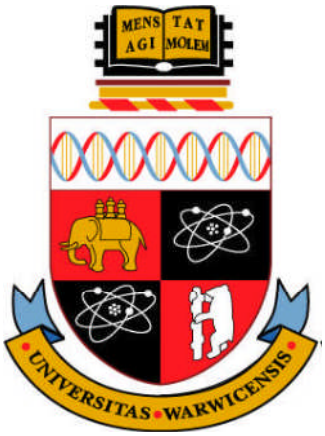
The Twin Arginine Translocation Pathway in *Escherichia coli*: Mechanism and Quality Control

Chao Ren

A thesis submitted for the degree of Doctor of Philosophy



University of Warwick

School of Life Sciences

December 2011

Table of Contents

List of Figures

List of Tables

Acknowledgments

At first, I would like to thank Professor Colin Robinson for giving me the opportunity to deeply study on the presented work, in addition, providing me the important support and suggestion during the supervision of research as well as the preparation of this manuscript.

Then, I appreciate Professor Frank Sargent (University of Dundee) for the gift of *E. coli* gene deletion strains in the proofreading study of Tat substrates which facilitated the recognition of specific chaperons associated with proteins.

Also, I am very grateful to Dr. Cristina Matos for technical advice and invaluable discussions during the undertaking of this work.

Many thanks go to Dr. James Barnett (analysis of metal content in proteins by ICP-MS, Chapter 4), Michael Caves (guidance of circular dichroism for protein secondary structure, Chapter 4), Nishi Vashist (suggestion of protein purification), Anna Albiniak (advice of confocal microscopy work, Chapter 6), Jacopo Baglieri, Daniel Beck and Roshani Patel for all the help and encouragement during my work.

I thank all of members in the Robinson group to offer me the friendly, happy and memorable time in the last three years, and lots of chocolates and biscuits always give me the energy to continue the work!

At last, I extremely thank my family, particularly my parents, for their patience, support and love throughout my life. And I would like to dedicate this thesis to my grandmother who sadly passed away a few months ago for her encouragement.

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 (University of Warwick). It has not previously been presented for another degree.

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

ICP-MS analysis was carried out by Dr. James Barnett and technical support by Dr. Lijiang Song and Mr. Philip Aston (University of Warwick).

nanoLC-ESI-MS/MS was performed by the Warwick/Waters Centre for BioMedical Mass Spectrometry and Proteomics in the School of Life Sciences, University of Warwick.

Summary

The bacterial twin-arginine translocation (Tat) pathway has the unique ability to export pre-folded proteins across the cytoplasmic membrane. Its name came from the almost invariant twin-arginine motif in the signal peptide of Tat substrates. *Escherichia coli*, a Gram-negative bacterium, is typically used to understand Tat function in bacteria. Until now, evidence has shown that TatA, TatB and TatC comprise the minimally functional unit, moreover, a quality control system exists to monitor the assembly of cofactors and the correctly folding state of proteins which avoids the futile export and initiates the degradation of rejected molecules.

The research presented in this thesis sought to gain insight into the quality control mechanism of the Tat pathway in *E. coli*, and also study the relationship between transport and maturation of substrates.

In the first place, a novel Tat substrate, YedY, was used to analyse the nature and variety of proofreading functions operating in conjunction with the Tat pathway. The single substitutions in three predicted ligands for the YedY molybdopterin centre led to complete inhibition of export and variable degradation of mutated YedY forms, indicating an effective proofreading activity. Circular dichroism spectroscopy and inductively coupled plasma mass spectrometry of purified proteins demonstrated the change of secondary structures between YedY and mutated variants, and also indicated the content of Mo in molybdopterin cofactor within proteins. The data suggested that the three mutated forms failed to correctly assemble cofactor which resulted in rejection by the Tat export pathway on the basis of the different changes of secondary structures. Further analysis shows that none of the known export chaperones for molybdenum cofactor-containing Tat substrates is required for YedY biogenesis; export is unaffected in cells lacking DmsD and TorD.

In the second place, maturation of pre-YedY was blocked when an Ala>Leu substitution was made at the -1 position of the signal peptide, and a membrane-bound precursor form accumulated in the membrane. However, the mature domain had been transferred to the periplasm. The accumulation did not block transport of other Tat substrates, indicating the precursor exited from the translocation channel and integrated into the membrane bilayer. Since the precursor was not detected in the periplasm, it was suggested that the precursor has undergone lateral transfer into the bilayer during translocation.

These results are discussed in relation to the overall mechanism of translocation and proofreading by the Tat pathway in *E. coli*.

List of Abbreviations

Amp	Ampicillin
APS	Ammonium persulfate
ATP	Adenosine triphosphate
AU	Absorbance units
BCA	Bicinchoninic acid
BN PAGE	Blue-native PAGE
BSA	Bovine serum albumin
c	Concentration
C	Cytoplasm
C-	Carboxyl-
CD	Circular dichroism
dH2O	Distilled water
dNTP	Deoxynucleotide triphosphate
DMSO	Dimethyl sulphoxide
DmsA	Dimethyl sulphoxide reductase
E	Elution fraction
ECL ™	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
ESI-MS	Electrospray ionisation mass spectrometry
FeS	Iron-sulphur
FT	Flow through fraction
GFP	Green fluorescent protein
GTP	Guanosine triphosphate
his/h	Hexahistidine tag
HRP	Horseradish peroxidase
IPTG	Isopropylthiogalactoside
kDa	kilo Dalton
LB	Luria-Bertani broth
M	Membrane
MBP	Maltose binding protein
Moco	Molybdopterin cofactor
MGD	Molybdopterin Guanine Dinucleotide
NapA	Nitrate reductase
N-	Amino-
OD	Optical density
P	Periplasm
PAGE	Poly-acrylamide gel electrophoresis
PBS(-T)	Phosphate buffered saline (plus tween-20)
PCR	Polymerase chain reaction
PMF	Proton motive force
PVDF	Polyvinylidene fluoride membrane

RNA	Ribonucleic acid
rpm	Revolutions per minute
RR-motif	Twin-arginine motif
SD	Standard deviation
SDS	Sodium dodecyl sulphate
Sec	Secretory pathway
SRP	Signal recognition particle
Strep/s	Strep-tag II
Tat	Twin-arginine translocase
TCA	Trichloroacetic acid
TEMED	Tetramethylethylenediamine
TM	Transmembrane
TMAO	Trimethylamine-N-oxide
TorA	Trimethylamine-N-oxide reductase
UV	Ultraviolet
W	Wash fraction
WT	Wild type
w/v	Weight per volume
Δ Delta	(gene knockout e.g. Δ*tatB*)
λ	Wavelength

List of amino acids

Amino acid	Single letter code	3 letter code
Alanine	A	Ala
Asparagine or Aspartic acid	B	Asx
Cysteine	C	Cys
Aspartic acid	D	Asp
Glutamic acid	E	Glu
Phenylalanine	F	Phe
Glycine	G	Gly
Histidine	H	His
Isoleucine	I	Ile
Lysine	K	Lys
Leucine	L	Leu
Methionine	M	Met
Asparagine	N	Asn
Proline	P	Pro
Glutamine	Q	Gln
Arginine	R	Arg
Serine	S	Ser
Threonine	T	Thr
Valine	V	Val
Tryptophan	W	Trp
Tyrosine	Y	Tyr
Glutamine or Glutamic acid	Z	Glx

Chapter 1 Introduction:

Critical theme in biology - protein targeting

Proteins are the most versatile macromolecules in living cells, performing essential biological processes of activities encoded by genes. Nearly all proteins are synthesized in cytosol but an estimated half of them have to cross lipid bilayer membranes to various subcellular membrane or aqueous spaces to carry out their proper functions in the correct location. However, lipid membranes are barriers that prevent unfavorable molecules passing through (such as ions), but they have to allow proteins to cross. Thus, protein targeting is essential for all cellular life and this field has been studied for nearly 40 years, starting with the work of George Palade (Nobel Laureate in Medicine in 1974) that defined the basic structure of the secretory pathway in eukaryotic cells (Dalbey and von Heijne, 2002). In the last 30 years, many diverse protein translocation systems have been identified and characterized in the endoplasmic reticulum (ER), mitochondria, peroxisomes, chloroplasts of eukaryotes, and the inner and outer membranes of prokaryotes (Holland, 2004; Jarvis and Robinson, 2004; Johnson and van Waes, 1999; Pugsley, 1993; Aldridge *et al.*, 2009; Cross *et al.*, 2009; Rapoport, 2007).

1.1 General characteristics in protein translocation machineries

Although fascinating differences are obvious between various translocation mechanisms, all of them share general principles including the presence of a signal peptide (targeting sequence/transit peptide/leader sequence/pre-sequence), a hydrophilic protein conduction channel or membrane translocase, and requirement for energy from nucleoside triphosphate (NTP) hydrolysis or proton motive force (Schatz and Dobberstein, 1996).

In the first place, protein translocation in archaeal, bacterial, and eukaryotic cells depends on the use of signal peptides to direct proteins to locations where their functions are required. Signal peptides are often N-terminal extensions but not always (e.g. type III secretion signals are located at the C-terminal end of proteins in

bacteria) and they are usually cleaved off by signal peptidases once targeting is completed.

In bacteria, the General Secretory pathway (Sec) is one of the well characterized systems to be responsible for the export of the majority of unfolded extracellular proteins or the insertion of membrane proteins into the cytoplasmic membrane (Xie and Dalbey, 2008). Sec signal peptides were the first studied and they have a common architecture in eukaryotic and prokaryotic organisms (von Heijne and Abrahmsén, 1989). The signal peptide is located at the N-terminus and is composed of three distinct regions: a short, positively charged N-terminal region (n-region) which contains a specific motif in some cases, a hydrophobic core (h-region) which forms an α-helical structure, and a polar conserved C-terminal region (c-region) which contains the signal peptidase cleavage site (Duffaud *et al*., 1985; McGeoch, 1985; von Heijne, 1990; Izard and Kendall, 1994). In eukaryotes, the first component of SRP was isolated in 1978/1980 as a binding factor which is a ribonucleoprotein particle, consisting of a RNA and six distinct proteins.

The second requirement for protein transport is energy consumption since protein transport across or into low dielectric biological membranes is a thermodynamically unfavorable process. Most translocases are driven by the expenditure of metabolic energy originating from nucleoside triphosphate (NTP)-binding and hydrolysis. For instance, ATPases mediate energy transduction and have been identified in various protein translocases including those in the endoplasmic reticulum (ER), mitochondria, chloroplasts and bacteria. Energy-transducing mechanisms can function on the *cis*-side of the membrane (pushing force) or the *trans*-side (pulling force); the best-characterized examples are SecA and Hsp70 respectively. SecA acts as the ATPase in the bacterial Sec pathway to promote protein transport by undergoing cycles of nucleotide-dependent conformational changes coupled to reversible binding to the SecYEG translocon (Schiebel *et al*., 1991; van derWolk *et al*., 1997; Alder and Theg, 2003). On the other hand, *trans*-side use of ATP drives protein import in

mitochondria/chloroplasts and post-translational translocation across the endoplasmic reticulum (ER) membrane. A member of the Hsp70 family of chaperones drives folding of polypeptides by undergoing iterative cycles of substrate binding and release along with a conformational change cycle from a low affinity Hsp70–ATP to a stable Hsp70–ADP by hydrolysis. This ATP-ADP cycle is modulated by DnaJ-type proteins (Bukau, 1998).

In addition, GTPases act as molecular binary switches. The intersecting GTPase cycles of the SRP and the SRP receptor are involved in the protein targeting. Examples include signal recognition particle (SRP)-mediated targeting of ribosome-nascent chain complexes (RNCs) in eukaryotes as well as participation of Ffh and FtsY in the cognate prokaryotic pathway. To be binary molecular switches, GTPases have the ability to impart unidirectionality and substrate–receptor binding specificity through GTP binding and irreversible hydrolysis (Keenan *et al*., 2001; Alder and Theg, 2003). GTPase-mediated substrate–receptor interactions are also identified in post-translational protein translocation into chloroplasts: Toc159 and Toc34 GTPases provide a pushing force to protein translocation via a membrane insertion or de-insertion step analogous to SecA activity (Schleiff *et al.*, 2003).

The proton motive force (PMF) is found to be the third energy source used to enhance or act sequentially with ATPase/GTPase-driven cycles in most protein transport systems. Strictly, proton motive force (PMF) comprises the thermodynamically equivalent electric potential ($\Delta\psi$) and H^+-concentration difference across the membrane (ΔpH) (Alder and Theg, 2003). Although the Sec translocon in bacteria is a true PMF-using apparatus, wherein the ΔpH and $\Delta\psi$ can contribute equivalently to the driving force, there is no mechanistic reason to expect that reactions depend on both. Mitchell first proposed that chemiosmotic systems operate through a ‘proton well’ to explain their functional interchangeability (Bakker and Randall, 1984; Mitchell, 1968; Griwatz and Junge, 1992). There is limited knowledge to characterise the machinery of PMF transduction in the Sec pathway,

however, several points have been studied including guidance of unidirectionality and enhancement of transport rates. For examples, PMF drives the completion of transport once ATP-dependent SecA binds to substrates and initiate translocation cycles (Schiebel *et al*., 1991). Also, the PMF supports the SecA ATPase reaction cycle by stimulating SecA de-insertion from SecYEG and promoting substrate release (Nishiyama *et al*., 1999; Mori and Ito, 2003). A new protein translocation pathway was found in nearly 20 years and termed twin-arginine (Tat) pathway, which uses the ΔpH as the sole energy source to transport folded proteins across membranes in most bacteria and plant thylakoids (details see chapters below).

To date, energy sources for protein translocation have been well characterized, but few direct measurements exist to work out total energy cost during these processes. It is possible to estimate the force generated by protein translocation ATPases (SecA or Hsp70) on the basis of the known free energy of ATP hydrolysis under cellular conditions (Ishijima *et al*., 1998). A recent study estimated that 3% of the total energy output of the chloroplast was used by the thylakoidal ΔpH-dependent cpTat pathway (Iuone, 2003). Nevertheless, energy sources cooperate with each other in most translocation systems and it is difficult to calculate the cost in these complicated processes without the development of further experimental methods.

Another feature common to most protein transport systems is that molecular chaperones/chaperonins in the cytosol have functions to assist protein folding, oligomeric assembly, rejected molecule degradation and even protein translocation. Hsp70 chaperones are a family of conserved ATPase molecules which stabilize newly synthesized polypeptides and avoid aggregation in a folding-competent state until all segments of the chains are ready for folding (Hartl, 1996). For instance, DnaK and partner DnaJ is a homolgue of Hsp70 in *Escherichia coli* and they perform an effective cycle to fold proteins. Firstly, DnaJ interacts with unfolded polypeptide and targets DnaK. Secondly DnaJ stimulates ATP-bound DnaK hydrolysis to produce the ADP-bound stable state which forms a DnaK-DnaJ-substrate complex.

Then the complex nucleotide exchange factor GrpE promotes release of ADP from DnaK as well as DnaJ release from the complex. Finally, substrate is released to fold or repeat the cycle if necessary, or transfer to chaperonin GroEL to finish folding (Schmid *et al*., 1994; Palleros *et al*., 1993). GroEL is one of the chaperonin subgroups in eubacterial chloroplasts and mitochontria, another one, TRiC, is found in archaebacteria and eukaryotes. Chaperonin family members cooperate with molecular chaperones to assist in folding by forming cylindrical complexes that promote protein folding in the sequestered environment of their central cavity (Basha *et al*., 2004; Ben-Zvi and Goloubinoff, 2001). The Chaperonin GroEL and Cpn60 families seem more generally and crucially to play a role in efficient protein folding in prokaryotes and eubacteria, respectively.

Lastly, most protein translocation pathways contain membrane-bound subunits that form a channel-like complex which facilitates protein transport. The general secretory (Sec) pathway is the most common one in all organisms and a novel twin-arginine translocase (Tat) pathway has been recently identified to export folded or partially folded proteins across bacterial membranes. Both contain a specific organization to finish the task of protein targeting and it will be described in detail below.

1.2 Protein translocation in bacteria and plant chloroplasts

1.2.1 Bacterial protein translocation

Prokaryotic cells consist of simple units surrounded by the plasma membrane and they generally lack the defined nucleus and any other membrane-bound compartments. Bacteria are the mainly group of unicellular prokaryotes. They are broadly differentiated into two large groups (Gram-positive and Gram-negative) according to their physical properties of the cell walls by Gram-stain (Figure1.1).

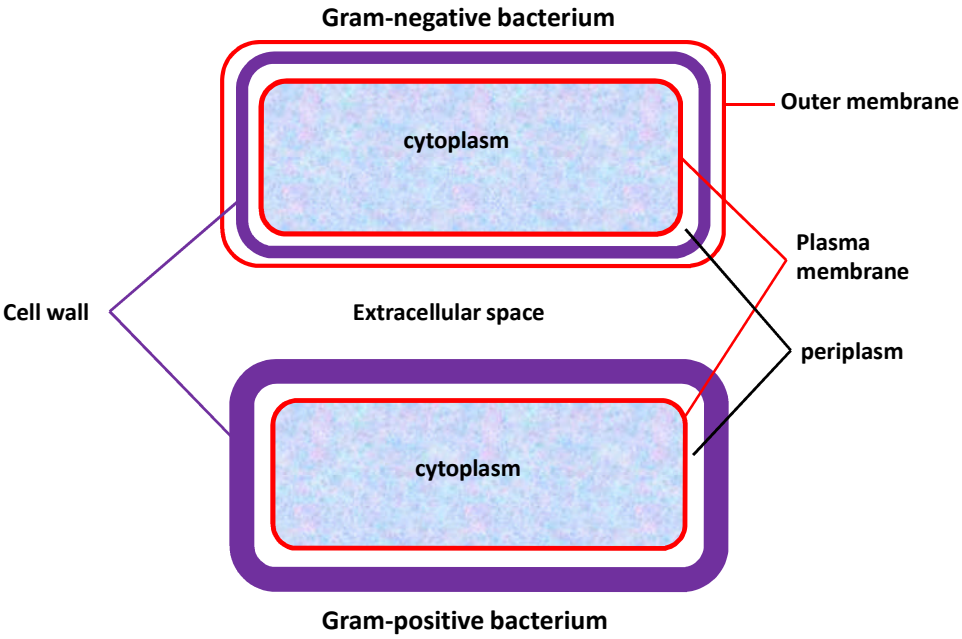


Figure 1.1 Schematic representation of the Gram-negative and Gram-positive bacterial cell The typical structure of a Gram-negative and a Gram-positive bacterial cell are shown in the diagram. A Gram-negative cell consists of four compartments which are a cytoplasmic membrane, the periplasmic space, an outer membrane and the cytoplasm. A Gram-positive cell contains a single plasma membrane surrounded by a thick cell wall.

The Gram-positive bacterial cell comprises three compartments: cytoplasm, plasma membrane and a thick cell wall layer of peptidoglycan and extracellular media. Proteins can be inserted into the cytoplasmic membrane or translocated across it. However, Gram-negative bacteria have a different and complex arrangement. The cell wall is composed of a thin layer of peptidoglycan and an outer membrane encompasses the cell. In this case, some proteins maybe further inserted into the outer membrane or secreted from the cell by specialized secretion systems such as type I–V protein secretion systems. There are also a few specialized translocation systems that enable proteins to move through both membranes into the extracellular space (Saier, 2006).

Approximately 20% of all proteins synthesized in the cytoplasm must be exported across the cytoplasmic membrane in bacteria (Pugsley, 1993). The major route of protein translocation is the so-called secretory pathway abbreviated as 'Sec pathway', but a smaller number of substrates use a more recently discovered Tat pathway. I will focus on the translocation of proteins by the Tat pathway in the Gram-negative bacterium *E. coli*.

1.2.2 Protein sorting in the chloroplast

Chloroplasts are a specialised plastid-type organelle that assimilates carbon dioxide and generates oxygen through photosynthesis in plant cells, and also participates in the biosynthesis of fatty acids, lipids and amino acids. The complex structure of chloroplasts includes six sub-compartments, including an envelope consisting of the outer and the inner lipid bilayer membrane, the inter-membrane space, the stroma, the internal thylakoid membranes and the enclosed soluble phase - the thylakoid lumen. The thylakoid network within the stroma is composed of stacks of disc-like membrane structure that directly perform the light-mediate reaction of photosynthesis.

Most proteins destined for the chloroplasts are made in the cytosol and need to be imported. Two large complexes, the translocases of the outer membrane (Toc) and inner membrane (Tic) were identified to be responsible for protein import (Jarvis and Robinson, 2004; Oreb *et al*., 2008). Proteins are synthesized bearing a cleavable, N-terminal presequence which can be targeted to translocons with the help of cytosolic chaperones. When proteins are successfully translocated into the chloroplast, the stromal processing peptidase (SPP) will remove the transit presequence of proteins. Some of them remain to function in the stroma, but others contain a second signal peptide to direct a subsequence routing pathway if their final destination is in the thylakoid lumen. Both Sec and Tat pathway utilise the second signal peptide to transport proteins into the thylakoid lumen but they have different translocation mechanisms. For proteins transported into the thylakoid membrane, two pathways were identified: the signal recognition particle (SRP) pathway and the 'spontaneous' pathway are involved (Robinson *et al*., 2001; Di Cola *et al*., 2005).

1.3 The General Secretory (Sec) pathway

The mechanism of the Sec pathway and the structure of the major components of the Sec translocase in *E. coli* have been elucidated by genetic, biochemical and electron microscopy (EM) studies from 1970s. When the signal peptide of a protein is translated by the ribosome, it can be targeted to the Sec translocase by two different mechanisms: either co-translational or post-translational (Figure 1.2).

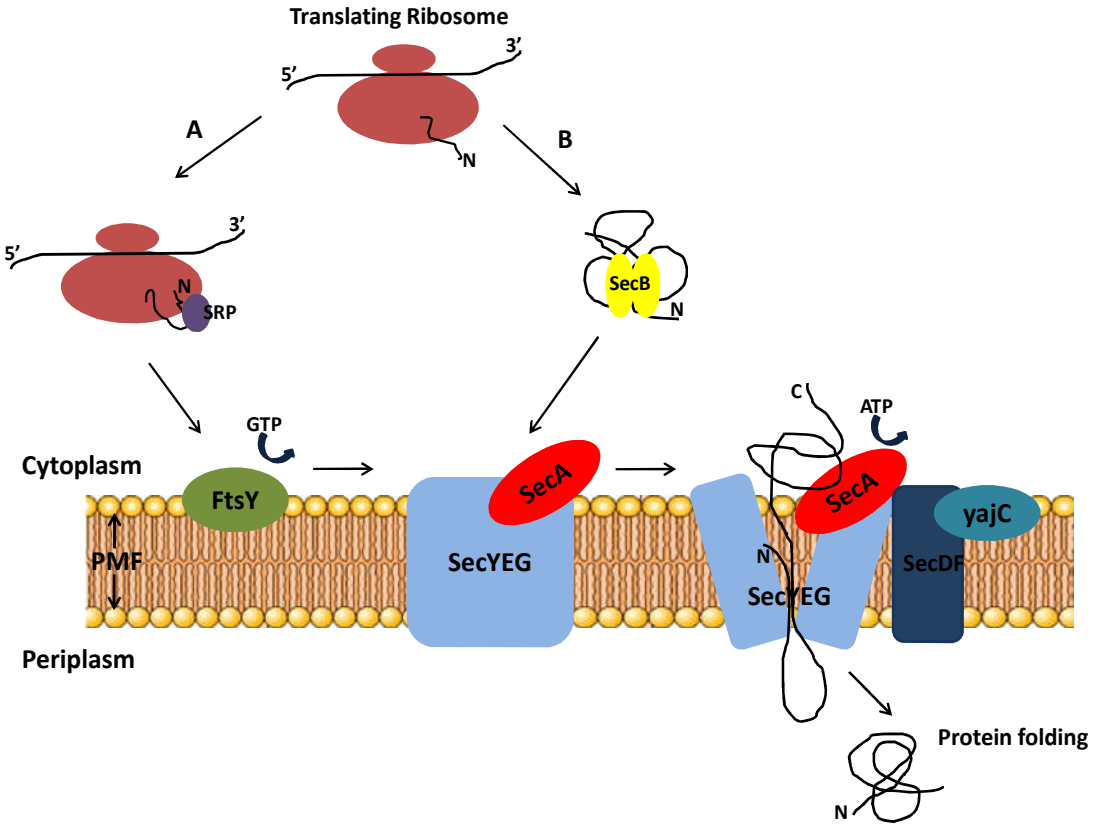


Figure 1.2 The protein translocation by the Sec pathway in *E. coli* (Adapted from Driessen and Nouwen, 2008**).** Newly synthesised proteins are targeted to the Sec translocase by different routes. **A.** Co-translational and **B.** post-translational targeting routes and translocation of unfolded proteins by the Sec-translocase.

In bacteria, the co-translational targeting pathway is generally used for the integration of proteins into the cytoplasmic membrane, unlike ER, where only a few proteins are transported across the membrane by this route (Osborne *et al*., 2005). In the post-translational mode, the preprotein is synthesized to its full-length in the cytosol and kept in a loosely folded, translocation-competent state by the binding of the molecular chaperone SecB to multiple regions of the mature domain (Randall *et al*., 1990; Fekkes and Driessen, 1999; Kumamoto and Francetiç, 1993). Although general chaperones such as GroEL and DnaK can compensate for the stabilization of unfolded preproteins in the absence of SecB, they can not perform the specific targeting role in protein export which particular depends on the ability of SecB to associate with SecA (Hartl *et al*., 1990; Kumamoto, 1991; Wild *et al*., 1996). The SecB–preprotein complex is targeted to the SecYEG-bound SecA where the signal sequence of the preprotein is exposed to SecA during the SecB–SecA interaction (Fekkes *et al*., 1997). Then SecB is released from this ternary complex when SecA binds ATP to initiate translocation.

SecA is a central component of the bacterial Sec system as it is the only ATPase to couple hydrolysis of ATP to drive translocation of preproteins and it interacts with almost all other components of the translocase (Schiebel *et al*., 1991; Manting and Driessen, 2000).

The SecYEG complex forms a hydrophilic protein-conducting pore to direct the movement of secretory proteins across membrane in bacteria (Joly and Wickner, 1993; Cannon *et al*., 2005). In *E. coli*, SecY is a 48 kDa protein with 10 membrane-spanning α-helices, while SecE is 15 kDa with three membrane-spanning α-helices. SecY and SecE are the central components that form a stable complex. SecY is degraded by an ATP-dependent membrane-bound protease FtsH in the absence of SecE (Akiyama and Ito, 1987; Kihara *et al*., 1995). SecG, a third membrane protein of two spans that forms the heterotrimeric SecYEG complex, is not essential for viability but needed for efficient protein export, in particular at

lower temperature (Brundage *et al*., 1990; Nishiyama *et al*., 1993). The proposed translocation channel is formed by a SecYEG dimer and has the overall shape of an hourglass that is ~ 4 Å at it most constricted point and ~ 20-25 Å at its widest (Breyton *et al*., 2002; Mitra *et al*., 2005; Van den Berg *et al*., 2004). Another heterotrimeric integral membrane protein complex, SecDFYajC, is needed for efficient protein export by SecYEG (Duong and Wickner, 1997).

A mechanism of the Sec pathway has been proposed. After SecB passes the preproteins to SecA, SecA binds to ATP leading to a loop of the signal sequence as the preprotein inserts into the SecYEG. Upon the hydrolysis of ATP, preproteins are released from SecA but SecYEG-bound SecA rebinds ATP to finish the step-wise translocation (Mitra *et al*., 2006; van der Sluis and Driessen, 2006; van derWolk *et al*., 1997). In the late stage, The PMF can drive translocation when SecA is not associated with the preprotein (Schiebel *et al.*, 1991; Driessen, 1992).

1.4 The Twin-Arginine translocation (Tat) pathway

Although the Sec pathway is involved in the export of unfolded proteins across membranes in all domains of life, a specialized protein translocation pathway exists and acts in a different manner in archaea, bacteria and plant plastids. The pathway was first discovered in the thylakoid membrane of chloroplasts towards the early 1990s and initially designated the ΔpH pathway because of the reliance on the proton motive force as the sole energy resource (Cline *et al.*, 1992; Klosgen *et al.*, 1992; Mould & Robinson, 1991). Thereafter, a group of bacterial redox proteins were observed to export via the same pathway, and their signal peptides were analyzed to harbour a highly conserved twin arginine (RR) motif, which led to the pathway being termed the twin-arginine translocation (Tat) pathway (Berks, 1996; Sargent *et al*., 1998).

The Tat pathway possesses the unique and remarkable ability to transport partially or even fully folded proteins across energy-transducing membranes, and does not require nucleoside triphosphate (ATP) hydrolysis at any stage of the translocation process (reviewed by Robinson and Bolhuis, 2004; Muller, 2005). The composition of the Tat pathway has been identified but it is surprisingly varied in different species. In *E. coli*, it consists of three subunits TatA, TatB and TatC which are the minimal functional group. Homologues of each subunit are found in chloroplasts as Tha4, Hcf106 and cpTatC respectively. However, all Gram-positive bacteria except *Streptomyces* species and some archaea contain only TatA and TatC (Yen *et al*., 2002). TatD and TatE are also encoded by *tat* genes but they are not essential for the protein translocation. The *tatD* gene is co-expressed in an operon with *tatA/B/C* and encodes a protein with DNase activity (Wexler *et al*., 2000). The *tatE* gene is thought to be a gene duplication of *tatA* but is expressed at significantly lower levels, and can partially complement *E. coli ΔtatA* cells if overexpressed (Sargent *et al*., 1998; Jack *et al*., 2001). Surprisingly, recent researches have speculated that the involvement of the TatD and TatE in the quality-control of Tat substrates (Matos *et al*., 2009).

At present, the Tat pathway is still not completely understood and key aspects need to be studied. For instance, although other pathways are capable of transporting folded proteins, namely import pathways into peroxisomes and the nucleus in eukaryotes (Glover *et al.*, 1994; Walton *et al.*, 1995; Feldherr *et al*., 2002) as well as export across the Gram-negative outer membrane by Type II secretion, the Tat pathway is the only one that achieves this in a coupled membrane. This may provide a means for secretion of complex recombinant proteins which are not compatible with the Sec pathway (Brüser, 2007; Kim *et al.*, 2005). In addition, the Tat pathway is widespread within the microbial world and plants but absent in animal and fungal cells which makes it a novel target for antimicrobial drugs with low human toxicity (Lee *et al*, 2006).

1.4. 1 Components of the Tat pathway

The subunits of the Tat machinery in chloroplasts and Gram-negative bacteria share highly homologous sequences which help to identify the presence of Tat proteins in most bacteria. The maize *hcf106* gene was the first gene to be identified and characterized in the Tat pathway (Settles *et al.*, 1997; Voelker & Barkan, 1995). In *E. coli*, three inner membrane proteins TatA, TatB and TatC represent the minimal requirements for efficient translocation, and the countparts in plants are Tha4, Hcf106 and cpTatC respectively (Sargent *et al.*, 1998; Weiner *et al.*, 1998; Bogsch *et al*., 1998). However, TatB does not seem to be essential for export in some Gram-positive bacteria which contain only TatA and TatC-type proteins of the Tat apparatus (Chan *et al*., 2009; Robinson *et al.*, 2011). The fouth gene *tatD* located in an operon of *tatABC* and the fifth monocistronic gene *tatE* are thought to play an role in quality control of Tat system. The organisation of the *tat* genes of *E. coli* is shown schematically in Figure 1.3.

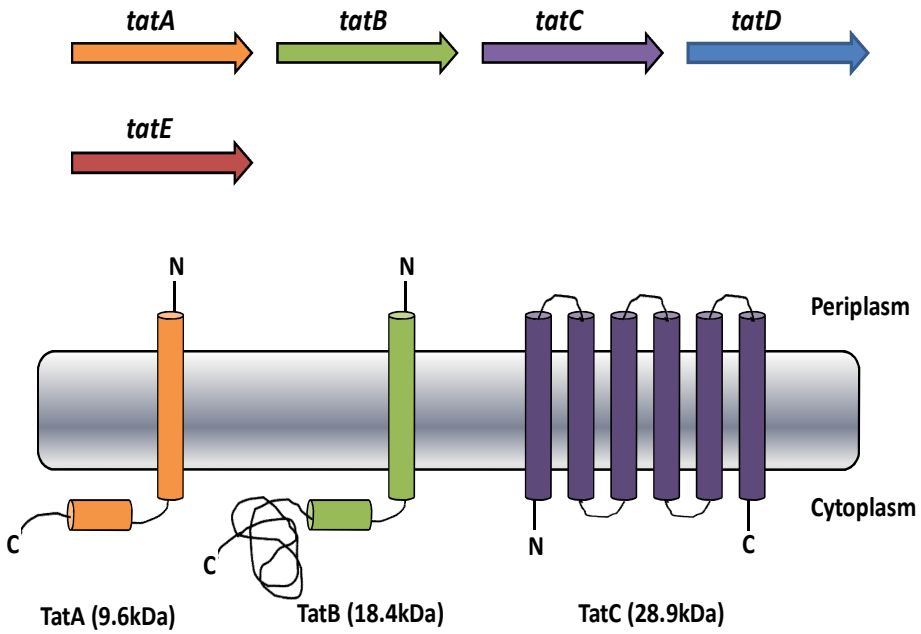


Figure 1.3 The *tat* genes and the overall structure of three essential Tat subunits of *E. coli* (Adapted from Robinson *et al.*, 2011) In *E. coli*, three essential *tat* genes are constitutively expressed alongside the *tatD* gene in an operon, also *tatE* which is the paralogue of *tatA* is located elsewhere in the genome.The picture shows the proposed overall structure for the subunits of the TatABC system: the TatA and TatB have a single TM spanning domain following an amphipathic helix that lies along the cytoplasmic face of the plasma membrane. TatC contains 6 TM spans with the N- and C-termini in the cytoplasm.

1.4.1.1 TatA subunit

TatA is the most abundant Tat component and is present at around 25- and 50-times more than TatB and TatC, respectively (Jack *et al.*, 2001; Sargent *et al*., 2001), and it is suggested to form large homooligomeric complexes of differing diameter as the translocation channel (Gohlke *et al*., 2005; Leake *et al*., 2008).

The *tatA* gene encodes an 89 amino-acid integral membrane protein of 9.6 kDa which is predicted to contain an N-terminal α-helical transmembrane (TM) domain followed by a short 'hinge' region leading to an amphipathic helix (APH) and a largely unstructured C-terminal tail (Settles *et al.*, 1997; Sargent *et al.*, 1998). Studies of TatA secondary structure by circular dichroism (CD) and oriented CD (OCD) provided the evidence that the N-terminal TM domain does traverse the membrane with the APH lying along the surface of the membrane (Lange *et al*., 2007; White *et al*., 2010). Also, the study using fusion of TatA and YFP showed TatA protein has an even distribution around the periphery of the cell (Berthelmann and Bruser, 2004; Leake *et al*., 2008; Ray *et al*., 2005). TatA homologs in *B. subtilis and H. volcanii,* even Tha4 in plants, are partially present in a soluble cytoplasmic form and the role of this form in the translocation process is still ambiguous (Dilks *et al*., 2005; Pop *et al*., 2003; Frielingsdorf *et al*., 2008). The TM domain of TatA is essential for export, (De Leeuw *et al.*, 2001), and the TatA TM domain is also thought to be involved in interactions with TatB (Barrett and Robinson, 2005; Lee *et al*., 2002). Furthermore, point mutations located in the hinge region and APH severely affect translocation activity indicating a functional significance (Hicks *et al*., 2003).

A dual topology of *E. coli* TatA has been proposed in terms of two studies: localization analysis with fusion of TatA and YFP as well as accessibility of single cysteine mutants to thiol-reactive components. It is speculated that lipid interactions of the APH of TatA allow a temporary insertion into the membrane during the translocation process (Gouffi *et al.*, 2004; Porcelli *et al.*, 2002). In contrast, the

C-terminus of TatA appears to be non-essential for function. Truncations by up to 40 amino acids from the C-terminus of *E. coli* TatA are still able to function (Lee *et al*., 2002). Recently, a study proposed that TatA, in conjunction with TatE, might be involved in the proofreading of rejected Tat substrates (Matos *et al*., 2008), the detail will be discussed.

1.4.1.2 TatB subunit

TatB and TatA share 25% homology in *E. coli*, and are predicted to have a similar secondary structure as well as membrane topology. TatB can interact with TatC to form the translocation receptor complex.

TatB is made up of 171 amino acids and is 18.4kDa; it contains an N-terminal TM domain followed by a small hinge region and amphipathic helix (APH) that lies along the cytoplasmic side of the membrane (De Leeuw *et al*., 2001). TatB has a long extended C-terminal tail that was shown to be nonessential for Tat translocation activity through truncation analysis (Lee *et al*., 2002). Furthermore, mutations located in the TM domian of TatB did not have an effect on transport and complex formation, even when the domain was substituted by the homologous TM domain from TatA (Lee *et al.*, 2002). However, similar to TatA, various mutagenesis studies have indicated that single amino acid substitutions in the hinge region and APH dramatically affect activity (Barrett *et al*., 2003; Hicks *et al*., 2003; Lee *et al*., 2006). Recent two hybrid analysis demonstrated no specific interaction between the membrane-extrinsic domains of TatA and TatB, even thought the stronge homodimerization of them (Maldonado *et al*., 2011). This is consistent with earlier chemical crosslinking studies which did not detect contacts between TatA and TatB (De Leeuw *et al*., 2001).

Recent study of the membrane-extrinsic domain of TatB has shown that the first 55 or 101 amino acids of TatB is efficient of TatBC complex formation or stability, and

this implied that only the TM and/or APH is required for complex formation with TatC (Maldonado *et al*., 2011). A model proposed that the TatBC receptor consists of four TatBC protomers, in which TatB is able to transiently accommodate large parts of a folded Tat precursor and TatC is located on the periphery of the complex (Lee *et al*., 2006; Maurer *et al.*, 2010). Indeed, the interaction between TatB and Tat precursor occurs at multiple contact sites along the TM and amphipathic helix, with two Tat precursors able to simultaneously bind to a TatB subunit (Tarry *et al*., 2009; Ma and Cline, 2010).

Despite the similarity of TatA and TatB, the two subunits can not substitute for each other, even when over-expressed (Sargent *et al*., 1999). TatB is a prerequisite for the translocation of endogenous *E. coli* Tat substrates, but is not always essential for transport of all substrates. For instance, Colicin V was translocated in the *tatB* deletion strain of *E. coli* (Ize *et al*., 2002) and export of a reporter protein fusing the signal peptide of TMAO reductase (TorA) with MalE was observed in the absence of TatB (Blaudeck *et al*., 2005).

In most Gram-positive bacterial systems a TatB homologue is lacking and the Tat pathway comprises a TatAC system (Dilks *et al.*, 2003; Yen *et al.*, 2002). This has led to the hypothesis that in these systems the TatA/B homologue may be bifunctional and takes the role of both TatA and TatB. This has been demonstrated for the TatAdCd system from *B. subtilis* (Barnett *et al.*, 2008).

1.4.1.3 TatC subunit

TatC is the largest and most highly conserved Tat subunit in *E. coli*. It is 28.9kDa and is predicted to have six TM helices and both N- and C-termini protruding into the cytosol by sequence analysis as well as gene fusions (Behrendt *et al*., 2004; Jeong *et al*., 2004). Some studies suggested that only four TM helices is present in membrane and two may lie on the surface of membrane (Gouffi *et al*., 2002; Nolandt

et al., 2009). However, the most likely scenario is the presence of six TM helices which has been proved by crosslinking (Punginelli *et al*., 2007). The cytoplasmic side of TatC is particularly important for function. Several site-directed mutagenesis studies of conserved residues in cytosolic loops have revealed that even single amino acid substitutions can have a severe effect on translocation activity in prokaryotes and plant plastids. The data also suggested two residues, Phe94 in the second TM domain and Pro48 in the first periplasmic loop, are critical for the assembly of Tat complexes (Allen *et al*., 2002; Barrett and Robinson, 2005; Buchanan *et al.,* 2002). TatC either alone or in a complex with TatB initially interacts with the signal peptide of Tat substrates, and this interaction requires the twin arginine motif in *in vitro* biochemical studies (Alami *et al*., 2003; Cline and Mori, 2001; De Leeuw *et al*., 2002).

1.4.1.4 TatD subunit

The *tat* operon encodes the fourth gene *tatD* which is a soluble protein with DNase activity. TatD was originally thought to play no apparent role in Tat-dependent translocation, because the translocation of Tat substrates was not affected in TatD deletion strains (Wexler *et al*., 2000; Robinson and Bolhuis, 2004). However, it is proposed recently that TatD plays a role in the proofreading of Tat pathway which can stabilize the correctly assembled proteins in wild type cells but turn over rejected molecules (Matos *et al*., 2009). So far, the exact role of TatD is still uncertain and more work will be done in the future.

1.4.1.5 TatE and TatF subunit

The *E. coli* chromosome encodes a monocistronic gene outside the *tatABC* operon, *tatE,* which is the fifth subunit of the Tat machinery. TatE shares up to 50% amino acid sequence homology with TatA, resulting in the possibility that TatE seems to have overlapping functions with TatA and it can partially complement tatA null

mutants (Jack *et al*., 2001). TatE is produced at much lower levels than TatA and does not have any significant role in Tat function (Sargent *et al*., 1998; Sargent *et al*., 1999). Nevertheless, *E. coli* TatE is induced in biofilms, and in *Pseudomonas stutzeri* TatE expression is associated with denitrification, pointing to a more important role of TatE proteins under specific growth conditions—possibly as an adaptation to specific substrates (Beloin *et al*., 2004; Heikkilä *et al*., 2001). More importantly, the degradation of FeS deficient Tat substrates in previous studies suggested TatE alone or accompanied by TatA may be involved in the proofreading/turnover process of Tat system (Matos *et al*., 2008).

No evidence demonstrated that TatF is a Tat subunit participating in the Tat translocation pathway. But it is related to the AAA family of ATPases that often have chaperone and/or protease functions (Lindenstrauss and Brüser, 2006).

1.4.2 Structural organization of Tat complexes

1.4.2.1 Tat complexes in Gram-negative bacteria

In Gram-negative bacteria, three subunits are required to come together and form a large hetero-oligomeric TatABC complex to translocate large folded proteins. On the basis of the key points, some Tat mechanistic models have been built in recent years.

Initially two independent studies of isolated complexes showed different results: one found that the TatAB complex contained oligomeric TatA complex and a small amount of TatB associated with TatA, and the isolated single particle was analysed to form a double-layered ring-like structure by electron microscopy (Bolhuis *et al*., 2001; De Leeuw *et al*., 2001; Sargent *et al*., 2001); another was that only TatA formed homo-oligomeric complexes of variable size but mainly a discrete complex of ~ 460 kDa (Porcelli *et al*., 2002).

However, it is now widely accepted that two distinct types of complexes are present: a TatABC complex and homo-oligomeric TatA complex (Bolhuis *et al*., 2001; Lee *et al*., 2006; Robinson *et al*., 2011). The TatABC complex was purified from *E. coli* and estimated to be ~600 kDa at the beginning by gel filtration chromatography (Bolhuis *et al*., 2001), thereafter, the accurate complex size was determined as 370 kDa according to blue native PAGE (Oates *et al*., 2005). Within this complex TatB and TatC interact with a 1:1 stochiometry to form a functional unit with slightly variable amounts of TatA presents (Bolhuis *et al*., 2001). TatBC is essential to recognize and bind to substrates before export. Surprisingly, there is evidence that TatB and TatC assemble into autonomous units in the absence of each other (Orriss *et al*., 2007). Furthermore, TatC is unstable in the absence of TatB, suggesting a role of TatB in stabilizing the TatBC complex (Sargent *et al*., 1999), and TatA has no effect on either the formation or stability of the TatBC complex (Orriss *et al*., 2007).

The majority of TatA is found in separate homo-oligomeric complexes that co-purify with the TatABC complex and form a ladder of bands on blue native gels ranging in size from 50 kDa to well over 500 kDa (Gohlke *et al.*, 2005; Oates *et al.*, 2005). The formation of the TatA complex is independent of expression level and does not require the presence of TatB/C (Muller and Klosgen, 2005). Electron microscopy and random conical tilt reconstructions found TatA complexes form channels with a possible lid structure and variable diameter (Gohlke *et al.*, 2005). A recent study using *in vivo* imaging of TatA-YFP at native levels demonstrated that approximately 15 TatA complexes were mobile in the membrane with some disperse TatA molecules present (Leake *et al*., 2008). The cross-linking studies suggested TatA form trimers and tetramers in *E. coli* membranes (De Leeuw *et al.*, 2001; Oates *et al.*, 2005; Leake *et al.*, 2008). Most of the studies mentioned above have performed under the situation in absence of substrate and over-expression of TatA/B/C, it is hard to know exact complex organization when folded Tat substrates of vastly differing sizes are translocated.

1.4.2.2 Tat complexes in Gram-positive bacteria

The minimal Tat translocation systems are comprised of only TatA and TatC-type proteins in most Gram-positive bacteria (Jongbloed *et al*., 2004; Barnett *et al*., 2008). *Bacillus subtilis* is used as a model organism to study the Tat pathway and it has two TatAC-type pathways (TatAdCd and TatAyCy) that operate in parallel with differing substrate specificities (Jongbloed *et al*., 2004). Analysis of the TatAdCd complexes showed TatAdCd forms ~ 230 kDa complexes and a separate, rather discrete ~270 kDa TatAd complex (Barnett *et al*., 2008). A soluble cytosolic population of TatAd was found alongside the membrane localized form (Pop *et al*., 2003). This soluble TatAd and the membrane localised TatCd led to a proposed distinct model of Tat pathway with TatAd acting as the initial receptor for Tat substrates (Pop *et al*., 2003). Studies also indicated that TatAd was capable of complementing *E. coli* tatA and tatB null mutants, resulting in the truly bifunction of TatAd in lack of TatB (Jongbloed *et al*., 2006). The overall structure comprising the TatAdCd complex and the separate TatAd complex has been confirmed in solution NMR studies on detergent-solubilised TatAd in which the protein was monomeric under the conditions used, but it is probably part of an oligomeric complex in nature (Hu *et al*., 2010). Further investigation of TatAdCd and Tat AyCy in the Tat system is needed in view of the limited knowledge.

1.4.2.3 Tat complexes in plants

Similar to Tat components in *E. coli*, three subunits Tha4, Hcf106 and cpTatC make up of the Tat translocation pathway in the plant. In contrast to bacterial TatABC organisation, Tha4 (TatA) is not found in a complex with hcf106 and cpTatC (TatBC). It is shown that hcf106-cpTatC form a stable 700 kDa complex, while Tha4 forms a distinctly discrete homo-oligomeric complex of 400 kDa (Cline and Mori, 2001). A recent study in thylakoids used cross-linking to show that Tha4 oligomers associate through their TM domains in unstimulated membranes, and formed higher

order oligomers when stimulated by the addition of a Tat substrate (Dabney-Smith and Cline, 2009). Moreover, the separable hcf106–cpTatC and Tha4 subcomplexes located in the thylakoid membranes of plant chloroplsts only associated in the presence of substrate with ΔpH as energy (Cline and Mori, 2002).

1.4.3 Tat signal peptides

Proteins transported by the Tat pathway are targeted by N-terminal signal peptides that contain three distinct regions: a positively charged N-terminal region, a hydrophobic region and a polar C-terminal region ending with an Ala-Xaa-Ala cleavage motif which is removed after translocation (von Heijne, 1995). Bacterial Tat substrates display a highly conserved consensus motif T/S-R-R-x-F-L-K (where x is an apolar amino acid) between the N and H regions in which the invariant twin-arginine residues result in the name of 'Twin arginine translocation pathway' (Figure 1.4) (Berks, 1996; Crsitobal *et al*., 1999). But chloroplast Tat signals show a slight alteration as R-R-xx-Φ, where Φ is leucine, phenylalanine, valine or methionine (Brink *et al.*, 1998). RR is essential in the Tat pathway of plants since the substitution of either Arg to Lys completely abolishes transport (Chaddock *et al*., 1995). In contrast, in bacteria substitution of a single Arg is tolerated or only affects the translocation activity whereas mutation of both Arg blocks Tat-specific export, implying the involvement of other factors (Buchanan *et al*., 2001; De Lisa *et al*., 2002; Mendel *et al*., 2008). Although the RR-motif is almost invariant in all Tat substrates, a minority of exceptions have been identified to date with a KR motif in the signal peptide of TtrB subunit of *Salmonella enterica* and the RNR motif of *E. coli* penicillin amidase (Hinsley *et al*., 2001; Ignatova *et al*., 2002).

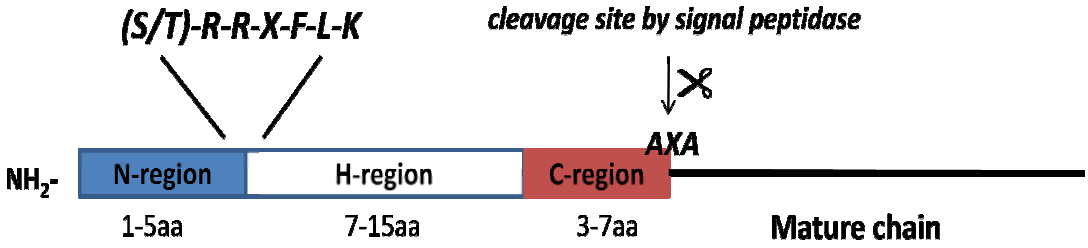


Figure 1.4 Tat signal peptide A typical tripartite signal peptide structure applies to Tat signal peptides; a polar amino domain (N-region), hydrophobic core (H-region), and polar carboxyl domain (C-region). Tat signal peptides are distinguished by the presence of a specific twin-arginine containing motif at the end of the N-region. They are also typically longer and less hydrophobic than Sec signal peptides. The cleavage motif at the end of the C-region, A-x-A, where the signal peptide is cleaved off to form the mature protein is also shown.

The RR motif is not the only important determinant in Tat signals, other factors also appear to be critical. Some residues around RR motif, especially position -1 and +2/+3 appear to be important for translocation in bacteria. Studies have shown the Serine residue before the RR-pair seems to be important in some substrates because the substitution of this residue of *E. coli* TorA and DmsA led to a large reduction in translocation activity yet had no effect on SufI transport (Stanley *et al*., 2000; Mendel *et al*., 2008). Similarly, the +2 residue in front of RR-pair that is usually either a Phenylalanine or a Leucine is significant in different substrates but an acidic residue at this position cannot be tolerated. For instance, substitution of this residue with either Alanine or Serine had no effect on translocation in *E. coli* of TorA-GFP, whilst substitution by Aspartic acid caused a complete block in transport. In DmsA, substitution of Leucine with either Alanine or Aspartic acid caused a complete block as well (Mendel *et al*., 2008).

Moreover, basic residues in the C-region and a proline residue at the end of the H-region have been identified as playing a role in translocation activity (Ize *et al*., 2002; Bogsch *et al*., 1997; Cristobal *et al*., 1999). In chloroplasts the basic C-domain is recognized as a 'Sec-avoidance motif' (Bogsch *et al.*, 1997). In addition, Tat signal peptides are less hydrophobic than Sec-type signal peptides (Cristobal *et al.*, 1999; Ize *et al.*, 2002).

Also, the N-termini of Tat signal peptides are generally longer than Sec counterparts with an average length of 38 and 24 residues, respectively (Cristobal *et al*., 1999). Although there is the difference of Tat RR- motif significance between the plants and bacteria, the high conservation leads to efficient direction of transport when expressed heterologously in distinct species, for example a bacterial signal peptide fused GFP can be transported in plants (Spence *et al.*, 2003). Nevertheless, at least eight proteins in *E. coli* lacking a signal peptide seem to be exported in a 'hitchhiker mechanism' which means they interact with another protein that has a Tat signal peptide and transport together via the Tat pathway, such as DmsB (Rodrigue *et al*.,

1999). Apart from the targeting of proteins to Tat translocase, more recent studies have indicated that signal peptides that bind to cytosolic chaperones in order that the correct assembly of substrates can be finished, thus performing a valuable function in proofreading (Buchanan *et al*., 2008; Robinson *et al*., 2011).

1.4.4 Tat substrates

At the beginning, Tat pathway was discovered to translocate complex cofactor containing redox proteins in bacteria. It was shown that 6% of secreted proteins are predicted to be Tat-dependent and the majority are cofactor-containing redox proteins in *E. coli* including hydrogenases, formate dehydrogenases, nitrate reductases, trimethylamine *N*-oxide (TMAO) reductases, and dimethyl sulfoxide (DMSO) reductases, all of which function in anaerobic respiration (Berks, 1996; Gennis and Stewart, 1996; Palmer *et al*., 2005; Rose *et al*., 2002). These substrates require the insertion of metal ion cofactors such as FeS, NiFe, copper and molybdopterin centres within the cytosol prior to translocation (Berks, 1996; Santini *et al.*, 1998; Weiner *et al.*, 1998). Besides, the Tat pathway has the instinct to discriminate against misfolded proteins (Delisa *et al*., 2003). Since then several Tat substrates have been identified that do not have a cofactor, such as AmiA and AmiC in *E. coli* that are required for cell wall biogenesis. Therefore, a widely accepted hypothesis is that rapid folding kinetics for these proteins may not be compatible with the Sec pathway (Ize *et al*., 2003; Bernhardt and de Boer, 2003).

Another hypothesis is that some Tat substrates are unable to attain their biologically active state within the periplasmic milieu, so they have to properly assemble to be active in the cytoplasm, in the same way that green fluorescent protein (GFP) is only fluorescent in the periplasm when exported by the Tat pathway (Feilmeier *et al*., 2000; Thomas *et al*., 2001). Some organisms favor the Tat pathway due to their harsh external environments. *Halophilic archaea* are thought to use the Tat pathway as the

high salt concentrations in the external environment would cause aggregation of unfolded proteins (Rose *et al*., 2002).

The Tat pathway not only exports soluble proteins across the membrane, but also targets integral membrane proteins in some cases. However, it is not known whether integral membrane proteins are inserted directly from the Tat translocase or inserted by other mechanisms following their translocation to the periplasm. There is also some evidence to suggest that the Tat pathway is capable of translocating lipoproteins that are attached to the membrane via lipidation of the N terminus (Dilks *et al*., 2005; Li *et al*., 2005). The Tat system is identified in most bacteria but is absent in some. *E. coli* has ~ 30 Tat substrates and *Streptomyces coelicolor* about 25 proteins (Widdick *et al*., 2006). In contrast the Gram-positive bacterium *B. subtilis* has just 4 confirmed Tat substrates (Widdick *et al*., 2008). Importantly, several virulently pathogenic bacteria also rely heavily on the Tat pathway. By contrast some bacteria such as *Lactococcus lactis* do not possess any Tat protein homologues (Dilks *et al*., 2003). In summary, the Tat pathway has the unique ability to transport a certain degree of folded proteins through tightly sealed membranes, but different bacterial species utilise the Tat pathway to considerably extents.

1.4.5 Proposed mechanism of Tat translocation

The Tat components and their organization have been identified, whereas the novel mechanism of export folded proteins across energetically sealed membranes in Tat pathway is still unraveled. Based on most studies in Gram-negative bacteria and plants, a compelling model has been proposed as ‘Tat-transport cycle’ (Figure 1.5).

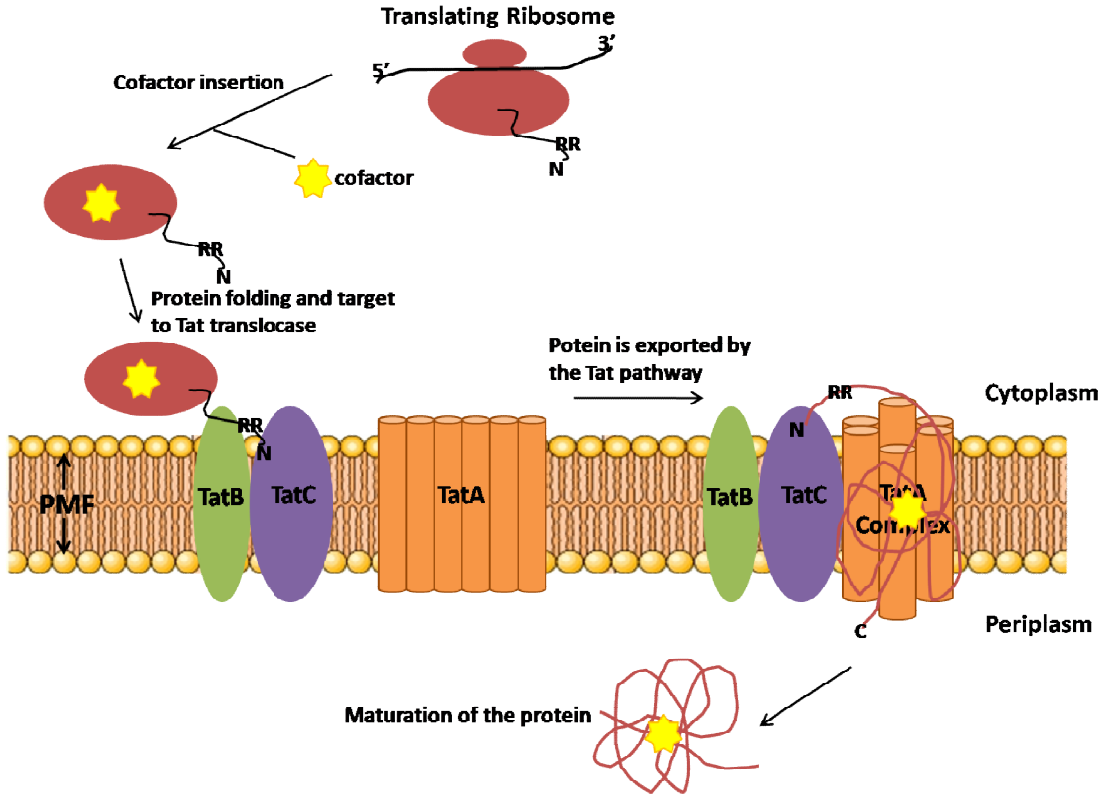


Figure 1.5 Proposed mechanism of Tat translocation in *E. coli* Firstly, Tat substrates are synthesised with the signal peptide bearing a twin arginine motif in the cytoplasm. Then cofactor is inserted and the protein is folded. Secondly, TatBC complex recognize and target the protein to the Tat translocon which TatA is recruited to form a channel relying on the membrane proton motive force (PMF). Lastly, translocation subsequently occurs. Upon translocation the signal peptide is cleaved and the mature protein is released into the periplasm.

Two distinct TatABC and TatA complexes were purified and characterized, and they are capable of recognizing the substrates and forming a variable translocation channel to export. Tat substrates are synthesized by transcription/translation and assembled correctly including the insertion of cofactors, and then their signal peptides are present in α-helical to interact with the membrane (Hou *et al.*, 2006; Shanmugham *et al.*, 2006). The translocation process is roughly divided into three steps. In the first step substrates are targeted to the Tat translocon under an energy-depleted condition (Alami *et al.*, 2002; Mould & Robinson, 1991). TatB and TatC in the TatABC complex are critical for the substrate-Tat association in which TatC initially recognizes the twin arginine motif of the signal peptide and TatB interacts more extensively with the substrate, resulting in the recruitment of TatA complexes to TatABC complex in a PMF-dependent manner (Alami *et al*., 2002; Alami *et al*., 2003; Mori and Cline, 2001, Mori & Cline, 2002). The electron microscopic (EM) study using over-expressed TatBC interacting with Tat substrate SufI in a *tatA* deletion stain provided evidence that one or two substrate molecules bind the periphery of TatBC to form TatBC-SufI complexes. Also the TatBC complex appears roughly spherical with a small central cavity which is unlike TatA complexes which form a size-variable channel with a lid on the cytoplasmic face (Tarry *et al*., 2009; Gohlke *et al*., 2005; Porcelli *et al*., 2002). It is suggested that TatB and TatC protomers can associate to be homo-polymers and the TatBC heterodimer contains at least one substrate-binding site (probably more). Furthermore, the study of chloroplasts has shown that the Tha4 (TatA in bacteria) complex is only recruited to compose the channel when the substrate binds to Hcf106-cpTatC (TatBC in bacteria) in the presence of the proton motive force (Mori and Cline, 2002). The next step is that TatA complexes combine to generate the differently-sized channels that could accommodate a diverse array of folded proteins across membrane (Gohlke *et al*., 2005; Oates *et al*., 2005). The topological evidence suggested the C-terminus of TatA is flipped across the membrane to give a hydrophilic-lined channel composed of the amphipathic domain (Gouffi *et al*., 2004) and the N-terminus remains on the *cis*-side of the membrane (Fincher *et al*., 1998).

At last, while the substrate is exported to the periplasm, TatA subsequently disassociates from the TatBC units and reset the system to the beginning ready for another round of translocation. Because the transport cycle is transient and efficient, it could provide an explanation of why Tat pathway is able to transport the largely folded proteins without effect on the membrane integrity.

Early work focused on the thylakoid system showed that translocation solely depends on the pH gradient of the PMF (Alder and Theg, 2003; Mori and Cline, 2002; Mould and Robinson, 1991). Bacterial Tat transport is also PMF-dependent since some in vitro assays have been developed using inverted *E. coli* membrane vesicles (Alami *et al*., 2002; Alami *et al*., 2003; Yahr and Wickner, 2001). However, observed that the thylakoid Tat pathway could be driven by the $\Delta\psi$ in vivo (Finazzi *et al.*, 2003), and it has to be pointed out that the membrane electrical potential is large in other organisms (Alder and Theg, 2003). For instance, *E. coli* SufI could be imported into IMVs in the absence of a ΔpH and it was demonstrated that two electric potentials were required in the translocation process (Bageshwar and Musser, 2007). Surprisingly, a recent study indicated that Tat dependent export of α-amylase in the haloarchaeon *Haloarcula hispanica* is driven by the sodium motive force (Kwan *et al.*, 2008). Thus, the energy resource of the Tat pathway is an essential point to study which remains poorly understood now and to a certain extent controversial.

1.4.6 Proofreading and quality control system in Tat pathway

The Tat system is specifically dedicated to the translocation of fully folded proteins and most of them require cofactor insertion or oligomerization with partner proteins in the cytosol prior to export, this raises a question how do cells ensure that proteins are correctly folded to be competent for transport? The majority of studies in both bacteria and chloroplasts have been performed to understand the mystery, resulting in a proposed proofreading or quality control mechanism that prevents the futile export of apo-proteins in Tat pathway. Generally speaking, it can be separated into

two aspects including the inherent ability of Tat system and the extrinsic assistance of other factors.

On the one hand, several studies have shown that the Tat system preferentially exports folded proteins even they are heterologous. For instance, Tat-mediated export of *cytochrome c* into the periplasm occurs only if maturation and folding in the cytoplasm are allowed (DeLisa *et al*., 2003) and *PhoD*, heterologously expressed with a Tat signal peptide, could only be exported when correctly folded (DeLisa *et al*., 2003), also a disulfide bond-containing protein PhoA was only translocated if folding of oxidative state and disulfide bond formation could take place in the cytoplasm (Richter and Bruser, 2005). In addition, the Tat pathway tends to recognise proteins containing weak hydrophobic signal peptides or hydrophilic domains (DeLisa *et al.*, 2003). It is supported by research which demonstrated that a small, unfolded hydrophilic polypeptide fused to Tat signal sequences was translocated (Panahandeh *et al*., 2008) and insertion of hydrophobic patches onto the surface of a globular protein abolished export via the Tat pathway (Richter *et al.*, 2007). Thus, the Tat pathway itself can proofread the passenger proteins and is predisposed to recognize and export folded proteins.

On the other hand, there are several surveillance factors to monitor the correctly assembly of Tat substrates in the cytoplasm (Fisher and Delisa, 2004). Firstly, it is speculated that pre-proteins require the assistance of general molecular chaperone systems such as GroELS and DnaK/Hsp70 for correct folding. All Tat leader peptides were observed to contain DnaK binding sites in a study (Rudiger *et al*., 1997). Indeed, it was discovered in plants that the Tat-mediated Rieske Fe/S protein makes contacts with both Cpn60 (GroEL in *E. coli*) and the DnaK/Hsp70 homolog prior to membrane insertion (Madueno *et al*., 1993; Molik *et al*., 2001). However, there was evidence for exceptions. For example, DnaK displays affinity for Tat leader peptides in vitro but is not required for the transport of the high-potential

iron-sulfur proteins in vivo (Oresnik *et al*., 2001; Bruser and Sanders, 2003). So far, the roles of generalized molecular chaperones in Tat export require more studies.

Secondly, a subset of specific cytosolic chaperones termed redox enzyme maturation proteins (REMPs) help Tat substrates to insert cofactor, fold properly and oligomerize with partner proteins, and they are often found in the same operon as their corresponding gene-encoded Tat substrates (Turner *et al*., 2004; Jack *et al*., 2004). Previous studies have examined two well-characterized REMPs, TorD and DmsD, which interact with trimethylamine *N*-oxide (TMAO) reductase (TorA) and dimethyl sulfoxide (DMSO) reductase (DmsA) respectively.

The *E. coli* DmsD was the first REMP analysed and it plays an important role in the DmsA localization and maturation (Weiner *et al*., 1992; Li *et al*., 2010). Studies have indicated that DmsD recognizes both N-terminal and C-terminal regions of the DmsA signal peptide (Chan *et al*., 2009; Li *et al*., 2010). The data show a clear interaction between DmsD and the ribosome-associated trigger factor (Tig), which suggested DmsD may interact with DmsA by attaching to its RR-leader immediately after the DmsA polypeptide leaves the ribosome and acts in the concert with Tig to protect it from untimely degradation or aggregation processes (Li *et al*., 2010). More recently, DmsD was shown to bind general chaperones (such as DnaK, DnaJ and GroEL) as well as proteins involved in the biosynthesis pathway of MGD cofactor in DmsA (Buchanan *et al*., 2008; Robinson *et al*., 2011). In addition, it was found that DmsD interacts with TatBC complex, implying DmsD may assist the localization of DmsA in Tat pathway (Ray *et al*., 2003). Surprisingly, DmsD is not only a REMP for DmsA, but is also suggested to interact with TorA and other putative DMSO reductases such as YnfE and YnfF (Chan *et al*., 2009).

In *E. coli* TorD is the most well-known REMP and this protein is required for the maturation of the Tat substrate TorA prior to export (Potter and Cole, 1999). So far, several roles have been experimentally attributed to TorD. First, Tat proofreading

involves a direct, specific interaction between TorD and N-terminal/ C-terminal regions of the signal peptide to stabilize and protect apo-TorA (Buchanan *et al*., 2008; Genest *et al*., 2008; Li *et al*., 2010). Furthermore, TorD binds the core of apo-TorA and contributes to the maturation of TorA. Structural analysis has predicted the very high α-helix content in TorD, in addition, study also has established that TorD α-helix 5 is crucial for the interaction with the core of TorA and this interaction is required for bis-MGD insertion (Buchanan *et al*., 2008; Genest *et al*., 2008). Recent studies revealed that TorD proofreads in a GTP-dependent way the cofactor insertion of TorA precursor and promotes its Tat translocation pathway (Genest *et al*., 2008). The study demonstrated that TorD exhibits a magnesium-dependent GTP hydrolytic activity and the GTPase activity is shown to be present only in the domain-swapped homodimeric form of the protein. GTP hydrolysis of TorD is used to regulate the maturation of TorA which involves the cooperation of two aspects of TorD with the folding of TorA and possible involvement of the MGD biosynthetic process. In conclusion, TorD takes part in all stages of TorA transportation from synthesis of apoprotein to translocation.

Another well-studied group of REMPs, HyaE and HybE, are required for the assembly of the hydrogenase components HyaA-HyaB and HybO-HybC respectively. These are multimeric but only one subunit bears a Tat signal peptide and the complexes are transported in dimeric form. For example, HyaE interacts with the signal peptide of HyaA but not with the mature form and HyaA form a dimer with HyaB in the cytoplasm prior to export (Dubini and Sargent, 2003; Rodrigue *et al*., 1999).

Although Tat substrates discussed above have the specific chaperones to assist in the folding of the protein and insertion of cofactors, this is not the case for all Tat passenger proteins. In addition, while homologs of the *E. coli* REMPs are found in many bacteria, they are still unknown in plants (Turner *et al*., 2004). Therefore, it

must be considered the general folding machinery of the cytosol plays a role in the transport of most substrates.

Most studies of proofreading were concentrated on bacterial systems. Recently, it was reported that a translocation reversal event occurres in the cpTat pathway. The study described that signal peptides were removed while a natural substrate pre-23K and a fused 23K-GFP traversed the thylakoid membrane, but they were somehow rejected by the translocase and returned to the stroma in many cases (Di Cola and Robinson, 2005). This suggests that even at a late stage the Tat pathway can sense and reject an incompatible substrate in plants. However, no significant Tat dependent retrograde transport has been reported in bacteria or archaea and the mechanism for this rejection is not clear. Overall, it is widely accepted that a folding quality control or proofreading mechanism monitors the "foldedness" of a Tat preprotein prior to transport, but it is unknown how such a process operates.

Either the Tat pathway itself or cytoplasmic chaperones participate in the proofreading processing, it is necessary to have a salvage pathway to refold or degraded rejected proteins. Indeed, a recent report provided evidence that two *E. coli* Tat substrates NrfC and NapG are blocked in export by the Tat pathway if FeS centre ligands were mutated, and mutant versions of the proteins were directed for degradation (Matos *et al*., 2008). There is evidence that accumulation of nontransported Tat preproteins that arise either from misfolding in the cytoplasm or from depletion of the *tat* genes often results in inactivation and degradation in the cytoplasm (Chanal *et al*., 1998; Delisa *et al*., 2003). The precise players in this degradation process are not currently known, although likely candidates include the FtsH protease (Bruser and Sanders, 2003) and the Clp machinery. The issue of proofreading and quality control on the Tat pathway is still a controversial one and remains undetermined now. The proposed model of proofreading in the Tat system is depicted in the Figure 1.6.

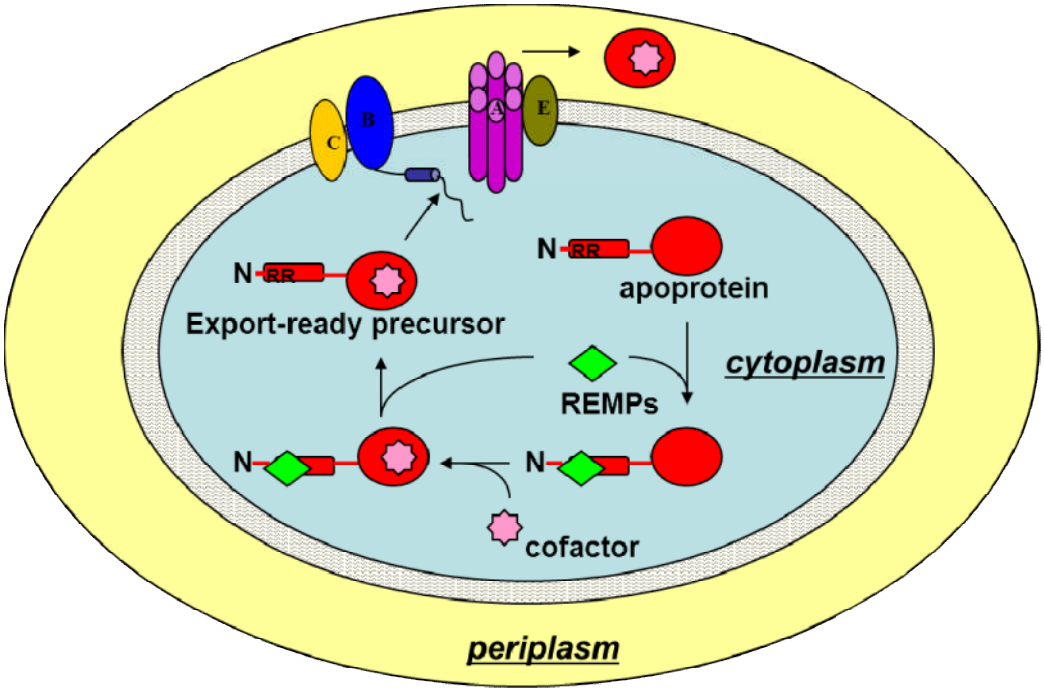


Figure 1.6 The model of the mechanism of proofreading by the Tat pathway in *E. coli* After the synthesis of proteins in the cytosol, REMPs (Redox Enzyme Maturation Proteins) can recognize and associate with the signal peptide of proteins which help the assembly of cofactor and correct folding of proteins. Then folded proteins are exported to the periplasm by the Tat patway. If the cofactor fails to insert into proteins or proteins are misfolded, they will be rejected to export by Tat pathway and can be degraded by the proteopysis.

1.4.7 Applications of Tat pathway in the future

Since the Tat translocation pathway is a potential target for novel antimicrobial compounds. For example, *E. coli* cells are more susceptible to lysis and sensitive to compounds such as the detergent SDS if two cell wall amidases AmiA and AmiC are mislocalised in the absence of the Tat translocase (Ize *et al*., 2003; Bernhardt and de Boer, 2003). Indeed, the ability of *E. coli* to infect host organisms can be abolished in *tat* mutants. Also, the Tat pathway is of importance in plant pathogens due to the case that it plays a role in establishment of plant infections by *Agrobacterium tumefacians* (Ding and Christie, 2003). Importantly, the absence of any Tat homologues in mammalian cells makes the Tat pathway of human pathogens a potential target for the development of novel antimicrobial drugs. Furthermore, the Tat pathway is attractive to biotechnological exploitations because it can be used to produce commercially important protein products that are not compatible with translocation by the Sec pathway. For reviews of the application of Tat pathway see Bruser, 2007 and De Buck *et al*., 2008.

1.5 The aims of project

The proofreading of substrates is an essential aspect in the Tat pathway and the overall aim of the project was to investigate and elucidate the proofreading of Tat substrates in *E. coli*. In order to realise it, work was undertaken to answer the following questions:

1. In *E. coli*, how are Tat substrates degraded if they can not be properly folded before export? The degradation is the last stage in the proofreading if proteins fail to insert cofactors and assemble correctly, otherwise these rejected macro-molecules will accumulate in cells.

2. In *E. coli*, how is this degradation monitored by Tat-dependent systems or other proteolysis programs? The majority of proteases are produced by cells and some protein degradation systems exist in the cytosol such as FtsH protease and the Clp machinery respectively which are important for degradation of rejected macromolecules. The possible role of the Tat pathway in proofreading was also considered to be important.

3. Proofreading/quality control systems are important in the Tat pathway, but how do these mechanisms work? One of the aims was to identify novel components of the quality control machinery.

4. Signal peptide cleavage is essential for completion of the overall translocation pathway. Also the cleavage of the signal peptide after translocation is important for the maturation of proteins. Here, I studied how the Tat translocation process is related to the maturation of proteins, and whether maturation is linked to proofreading of substrates and possible reverse translocation.

Chapter 2

Materials and Methods

2.1 Suppliers of Materials

2.1.1 Antibodies

Strep-tactin™ HRP-conjugate (**IBA, Germany**)
10× Buffer E (**IBA, Germany**)
Anti-His (C-terminal) antibody (**Invitrogen, USA**)
Anti-mouse-HRP conjugates (**Promega, UK**)
Rabbit Anti-GFP (**Invitrogen, USA**)

2.1.2 Kits

ECL™ Western Blotting detection reagents (**Amersham Pharmacia Biotech, UK**);
EZ-ECL Chemiluminescence Detection kit for HRP (**Geneflow Limited, UK**)
QIAprep miniprep kit **(Qiagen, Germany)**
DNeasy® blood & tissue kit (**Qiagen, Germany**)
Sulfo-SBED Biotin label transfer kit-western blot application (**Thermo scientific, USA**)

2.1.3 Molecular weight markers

GeneRuler™ 1kb DNA ladder (**Fermentas, UK**)
Pre-stained broad range protein markers (**New England Biolabs, USA**)

2.1.4 Enzymes

Pfu DNA polymerase (**Stratagene, USA**)
Platinum® *Pfx* DNA polymerase (**Invitrogen, USA**)
Platinum® *Taq* DNA polymerase (**Invitrogen, USA**)
FastAP™ Thermosensitive Alkaline Phosphatase (**Fermentas, UK**)
T4 DNA ligase (**Fermentas, UK**)
DpnI, NcoI, EcoI, XbaI, NheI, KpnI restriction enzyme (**Invitrogen, USA**)

2.1.5 Affinity/ ion exchange Resin

Talon™ affinity chromatography resin (**BD Clontech, USA**)
Q-sepharose anion exchange resin (**Amersham Pharmacia Biotech, UK**)
Profinity™ IMAC Uncharged Resin (**Bio-Rad, UK**)
ABT® Affinity his-tag purification low density chelate kit (**Webscientific, UK**)
His-Select® Cobalt Affinity Gel (**Sigma, UK**)
Ni-NTA Agarose (**Qiagen, Germany**)
Chelex® 100 Ion Exchange Resin (**Bio-Rad, UK**)

2.1.6 Chemicals

All general laboratory reagents used in this study were of the highest grade available and were supplied by Fisher Scientific (UK), Sigma (UK) or VWR (UK) unless otherwise stated below.
dNTPs Mix (**Promega, UK**)
Bradford Reagent (**Fermentas, UK**)
Instant Blue stain (**Stratech, UK**)
CompleteTM protease inhibitor cocktail tablets (**Roche applied science, UK**)
Dried skimmed milk power (**ALDI, UK**)

2.1.7 Additional materials

Hybond™-P PVDF membrane (**Amersham Pharmacia Biotech, UK**);
PVDF blotting membrane (**peQlab, Germany**)
Super RX film (**Fuji, Japan**)
Novagen® D-Tubet™ Dialyzer Maxi (**MERCK, Germany**)
Slide-A-Lyzer® Cassette (**Pierce, USA**)
Zeba™ Desalt Spin Columns (**Thermo scientific, UK**)
Disposable 5ml/10ml Polypropylene Columns (**Thermo scientific, UK**)
Vivaspin concentrator (**Sartovius stedin biotech, UK**)

2.2 Growth and storage of *E. coli* culture

2.2.1 Strains of *E. coli* used

Table 2.1 List of the *E. coli* strains used in this study.

Strain	Relevant properties	Reference/ source
DH5α	*F_80lacZ_M15_(lacZYA-argF) U169 deoR recA1 endA1 hsdR17 (rk_, mk_) gal-phoA supE44 thi-1 gyrA96 relA1*	Invitrogen, Carlsbad, CA, USA
MC4100AR	*AraR, FaraD139DlacU169 rpsL150 relA1 flB5301 deoC1 ptsF25 rbsR*	Casadaban *et al* (1979); Bolhuis *et al* (2000)
ΔtatABCDE	As MC4100AR; *ΔtatABCDE*	Wexler *et al* (2000); Bolhuis *et al* (2000)
ΔtatA	As MC4100AR; *ΔtatA*	Sargent *et al* (1998); Bolhuis *et al* (2000)
ΔtatB	As MC4100AR; *ΔtatB*	Sargent *et al* (1998); Bolhuis *et al* (2000)
ΔtatC	As MC4100AR; *ΔtatC*	Sargent *et al* (1998); Bolhuis *et al* (2000)
ΔtatD	As MC4100AR; *ΔtatD*	Englesberg *et al* (1962); Wexler *et al* (2000)
ΔtatE	As MC4100AR; *ΔtatE*	Sargent *et al* (1998); Bolhuis *et al* (2000)
ΔtatA/E	As MC4100AR; *ΔtatA/E*	Sargent *et al* (1998); Bolhuis *et al* (2000)
ΔtorD	*torD* deletion strain	FTD100/ Sargent *et al* (2004)
ΔdmsD	*dmsD* deletion strain	FTD102/ Sargent *et al* (2004)
ΔnapD	*napD* deletion strain	SCD1/Sargent *et al* (unpublished)

2.2.2 Aerobic growth condition

E. coli cells were generally cultured in Luria-Bertani (LB) liquid medium (10 g/L sodium chloride, 10 g/L bacto-tryptone, 5 g/L yeast extract). Growth of culture was achieved by inoculating the pre-warmed LB medium with a single bacterial colony taken from an agar plate. The stated liquid cultures were grown at 37°C in a shaking incubator at 200 rpm for approximately 16 hrs under aerobic conditions. LB medium with 16 g/l bacto-agar was used for the aerobic growth of *E. coli* on plates, followed by incubation at 37 °C oven overnight (~16 hours).

2.2.3 Antibiotic supplements

Following concentration antibiotics were used if necessary: Ampicillin(100 μg/ml) for pBAD24 vector and Kanamycin (50 μg/ml) for pEXT22 vector.

2.2.4 Storage of *E. coli* strains

E. coli strains were maintained on LB-agar plates at 4°C for a maximum of 6 weeks; and were re-plated every two months.

Glycerol stocks of *E. coli* strains were prepared for long-term storage by the addition of 2 parts of liquid culture with 1 part of 50% (v/v) glycerol. Cultures were frozen rapidly on dry ice and stored at -70°C.

2.3 Preparation and transformation of competent *E. coli* cells

2.3.1 Preparation of competent *E. coli* cells

A starter culture of *E. coli* cells was inoculated in 5mL LB with antibiotics and grown at 37°C overnight. 1:100 diluted in 10 mL LB, and the culture was grown at 37°C with shaking at 200 rpm until OD600 = 0.3–0.4. Cells were centrifuged at 3000 rpm (Beckman Jouan GR422) at 4°C for 10 min, and pellets re-suspended in 10 mL of ice cold 0.1M $MgCl_2$ were incubated on ice for 5 min. Cells were collected again

and re-suspended in 1.0 mL ice-cold 0.1M $CaCl_2$. The prepared cells were incubated on ice for 2–24 hours before use.

2.3.2 Preparation of super competent *E. coli* cells

The culture was prepared by inoculating 200 mL LB with *E. coli* DH5α cells in a 2L flask, then grown at 25°C with shaking at 200 rpm overnight until OD600 = 0.3–0.4. Cells were incubated in the cold room for 30 min before being pelleted by centrifugation at 3000 rpm (Beckman Jouan GR422) at 4°C for 10 min, re-suspended in 10 ml ice cold 0.1 M $CaCl_2$ and incubated on ice for 30 min. Cells were pelleted, resuspended in 8 mL ice cold 50 mM $CaCl_2$ plus 20% glycerol, split into 100 µL aliquots and frozen immediately on dry ice. Aliquots were stored at -80°C until needed.

2.3.3 Preparation of plasmid DNA from *E. coli* cells

Plasmid DNA was isolated from 5 ml overnight culture using a 'QIAprep Mini-Prep' kit (Qiagen, Germany). The method was carried out exactly according to the manufacturer's instructions, plasmid DNA was recovered in 50 µL dH_2O and stored at -20°C.

2.3.4 Transformation of competent *E. coli* cells

2-3µl plasmid DNA or ligation mix was incubated with 50µl competent cells, and incubated on ice for 30 mins. The cells were transferred to a 42°C heat block for 2 mins, and then put back on ice for 2 mins, after which 0.5 ml of LB was added subsequently to incubate at 37°C with 200rpm shaking for 1 hour. The cells were pelleted by centrifugation at 4000rpm for 5 mins, 400 µl supernatant was removed, and pellets were resuspended in the remaining liquid and spread on to agar plates containing the appropriate antibiotics. Plates were incubated at 37°C oven overnight.

2.3.5 Bacterial plasmids used in this work

Table 2.2 Plasmids used in this study (see appendix for plasmid maps of pBAD24 and pEXT22).

Plasmid	Details	Reference/ source
pBAD24	*araC, Amp*^{r}	Guzman *et al*., 1992
pEXT22	*LacI, Kan*^{r}	Dykxhoorn *et al.*, 1996
pBLA	pBAD24 expressing pre-LolA with a 6-his-tag	This study
pXLA	pEXT22 expressing pre-LolA with a 6-his-tag	This study
pBTA	pBAD24 expressing pre-TorA with a 6-his-tag	This study
pXTA	pEXT22 expressing pre-TorA with a Strep-tag	This study
pXTAM	pEXT22 expressing mature TorA with a Strep-tag	This study
pXTA_KK	pEXT22 expressing pre-TorA with a Strep-tag but RR in signal peptide are mutated to KK	This study
pJDT1	pBAD24 + *torA-GFP*	Thomas *et al*., 2001
pBNA	pBAD24 expressing pre-NapA with a 6-his-tag	This study
pBDA	pBAD24 expressing pre-DmsA with a 6-his-tag	This study
pCRY1	pBAD24 expressing pre-YedY with a 6-his-tag	This study
pCRY2	As pCRY1 but point mutations C146A in the yedY gene	This study
pCRY3	As pCRY1 but point mutations E92D in the yedY gene	This study
pCRY4	As pCRY1 but point mutations R240K in the yedY gene	This study
pCRY1_KR	As pCRY1 but RR in signal peptide is mutated to KR	This study
pCRY1_RK	As pCRY1 but RR in signal peptide is mutated to RK	This study
pCRY4_KR	As pCRY4 but RR in signal peptide is mutated to KR	This study
pCRY4_RK	As pCRY4 but RR in signal peptide is mutated to RK	This study
pCRY1_A44L	As pCRY1 but the SP cleavage site AxA is mutated to AxL	This study

Plasmid	Details	Reference/ source
pCRYG	pBAD24 + yedY-GFP	This study
pCRYG_A44L	As pCRYG but the SP cleavage site AxA is mutated to AxL	This study
pCRY5	pEXT22 expressing pre-YedY with a 6-His-tag	This study
pCRY6	As pCRY5 but point mutations C146A in the yedY gene	This study
pCRY7	As pCRY5 but point mutations C146A in the yedY gene	This study
pCRY8	As pCRY5 but point mutations C146A in the yedY gene	This study
pCRY5_KR	As pCRY5 but RR in signal peptide is mutated to KR	This study
pCRY5_RK	As pCRY5 but RR in signal peptide is mutated to RK	This study
pCRY8_KR	As pCRY8 but RR in signal peptide is mutated to KR	This study
pCRY8_RK	As pCRY8 but RR in signal peptide is mutated to RK	This study
pCRY5_A44L	As pCRY5 but the SP cleavage site AxA is mutated to AxL	This study

2.3.6 Affinity tags used

A C-terminal hexa-histidine (*his*) tag or ten- amino acid *Strep*-tag II (IBA, Germany) was incorporated into constructs, where indicated as 6-his-tag or strep-tag, to facilitate protein detection and purification.

2.3.7 Primers used for the constructs generated

Oligonucleotides were supplied by Invitrogen (UK) and used for molecular cloning.

Table 2.3 DNA primers used for the preparation of *E. coli* substrates.

Name	Sequence 5' to 3'
NcoI_TorA_F	CGTAGCCACCATGGCCAACAATAACGATCTCTTTCAGG CATCACGTCGGC
TorA_His_XbaI_R	TGGCTACGTCTAGATCAGTGATGGTGATGGTGATGTGA TTTCACCTGCG
NcoI_NapA_F	CGTAGCCACCATGGCCAAACTCAGTCGTCGTAGCTTTA TGAAAGCTAACGCCGTTGC
NapA_His_XbaI_R	TGGCTACGTCTAGATTAGTGATGGTGATGGTGATGCAC CTTCTCCAGTTTGACCG
NcoI_DmsA_F	CGTAGCCACCATGGCCAAAACGAAAATCCCTGATGCGG TATTGGCTGC
DmsA_His_XbaI_R	TGGCTACGTCTAGATTAGTGATGGTGATGGTGATGCAC CTTTTCAACCTGAACAAGGTTTGTATGTGACG
NheI_RBS_YedY_F	CGTAGCCAGCTAGCAGGAGGAATTCAATGAAAAAGAA TCAATTTTTAAAAGAATCAGATGTTACG
YedY _His_XbaI_R	TGGCTACGTCTAGATTAGTGATGGTGATGGTGATGGAA ATTCTCCCG
X_KnpI_RBS_TorA_F	CGTAGCCAGGTACCAGGAGGAATTCAATGAACAATAA CGATCTCTTTCAGGC
X_MATURE_TorA_F	CGTAGCCAGGTACCAGGAGGAATTCAATGGCGCAAGC GGCGACTGACG
X_TorA_Strep_ XbaI_R	TGGCTACGTCTAGATCATTTTTCAAACTGTGGGTGCGAC CAATTCGATGATTTCACCTGCG
B_ NcoI_LolA_F	CGTAGCCACCATGGATGATGAAAAAAATTGCCATCACC TGTGC
X_KnpI_RBS_ LolA_F	CGTAGCCAGGTACCCTACCACAGAGGAACATGTATGAT GAAAAAAATTGCCATCACCTGTGC
BX_LolA_His_ XbaI_R	TGGCTACGTCTAGACTAGTGATGGTGATGGTGATGCTT ACGTTGATCATCTACCG

2.4 DNA manipulation and cloning techniques

2.4.1 Preparation of Genome DNA from *E. coli*

Genome DNA of *E. coli* was isolated using a DNeasy® blood & tissue kit (Qiagen, Germany). The kit was used exactly as the manufacturer's instructions. The extracted genome DNA was recovered in 50 µL dH_2O and used the template of PCR.

2.4.2 PCR

2.4.2.1 Typical PCR using *Pfu* DNA polymerase

PCR amplification was generally performed with 5 µl 10× buffer, 1 µl 10 mM dNTPs mix, 1 µl of each primer (from 20 µM stock), 1 µl template DNA, 1 µl *Pfu* DNA polymerase (5 units/µL) and dH_2O to a final volume of 50µl.

The reactions took place in a Biometra T3 Thermocycler with the following cycling conditions:

94°C 4 min
94°C 15 sec
52°C 45 sec
72°C X mins (2 mins/per kb) } 34 cycles
72°C 10 min
4 °C pause

2.4.2.2 PCR using Platinum® *Pfx* DNA polymerase

PCR amplification was carried out using 2.5 µl 10× amplification buffer, 3 µl 10× Enhance solution, 1 µl 10 mM dNTP mix, 1 µl of each primer (from 20 µM stock), 1 µl template DNA, 1 µl *Pfx* DNA polymerase (2.5 units/µL), gradient 1.5-5 mM $MgSO_4$ and dH_2O to a final volume of 25µl.

The reactions took place in a Biometra T3 Thermocycler with the following cycling conditions:

94°C 2 min
94°C 30 sec
55°C 30 sec
68°C X mins (2 mins/per kb)
} 20 cycles
68°C 10 min
4 °C pause

2.4.2.3 PCR using Platinum® *Taq* DNA polymerase

PCR amplification was carried out using 2.5 µl 10× PCR buffer, 1 µl 10 mM dNTPs mix, 1 µl of each primer (from 20 µM stock), 1 µl template DNA, 1 µl *Taq* DNA polymerase (2.5 units/µL), gradient 2-5 mM $MgCl_2$ and dH_2O to a final volume of 25µl. The reactions worked as *Pfx* DNA polymerase above.

2.4.3 Agarose gel electrophoresis

Agarose gels were prepared by dissolving 1–1.5% (w/v) agarose in TBE buffer (89 mM Tris; 89 mM H_3BO_3; 20 mM EDTA, pH 8.0) with heating to boiling. Ethidium bromide was added (0.5 µg/ml final concentration) and set the gel was submerged in TBE buffer immediately.

DNA samples for analysis were prepared by mixing 2 µl loading dye (0.25% bromophenol blue, 40% w/v sucrose). Electrophoresis was performed at 100–150 V for 45 mins until the dye front was at least half way through the gel. DNA bands were visualised under a UV transilluminator and image of gels were taken on a gel documentation system.

2.4.4 Purification of DNA by gel extraction

The DNA of interest was excised from agarose gels under a UV transilluminator and extracted using a QIAprep Gel Extraction kit (Qiagen, Germany). Kit was used according to the manufacturer's instructions supplied.

2.4.5 Restriction endonuclease digestion of DNA

Restriction endonuclease digestion of DNA was carried out as the enzyme manufacturer's guidelines. Reactions were typically incubated for 1-2 hours in a 37°C water bath, before being subjected to gel electrophoresis and purified using the procedure described above.

2.4.6 Dephosphorylation of vector DNA

The Digested vector DNA was added 2 µl Alkaline Phosphatase (Fermentas, UK) and incubated for 20 mins in a 37°C water bath before performing to gel electrophoresis and purified using the procedure described above.

2.4.7 Ligation of digested vector and insert DNA fragments

Ligation reactions were performed using T4 DNA ligase (Fermentas, UK) in the light of the manufacturer's guidelines. Vector and insert DNA fragments were added in 1:3 ratio of their concentration and incubated at room temperature for 1 hour.

2.4.8 Determination of DNA concentration

The concentration of DNA was measured by NanoDrop ND1000 spectrophotometer.

2.4.9 Sequencing of plasmid DNA

All DNA Sequencing was conducted by the Molecular Biology Service, University of Warwick, using Applied Biosystem Big Dye Terminator v3.1 chemistry and an ABI 3100 Genetic Analyser^{TM} (Applied Biosystems, UK). Meglign and Chromas software were used to analyse DNA sequence.

2.4.10 Sequencing primers used in this study

Table 2.4 Sequencing primers used in this study.

Name	Sequencing Target	Sequence 5' to 3'	Source
BADSEQ_F	pBAD24	TATTTGCACGGCGTCACA	C. Matos (School of Life Sciences, University of Warwick)
BADSEQ_R	pBAD24	AAAACTCTCAAGGATCTT	C. Matos (School of Life Sciences, University of Warwick)
pEXT22SEQ_F	pEXT22	CGCTCAAGGCGCACT CCCG	C. Matos (School of Life Sciences, University of Warwick)
pEXT22SEQ_R	pEXT22	CCGCGCACATTTCCCCG	C. Matos (School of Life Sciences, University of Warwick)
Front_TorA	TorA	GGCTCGCAAAAGTATCCG	This study
TorA_S1		GCTGGTATTGCAGAACAGC	This study
TorA_S2		CGATTCCGCCTGTTCACG	This study

2.4.11 DNA site-specific mutagenesis of plasmid DNA

Point mutations were introduced into *E. coli* plasmid DNA using *Pfu* DNA polymerase and a QuikChange™ Site-Directed Mutagenesis kit (Stratagene, USA), following to the manufacturer's instructions. PCR reactions were carried out using a Biometra T3 thermocycler. The amplified PCR product was transformed into super competent *DHα5* cells as described (§2.3.4)

2.4.12 Primers used for mutagenesis of plasmid DNA

Table 2.5 Primers used for mutagenesis of plasmid DNA.

Name	Sequence 5' to 3'	Substitution
YedY_C146A_F	CGTATTTATCGTATGCGCGCCGTGGAAGC GTGGTCG	Cys146 → Ala
YedY_C146A_R	CGACCACGCTTCCACGGCGCGCATACGAT AAATACG	
YedY_E92D_F	CGGTTATAATAACTTCTATGATTTCGGGCT GGATAAAGC	Glu92 → Asp
YedY_E92D_R	GCTTTATCCAGCCCGAAATCATAGAAGTT ATTATAACCG	
YedY_R240K_F	CGCGCCGGTGAAACTGATTGTGCCG	Arg240 → Lys
YedY_R240K_R	CGGCACAATCAGTTTCACCGGCGCG	
YedY_RK_F	CGGTATTCTTTATGAAGCGTAAGCAGGTG TTAAAAGC	Arg23 → Lys
YedY_RK_R	GCTTTTAACACCTGCTTACGCTTCATAAAG AATACCG	
YedY_KR_F	CGGTATTCTTTATGAAGAAACGGCAGGTG TTAAAAGC	Arg23 → Lys
YedY_KR_R	GCTTTTAACACCTGCCGTTTCTTCATAAAG AATACCG	
X_TorA_KK_F	CGATCTCTTTCAGGCATCAAAAAAACGTTT TCTGGCACAACTCG	Arg11/12 → Lys
X_TorA_KK_R	CGAGTTGTGCCAGAAAACGTTTTTTGATG CCTGAAAGAGATCG	
YedY_A44L_F	GCCTCACGCTGCGCATCTCGATCTGCTTAG C	Ala44 → Lys
YedY_A44L_R	GCTAAGCAGATCGAGATGCGCAGCGTGAG GC	

2.5 Methods concerning preparation of proteins

2.5.1 Plasmid induction

Typically, induction of pBAD24 or pEXT22 based plasmids were carried out by the addition of 0.1M arabinose or 5 mM IPTG to the culture medium for 3-4 hrs (until mid-exponential phase).

2.5.2 Preparation of whole cell fractions

Preparation of whole cell fractions was performed by harvesting 0.5 mL of a mid-exponential phase culture and resuspending in 80 µl SDS-PAGE sample buffer. Samples were incubated at 50°C for 10 min before analysis by SDS-PAGE or frozen -20°C until required.

2.5.3 Fractionation of *E. coli* cells

E. coli cells were separated into periplasm, cytoplasm, and membrane fractions based on a previously described lysozyme/cold osmotic shock method (Randall and Hardy, 1986). Briefly 1:20 diluted overnight cell culture of *E. coli* in 10 ml LB plus appropriate antibiotics and inducer to grow 3 hours until OD_{600} = 0.6. The culture was centrifugated and re-suspended in 0.5 ml of pre-chilled buffer 1 (100 mM tris-acetate pH 8.2, 0.5 M sucrose and 5 mM EDTA). 40 µl lysozyme (2 mg/mL) and 0.5 ml of H_2O were added simultaneously before incubation on ice for 5 mins. The resulting spheroplasts were treated with 20 µl of 1.0 M $MgSO_4$ and pelleted by centrifugation at 14,000 rpm (Eppendorf 5417R). The supernatant was collected as the periplasmic fraction. Spheroplasts were washed using 1 ml buffer 2 (50 mM tris-acetate pH 8.2; 0.25 mM sucrose; 10 mM $MgSO_4$) and pelleted by centrifugation as above. The supernatant was discarded and the spheroplasts were re-suspended in 1 ml pre-chilled buffer 3 (50 mM tris-acetate pH 8.2, and 2.5 mM EDTA). Spheroplasts pre-cooled, and briefly sonicated (3× 10 seconds at ~8 microns) to break the membranes. The membrane was separated from cytoplasmic fraction by ultracentrifugation at 70,000 rpm for 30 minutes at 4°C (Beckman TL100.3 rotor).

Membranes were solubilised in 1.0 ml buffer 3 containing 1 % (w/v) SDS unless otherwise stated.

2.5.4 Proteolysis of spheroplasts

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

2.5.5 Carbonate extraction

Spheroplasts were prepared as described and resuspended in 1.0 ml ice-cold 0.1 M pH 11.5 Na_2CO_3. Samples were incubated on ice for 30 min and pelleted by ultracentrifugation at 50,000 rpm for 30 minutes at 4°C (Beckman TL100.3 rotor). Supernatant was collected and precipitated with 10% TCA, and pellet was dried after the rinse with ice-cold water. The supernatant and pellet were dissolved in 80 µl SDS sample loading buffer.

2.5.6 Trichloroacetic acid (TCA) precipitation

Protein samples were concentrated for gel electrophoresis by TCA precipitation. 1 ml 10% TCA was added and samples were incubated on ice for 30 min before centrifugation at 14,000 rpm for 15 min (Eppendorf 5417R). The supernatant was removed and 1 ml acetone was added and incubated on ice for 5 min before centrifugation for 10 min. The supernatant was removed and samples were allowed to air dry at room temperature. Samples were resuspended in 80 µL SDS sample loading buffer.

2.5.7 Isolation of *E. coli* cytoplasmic fraction, membrane fraction and periplasmic fraction for protein purification and protein interaction assay

For achievement of proteins from the cytoplasmic fraction, membrane fraction and peripasmic fraction, samples were prepared from a 1 L culture of *E. coli*. 1 L of cell cutlure were centrifugated at 6000 rpm for 15 mins at 4°C (Beckman TLA8.1 rotor) and re-suspended in 48 ml of pre-chilled disruption buffer (100 mM tris-acetate pH 8.2 and 0.5 M sucrose) which added 1 complete protease inhibitor tablet. 2 ml lysozyme (2 mg/mL) was added before incubation on ice for 5 mins followed by the addition of 48 ml ice cold dH_2O. The resulting spheroplasts were treated with 2 ml of 1.0 M $MgSO_4$ and pelleted by centrifugation at 4000g for 15 min at 4°C (Beckman Jouan). The supernatant was collected as the periplasmic fraction. The pellet was re-suspended in 18 ml pre-chilled sonication buffer (50 mM tris-acetate pH 8.2) which added 1 complete protease inhibitor tablet. Spheroplasts were performed by sonication (3× 30 seconds at ~10 microns) to break the membranes. Unbroken cells and debris were removed by centrifugation at 4000g for 10 min at 4°C (Beckman Jouan). The cytoplasmic fraction was separated from membrane by ultracentrifugation at 70,000 rpm for 30 minutes at 4°C (Beckman TL100.3 rotor). Membranes were solubilised in 20 ml Solubilisation Buffer (20 mM Tris-HCl pH 8.0, 50 mM NaCl, 10% Glycerol, 1% Digitonin) overnignt, and membrane protein solution was clarified by centrifugation at 4000g for 30 min at 4°C (Beckman Jouan).

2.6 Protein chromatography methods

2.6.1 Q-Sepharose anion exchange chromatography

E. coli cytoplasmic fraction was applied to 10 ml Q-Sepharose anion exchange column (Amersham Biosciences, UK) that had been pre-washed with 2 CVs (column volumes) Equilibration buffer (20 mM Tris-HCl pH 8.0, 50 mM NaCl). The column was subjected to washing with 2 CVs of wash buffer (20 mM Tris-HCl pH 8.0, 100 mM NaCl) to remove weakly bound proteins from the column. Strongly bound proteins were eluted from the column by the addition of 150-300 mM gradient NaCl with 20 mM Tris-HCl pH 8.0.

2.6.2 Affinity chromatography

2.6.2.1 Immobilized metal ion affinity chromatography (IMAC)

1 ml of Profinity™ IMAC Uncharged Resin (Bio-Rad, UK) was prepared in a column, and 5ml buffer (50mM sodium acetate pH 4.0, 300 mM NaCl) was added to flow through. 5ml 0.1-0.3 M solution of the metal ion as Zn, Cu, Ni was applied to the column and 5ml buffer described above used to wash the column before 10ml dH_2O flowing through. When it was ready, the column was pre-equilibrated in 5ml of buffer (20 mM Tris-HCl pH 5.0, 400 mM NaCl, 10% Glycerol and 5 mM imidazole). The proteins were added to the column and the flow through collected. The column was washed using 100 ml buffer (20 mM Tris-HCl pH 5.0, 400 mM NaCl, 10% Glycerol and 5 mM imidazole). Bound proteins were eluted with elution buffer containing 50-500 mM gradient imidazole.

2.6.2.2 Talon™ metal affinity chromatography

2.5 ml of Talon affinity resin (BD biosciences, UK) was prepared in a colomn and pre-equilibrated in 25ml of buffer (20 mM Tris-HCl pH 6.0, 300 mM NaCl, and 5 mM imidazole). Partially purified proteins or cytoplasmic fraction/periplasmic fraction were added to the colomn and the flow through collected. The column was washed using 100 ml buffer (20 mM Tris-HCl pH 6.0, 300 mM NaCl, and 5 mM imidazole) to remove weakly bound proteins. Bound proteins were eluted with elution buffer containing 40-500 mM gradient imidazole.

2.6.2.3 His-Select® Cobalt Affinity chromatography

His-Select® Cobalt Affinity resin (Sigma, UK) has used to purify the protein from cytoplasmic fraction. Samples were poured into 2.5 ml of cobalt affinity coloumn which has been pre-equilibrated in 25ml of buffer (50 mM NaH_2PO_4 pH 8.0, 300 mM NaCl, and 10 mM imidazole) and the flow through collected. The column was washed using 100 ml buffer (50 mM NaH_2PO_4 pH 8.0, 300 mM NaCl, and 10 mM

imidazole) and bound proteins were eluted with elution buffer containing 40-250 mM gradient imidazole.

2.6.2.4 Ni-NTA Affinity chromatography

Buffers were super-cleaned using Chelex® 100 Ion Exchange Resin (Bio-Rad, UK) to remove all possible metal ions before purification.

The cytoplasmic fraction/periplasmic fraction were added and incubated with 2.5 ml of Ni-NTA affinity resin (Qiagen, Germany) which has been pre-equilibrated in 25ml of buffer (50 mM NaH_2PO_4 pH 7.0, 300 mM NaCl, and 10 mM imidazole) for 2-3 hours at 4°C. The resulting mixture was poured into a column and the flow through collected. The column was washed using 100 ml buffer (50 mM NaH_2PO_4 pH 7.0, 300 mM NaCl, and 10 mM imidazole) to remove weakly bound proteins. Bound proteins were eluted with elution buffer containing 40-250 mM gradient imidazole.

2.6.3 Determination the concentration of purified protein

To determine the concentration of protein, Bradford protein assay was used by Bradford Reagent (Fermentas, UK) according to the manufacturers' recommendation. The data were measured spectrophotometrically (Jenway 6505 UV/Vis spectrophotometer) and compared to the standard curve.

2.7 Methods of protein interaction

2.7.1 Protein interaction assay with Ni^{2+} affinity chromatography

Ni^{2+} affinity chromatography is one of the useful approached to carry out protein interaction research. Typically the Ni^{2+}-NTA resin (Qiagen, Germany) contained column was prepared and rinsed with 12 CV buffer (50 mM NaH_2PO_4 pH 7.0, 300 mM NaCl, and 10 mM imidazole). The purified bait protein was bulk incubated with column end-over-end for at least 1 hour before 3-4 CV cytoplasmic fraction or membrane fraction was routinely applied to incubate for 3 hours. Column was then

washed with buffer (50 mM NaH_2PO_4 pH 7.0, 300 mM NaCl, and 25 mM Imidazole) until wash fractions gave A_{280} less than 0.01. Bound proteins were eluted with 1 CV 100 mM glycine pH 2.8.

2.7.2 Formaldehyde cross-linking in vivo and in vitro

In vivo cross-linking studies were performed as described by Prossnitz *et al*. Briefly, cells were grown at 37°C to an OD600 of 0.6 and pelleted by centrifugation at 4000g for 15 min at 4°C (Beckman Jouan). The pellets were washed and resuspended at OD600 of 0.5 in ice-cold 10mM KPi pH 6.8, and incubated further for 20 min with the addition of final concentration of 1% Formaldehyde at the room temperature. The cross-linked samples were washed and resuspended with ice-cold 10mM KPi pH 6.8, then either heated at 37 °C for 10 min to maintain the cross-links or at 96 °C for 20 min to break the cross-links.

In vitro, purified protein and cytoplasmic fraction were mixed and incubated for 30 min at room temperature. Final concentration 1% formaldehyde was added, and the mixture was further incubated for 20 min. The cross-linked samples were heated at 37 °C for 10 min to maintain the cross-links or at 96 °C for 20 min to break the cross-links, and analysed by SDS-PAGE.

2.7.3 Sulfo-SBED Biotin label transfer

The experiment was performed as the description of protocol in the Sulfo-SBED Biotin label transfer Kit-Western blot application (Thermo scientific, USA).

2.8 Electrophoresis of proteins

2.8.1 SDS poly-acrylamide gel electrophoresis (PAGE)

Gel electrophoresis of proteins was carried out using vertical gel system from C.B.S. (USA) and based on the method described by Laemmli, (1970). A 15% separation gel was made which was composed of: 15% acrylamide gel; 375 mM Tris-HCl pH 8.8; 0.1% SDS; 0.1% APS; 0.06% TEMED. A 5% stacking gel was overlaid when it

was solidified, and the stacking gel composed of: 5% acrylamide gel; 125 mM Tris-HCl pH 6.8; 0.1% SDS; 0.6% APS, 0.06% TEMED.
Electrophoresis was performed with a uniform running buffer (25 mM tris; 250 mM glycine; 0.1% SDS), at typically 30 mA for 3 hours or 3 mA overnight until the dye front had migrated off the bottom of the gel.

2.8.2 Blue-native poly-acrylamide gel electrophoresis

All electrophoresis using blue-native gels was performed using the C.B.S. (USA) vertical gel system as instructed by the manufacturer. The method used was based on one originally described by Schagger and von Jagow (1991).

Table 2.6 A 1.5 mm gradient gel was prepared in this study:

Component	Stacker gel (4%)	Gradient separation gel	
		5%	13.5%
Acrylamide (40%)	0.5ml	1.37ml	3.72ml
Gel buffer (3×)	1.65ml	3.64ml	3.64ml
Glycerol (50%)	-	170ul	1.1ml
Water	2.85ml	5.82ml	2.53ml
APS (10%)	35ul	40ul	35ul
TEMED	3.5ul	4ul	3.5ul

Samples were prepared by mixing 10 µl of solubilised membranes with 1 µl of sample loading buffer (5% Serva G Coomassie blue and 750 mM aminocaproic acid).

Gradient gels were equilibrated at 4ºC before loading samples. Gels were initially run at 70 V at 4ºC with coloured cathode buffer (50 mM Tricine pH 7.0, 15 mM Bis-Tris-HCl pH 7.0, 0.02% Serva G Coomassie blue) and anode buffer (50 mM Bis-Tris-HCl pH 7.0). Once the blue colour had ran two thirds of the way down the gel, the cathode buffer was replaced with clear cathode buffer (50 mM Tricine pH 7.0, 15 mM Bis-Tris-HCl pH 7.0). The gel was then allowed to run until all of the coomassie dye had run off the bottom of the gel. After electrophoresis the gel was equilibrated

for 15 min in clear cathode buffer containing 0.05% SDS before transfer to PVDF membrane.

2.8.3 Sample preparation for SDS-PAGE

Samples were prepared by mixing 1:1 in SDS sample loading buffer (125 mM Tris-HCl pH 6.8, 20% glycerol, 4% SDS, 0.02% bromophenol blue, 5% *ß*-mercaptoethanol) and were heated to 50 °C for 10 minutes before loading on to the gel.

2.9 Protein detection

2.9.1 Coomassie staining

After electrophoresis, the gel was placed in fixing solution (50% methanol, 10% acetic acid) for 30 mins, then in stain solution (10% acetic acid, 0.025% Coomassie G250) for at least 1hour, followed by de-stain (20% methanol, 7% acetic acid) until the protein bands became visible. The gels were subsequently dried for storage. Alternatively, the gel was stained with Instant Blue^{TM} Coomassie stain according to the manufacturer's instructions.

2.9.2 Silver staining

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

2.9.3 Transfer of proteins to PVDF membrane

Proteins were transferred from an acrylamide gel to PVDF membrane (Amersham Biosciences, UK) was achieved by a semi-dry western blotting system (Sigma, UK). Towbin transfer buffer was prepared (25 mM Tris, 192 mM glycine, and 20 % methanol) (Towbin *et al*., 1979). Briefly, two sheets of Whatman paper soaked in transfer buffer and the PVDF membrane soaked in methanol, following a paper, the membrane, the gel and a paper placed in order. Transfer was achieved as a constant current of ~ 200 mA for 90 mins or ~ 20 mA overnight.

2.9.4 Detection of proteins by immuno-blotting

After transfer, the PVDF membranes were immunoblotted in PBS-T containing 5% (w/v) dried skimmed milk for at least 1 hour. The membranes were washed 3 x 10 mins in PBS-T before incubation with 20 ml of PBS-T containing the primary antibody for 1 hour. After at least 3 x 10 minutes washes in PBS-T, the secondary antibody diluted in 20 ml of PBS-T was added for 1 hour. PVDF membranes to be immunoblotted with antibodies to the StrepII™ tag were blocked with PBS-T containing 3% BSA for at least 1 hour. The membranes were washed 3 x 10 minutes in PBS-T before incubation in PBS-T with 6 μg/mL avidin for 10 min. The Streptactin-horseradish peroxidise (HRP) conjugate antibody (IBA) was added directly to the avidin solution and incubated for 2 hours. Finally the membrane was washed for at least 6 x 10 minutes in PBS-T and immunoreactive bands were detected using ECL™ detection reagents according to manufacturer's instructions. X-ray films were developed in an AGFA Curix 60 automatic developer as directed.

2.10 Analysis of proteins

2.10.1 UV Absorption

Purified protein solutions were used for buffer exchange (Thermo Desalt Spin Column, UK) to get rid of salt and Imidazole. Then Samples were measured UV absorption from 240nm to 800nm in a Cary 100 bio UV-visible spectrophotometer according to the guideline of manufacturer.

2.10.2 Circular Dichroism (CD) spectroscopy

Purified proteins were dialyzed using 100 mM K_2HPO_4/KH_2PO_4 at 4°C overnight in order to fit the condition of Circular Dichroism spectrophotometry. Samples were performed using a 1 mm path-length quartz cuvette (Starna; Optiglass Ltd., Hainhault, UK) and measured in a Jasco J-815 spectrophotometer exactly as instructed by the manufacturer. Scans were taken across a wavelength range of 190-260 nm at a constant temperature of 20 °C. All spectra shown were averaged from 16 successive scans.

2.10.3 Mass spectroscopy

Samples were ran the SDS-PAGE and detected by Coomassie blue. The unknown proteins were cut from gel and digested by Trypsin. Samples were analysed using electrospray ionisation mass spectrometry (ESI-MS) on a Bruker MicrOTOF. The positive ionisation mode was used and detection was between 50 and 3000 m/z. Spectra were typically recorded for 2 minutes, averaged and deconvoluted to determine the mass of the main species.

2.10.4 Molybdenum quantification

20 ug of purified protein was used for analysis. Protein was diluted to 3 ml in 0.1 M trace metal grade nitric acid (prepared in house). Samples were analyzed for ^{95}Mo content using a 7500 series inductively coupled plasma mass spectrometer (Agilent Technologies, USA) equipped with a cross flow nebulizer and a quartz spray chamber. Calibration was achieved using external Molybdenum ICP-MS standards (Sigma, UK) and ^{166}Er as an internal standard (Agilent, USA). To ensure reproducibility each sample was measured in triplicate and results from two independent experiments compared. As a blank sample buffer was also measured and found to be free of any contaminating molybdenum.

2.11 Confocal microscopy

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

2.12 Specialist software

Chromas: Analysis of sequencing data.

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

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

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

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

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

Matrix Science: A powerful search engine featured as Mascot that uses mass spectrometry data to identify proteins from primary sequence databases.

http://www.matrixscience.com/

ProteinLynx Global Server v2.4: A fully integrated Mass-Informatics™ platform for quantitative and qualitative proteomics research.

Integr8 web: It provides easy access to integrated information about deciphered genomes and their corresponding proteomes.

http://www.ebi.ac.uk/integr8

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

Chapter 3

Translocation and proofreading of molybdopterin proteins by the Escherichia coli Tat pathway

3.1 Introduction

The Twin arginine translocation (Tat) pathway exports folded proteins across tightly sealed membranes in bacteria (Robinson and Bolhuis, 2004). So far, approximately 30 Tat substrates in *Escherichia coli* have been identified and most of them are redox enzymes functioning in anaerobic respiration, containing cofactors such as FeS, NiFe, copper and molybdopterin centres (Berks *et al.*, 2005; Tullman-Ercek *et al.*, 2007). In this case, the efficient translocation of proteins by the Tat pathway requires the insertion of cofactor and adoption of a folded conformation within the cytosol.

The question is how the Tat pathway determines whether a substrate is correctly folded with any cofactors in place. It has been addressed that the Tat system seems to have an inherent ability to preferentially transport folded proteins. Furthermore, DeLisa *et al* suggest that the bacterial Tat translocase may be able to sense regions of unfolded protein, much like a housekeeping chaperonin senses the exposed hydrophobic core of a misfolded protein, and reject any immature proteins attempting export— the chaperone-mediated proofreading is an alternative concept of 'quality control' (DeLisa *et al*., 2003). Several studies for chaperone-mediated proofreading activity have related to trimethylamine N-oxide reductase (TorA) and DMSO reductase (DmsA) and have shown that the chaperone molecules, TorD and DmsD, are required for efficient cofactor insertion into pre-TorA and pre- DmsA. These REMPS specifically prevent the futile export of apo-protein forms (Oresnik *et al.*, 2001; Sargent, 2007). More details of chaperone-mediated proofreading will be described in Chapter 5. Indeed, the inaccurate assembly of cofactor and improper folding of proteins partially affect or completely abolish translocation by the Tat pathway.

Recently, two *E. coli* Tat substrates, NrfC and NapG, have been investigated to test how the pathway senses their folded state. The data showed that the two substrates are completely blocked in export if FeS clusters are mutated, and mutant versions were rapidly degraded by the cells (Matos *et al*., 2008). In summary, previous studies have proposed a proofreading or quality control mechanism that can determine the folding state of proteins and avoid the export of rejected precursors in Tat pathway.

A large number of Tat substrates are molybdenum-containing enzymes which are ubiquitous in many bacteria where they are normally involved in a wide range of redox reactions including those essential for the global carbon, nitrogen, and sulfur cycles. In general, molybdenum in enzymes is found coordinated to the dithioline group and made up the most simple cofactor molybdopterin (MPT). Additionally, the majority of well-known periplasmic bacterial molybdoenzymes commonly utilize molybdopterin guanine dinucleotide (MGD) or *bis* MGD as cofactor which is the variant form of MPT modified by attachment of GMP (Berks *et al.*, 2003). The Figure 3.1 depicts the structure of the molybdenum-cofactors MPT and MGD.

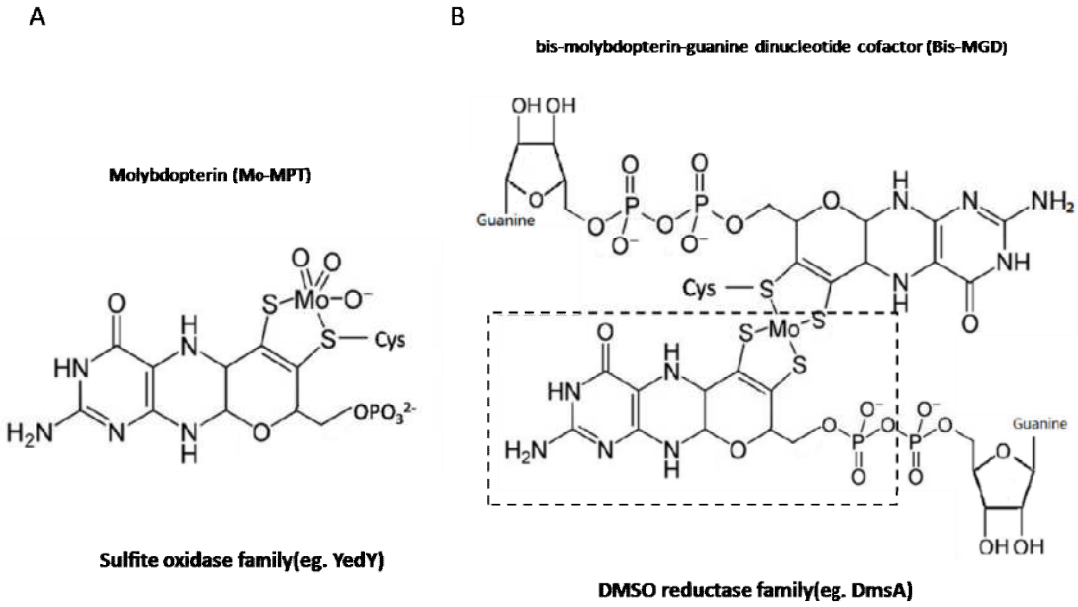


Figure 3.1 Structure of the pterin cofactor (A) Proposed structure of the molybdopterin cofactor (Mo-MPT) as present in Sulfite oxidase family (B) Structure of the molybdopterin guanosine dinucleotide cofactor (MGD) as present in DMSO reductase family.

To further study the quality control system in Tat pathway, molybdoproteins in *E. coli* were analysed. For instance, TorA is a water-soluble periplasmic protein that binds MGD as the single cofactor (Mejean *et al.*, 1994) and is synthesized with a twin-arginine signal peptide. Also, nitrate reductase (NapA) and DMSO reductase (DmsA) bind MGD as well as FeS cluster as cofactors and export by Tat pathway (Kisker *et al*., 1997; Papish *et al*., 2003). Importantly, the predicted Tat substrate, YedY, is a sulfite oxidase that binds molybdopterin (MPT) as a single cofactor, and it represents the first structural characterization of this form of cofactor in prokaryotes (Loshi *et al.*, 2004).

In this chapter, four substrates in *E. coli* were examined for translocation by Tat pathway and YedY was chosen as the best candidate to be characterized for proofreading.

3.2 Results

3.2.1 Translocation of four molybdopterin proteins by the Tat pathway

NrfC and NapG are Tat substrates containing four separate [4Fe-4S] centres (reviewed by Cole, 1996). The study of the proofreading in the Tat pathway using NrfC and NapG has demonstrated the completely block in export if one of four predicted FeS centres were mutated and the rejected molecules were disposed (Matos *et al.*, 2008).

Since the effective proofreading of FeS substrates by the Tat pathway has been revealed, it is of interest to further investigate whether this proofreading system equally applies to molybdopterin substrates, or whether there is another novel process to monitor the proofreading.

Molybdopterin proteins represent a large proportion of Tat substrates. Some of them have been well studied such as TorA, DmsA and NapA which are ~90 kDa proteins in *E. coli*. They contain the molybdopterin centre MGD that inserts into the protein in the cytoplasm and they are believed to be exported to the periplasm in a correctly folded form. Moreover, the predicted Tat substrate, YedY, is a more novel molybdopterin protein and a potential candidate of interest in my study because it only contains MPT (molybdopterin) as the cofactor which is the simplest form of molybdopterin centre (Sargent *et al.*, 1998; Weiner *et al.*, 1998). Table 3.1 briefly summarizes four substrates.

Substrate	Cofactor	Type
TorA	MGD	perplasmic
NapA	MGD, Fe-S cluster	perplasmic
DmsA	MGD, Fe-S cluster	membrane
YedY	MPT	perplasmic

MGD: molybdopterin guanine dinucleotide:
MPT: molydopterin

Table 3.1 List of four known or predicted *E. coli* Tat substrates including names, cofactor(s) inserted and location to be functioned.

At the beginning, a Sec substrate LolA was used as a negative control to confirm the proper functioning of the Tat pathway in MC4100 and Δ*tatABCDE* strains. LolA was amplified appending C-terminal his-tag into plasmid pBAD24, and expressed over 5-h (after addition of arabinose to induce synthesis) to check the expression of LolA during time. Additionally, LolA was over-expressed in wild-type MC4100 and Δ*tatABCDE* cells which were then fractionated into cytoplasm (C), membrane (M) and periplasm (P) compartments. Protein samples were analysed by SDS-PAGE gel and subsequent immunoblotting. Figure 3.2.1 show LolA is detected after 3-h in MC4100 and Δ*tatABCDE* cells and stable throughout the time period. Also, LolA is exported to the periplasm in two cells. The data indicate the Sec pathway in MC4100 cells works well.

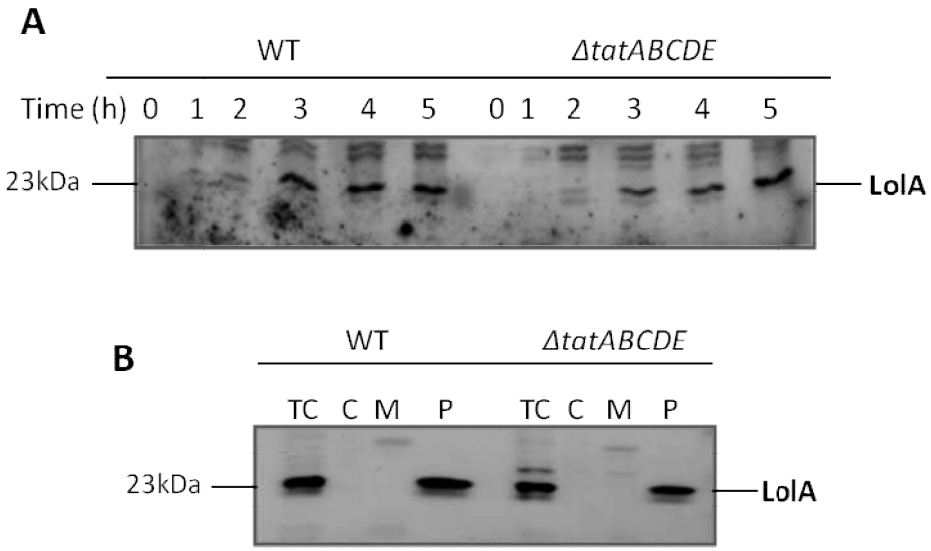


Figure 3.2.1 Translocation of Sec substrate LolA in wild-type MC4100 and Δ*tatABCDE* cells LolA (from plasmid pBAD24) was expressed in *E. coli* wild-type MC4100 cells (WT) and Δ*tatABCDE* cells. (A) Samples were taken at the indicated times after induction of synthesis with arabinose and immunoblotted by anti-his antobody. (B) Cells were taken after 2.5 h induction, total cell contents (TC) were collected and fractionated into membrane, cytoplasm, and periplasm (M, C, P). The location of LolA was detected by immunoblotting with anti-his antibody.

Next, to test the efficiency of translocation of four substrates by the Tat pathway, TorA, DmsA, NapA and YedY were cloned and amplified with C-terminal his-tags into the arabinose inducible plasmid pBAD24 and over-expressed in *E. coli* wild-type MC4100 and Δ*tatABCDE* cells. Total cell extracts (TC) were collected and cells were separated into cytoplasm (C), the membrane (M) and the periplasm (P) compartments. Each sample was measured to the same amount and run on an SDS-PAGE gel to analyse by immunoblotting with anti-his antibody. The result in the top panel of Figure 3.2.2 (A) indicates that some TorA is exported to the periplasm when expressed in wild-type MC4100 cells (WT), very little precursor is detected in the cytoplasm, and a lot of the precursor accumulates in the membrane. The second panel in Figure 3.2.2 (A) shows that that NapA is translocated to the periplasm in MC4100 and with some protein remaining in the cytoplasm. Importantly, YedY is efficiently exported and processed to the mature size in the periplasm, with some of precursor present in the membrane in MC4100 cells. DmsA is only detected in the membrane in wild-type cells because it is thought to be a membrane-anchored protein with the active site on the perplasmic face (The data in Figure 3.2.2 (B) shows DmsA is degraded by proteinase K in MC4100 but protected by membrane in Δ*tatABCDE* cells). The complete absence of export in the *tat* mutant cells confirms that TorA, DmsA, NapA and YedY are Tat substrates.

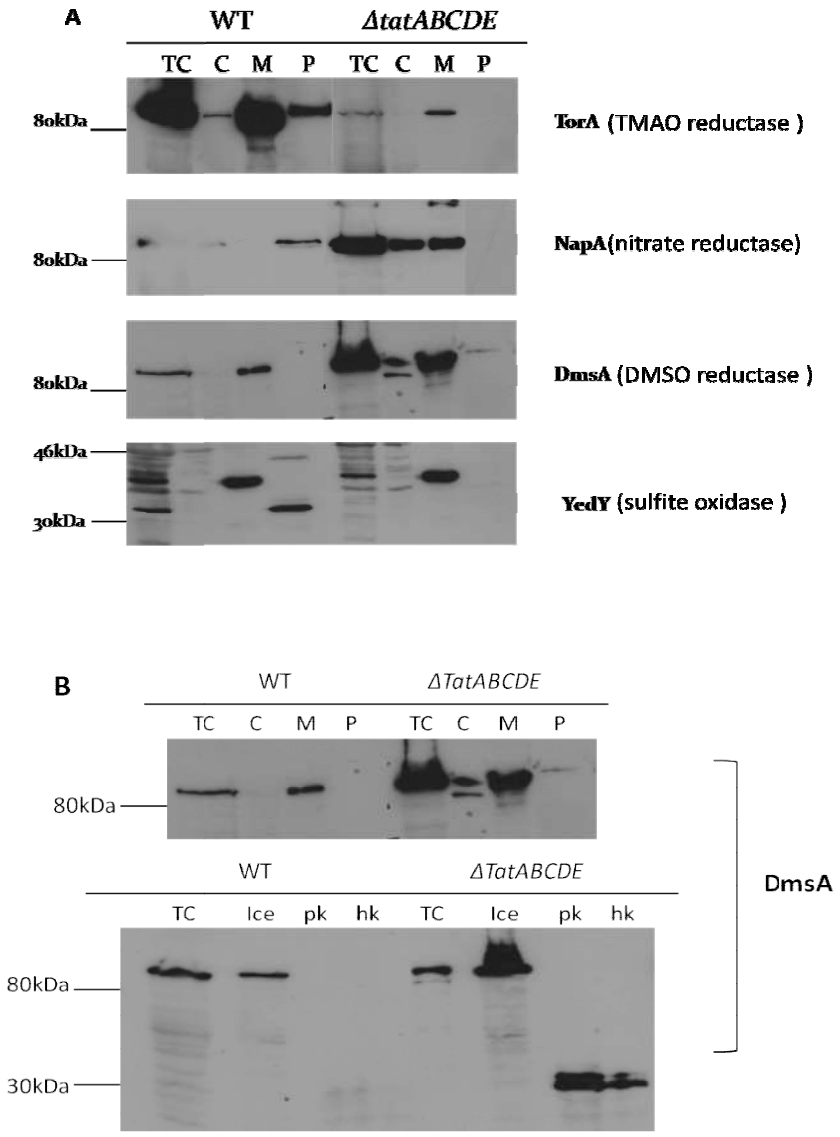


Figure 3.2.2 Translocation of Four molybdopterin proteins TorA, NapA, DmsA and YedY by the Tat pathway TorA, NapA, DmsA and YedY (from plasmid pBAD24) were expressed in *E. coli* wild-type MC4100 cells (WT) and Δ*tatABCDE* cells respectively. (A) Cells were taken after 3 h induction (3 h is the optimized time point for efficient export of proteins by the Tat pathway on basis of time course study) and total cell contents (TC) with separated membrane, cytoplasm, and periplasm (M, C, P) compartments were analysed by immunoblotting using anti-his antibody. (B) Cells producing DmsA were harvested for 3 h and spheroplasts extracted prior to proteolysis by proteinase K. Total cell contents (TC) with spheroplast untreated (ice), treated by proteinase K (pk) or treated by 5×freeze-thaw cycles following proteinase K (hk) were analysed by immunoblotting using anti-his antibody.

Among the four molybdopterin substrates exported by the Tat pathway in *E. coli*, YedY is the best candidate to study the proofreading process in Tat system. The reason includes three aspects. Firstly, YedY is shown to be a Tat substrate for the first time. Secondly, YedY is effectively exported and processed from a precursor of 37 kDa to the mature form of 32 kDa by Tat pathway. Lastly, YedY only binds the cofactor MPT which is the simplest form of molybdopterin centre and it is convenient to alter the insertion of cofactor in the study.

3.2.2 Translocation and proofreading of YedY MPT mutants by the Tat pathway

YedY, the catalytic subunit of the heterodimer oxidoreductase YedYZ from *E. coli*, is a soluble periplasmic protein which has a twin-arginine signal peptide and is exported by the Tat translocation system (Sargent *et al*, 1998; Weiner *et al*, 1998). *yedY* encodes a pre-protein with a molecular weight of 37 kDa, and it leads to a mature protein of 32 kDa after cleavage of signal peptide. YedY binds molybdopterin (MPT) as a single cofactor, and it represents the first structural characterization of this form of cofactor in prokaryotes (Loschi *et al.*, 2004). According to study of overall architecture of YedY, the active site of YedY has specific characters. In YedY, the molybdenum ion is coordinated by three sulphur ligands with one contributed by Cys102. Moreover, the pterin within MPT interacts with protein by seven direct hydrogen bonds including Try47, Glu48, Thr137, Lys207, Gly205, Arg194, and importantly, Arg194 also forms hydrogen bonds with the terminal phosphate group (Loschi *et al.*, 2004). These ligands conjugated on single molybdopterin are believed to be highly important to insert cofactor into protein and keep the correct folding state. Figure 3.2.3 is a schematic representation of the interaction between Moco and YedY which adopted from Loschi *et al.*, 2004.

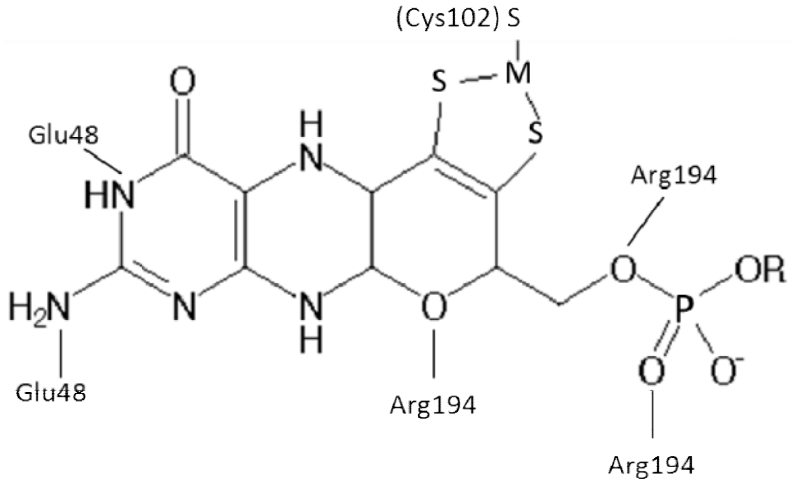


Figure 3.2.3 Schematic representation of the interactions between molybdenum-cofactor and YedY The interactions that anchor the molybdopterin cofactor tightly to YedY are highly important to the insertion of cofactor and correct folding of YedY. The picture shows one of three sulfur ligands, Cys102 (numbering refers to the processed YedY, starting at Asp-1), which is coordinated to the molybdenum ion. Also, Glu48 and Arg194 are essential ligands to interact with the pterin by direct hydrogen bonds (Adopted from Loschi *et al.*, 2004).

The aim of this study was to analyse the relationship between molybdopterin assembly and protein export in YedY. Three mutated versions of pre-YedY were designed to affect the insertion of MPT resulting in addressing the question of proofreading of Tat pathway. In the three mutations, the 102 Cys residue was substituted by Ala, the 48 Glu and the 194 Arg residues were substituted by Asp and Lys, respectively. Expression of these mutants used the arabinose-inducible pBAD24 vector, with a C-terminal his-tag appended to aid the identification of the proteins. They were expressed in wild-type MC4100 cells and a *tatABCDE* strain, and the cells were subsequently separated into cytoplasm, periplasm and membrane fractions (C, P, M) as shown in Figure 3.2.4. The top panel shows the data of non-mutated YedY which has been described above. The results with the three mutants are shown in the remaining panels. All of the substitutions have a dramatic inhibitory effect on export in wild-type MC4100 cells, with no periplasmic mature YedY apparent. Thus, the data provide strong evidence that a proofreading system and/or the Tat pathway effectively evaluates YedY's folding and assembly state and completely blocks the export of the misfolded/misassembled forms.

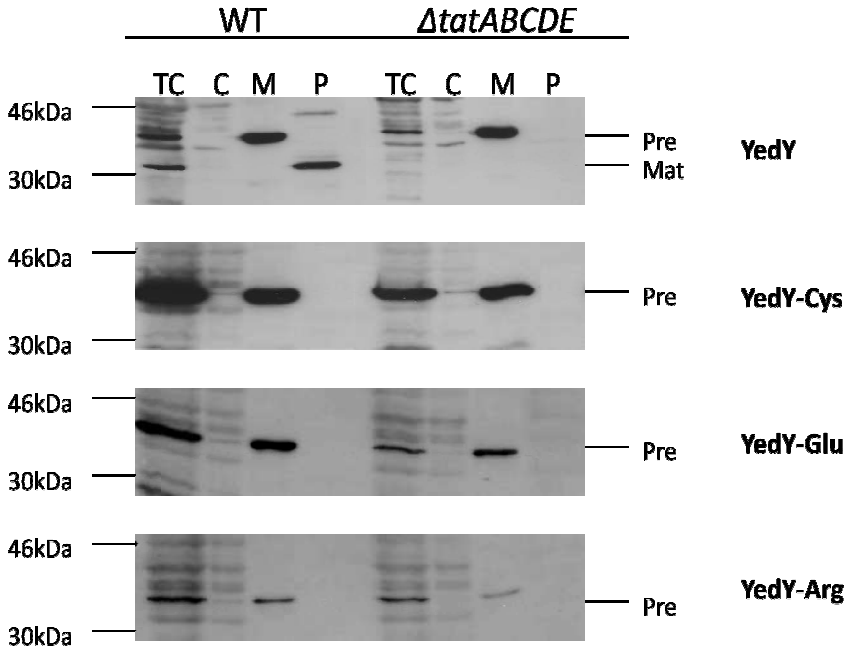


Figure 3.2.4 Translocation of YedY and three MPT mutants by the Tat pathway YedY is expressed in *E. coli* wild-type MC4100 cells (WT) or Δ*tatABCDE* cells. After induction with arabinose for 2.5 h, cells are harvested and total cell contents (TC) are analysed together with cytoplasm, membrane and periplasm samples (C, M, P) after osmotic shock and fractionation. Precursor and mature forms of YedY are indicated (Pre, Mat.). Mutated versions of YedY (Cys, Glu and Arg) are analysed in an identical manner (as indicated). The immunoblots detected by anti-his antibody show precursor and mature forms of proteins.

3.2.3 YedY molybdopterin (MPT) mutants are degraded at varying rates in wild-type cells but fully stable in Δ*tatABCDE* strain

The data shown above in Figure 3.2.4 were obtained from cells in which YedY expression had been induced with arabinose for 2.5 h. In order to analyse the mutants' stabilities in more detail, we conducted time-course studies in which samples were taken over a longer period; the data are shown in Figure 3.2.5.

The non-mutated YedY was expressed in wild-type and Δ*tatABCDE* cells over a 5-h period after addition of arabinose to induce synthesis. In wild-type cells, the precursor protein is first detected after 2 h and mature YedY is also detected (at low levels) after 2 h. After this point, the mature protein gradually becomes more prominent as exported proceeds. The translocation of YedY is a relatively slow process which presumably reflects the time is necessary to assemble and insert the MPT before export. Interestingly, the expression of the three mutated variants peaks at 2-h in wild-type cells, and they show differing stabilities over the time. Cys102 is relatively stable over the entire time-course and the overall accumulation of protein is similar to that of native YedY, with the exception that no export and processing occur. In contrast, the Glu48 mutant is much less stable, and is effectively absent after about 4 h. The Arg194 mutant is even less stable and is only apparent at the 2 h time point; after this, it is rapidly degraded. However, YedY and three variants all accumulate stably throughout the induction period when expressed in the *ΔtatABCDE* background. These data suggest that the Tat system not only senses the folding state of proteins, but may be involved in the degradation of mutated forms indirectly.

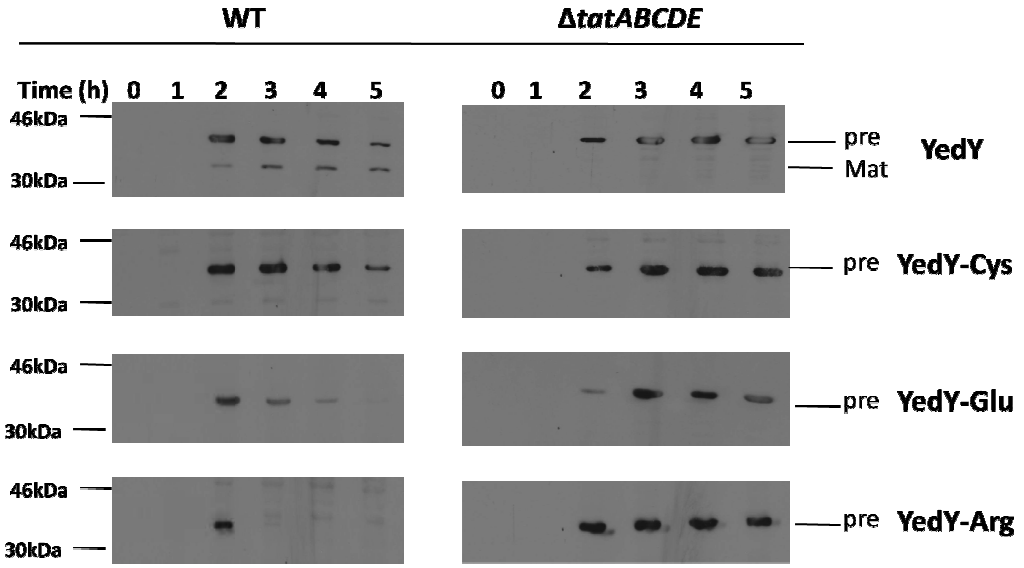


Figure 3.2.5 Translocation of YedY and molybdopterin(MPT) mutants by the Tat pathway YedY and mutants YedY-Cys, YedY-Glu, YedY-Arg (from plasmid pBAD24) was expressed in *E. coli* wild-type MC4100 cells (WT) and Δ*tatABCDE* cells. Cells are analysed at times (in hours) after induction of synthesis with arabinose. The immunoblots are detected by anti-his antibody, with precursor and mature forms of the protein indicated (Pre, Mat).

3.2.4 Translocation of YedY and three molybdopterin (MPT) mutants in a range of *tat* deleted strains

The question is what machineries in the cells are essential for the sensing and degradation of YedY mutants? If the Tat pathway is involved in the turnover, which components of the Tat apparatus are concerned? In *E. coli*, TatA,TatB, TatC consist of the minimal unit to carry out the translocation by Tat pathway. It is proposed that TatBC recognises substrates which are then exported through a channel which is produced by different numbers of TatA protomers. TatE, a TatA paralogue, which is expressed at much lower levels than TatA, and sometimes can partially complement the function of TatA when *tatA* is deleted. But most studies indicate that Δ*tatA* cells exhibit a much more severe export defect than do Δ*tatE* cells. Also, a complete block in translocation is observed only when both *tatA* and *tatE* are disrupted, and disruption of either *tatB* or *tatC* also results in a complete block in translocation (Sargent *et al.*, 1998, 1999).

We expressed YedY and three variants in a range of *tat* mutant strains and the data are shown in Figure 3.2.6. YedY is blocked in export in *ΔtatA, ΔtatB, ΔtatC, ΔtatA/E* cells, but still translocated and processed to the mature form in *ΔtatD, ΔtatE* cells. Three mutant variants Cys102, Glu48, Arg194 are prevented in export in all cells. The results display that Cys102 is relatively stable in all strains. Additionally, Glu48 levels seem to slightly decrease in the Δ*tatE* strain, but are stable in others which is completely different to the expression in wild-type cells. Arg194 is rapidly degraded in Δ*tatA,* Δ*tatE*, Δ*tatA/E* cells which is consistent with the observation in wild-type cells. Compared YedY and three variants expressed in wild-type cells with those in *ΔtatD* cells, the result suggest that proteins are more stable in *ΔtatD* cells.

Surprisingly, there is little detection of the proteins in Δ*tatB* cells, and the mutants are not stable in *ΔtatC* cells. It is suggested that TatBC is the binding site for proteins targeted to Tat translocon. TatB is proposed to be the first association with newly synthesized proteins and deliver to TatC to binding. The data suggest that if proteins fail to initiate binding and translocation, they will be subjected to proteolysis in the cytosol.

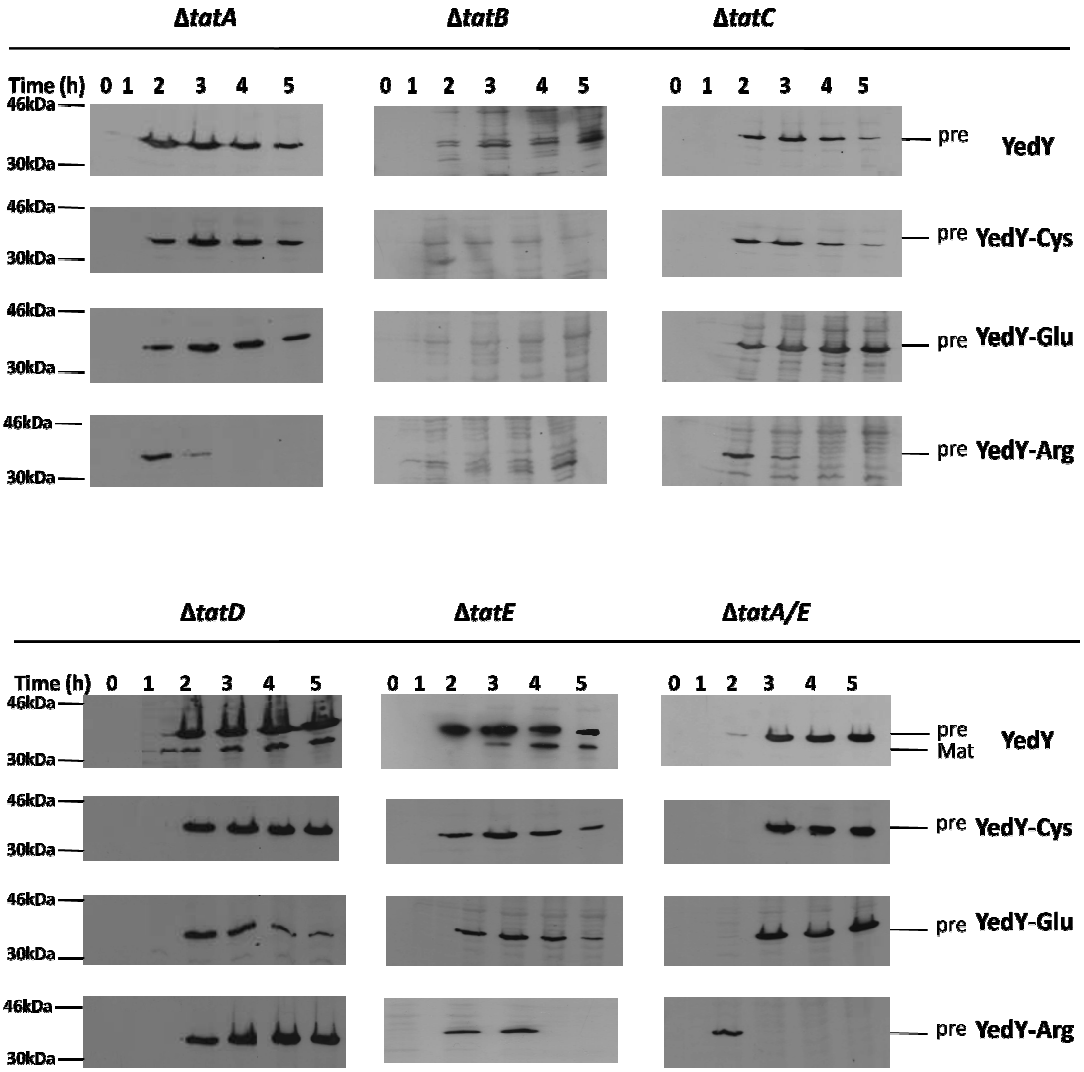


Figure 3.2.6 Translocation of YedY and molybdopterin (MPT) mutants in a range of *tat* deleted strains. YedY and molybdopterin (MPT) mutants YedY-Cys, YedY-Glu, YedY-Arg (from plasmid pBAD24) was expressed in a range of *tat* deleted strains including *ΔtatA, ΔtatB, ΔtatC, ΔtatD, ΔtatE* and *ΔtatA/E* strains. Each sample expressed in the same strain but collected at every time points (in hours) are measured as the same amount cells according to OD_{600}. Cells are analysed after induction of synthesis with arabinose. The immunoblots are shown in the insets and detected by anti-his antibody, with precursor and mature forms of the protein indicated (Pre, Mat).

Thus, there is a clear implication that TatA/TatE/TatAE perhaps plays a role in degradation of rejected substrates. Since it is different when proteins are amplified into the IPTG inducible vector pEXT22 (results are not shown here), we postulate that YedY molybdopterin (MPT) mutants are proofread and degraded by a proteolysis system in which the Tat pathway may be involved We thereafter studied the chaperone-mediated proofreading referring to the redox enzyme maturation proteins (REMPs) which have a function to monitor the assembly of cofactor, ensure the correctly folding of protein, and initiate the proteolysis for rejected molecules.

3.3 Discussion

In this chapter I studied the translocation and proofreading of the molybdopterin substrate YedY by the Tat pathway in *E. coli*. The data reveal that the YedY molybdopterin (MPT) mutants are not only completely blocked to export in the wild-type cells, but also degraded at varying rates by the unknown proteolysis system which the Tat pathway may be involved in.

Proofreading is an important aspect of Tat functioning because it prevents the wasteful exporting of substrates until all folding and assembly processes have been completed before transportation which is the point-of-no-return. Studies on a variety of heterologous proteins have suggested that the Tat export system has an inbuilt ability to preferentially export correctly folded proteins (Sanders *et al*., 2001). In one case, a heterologous protein, PhoA, was only transported if the native disulphide bonds had been formed to generate the correctly folded molecule (DeLisa *et al.*, 2003).

However, the pathway appears to use additional quality control system for at least some of the more complex, cofactor-containing substrates. Studies on molybdoproteins such as TorA and DmsA provide evidence to indicate that soluble, substrate-specific chaperones, or REMPs, have an important function to monitor the cofactor-insertion and export processes (reviewed by Sargent, 2007). Recently, the study of FeS substrates implicated that proofreading in the Tat pathway involves much more than monitoring the status of FeS cluster assembly within the protein folding process (Matos *et al*., 2008).

In *E. coli*, most Tat substrates bind redox cofactors including FeS, NiFe and molybdopterin centres (Berks, 1996) and molybdopterin proteins make up a large part. A major aim of the work was to characterise the proofreading of molybdoproteins by the Tat pathway, and elucidate the novel characters in this proofreading system.

YedY is the only Tat substrate which binds molybdopterin (MPT) as a single cofactor in *E. coli*, and I have studied its overall export process. Here, the data show

that substitution of any ligands coordinated with MPT is sufficient to completely block export of YedY, and these findings are clear evidence for the operation of a quality control system that operates during YedY export. Moreover, the study not only revealed the proofreading of the status of MPT assembly, but also confirmed the rapid degradation of MPT-deficient molecules by the cell. Although the degradation of mutant Glu48 is nearly abolished in some of the *tat* deleted strains, which implies the Tat pathway is involved in the turnover of the mutant, and mutant Arg194 is also degraded rapidly in *tatA/tatE/tatAE* deleted strains. It is therefore possible, but not certain, that the Tat pathway is directly involved in the degradation of the mutant YedY proteins.

It is common in biological systems that rejected substrates are degraded but highly unusual for a Tat translocation pathway to trigger the proteolysis if the observations are correct. Some studies have suggested that the Tat apparatus is not involved in the degradation of substrates. For examples, a few Tat substrate (whether native or misfolded) are destabilised, rather than stablilised, when transport is prevented by the absence of translocase components or mutagenesis of the signal peptide (Santini *et al.*, 2001; Barrett *et al.*, 2003; DeLisa *et al.*, 2003; Fisher *et al.*, 2006). However, further studies are required to determine whether the Tat pathway involves in the proteolysis indeed, whether other factors are necessary to perform the proofreading process.

To address these questions, one possibility is that 'proofreading' factors such as Tat-associated chaperones (also termed REMPs) remain associated with the substrate and prevent its correct targeting to the TatABC complex. Thereafter, the translocation event may actually serve to partially unfold the substrate and render it protease sensitive. Alternatively, the substrate has interacted with translocon in a transport-incompetent state and the Tat system may then stimulate the proteolysis machineries to degrade misfolded proteins.

Chapter 4

Purification and structural study of YedY and three mutations affecting cofactor insertion

4.1 Introduction

YedY, the soluble periplasmic catalytic subunit of the heterodimer oxidoreductase YedYZ, contains a RR motif in the signal peptide for targeting by the twin-arginine translocation (Tat) system (Sargent *et al.*, 1998, Weiner *et al.*, 1998). YedY is the only molybdoenzyme isolated from *E. coli* with the molybdopterin-type cofactor which is not conjugated by an additional nucleotide.

Several years ago, YedY from *E. coli* was purified to determine the crystal structure at 2.5-Å resolution which is shown in Figure 4.1. According to the study using MAD phasing from incorporated selenomethionine, YedY is a spherical molecule of mixed $\alpha + \beta$ structure with overall dimensions $50 \times 45 \times 40$ Å, and the overall fold of YedY consists of 10 β-strands, organized into two β-sheets, and 12 α-helices (Loschi *et al.*, 2004). The data in PDB (codes *1XDQ*) reveals that it has five chains in YedY, four of them are topologically aligned in a similar manner but one is sequence-unique. Generally, the Moco containing region in five chains lies on the same face, and five conserved hydrophobic regions form an extended hydrophobic surface that could play a role in membrane association with the redox partner YedZ.

The previous study also shows the active site of YedY possesses several novel features. In the first place, the molybdenum atom in the single molybdopterin is localized ~16 Å from the enzyme surface. Additionally, the molybdenum cofactor forms numerous hydrogen bonds with main chain and side chain atoms in YedY, which are strictly conserved across the family of bacterial YedY-related proteins. The pterin interacts with the protein by seven direct hydrogen bonds involving Tyr-47, Glu-48, Thr-137, Lys-207, Gly-205, Arg-194, and the terminal phosphate group is stabilized by six additional hydrogen bonds formed by Arg-194, Asn-189, Lys-207, and Asn-44. Furthermore, the molybdenum ion is coordinated by three sulfur ligands, two contributes from the dithiolene sulfurs of the molybdopterin with a Mo-S distance of 2.4 Å, and the third sulphur ligand is Sγ of Cys-102 at a distance of 2.4 Å (Loschi *et al.*, 2004).

To study the proofreading of the Tat pathway in chapter 3, three residues coordinated to molybdenum atom were substituted to affect the molybdopterin (MPT) assembly

during YedY folding process. Two mutations are important ligand Glu-48 and Arg-194 that interact with pterin, and one is the sulphur ligand Cys-102. Although the results indicated that the Tat pathway recognised the misfolded proteins and refused the export, it is still unknown that how the Tat pathway recognizes Cys/Glu/Arg mutants as incorrectly folded. It was thus of interest to determine how these mutations had affected the overall structure of YedY, and this would indicate how sensitive the Tat machinery is in terms of recognizing structural defects. Here, I purified the YedY and three variants to analyse the secondary structure and measure whether the cofactor had been inserted.

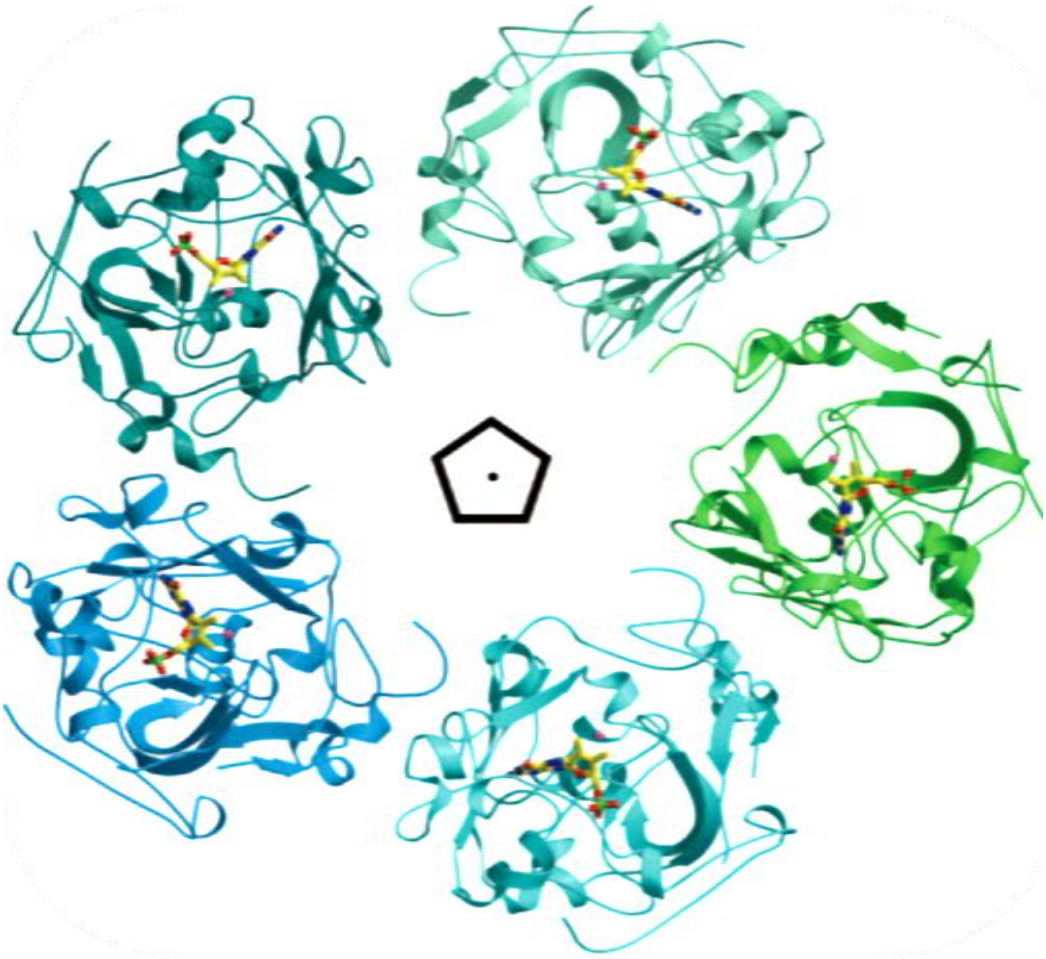


Figure 4.1 Overall architecture of YedY in *E. coli* (PDB code 1XDQ) Ribbon representation of the YedY pentamer. The five monomers are color-coded differentially. The 5-fold axis perpendicular to the plane is indicated. The molybdenum cofactors are in *ball* and *stick* representation with the molybdenum ions in *pink.* The figure was adapted from Loschi *et al.*, 2004

4.2 Results

4.2.1 Optimal growth temperature to express YedY in *E. coli*

E. coli can maintain balanced growth in rich medium over the temperature range from 10 to 49°C, but its metabolic state only can stabilize in a maximize growth rate between 20 and 37°C (Marr *et al.*, 1964). It is suggested that the synthesis and activity of proteins are affected by the inability of the cell which can not compensate for thermally (hot or cold) induced changes of reactions above 37°C (Patterson and Gillespie, 1972) or below 20°C (Broeze *et al*., 1978; Herendeen *et al*., 1979).

Here, in order to examine the optimal growth temperature to express YedY in *E. coli*, I tested the expression level of YedY at 30 °C and 37 °C. YedY was amplified with C-terminal his-tag into arabinose inducible pBAD24. In Figure 4.2.1, YedY was expressed in *E. coli* wild-type MC4100 and Δ*tatABCDE* cells at 30 °C and 37 °C. Samples were collected in terms of the same amount of cells measured by OD_{600} and the same amount of protein measured by BCA assay at two temperatures, and detected by immunoblot. The results have shown that pre-YedY was not detected at 30 °C in the wild-type cells, and the expression level of pre-YedY at 37 °C seems to be better than 30 °C in Δ*tatABCDE* cells. Thus, 37 °C is the optimal growth temperature to express YedY in both wild-type and Δ*tatABCDE* cells.

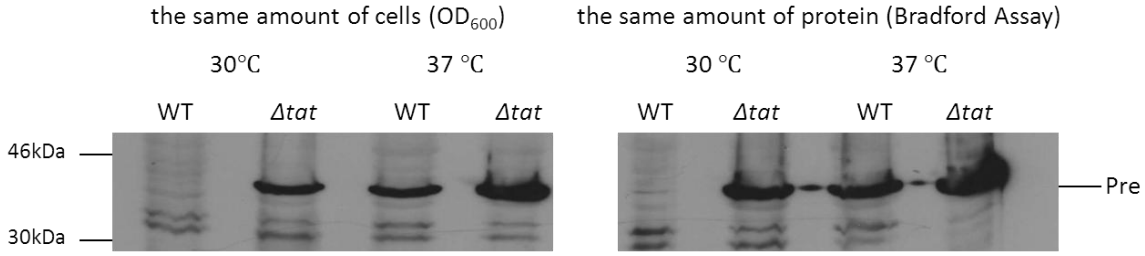


Figure 4.2.1 Optimal growth temperature to express YedY in *E. coli* YedY was expressed in *E. coli* wild-type MC4100 cells (WT) and Δ*tatABCDE* cells (Δ*tat*). Cells were grown at 30 °C and 37 °C respectively. Samples are collected by the same amount of cells determined by OD_{600} and by the same amount of proteins determined by Bradford assay. The expression level of YedY in both temperatures is detected by immunoblotting using anti-his antibody, with precursor as indicated (Pre).

4.2.2 Purification of pre-YedY and three cofactor mutants in *E. coli*

In *E. coli*, YedY has been isolated to identify and characterize the structure nearly 7 years ago. In that time, the mature form of YedY was firstly purified from periplasm by size exclusion chromatography and the crystal structure determined at 2.5-Å resolution (Lodovica *et al*., 2004). In order to elucidate how the Tat system might proofread the assembly of MPT and folding of pre-YedY in the cytosol, affinity chromatography was used to purify the proteins and analyze the difference in structure between pre-YedY and three mutants by circular dichroism as well as inductively coupled plasma mass spectrometry (ICP-MS).

In general, all purification studies were performed using proteins with a C-terminal 6×his-tag on YedY and three mutations (Cys102, Glu48, Arg194). *E. coli ΔtatABCDE* cells expressing YedY and three mutants from plasmid pBAD24 were fractionated and the proteins purified from the cytoplasmic fraction.

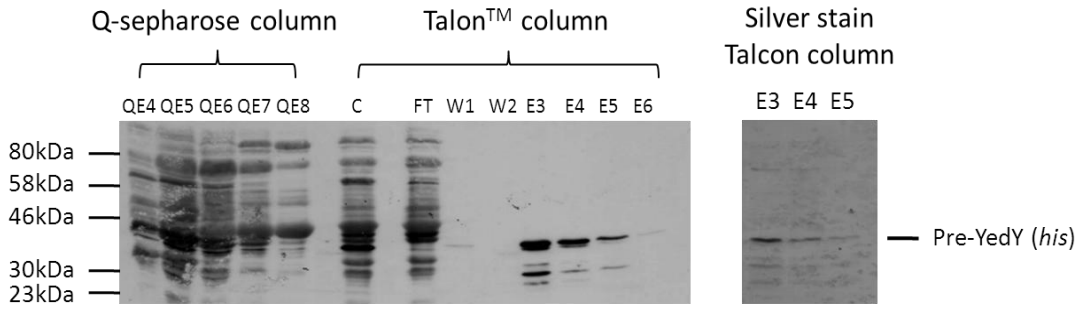
4.2.2.1 Two-step purification including anion exchange chromatography and affinity chromatography

The cytoplasmic fraction was firstly applied to an equilibrated 4 ml Q-Sepharose anion exchange column run with the optimal salt concentration to partially separate YedY from other proteins. The column was washed with 2 x 20 ml of buffer and protein subsequently eluted with 8 x 4 ml of buffer supplemented with 100-500 mM sodium chloride (NaCl). All eluted fractions were resolved on an SDS-PAGE gel and analyzed by the immunoblot with anti-*his* antibody (Figure 4.2.2). The result showed that Elution 6 was the optimal concentration of salt which contained 300 mM NaCl to isolate YedY.

Thereafter, this fraction was loaded to an equilibrated 2 ml Talon^{TM} affinity column that further purified his-tagged YedY. To remove unbound proteins, the column was washed using 2 x 20 ml of equilibration buffer. Tightly bound protein was eluted in equilibration buffer supplemented with 150-1000 mM imidazole in 8 x 2 ml fractions. All affinity column fractions were analysed by SDS-PAGE and detected by immunoblot using anti-his antibody as well as silver stain.The results presented in

Figure 4.2.2 demonstrated that most of the protein did not bound to the column and was detectable in the flow-through (FT). The majority of YedY can be eluted from the column over fractions 3-5 with a peak in fraction 3. However, it showed that a few small contaminants co-elute with YedY.

Pre-YedY (pBAD24) in the cytoplasm



1L culture, Elution 6 using 300mM NaCl from Q-Sepharose column (QE6) was chosen to process the Talon affinity chromatography. 2.5ml Talon column, Each elution=1ml, E1-E6 ~ 150mM – 1M Imidazole		
A_{595}	Elution	Concentration(ug/ml)
0.045	E3	90
0.033	E4	66
0.025	E5	50
0.011	E6	22
0.006	E7	12
-	total	240

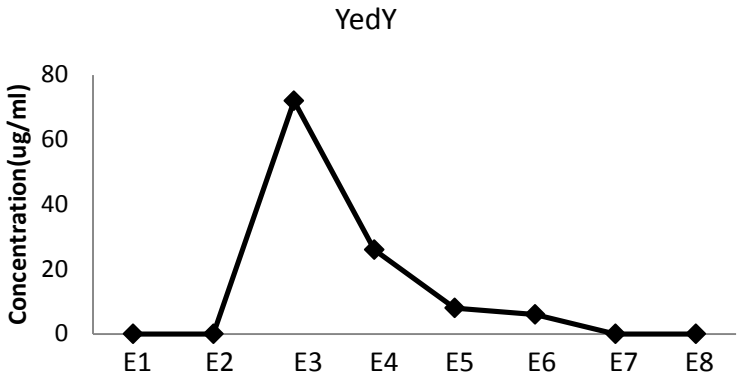


Figure 4.2.2 Two-step purification of pre-YedY YedY was amplified with C-terminal his-tag into arabinose inducible pBAD24. It started the 1L culture to over-express YedY in Δ*tatABCDE* cells. *E. coli* cytoplasmic fraction (C) was isolated from cells and firstly applied to Q-sepharose anion exchange column, and chose a half-purified elution (QE) to load into Talon affinity column. Thereafter, unbound proteins were washed a couple of times (W) and pre-YedY was eluted by elution buffer. Elution fractions from Talon column (E) were measured by Bradford assay. The collection of all samples were analysed by silver stain as well as immunoblot with anti-his antibody.

4.2.2.2 Purification with Talon™ affinity chromatography

With affinity chromatography, many factors influence the quantity and quality of purification, such as the scale of crude extract including expressed proteins, the optimal environment for isolating and stabilizing the objective protein (salt concentration, pH *etc*). With the ion exchange chromatography, 300 mM NaCl was used in all solutions during the purification. Assays to test optimal pH and whether glycerol is helpful were performed as described below.

Briefly, the cytoplasmic fraction isolated from *E. coli ΔtatABCDE* cells was applied to an equilibrated 2.5 ml Talon™ affinity column that specifically bound to *his*-tagged YedY. The column was washed using equilibration buffer to remove unbound protein, and eluted with equilibration buffer supplemented with imidazole. Protein presented in column fractions was resolved on SDS-PAGE, and the gel was immunoblotted with anti-*his* antibody or stained by silver stain.

Figure 4.2.3 shows the purification performed at pH 8.0. The anti-his immunoblot has shown that not all YedY bound to the column and protein was detected in the flow-through (FT) as well as wash fractions. $YedY_{his}$ that did bind specifically eluted from the column with a clear peak in the elution 4 and 5. However, some small proteolytic clipping of the protein was also found in two elutions. In terms of silver stain of eluted fractions, higher molecular weight bands were present which were absent in the immunoblot implying the presence of unspecific proteins co-eluted from the column. In short, the environment of pH 8.0 is not suitable to purify and stabilize YedY using a Talon affinity column.

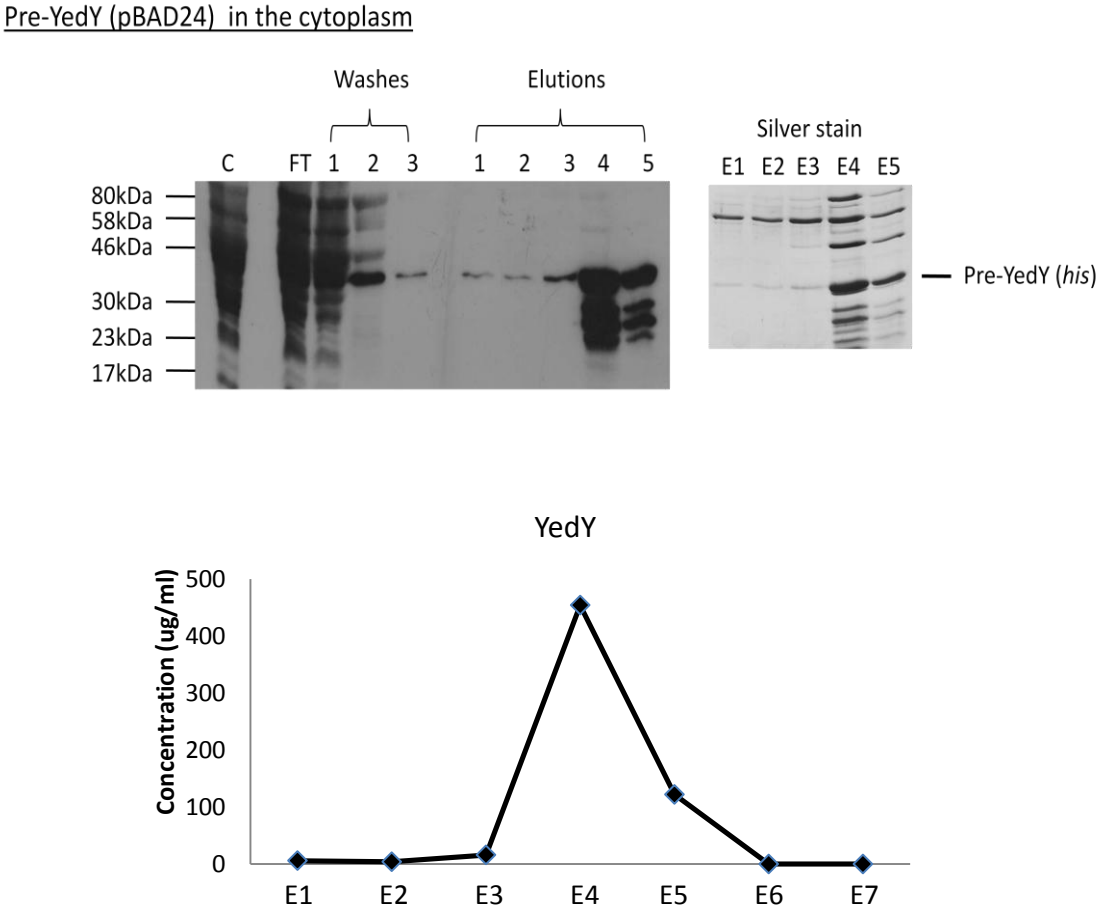


Figure 4.2.3 Talon affinity chromatography of pre-YedY with pH 8.0 YedY was amplified with C-terminal his-tag into arabinose inducible pBAD24. It started the 3L culture to over-express YedY in Δ*tatABCDE* cells. *E. coli* cytoplasmic fraction (C) was isolated from cells and directly applied to Talon affinity column. The flow-through (FT) sample collected and unbound proteins were washed (W). Bound proteins were eluted by elution buffer. All elution fractions (E) were measured by Bradford assay and analysed by silver stain as well as immunoblot with anti-his antibody.

The purification carried out with solutions at pH 5.0 is shown in Figure 4.2.4. The immunoblot shows that some YedY was detected in the flow-through (FT) indicating protein that did not bind to the column effectively. $YedY_{his}$ bound was then specifically eluted from the column and seemed to be very pure in the immunoblot. However, there is still some contaminations when purity was assessed by silver stain of SDS-PAGE. Moreover, the quantity of relatively pure YedY was limited (based on the BCA assay). Thus, pH 5.0 can improve the purity of protein, but fails to increase the amount of protein. In order to get enough purest protein to analyze the structure, more conditions need to be changed and examined such as adding glycerol in buffers which was suggested to remove unspecific proteins and improve the purity.

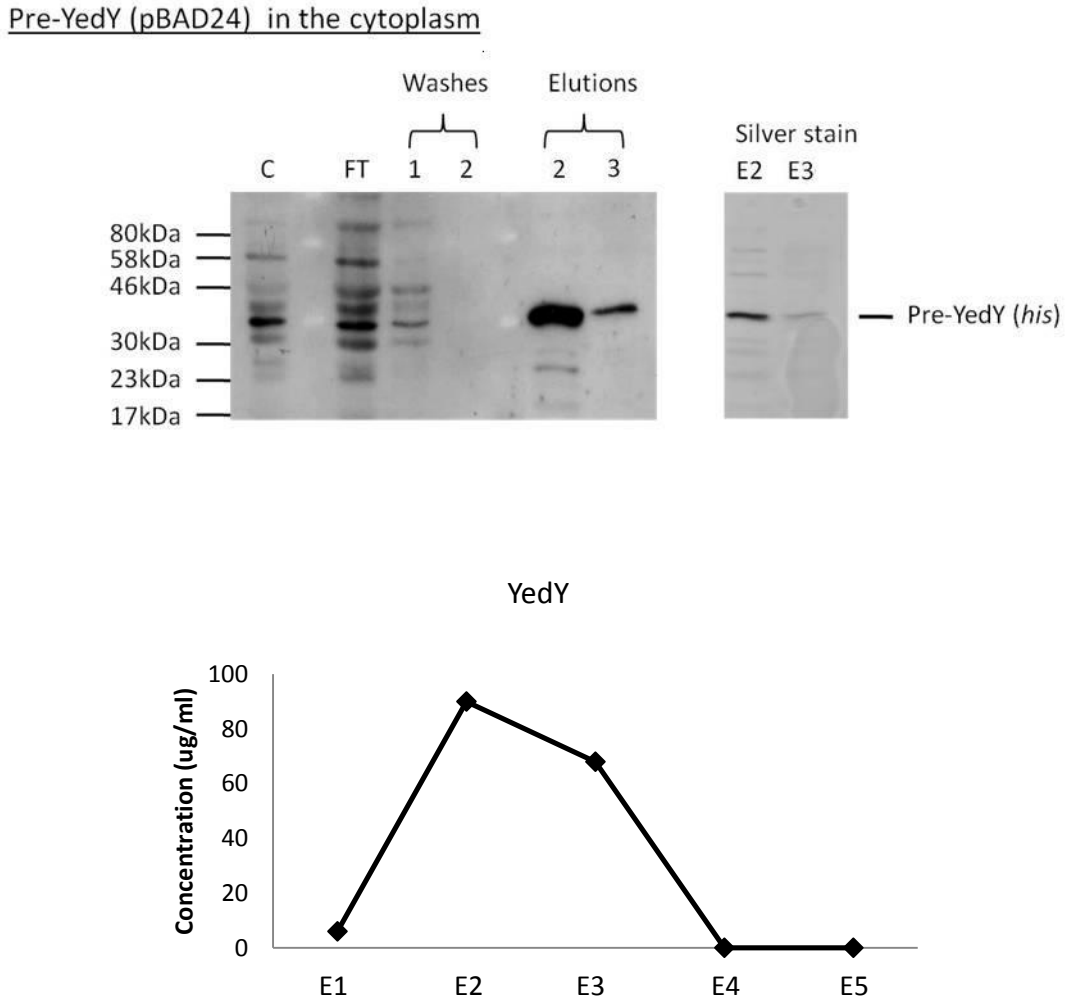


Figure 4.2.4 Talon affinity chromatography of pre-YedY with pH 5.0 YedY was amplified with C-terminal his-tag into arabinose inducible pBAD24. It started the 1L culture to over-express YedY in Δ*tatABCDE* cells. *E. coli* cytoplasmic fraction (C) was isolated from cells and directly applied to Talon affinity column. The flow-through (FT) sample collected and unbound proteins were washed (W). Bound proteins were eluted by elution buffer. All elution fractions (E) were measured by Bradford assay and analysed by silver stain as well as immunoblot with anti-his antibody.

Based on the experiment carried out at pH 5.0, 5% glycerol was added to help the purification. The immunoblot in Figure 4.2.5 shows that some YedY was lost in the flow-through (FT) indicating protein that still was not effectively bound to the column. Importantly, the specifically bound $YedY_{his}$ was eluted in one elution. It was shown that this YedY was pure, even when concentrated to a small volume, but a higher molecular weight band after concentration was present in the silver stain of SDS-PAGE. The purification of YedY was not improved by adding glycerol, because the quantity of protein was still a key problem to solve. Hence, it was important to increase the binding efficiency of protein using the affinity chromatography.

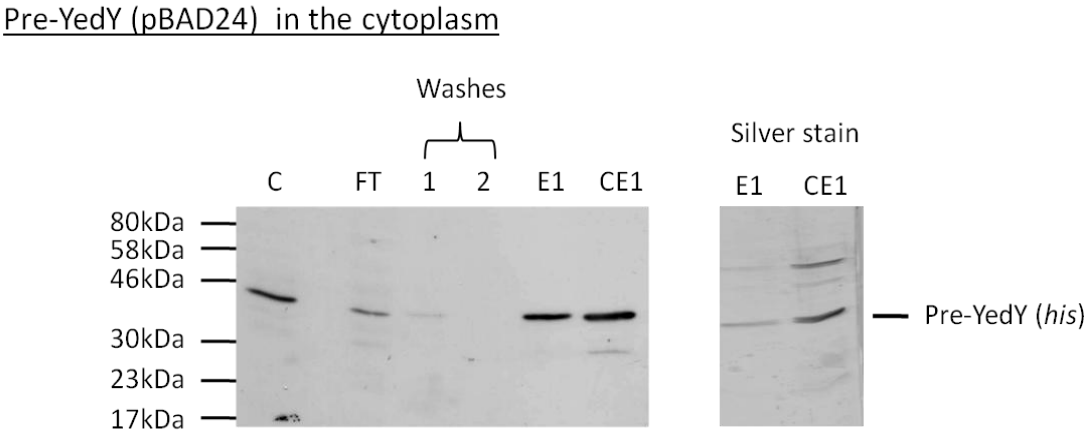


Figure 4.2.5 Talon affinity chromatography of pre-YedY with pH 5.0 as well as adding of Glycerol YedY was amplified with C-terminal his-tag into arabinose inducible pBAD24. It started the 0.5L culture to over-express YedY in Δ*tatABCDE* cells. *E.coli* cytoplasmic fraction (C) was isolated from cells and directly applied to Talon affinity column. The flow-through (FT) sample collected and unbound proteins were washed (W). Bound proteins were eluted by elution buffer. Elution fractions (E) and concentrated elution fraction (CE) were measured by Bradford assay and analysed by silver stain as well as immunoblot with anti-his antibody.

4.2.2.3 Purification with Immobilized metal ion affinity chromatography (IMAC) charged with Cu^{2+}, Zn^{2+} or Ni^{2+}

Immobilized metal ion affinity chromatography (IMAC) relies on the specific interaction between certain exposed amino acids, particularly histidine, and transition metal ions such as Zn^{2+}, Ni^{2+}, Cu^{2+} or Co^{2+} to widely purify a broad range of proteins and peptides. The IMAC sepharose will be charged with the most suitable metal ion to achieve the high performance of purification before use.The choice of metal ion is dependent on the type of application and the specific protein to be purified. For instance, Ni^{2+} is usually the first choice when purifying most histidine-tagged recombinant proteins. Furthermore, the strength of protein binding with immobilized metal ions is affected by several factors, including the number and spatial distribution of the affinity tag on the protein, the nature of metal ion used and the pH of buffers. Some histidine-tagged proteins could be easier to purify with metal ions other than Ni^{2+}, e.g. Zn^{2+}, Cu^{2+} or Co^{2+}. The purification of $YedY_{his}$ using Talon affnity chromatograpghy have been tested under vary conditions which were displayed the problems of quantity and quality to get the purified YedY. Here, we examined the efficiency of purification of $YedY_{his}$ with IMAC charged with Zn^{2+}, Ni^{2+} and Cu^{2+} other than Co^{2+} of Talon resin.

3 ml uncharged IMAC sepharose was separated into 3 columns and charged with Zn^{2+}, Ni^{2+} and Cu^{2+} respectively to ready for purification. In order to compare the performance, the purification used an optimal choice of pH 7.0 and 300 mM NaCl. The same amount of crude cytoplasmic extract isolated from *E. coli ΔtatABCDE* cells was applied to each equilibrated charged IMAC column. The column was washed using equilibration buffer to remove unbound protein, and eluted tightly bound $YedY_{his}$ with equilibration buffer supplemented with imidazole. Protein presented in elutions was resolved on SDS-PAGE, and the gel was stained by silver stain.

The result in Figure 4.2.6 shows that the purification of pre-YedY only worked using IMAC charged with Cu^{2+} but Zn^{2+} and Ni^{2+} did not bind any $YedY_{his}$. In terms of the IMAC charged Cu^{2+}, not all YedY bound to the column and some protein was detected in the wash fractions. $YedY_{his}$ eluted from the column was not pure since

some higher molecular weight bands were present. Thus, it was shown that $YedY_{his}$ may prefer the Cu^{2+} charged IMAC but the metal ion to be charged in the column can affect the performance of purification. However, compared with Talon resin, the Co^{2+} charged Talon is better than Cu^{2+} charged IMAC, and Ni^{2+} is generally the first choice to purify most his-tagged recombinant proteins, I then examined the commercial charged IMAC, His-Select Cobalt affinity chromatography (Sigma) and Ni-NTA affinity chromatography (Qiagen), to optimize the purification.

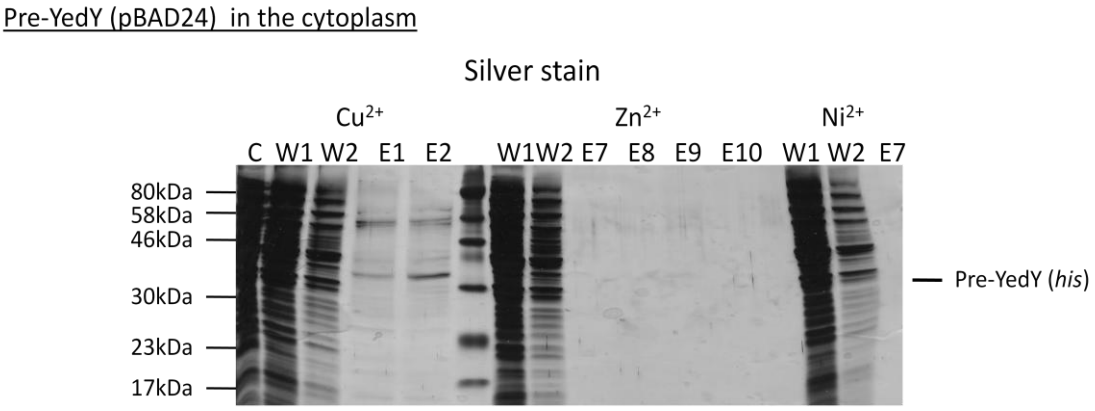


Figure 4.2.6 Purify pre-YedY by Immobilized metal ion affinity resin (IMAC) charged with Cu^{2+}, Zn^{2+} or Ni^{2+} YedY was amplified with C-terminal his-tag into arabinose inducible pBAD24. It started the 1L culture to over-express YedY in Δ*tatABCDE* cells. *E. coli* cytoplasmic fraction (C) was isolated from cells and separated to 3 parts. Each fraction applied to IMAC-Cu^{2+}, IMAC- Zn^{2+} and IMAC-Ni^{2+}column respectively. The flow-through (FT) sample collected and unbound proteins were washed (W). Bound proteins were eluted by elution buffer. Elution fractions (E) were measured by Bradford assay and analysed by silver stain.

4.2.2.4 Purification with His-Select Cobalt affinity chromatography (Sigma) and Ni-NTA affinity chromatography (Qiagen)

According to the product information, HIS-Select Cobalt Affinity Gel from Sigma allows for high purity, low non-specific binding and high binding capacity for his-tagged proteins. On the other hand, NTA (nitrilotriacetic acid) provides four chelation sites for nickel ions resulting in more tightly bound nickel than normal metal-chelating purification systems of three binding sites. The QIAexpress Ni-NTA Protein Purification System has the lowest nickel-ion leaching and a greater binding capacity intended to obtain the higher purity of proteins.

Briefly, I prepared 2×1 ml column with His-Select Cobalt affinity gel (Sigma) and Ni-NTA affinity resin (Qiagen), and the purification of $YedY_{his}$ was carried out using the optimal choice of pH 7.0 and 300 mM NaCl. The same amount of crude cytoplasmic extract isolated from *E. coli ΔtatABCDE* cells expressing pre-YedY was loaded onto each equilibrated column. The column was washed using equilibration buffer to remove unbound protein, and tightly bound $YedY_{his}$ was eluted with equilibration buffer supplemented with imidazole. Protein present in elutions was resolved on SDS-PAGE, and the gel was stained by silver stain.

Figure 4.2.7 shows that a lot of YedY did not bind to the column and was detected in the flow-through (FT) and wash fractions when it was purified using His-Select Cobalt affinity column. $YedY_{his}$ that did bind specifically eluted from the column and the pure pre-YedY was detected in elution fraction 3. Figure 4.2.7 also depicts that some YedY was lost in the flow-through (FT) while pure pre-YedY was eluted in elution 3-5 with the Ni-NTA affinity column. Comparing the amount of pure pre-YedY using BCA assay, it was found that the Ni-NTA affinity resin was the better method to obtain a large amount of pure pre-YedY. In summary, Ni-NTA affinity chromatography was chosen to purify pre-YedY and three mutants for study of structure changes.

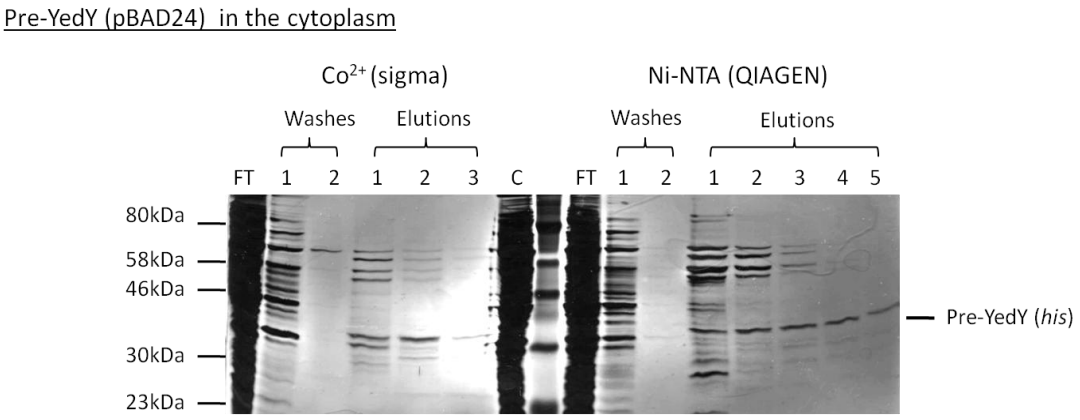


Figure 4.2.7 Comparison of the purification of pre-YedY by His-Select Cobalt affinity chromatography (Sigma) and Ni-NTA affinity chromatography (Qiagen) YedY was amplified with C-terminal his-tag into arabinose inducible pBAD24. It started the 1L culture to over-express YedY in Δ*tatABCDE* cells. *E. coli* cytoplasmic fraction (C) was isolated from cells and separated to 2 parts. Each fraction applied to Co^{2+} (Sigma) and Ni-NTA (Qiagen) column respectively. The flow-through (FT) sample collected and unbound proteins were washed. Bound proteins were eluted by elution buffer. Elution fractions were measured by Bradford assay and analysed by silver stain.

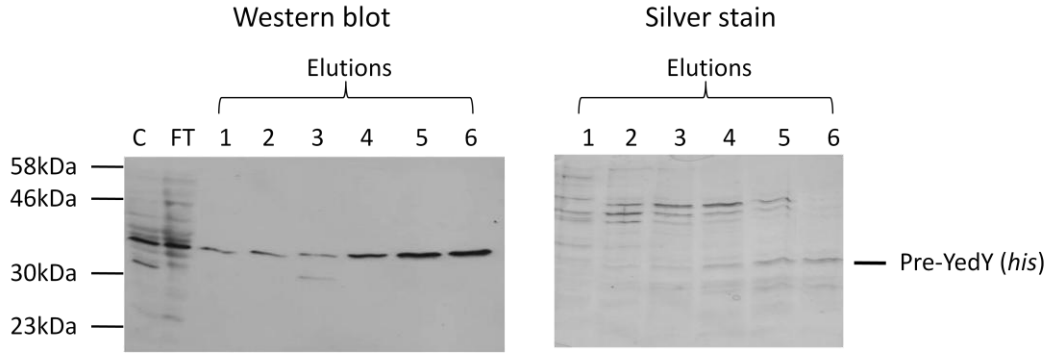
4.2.2.5 Purification of YedY (precursor and mature) and three mutants with Ni-NTA affinity chromatography (Qiagen)

For purification, YedY and three mutants were amplified with C-terminal his tags into IPTG inducible pEXT22. Pre-YedY was expressed for 3 h in 1 L Δ*tatABCDE* culture and cells were fractionated to obtain the crude cytoplasmic extract. The 2.5 ml Ni-NTA affinity column was prepared and equilibrated with buffer including 300 mM NaCl in pH 7.0. The isolated cytoplasmic fraction was applied to the column and the column was washed using equilibration buffer to remove unbound protein, and then eluted with equilibration buffer supplemented with imidazole. Protein present in column fractions was resolved on SDS-PAGE, and the gel was immunoblotted with anti-his antibody or stained by silver stain. Three mutants, Cys102, Glu48 and Arg194, were expressed and purified in the same manner. However, the mature form of YedY was isolated from the periplsam after expression in *E .coli* wild-type MC4100 cells, with the rest of the purification the same as for pre-YedY.

The results revealed in Figure 4.2.8 show that some of the YedY and mutant proteins were detected in the flow-through (FT) indicating protein that did not completely bind to the column. However, most of the YedY and the three mutants were found in the elution fractions where they were detected by immunoblot and the purity was determined by silver stain.

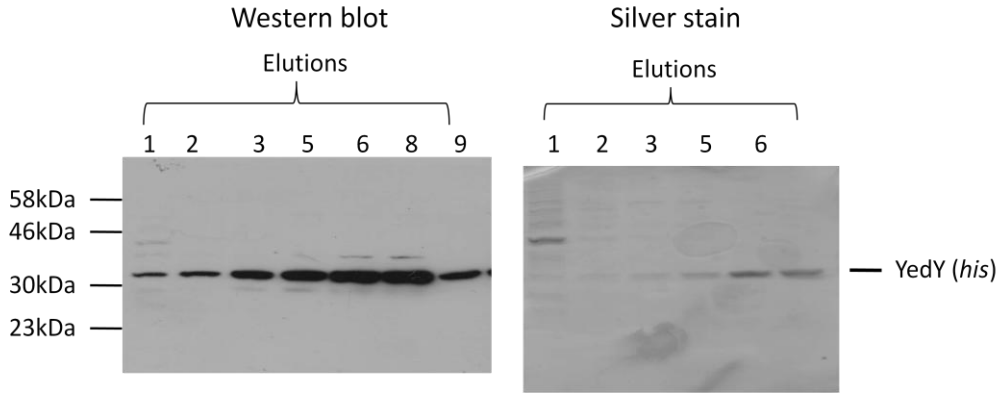
A

Pre-YedY (pEXT22) in the cytoplasm



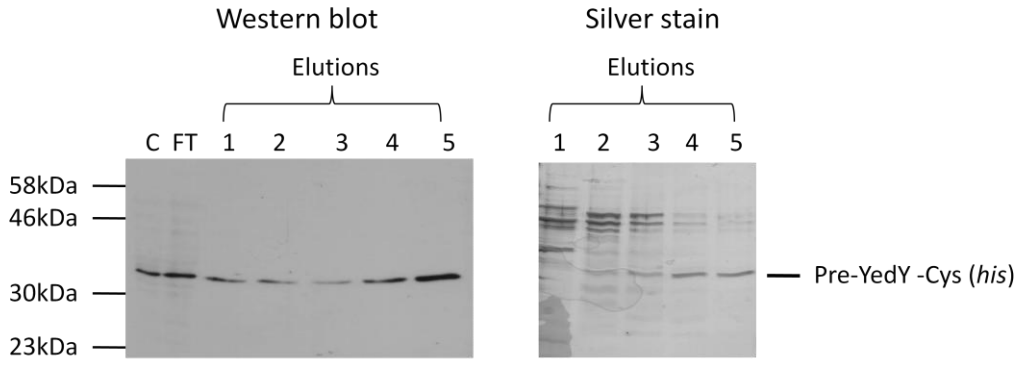
B

YedY (pEXT22) in the periplasm



C

Pre-YedY-Cys (pEXT22) in the cytoplasm



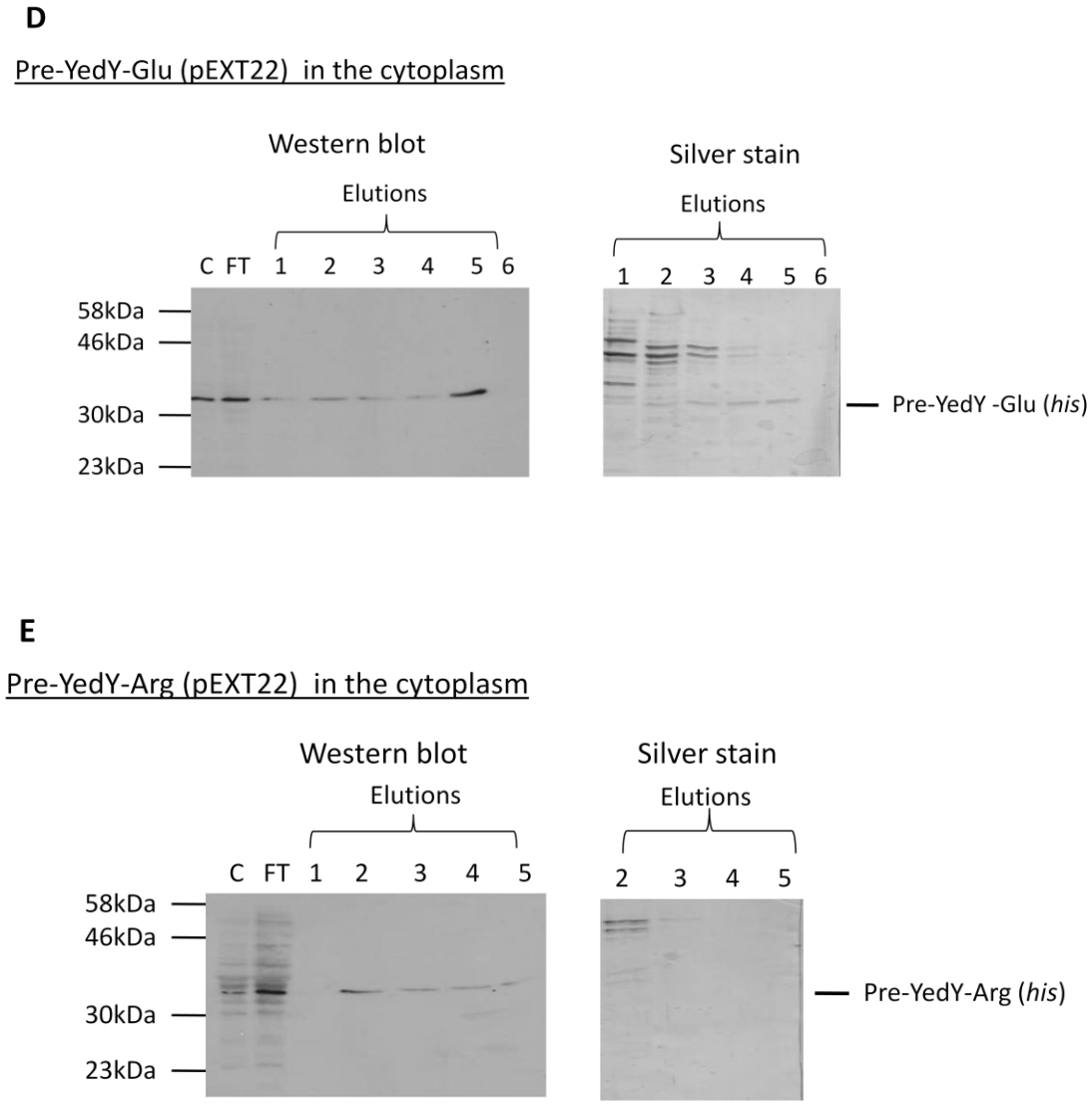


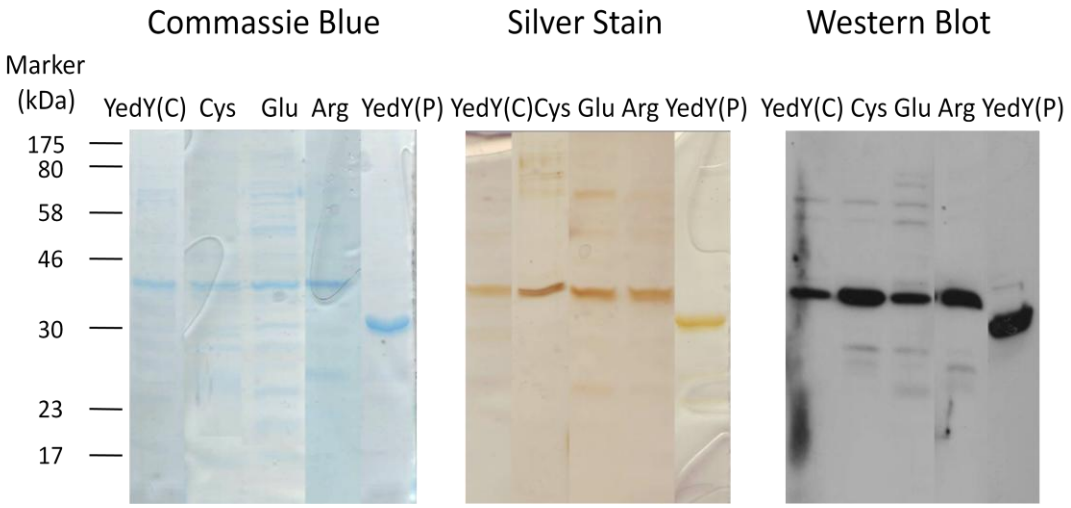
Figure 4.2.8 Ni-NTA affinity chromatography of pre-YedY (cytoplasm), YedY (periplasm), pre-YedY-Cys, pre-YedY-Glu and pre-YedY-Arg Pre-YedYand YedY, and precursor of three mutants Cys/Glu/Arg were amplified with C-terminal his-tag into IPTG inducible pEXT22. Precursor proteins all started the 1L culture to over-express in Δ*tatABCDE* cells and mature YedY was expressed in wild-type MC4100 cells. *E.coli* cytoplasmic fraction (C) or periplasmic fraction (P) was isolated from cells and applied to Ni-NTA affinity column. The flow-through (FT) sample collected and unbound proteins were washed. Bound proteins were eluted by elution buffer. All elution fractions were measured by Bradford assay and analysed by silver stain as well as immunoblot with anti-his antibody.

4.2.3 Study of the structures of pre-YedY and the three mutant forms

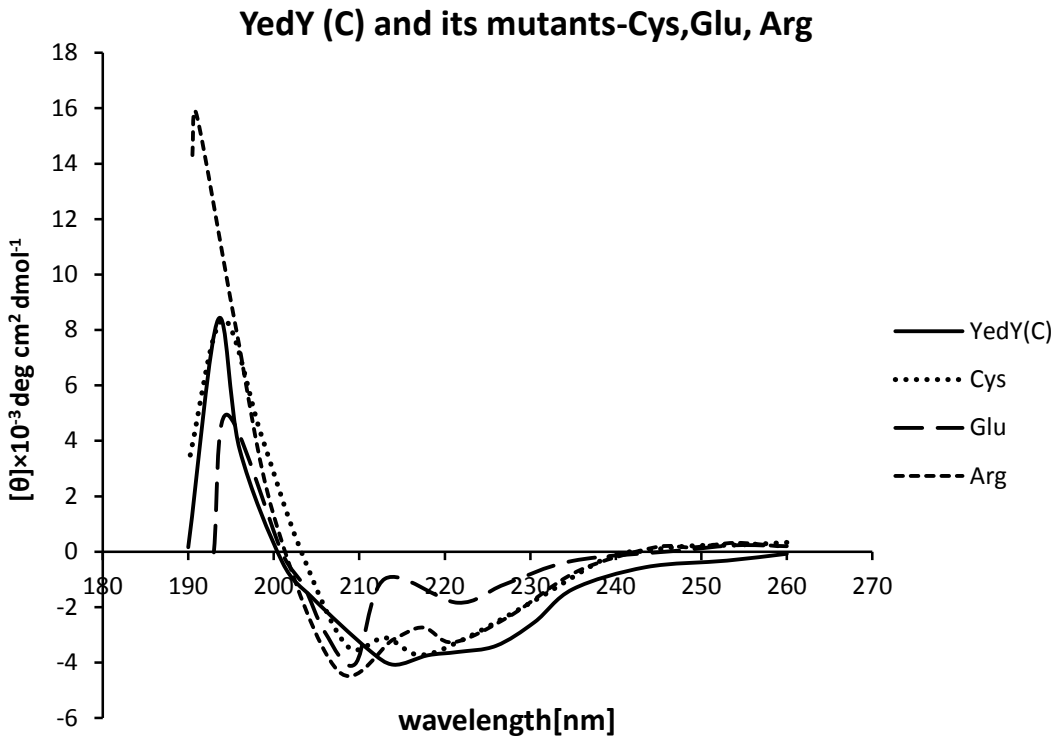
Purified pre-YedY and the three mutant forms were concentrated and detected by immunoblot, Commassie blue and silver stain as shown in Figure 4.2.9A which reveals the proteins to be highly pure. The purified proteins were then studied by circular dichroism (CD) spectroscopy to assess the secondary structure content of the proteins and inductively coupled plasma mass spectrometry (ICP-MS) to analyze the extent of cofactor insertion through measurement of the molybdenum content of the proteins.

As shown in Figure 4.2.9B and C, the non-mutated YedY spectrum shows negative maxima at 214 and 226 nm, corresponding to α-helical structure, and quantitation (using the fitting programme CDSSTR) gives an estimate of 47% α-helix and 24% β-strand. The α-helical content is higher than that estimated by Loschi *et al*. (2004) but this may reflect the presence of the N-terminal signal peptide in our tests, since we analysed the precursor form rather than mature-size YedY. The mutated YedY variants (again, all as precursor proteins) exhibit significantly different CD spectra, confirming that all three substitutions affect secondary structure content. Quantitation of the data (Figure 4.2.9C) shows that Cys102 contains more α-helix and slightly less β-strand than non-mutated pre-YedY whereas Glu48 and Arg194 contain lower levels of α-helix, slightly more β-strand, and greater levels of unordered structure. However, compared the overall folding state of wild-type YedY with three mutants, Cys102 seems to have the most similar secondary structure to pre-YedY.

A


Commassie Blue
Silver Stain
Western Blot
Marker
(kDa)
YedY(C) Cys Glu Arg YedY(P)
YedY(C)Cys Glu Arg YedY(P)
YedY(C) Cys Glu Arg YedY(P)
175
80
58
46
30
23
17

B


YedY (C) and its mutants-Cys,Glu, Arg
[θ]×10-3 deg cm2 dmol-1
wavelength[nm]
YedY(C)
Cys
Glu
Arg

C

	YedY(C)	Cys	Glu	Arg
Helix	0.47	0.66	0.18	0.19
strand	0.24	0.14	0.37	0.41
turns	0.09	0.11	0.13	0.07
unordered	0.2	0.1	0.29	0.24
total	1	1.01	0.97	1.01

D

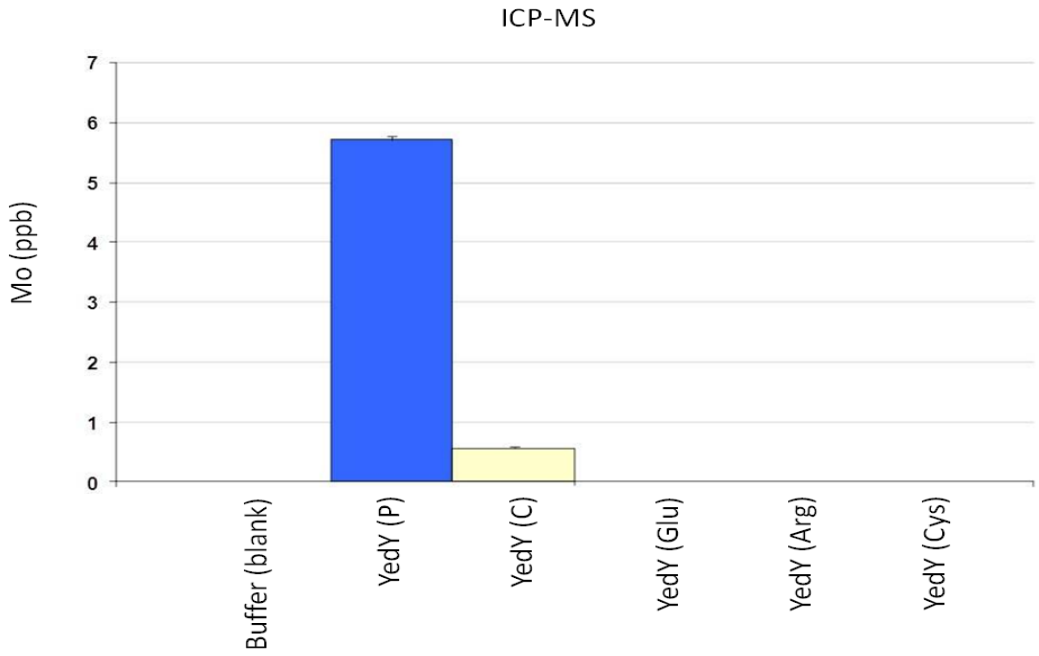


Figure 4.2.9 Study of secondary structures of purified *E. coli* YedY(C), YedY(P) and its mutations Cys/Glu/Arg by Circular Dichroism spectra and inductively coupled plasma mass spectrometry (ICP-MS). A. The elution fraction from the NI-NTA affinity column is dialysed and concentrated. The samples are checked with Commassie blue, silver stain and Western immunoblotted by anti-his antibody to determine purity (molecular weight markers are given to the left). **B**. Purified YedY (C) and three mutants are used to obtain CD spectra over 190-260 nm using a Jasco J-815 spectrophotometer. The spectra shown is an average taken from 16 consecutive scans. **C**. CD data analysed by Dichroweb is shown the structure alternation of YedY(C) and three mutants. **D**. Mo content of cofactor molybdopterin within purified YedY (C), YedY (P) and its mutations were measured by ICP-MS.

Consistent with these structural changes, analysis of metal content shows that while periplasmic mature YedY and cytoplasmic pre-YedY contain detectable levels of molybdenum (Figure 4.2.9D) no molybdenum can be detected in preparations of the mutated forms. The combined data show that the Cys, Arg and Glu substitutions block cofactor assembly and lead to significant, though different, structural changes in terms of overall secondary structure. It is interesting that the preparation of cytoplasmic pre-YedY contains less molybdenum than the periplasmic mature protein, and this finding is reproducible. It suggests that a substantial proportion of the pre-YedY is in the form of apo-protein, possibly reflecting a relatively slow cofactor insertion process.

4.3 Discussion

The Tat translocation pathway has the unique ability to export fully folded proteins across the bacterial plasma membrane. Most Tat substrates in *E. coli* contain redox cofactors such as FeS, copper, NiFe and molybdopterin centres which are required to insert before transportation (Palmer *et al.*, 2005; Sargent, 2007; Robinson *et al.*, 2011). In order to prevent the futile export of apo-proteins, effective proofreading systems are essential to monitor the assembly of cofactor and folding of proteins (Sargent *et al.*, 2005).

The Tat pathway has been demonstrated to preferentially export proteins if they are in a correctly folded state, for example, cytochrome c and disulphide protein PhoA (Delisa *et al.*, 2003; Richter and Bruser, 2005). Moreover, recent studies of two FeS Tat substrates NrfC and NapG has revealed that proofreading did not only describe the recognition of rejected proteins, but also involved degradation of malfolded proteins (Matos et *al.*, 2008).

In terms of the architecture of YedY and the analysis of molybdopterin (MPT) assembly, molybdopterin protein YedY and three mutants involved in the Moco-binding site have been studied in the Chapter 3. Cys102 is coordinated to the molybdenum ion by sulfur ligands and it is located between two β-strands. Glu48 and Arg194 interact with the pterin by direct hydrogen bonds which settle in α-helices (Loschi *et al.*, 2004). The data from the export assays show that mutated precursors are blocked in export and degraded at varying rates in wild type cells.

To elucidate how the Tat pathway recognizes the Cys/Glu/Arg mutants as incorrectly folded, and what the structural changes that are recognized by the proofreading system, YedY and three mutants were purified to analyze the secondary structure and cofactor insertion in this chapter. Purified proteins were examined using ion exchange chromatography and Immobilized metal affinity chromatography (IMAC) charged different metal ion to achieve the native and highly pure proteins.

The purified pre-YedY and three mutated precursor proteins were firstly studied by circular dichroism spectroscopy (CD) to assess the secondary structure. It was shown

that typical α-helical secondary structure which corresponds on the 214 and 226 nm negative maxima is present in pre-YedY. In contrast, the negative peaks and observed ellipticity (degree) at wavelength of the three mutants are shifted which indicates the proteins adopt different conformations. Pre-YedY comprises 47% α-helix and 24% β-strand, as determined by the fitting program CDSSTR using the SP175 reference set. Compared with the data from PBD (codes *1XDQ*) which predicted YedY has five chains consisted of nearly 22% α-helix and 24% β-sheet, the higher α-helical content is speculated to the presence of the N-terminal signal peptide of pre-YedY.

Of the three mutants, the structure of Cys102 appears to be only slightly changed since the α-helix content is slightly higher and β-strand content is slightly lower, however, the overall structure is still ordered. On the other hand, the structures of Glu48 and Arg194 are significantly altered including much less α-helix content, higher β-strand content and, respectively, 29% and 44% unordered content which suggests a much looser folding state.

Although the three mutants have different secondary structure contents, all of them are inhibited in export by the Tat pathway. Due to the relationship between the assembly of cofactor and the folding state of proteins, the metal content within purified proteins was analyzed by inductively coupled plasma mass spectrometry (ICP-MS) to test for insertion of molybdopterin cofactor. It was found that cytoplasmic pre-YedY and periplasmic mature YedY contain molybdenum, but the three mutants lack it. Importantly, the precursor of YedY has less molybdenum than the mature form which implies a proportion of pre-YedY is in the apo-protein form in the cytosol.

In short, three mutations, Cys102, Glu48 and Arg194, prevent the insertion of cofactor and lead to secondary structural changes. The structural change can be recognized by the Tat pathway which blocks their translocation, and also leads to the proteolysis of the unfolded/misfolded proteins. This elegant and complicated process is termed ‘proofreading’ or ‘quality control’ which play an essential role in the Tat translocation pathway. However, still some general principles relating to proofreading have yet to be properly elucidated, for instance, how does the system

monitors the folding state of cofactor-less Tat substrates? And whether specific cytosolic chaperones involve in the translocation of Tat-dependent redox proteins to assist in folding, cofactor insertion and oligomerization with partner proteins? To learn about proofreading for the redox proteins, the function of REMPs (redox enzymes maturation proteins) in the export of Tat substrates was studied in the Chapter 5.

Chapter 5

Interaction and the quality control of Tat substrates with REMPs (Redox Enzyme Maturation Proteins)

5.1 Introduction

Although the Tat pathway appears to have an inherent preference to transport folded proteins, a more complex quality control system has been shown to operate for a number of redox substrates. This involves cytoplasmic chaperones that associate with signal peptides to effectively act on the assembly of cofactor, folding of proteins and oligomerization with partner proteins to prevent substrates from targeting to the translocase until the entire process is complete (Robinson *et al.*, 2011). The genes encoding these chaperones are often found in the same operon as the Tat substrate. These chaperones also have been termed REMPs (redox enzyme maturation proteins) (Turner *et al*., 2004, Jack *et al*., 2004, Lee *et al*., 2006). Thus, this chaperone-mediated proofreading is an alternative concept of the stringent quality control system (DeLisa *et al.,* 2003).

The first practicable assay for chaperone-mediated proofreading activity was established in the trimethylamine N-oxide reductase (TorA) system. In *E. coli,* TorA is a water-soluble periplasmic Tat-targeted enzyme that contains bis-molybdopterin guanine dinucleotide (MGD) (Mejean *et al.*, 1994). TorD, which is encoded by the *torCAD* operon, helps to monitor the biosynthesis of TorA as well as the translocation of TorA by the Tat pathway. The direct, specific interaction between TorD and the signal peptide of TorA helps cofactor insertion and protects apo-TorA against proteolysis (Genest *et al*., 2006, Ilbert *et al*., 2003). In addition, TorD associates with the core of apo-TorA to facilitate the maturation of TorA, and is involved in the synthesis of bis-MGD (Guymer *et al*., 2010). Also, recent studies suggested that TorD proofreads TorA in a GTP-dependent way since GTP hydrolysis of TorD is used to regulate the maturation of TorA. This involves the cooperation of two aspects of TorD with the folding of TorA and possible involvement of the MGD biosynthetic process (Genest *et al*., 2008, Chan *et al*., 2009). In short, TorD takes part in all stages of TorA transport from synthesis of apoprotein to translocation.

Another well-known REMP is DmsD, a homologue of TorD, which performs proofreading in the biogenesis of the molybdopterin-containing *E. coli* enzyme dimethyl sulfoxide (DMSO) reductase. DMSO reductase is a trimeric complex comprising DmsA (the catalytic subunit), DmsB and an integral membrane protein

DmsC (Weiner *et al*., 1992; Simala-Grant and Weiner, 1996). DmsA is a Tat membrane-anchored substrate containing MGD and FeS clusters, and DmsD is the proofreading-chaperone to assist in the biogenesis of DmsA. Studies indicated that DmsD recognizes and interacts with the signal peptide of DmsA (Chan *et al*., 2009, Li *et al*., 2010). Moreover, DmsD is essential for biosynthesis and insertion of cofactor in DmsA. Importantly, DmsD interacts with the TatBC subunits and is proposed to escort fully folded protein to the Tat translocase (Ramasamy *et al.*, 2009).

In *E. coli*, a well-studied Tat substrate NapA is also subject to chaperone-mediated proofreading during the translocation process. Nitrate reductase (NapA) belongs to the DMSO reductase family of molybdenum-containing enzymes and is a periplasmic protein binding a FeS cluster and *bis*-MGD as redox cofactors (Kisker *et al*., 1997; Jepson *et al*., 2007). NapD encoded by the *nap* operon has been shown to play an important role in the NapA-dependent periplasmic nitrate reductase activity (Potter and Cole, 1999). Furthermore, NapD has been suggested to be the REMP of NapA and to function in a similar manner of TorD. However, NapD is structurally and functionally defined as the paradigm representative of a second family of 'Tat proofreading chaperones' because it forms a ferredoxin-type fold which is completely different from the TorD family (Maillard *et al*., 2007). Studies have demonstrated that NapD directly interacts with the NapA twin-arginine signal peptide to inhibit the premature export of NapA. Thus, NapD is essential for the translocation of NapA by the Tat pathway.

The interactions between REMPs and Tat signal peptides of known and predicted Tat-specific redox enzyme subunits were recently analyzed. The study demonstrated that some REMPs are specific to a redox enzyme(s) of similar function, whereas others are less specific and able to interact with signal peptides of related enzymes (Chan *et al*., 2009). For instance, DmsD is not only the REMP of DmsA, but also interacts with putative DMSO reductases YnfE and YnfF. The summary of Tat substrates and corresponding REMPs is listed in the Table 5.1.

Tat-dependent redox enzyme systems studied by Chan *et al.*, 2008[a]

Redox enzyme	RR-containing subunit(s)	Predicted REMP(s)	Catalytic cofactor
Biotin sulfoxide reductase 1	BisC	YcdY[b]	MoPt
Biotin/TMAO reductase	BisZ/TorZ	YcdY/TorD[b]	MoPt
TMAO reductase	TorA	TorD	MoPt
DMSO reductase	DmsA	DmsD	MoPt
Putative DMSO reductase	YnfE	DmsD[b]	MoPt[d]
Putative DMSO reductase	YnfF	DmsD[b]	MoPt[d]
DMSO/TMAO reductase	YedY	DmsD/TorD	MoPt[d]
Formate dehydrogenase	FdnG	FdhD/FdhE	MoPt
Formate dehydrogenase	FdoG	FdhD/FdhE	MoPt
Hydrogenase 1	HyaA	HyaE	Ni
Hydrogenase 2	HybO	HybE	Ni
Nitrate reductase (periplasmic)	NapA	NapD	MoPt
Nitrate reductase (cytoplasmic)	NarG	NarJ	MoPt
Nitrate reductase (cytoplasmic)	NarZ	NarWb	MoPt
Hypothetical protein	YfhG	?[c]	None predicted

a The predicted or known REMP(s) for each of the redox enzymes was investigated for interaction with peptides of the RR containing subunits.
b Predicted REMP based on the predicted function/role of the enzyme and its relatedness to other enzymes with an appropriate REMP.
c No predicted REMP due to limited knowledge of the function of the enzyme.
d Predicted catalytic cofactor of MoPt based on the presence of a MoPt domain in its sequence.

Table 5.1 Possible REMPs of Tat substates are summarized by Chan *et al.*, 2008

In the light of sequence alignments and other interaction experiments, DmsD and TorD are the best candidates as REMP for the sulfite oxidase YedY. The sequence alignment of the interaction between Tat substrates and REMPs is shown in Figure 5.1 (Chan *et al.*, 2008). While the specific chaperones discussed above generally bind the signal peptides of their target protein, in order to assist in the folding of the protein and insertion of cofactors, this is not the case for all Tat substrates. It must therefore be considered that, for Tat proteins which do not have specific chaperones, the general folding machinery of the cytosol plays a role.

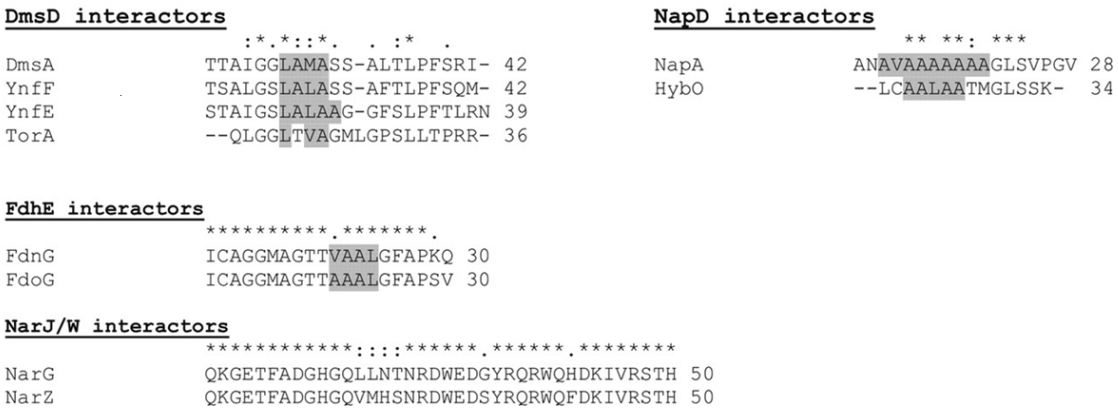


Figure 5.1 Sequence alignment of the interaction between Tat substrates and REMPs Alignment of the hydrophobic region immediately following the RR motif, based on RR peptides interacting with a common REMP. The continuous stretch of small hydrophobic residues that may play a role in recognition specificity by the REMPs is shaded. Residues in the alignment that are identical (*****), conserved substitutions (:), and semiconserved substitutions (.) are indicated. Numbering is based on the original positions of residues in the full-length RR peptide. (Figure is adapted from Chan *et al.*, 2008)

In this chapter, the translocation of YedY in the absence of some REMPs was tested to study the chaperone-mediated proofreading in Tat pathway, and identify the new possible chaperones which play an important role in the proofreading process. The importance of REMPS for export of TorA, DmsA and NapA were first studied as control tests.

5.2 Results

5.2.1 REMP (redox enzyme maturation protein) of TMAO reductase during the Tat translocation pathway

Many Tat substrates in *E. coli* are cofactor-containing redox enzymes that require REMPs to regulate folding and cofactor assembly before targeting to Tat translocation machinery (Turner *et al.*, 2002; Sargent, 2007). TMAO reductase (TorA) is a well-known Tat substrate, and previous studies have shown that TorD is the REMP that associates with both the signal peptide and the core of apo-TorA to monitor the assembly of the molybdopterin centre by the hydrolysis of GTP (Genest *et al.*, 2008; Guymer *et al.*, 2010). Furthermore, DmsD was also suggested to bind pre-TorA, an interaction between TorD and DmsA has not been shown (Oresnik *et al*., 2001; Ilbert *et al*., 2004). Here, the translocation of TorA by the Tat pathway in the absence of TorD and DmsD was studied to investigate the importance of REMPs.

TorA was amplified with a C-terminal strep-tag into the IPTG inducible pEXT22 and expressed in wild-type (WT), Δ*tatABCDE*, Δ*torD* and Δ*dmsD* cells. Figure 5.2.1 depicts the results obtained when cells were fractionated after the expression of TorA for 3 h. In wild-type cells (WT) the protein was detected in total cells (TC) and in the cytoplasm (C), membrane (M) as well as periplasm (P) which indicated export of TorA. No export was observed in Δ*tatABCDE* cells, as expected, and export was also blocked in the Δ*torD* strain, confirming the importance of TorD for the overall TorA export process. Interestingly, the TorA protein is found in the form of a smaller degradation fragment in the cytoplasm of Δ*torD* cells, together with a larger form (possibly an SDS-resistant dimer) in the membrane fraction. To a certain extent, the data is consistent with previous studies which suggested the interaction between TorD and TorA not only helps cofactor insertion, but also protects apo-TorA against proteolysis (Genest *et al*., 2006, Illbert *et al*., 2003). Nevertheless, in Δ*dmsD* cells, TorA was found to export to the periplasm confirming that DmsD is not required for the translocation of TorA by the Tat pathway. Thus, TorD is essential for the export of TorA by Tat pathway, but the absence of DmsD might partially affect the translocation.

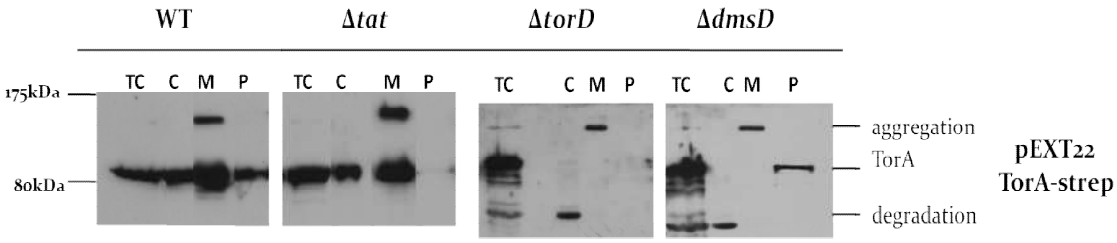


Figure 5.2.1 Translocation of TorA in wild-type, Δ*tatABCDE*, Δ*torD* and Δ*dmsD* cells TMAO reductase (TorA) was amplified with C-terminal *strep-tag* into IPTG inducible plasmid pEXT22. TorA was expressed in *E. coli* wild-type MC4100 (WT), Δ*tatABCED* (Δ*tat*), Δ*torD* and Δ*dmsD* cells. Cells were harvested in 3h (3 h is the optimized time point to efficient export of TorA) and total cell contents (TC) were analysed together with cytoplasm, membrane and periplasm samples (C, M, P) after osmotic shock and fractionation. TorA is detected by immunoblot using anti-strep antibody and indicated forms of aggregation or degradation. Mobilities of molecular mass markers (in kDa) are shown on the left.

5.2.2 REMP of nitrate reductase (NapA) during the Tat translocation pathway

Nitrate reductase (NapA) in *E. coli* is a periplasmic protein bearing a twin-arginine signal peptide, and the export involves a ‘‘Tat proofreading’’ process which requires the specific association between the Tat signal peptide and its chaperone NapD to monitor the cofactor insertion and protein folding (Maillard *et al*., 2007). In this study, a similar experiment to the TorA experiment above was conducted to study the importance of NapD for the export of NapA.

In Figure 5.2.2, NapA was amplified with a C-terminal his-tag into arabinose inducible pBAD24. When expressed in wild type cells, the appoximately 90 kDa mature NapA was detected in the periplasm, whereas export was blocked in Δ*tatABCDE* cells and the precursor protein was only found in the cytoplasmic (C) and membrane (M) fraction. Importantly, export is also retarded in Δ*napD* cells confirming a key role for the NapD as the REMP in the export of NapA. In addition, there was a larger form observed in the membrane fraction which may be an SDS-resistant dimer. The data indicate that NapD is essential for the biosynthesis and export of the Tat-dependent nitrate reductase NapA.

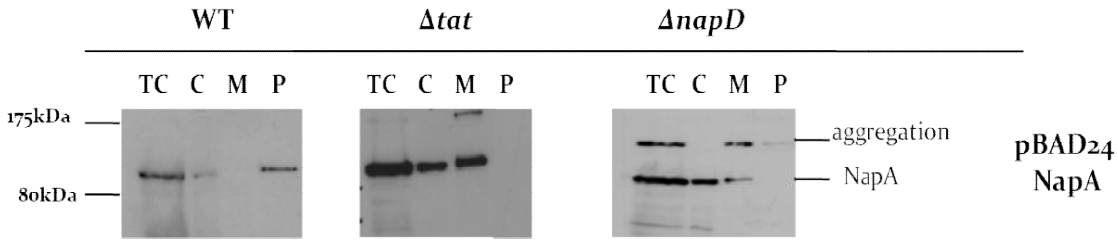


Figure 5.2.2 Translocation of NapA in wild-type, Δ*tatABCDE* and Δ*napD* cells
Nitrate reductase (NapA) was amplified with C-terminal *his-tag* into arabinose inducible plasmid pBAD24. NapA was expressed in *E. coli* wild-type MC4100 (WT), Δ*tatABCED* (Δ*tat*) and Δ*napD* cells. Cells were harvested in 3h (optimized time point) and total cell contents (TC) were analysed together with cytoplasm, membrane and periplasm samples (C, M, P) after osmotic shock and fractionation. NapA is detected by immunoblot using anti-his antibody and indicated forms of aggregation. Mobilities of molecular mass markers (in kDa) are shown on the left.

5.2.3 The export of YedY with respect to chaperone-mediated proofreading

It has been shown that TorD and NapD are the REMPs of TorA and NapA, respectively, playing a critical role in the export by the Tat pathway. It is still unclear whether all cofactor-containing Tat substrates require REMPs during the process of cofactor coordination and folding.

On the basis of sequence alignments and other interaction experiments, DmsD/TorD have been proposed as candidate REMPs for YedY assembly/export (Chan *et al.*, 2008), and experiments were carried out that directly address this question. Firstly, YedY export assays were performed in the *torD, dmsD* or *napD* genes deletion strains. Figure 5.2.3A shows time course studies of the export and processing of pre-YedY in wild-type cells and the three deletion strains. Expression of YedY in wild type cells led to the appearance of precursor and mature-size band, and a similar pattern of bands were observed when pre-YedY was expressed in Δ*torD,* Δ*dmsD* or Δ*napD* cells. Cells were then fractionated after 3 h expression confirming that mature-size YedY was exported to the periplasm in three deletion strains (Figure 5.2.3B). These data suggest that YedY export does not depend on any of these REMPs.

The YedY-Arg194 mutant also was expressed in *torD, dmsD or napD* genes deletion strains to study whether export of this protein may occur in the absence of REMPs (since the proofreading processes might be compromised). The result in Figure 5.2.3A displayed that the precursor of Arg194 mutant was only detected after 2h, and then degraded rapidly in Δ*torD,* Δ*dmsD* or Δ*napD* cells which was consistent with the expression in wild-type cells. Fractionation study again confirmed that Arg194 mutant was blocked to export in all strains (data not shown).

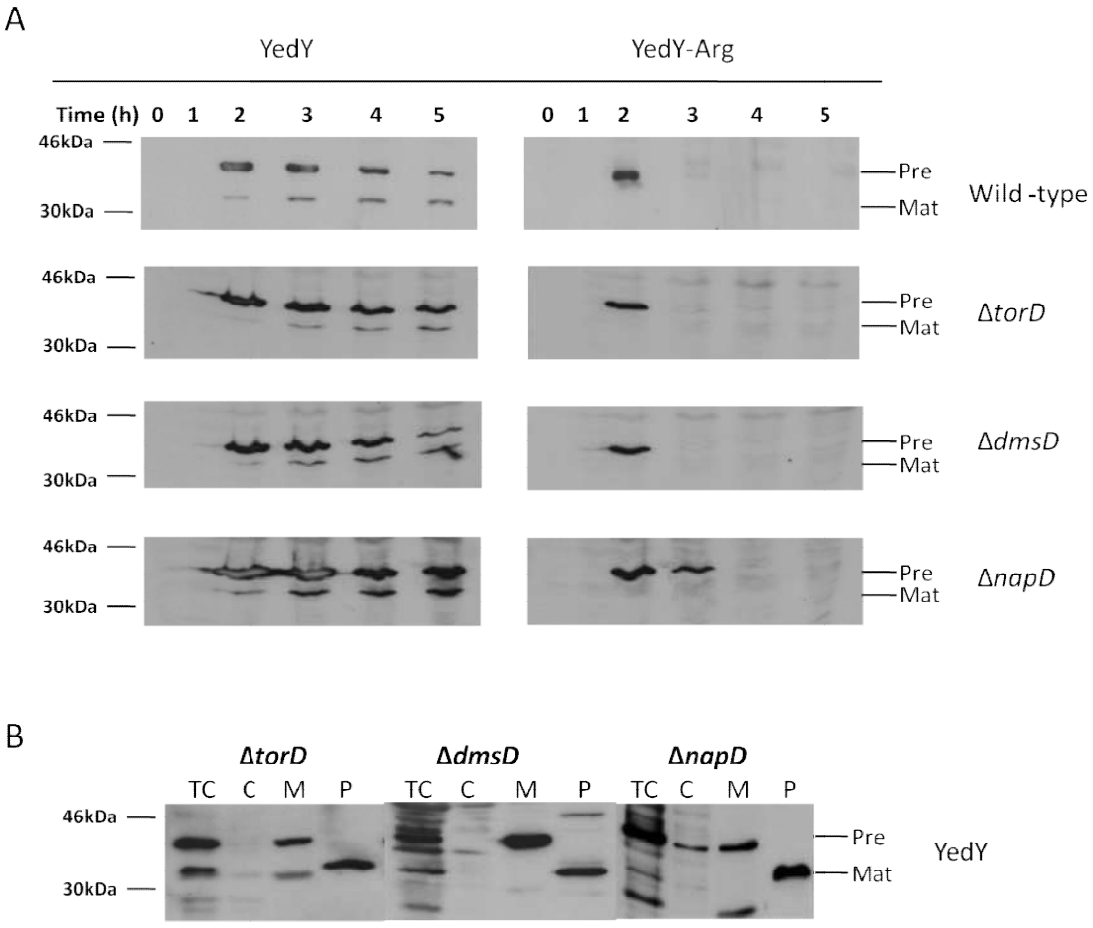


Figure 5.2.3 Translocation of YedY and a mutant YedY-Arg in wild-type, Δ*torD*, Δ*dmsD* and Δ*napD* cells A.YedY and a mutant YedY-Arg were expressed in wild-type (WT) and *torD/dmsD/napD* deletion cells. Cells were analysed at times (in hours) after induction of synthesis with arabinose and the immunoblots were shown in the inset, with precursor and mature forms of the protein indicated (Pre, Mat) **B**. YedY was expressed for 2.5 h after induction in *torD/dmsD/napD* deletion cells, and total cell contents (TC) were analysed together with cytoplasm, membrane and periplasm fractions (C, M, P). Samples were analysed by immunoblot using anti-his antibody, and Precursor and mature forms of YedY are indicated (Pre, Mat).

In conclusion, these data indicated that YedY does not rely on any of these REMPs (TorD, DmsD and NapD) for the cofactor assembly and export in *E. coli*. The results therefore suggest that YedY uses other cytosolic chaperones to regulate export, or possibly a simpler proofreading system that does not require substrate-specific REMPs at all. Such a system may involve novel proofreading components and experiments were carried out to identify these.

5.2.4 Interaction between pre-YedY and cytoplasmic extract to identify possible REMPs involved in the proofreading of YedY

Although TorD, DmsD and NapD are not REMPs of YedY in *E. coli*, it is possible to have unknown specific chaperones in the cytosol which play an essential role in the proofreading. In order to investigate the possible REMPs of YedY, protein-protein interaction studies were carried out by several methods.

The *in vitro* protein-protein interaction methods used here include Ni-NTA affinity chromatography, chemical cross-linking and label transfer. Each method has its own advantages. Affinity chromatography permits low-abundance bait proteins to be exposed to large volumes of cell-free extracts containing prey proteins, allowing for the specific interactions to occur while background is minimized (Howell *et al*., 2006). Besides, to detect transient or weak protein interactions, even the cases of protein complexes, chemical cross-linking strategies were developed (Fancy, 2000). Label transfer methods, as a subset of advanced protein cross-linking techniques, can be used for the identification of protein interactions (Phizicky and Fields, 1995; Fancy, 2000).

Purified pre-YedY and cytoplasmic extract containing prey proteins were prepared. As detailed in chapter 4, the tagged His_6-preYedY was purified from *E. coli* Δ*tatABCDE* cells. Also, the cell-free cytoplasmic extract was isolated from 1 L culture of pre-YedY expressed in Δ*tatABCDE* cells. The aim was to immobilise the His-tagged pre-YedY on affinity columns and identify proteins that specifally bound to it.

5.2.4.1 Protein interaction by Ni-NTA affinity chromatography

Ni-NTA affinity resin and purified His_6-preYedY were incubated for 3 h to prepare the special Ni-NTA affinity column. All experiments were performed using gravity flow at 4 °C. The cytoplasmic extract was loaded into the column and incubated end-over-end for 2 h following washes to remove unbound proteins. Then the column was eluted with buffer containing 25 mM imidazole to remove unspecific binding proteins. Finally the specific proteins associating with YedY were eluted by Glycine pH 2.8, and analyzed via SDS-PAGE. Figure 5.2.4 shows three bands (two large bands were faint and another one was strong) were detected by the silver stain of SDS-PAGE.

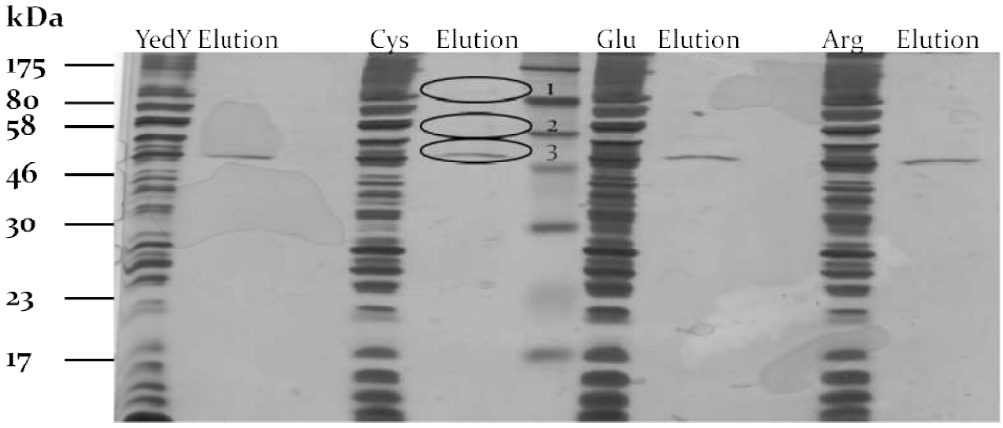


Figure 5.2.4 Silver stain of protein interaction between YedY (or mutants) and cytoplasmic extract by Ni-NTA affinity chromatography His-tagged purified pre-YedY (or mutants) used to prepare the column before loading the cytoplasmic extract. When the interaction between proteins has been made, equilibration buffer was added to remove unspecific binding proteins. Then the specific proteins associating with pre-YedY or mutants were eluted by Glycine pH 2.8. Cytoplasmic extracts from YedY, Cys, Glu, Arg and the specific binding proteins were analysed by SDS-PAGE and checked by silver stain.

The specific binding proteins were excised from instant Coomassie blue stained SDS-PAGE gels and treated by trypsin digestion. The identification of peptides was performed using mass spectrometry (MS) in the department of Chemistry, Warwick University. The data from MS was analyzed using the website 'Matrix Science'. It was found that three specific binding proteins were Pyruvate dehydrogenase E1 component, Dihydrolipoamide succinyltransferase and cytosolic chaperonin GroEL (Figure5.2.5). Pyruvate dehydrogenase E1 component and Dihydrolipoamide succinyltransferase are important component enzymes of the Pyruvate dehydrogenase complex (PDC) which works in the TCA cycle and there is no clear reason for them to interact with preYedY, implying they are contaminating bands. GroEL is a bacterial chaperone involved in protein folding in the cytosol. Previous studies have revealed that GroEL was found to associate with DmsD and NapD, the REMP for DMSO reductase (DmsA) and periplasm nitrate reductase (NapA) in *E. coli*, which implied GroEL could be involved in metalloenzyme maturation (Dionisi *et al*., 1998; Ribbe and Burgess, 2001; Butland *et al*., 2005; Li *et al*., 2010). Thus, although GroEL may play a role in the folding process of YedY, there is no specific protein found to be the REMP which interacts with YedY by the Ni-NTA affinity chromatography used in this study.

Mascot search parameters:

Database- NCBI (16/11/2010)
Enzyme – Trypsin
Missed cleavages allowed - 1
Max peptide mass tolerance 0.6 kDa
Taxonomy – *E. coli*

Sample	Estimated mass from gel	Protein name	Calculated mass	Number of matched peptides	Sequence coverage	MOWSE score
1	~ 80 kDa	Pyruvate dehydrogenase E1 component (ODP1_ECO57)	99.95 kDa	17	16 %	59 (>56 is significant)
3	~ 46 kDa	Dihydrolipoamide succinyltransferase (ODO2_ECO57)	44 kDa	17	40 %	103 (>56 is significant)
2	~ 58 kDa	60 kDa chaperonin – GroEL (CH601_ECOK1)	57.5 kDa	17	34 %	81 (>56 is significant)

Note:

- 1 and 3 are the important component enzyme of pyruvatedehydrogenase complex (PDC) which works in TCA cycle;
- 2-general folding chaperonin

Figure 5.2.5 Identification of the specific binding proteins of pre-YedY by Ni-NTA affinity chromatography to analyze with Mass spectrometry The specific binding proteins were analysed by SDS-PAGE and stained by instant coomassie blue. The samples were cut from gel and treated by Trypsin to analyse by mass spectrometry. The data was analysed by "Matrix Science".

5.2.4.2 Protein interaction studies using formaldehyde (1%) cross-linking

Formaldehyde cross-linking for the interaction between preYedY and the cytoplasmic extract containing prey proteins was carried out *in vivo* and *in vitro*. *In vivo* cross-linking experiment was performed according to the Derouiche *et al,* 1995 (see chapter 2 Methods). Briefly, *E. coli* Δ*tatABCDE* cells expressed preYedY$_{his}$ were grown at 37 °C to an OD_{600} of 0.6 and pelleted by centrifugation. The cell pellet was washed and resuspended at OD_{600} of 0.5 in 10 mM KP_i (pH 6.8) and incubated further for 20 min with the addition of a final concentration of 1% formaldehyde at the room temperature. The cross-linked samples were harvested and washed following the purification process by Ni-NTA affinity chromatography. The purified cross-linked samples were either heated at 37 °C for 10 min to maintain the formaldehyde cross-links or at 96 °C for 20 min to break the cross-links. Samples were analyzed by the SDS-PAGE and detected by instant Coomassie blue.

The result in Figure 5.2.6 displayed that there were less proteins detected in the flow through if compared with the purification of pre-YedY in chapter 4 implying lot of proteins bind to the column or they were degraded after cross-linking. Due to only the 37 kDa pre-YedY band showing in the elution lane, no interacting proteins can be detected. Thus, formaldehyde cross-linking in vivo was not suitable to study the protein interaction in my work.

For the *in vitro* experiment, purified pre-YedY (300μg/ml) and the cell-free cytoplasmic extract were gently mixed at a ratio of 1:3. The mixture was incubated for 2 h at room temperature and formaldehyde (1% final concentration) was added to further incubate for 20 min. Cross-linked samples were heated at 37°C for 10 min to stabilize the formaldehyde cross-links. Samples were analyzed by SDS-PAGE and detected by instant Coomassie blue as well as immunoblotting. There are a total of 4 bands in the immunoblot in Figure 5.2.6 which suggests that three bands could be possible proteins associated with pre-YedY in the cytosol and the band of 37 kDa should be pre-YedY. The three bands were excised to analyze by ESI-MS.

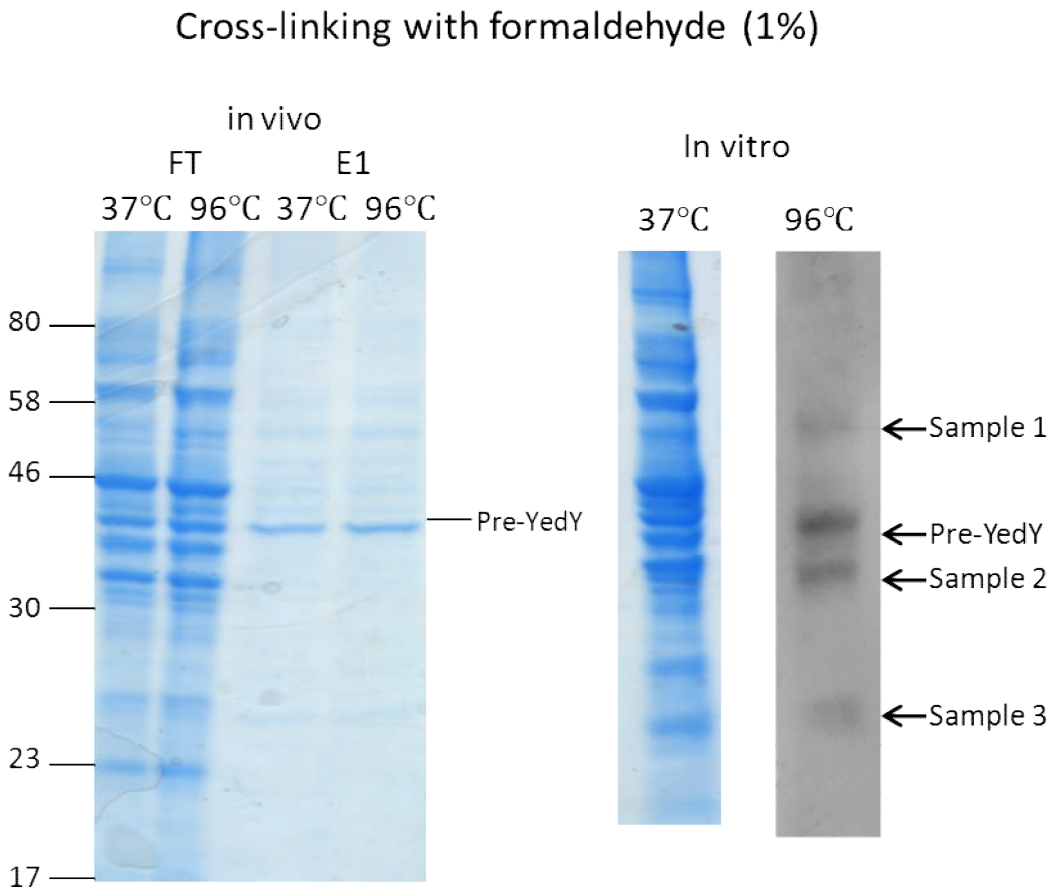


Figure 5.2.6 Commassie blue and immunoblot to detect proteins interacted between YedY and cytoplasmic extract by Formaldehyde cross-linking In vivo, 1% Formaldehyde adds to the culture which has been produced pre-YedY for 2.5 h and cells continue to grow for another 1h which facilitates cross-link between pre-YedY and proteins in cytosol. Then Pre-YedY$_{his}$ together with possible interactors is purified by Ni-NTA affinity chromatography. Flow through (FT) and Elution (E) are treated at 37 ℃ which maintains the crosslinking or 96 ℃ which breaks the crosslinking to release interactors. In vitro, His-tagged purified pre-YedY was mixed with cytoplasmic fraction to bind to interacting proteins. 1% formaldehyde was added to cross-link the proteins. The sample is treated at 37 ℃ or 96 ℃. Proteins were detected by Coomassie blue as well as analysed by immunoblotting.

According to the data from ESI-MS which were analyzed in the *E. coli* MC4100 database using ProteinLynx Global Server v2.4, the possible proteins interacting with pre-YedY are shown in Figure 5.2.7. The protein of approximately 43 kDa was found to be the translation elongation factor *tufA* which also was detected to associate with DmsD, the REMP of DmsA (Li *et al*., 2010). Elongation factor TufA promotes the GTP-dependent binding of aminoacyl-tRNA to the A-site of ribosomes during protein biosynthesis (Hwang and Miller, 1987; Jacquet and Parmeggiani, 1989), and plays an important regulatory role in cell growth and in the bacterial response to nutrient deprivation (Young and Bernlohr, 1991). Another interesting protein is GAPDH (Glyceraldehyde 3 phosphate dehydrogenase) which is involved in the metabolism of glucose and several non-metabolic processes including transcription activation. The last small protein is the Trp repressor-binding protein *WrbA* which enhances the formation and/or stability of noncovalent complexes between the trp repressor protein and operator-bearing DNA. In addition, *WrbA* can be an accessory element in blocking trpR-specific transcriptional processes that might be physiologically disadvantageous in the stationary phase of the bacterial life cycle (Yang *et al*., 1993; Grandori *et al*., 1998). Hence, the three proteins either play an important role in translation/transcription processes, or function in metabolism during the growth of bacteria. They are unlikely to be specifically involved in the folding process or Tat translation pathway of YedY depending on the interaction.

sample	Protein description	Calculated Mass (Da)	Number of matched peptides	Sequence coverage (%)
1	Elongation factor Tu 1 OS *Escherichia coli* strain K12 MC4100 BW2952 GN tufA	43256	25	66.75
2	Glyceraldehyde 3 phosphate dehydrogenase A OS *Escherichia coli* strain K12 MC4100 BW2952 GN gap	35510	19	59.21
3	Flavoprotein wrbA OS *Escherichia coli* strain K12 MC4100 BW2952 GN wrbA	20832	5	37.37

Figure 5.2.7 Identification of the specific binding proteins of pre-YedY by formaldehyde cross-linking to analyze with ESI-MS The specific binding proteins were analysed by SDS-PAGE and stained by instant Coomassie blue. The samples were cut from gel and digested to analyse by ESI-MS. The data was obtained using ProteinLynx Global Server v2.4.

5.2.4.3 Protein interaction by Sulfo-SBED Biotin label transfer

Label transfer study was carried out using the Sulfo-SBED Biotin label transfer kit (Thermo Scientific). Briefly, purified pre-YedY (500μg/ml) was reacted with Sulfo-SBED reagent at room temperature for 30 min in the dark to allow labelling (250μg pre-YedY along with 14.7μg reagent dissolved in DMSO). Labeled pre-YedY was dialyzed in the 1×label transfer buffer for at least 4 h to remove unreacted reagent. The labeled pre-YedY was incubated with cell-free cytoplasmic extract at the room temperature for 20 min in the dark. Samples were then exposed to 305 nm UV light for 20 min and 50 mM DTT was added to finish the label-transfer. The label-transferred biotinylated proteins were detected by immunoblot using Streptavidin-HRP and corresponding proteins shown in the SDS-PAGE gel were excised to analyze by ESI-MS. The Figure 5.2.8 shows that four bands may bind to the pre-YedY and Figure 5.2.9 is the data from ESI-MS to show the identification of the proteins.

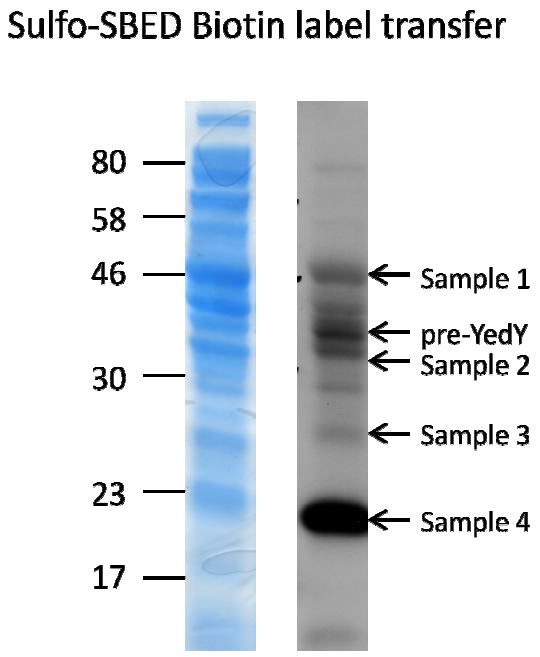


Figure 5.2.8 Commassie blue and immunoblot to detect proteins interacted between YedY in cytoplasmic extract by Sulfo-SBED Biotin label transfer His-tagged purified pre-YedY was labelled by Sulfo-SBED and mixed with the cytoplasmic fraction to carry out the interaction between proteins. The mixture was treated by UV to transfer the streptavidin-tag if the interaction happened. The interactors were detected by Commassie blue and immunoblot using anti-streptavidin.

The four bound proteins are GAPDH, Flavoprotein wrbA, Tryptophanase tnaA and 30S ribosomal protein rpsC, respectively. The first two proteins are detected in formaldehyde cross-linking as well, so the details are not discussed again. The other two proteins are Tryptophanase tnaA, which participates in tryptophan metabolism and nitrogen metabolism (Burns and Demoss, 1962), and 30S ribosomal protein rpsC that found to bind both of the lower part of the 30S subunit and mRNA in the 70S ribosome in order to facilitate translation (Takyar *et al*., 2005). Again, it seems that these proteins are unlikely to be specifically involved in proofreading during assembly process or the Tat translocation pathway of YedY.

sample	Protein description	Calculated Mass (Da)	Number of matched peptides	Sequence coverage (%)
1	Tryptophanase OS *Escherichia coli* strain K12 MC4100 BW2952 GN tnaA	52739	26	55.20
2	Glyceraldehyde 3 phosphate dehydrogenase A OS *Escherichia coli* strain K12 MC4100 BW2952 GN gap	35510	13	41.69
3	2 3 bisphosphoglycerate dependent phosphoglycerate mutase OS *Escherichia coli* strain K12 MC4100	28538	6	28.80
4	Flavoprotein wrbA OS *Escherichia coli* strain K12 MC4100 BW2952 GN wrbA	20832	3	26.77

Figure 5.2.9 Identification of the specific binding proteins of pre-YedY by Sulfo-SBED Biotin label transfer and analysis with ESI-MS The specific binding proteins were analysed by SDS-PAGE and stained by instant coomassie blue. The samples were cut from the gel and digested to analyse by ESI-MS. The data was obtained using ProteinLynx Global Server v2.4.

5.3 Discussion

The chaperone-mediated proofreading system is required to monitor the folding process of redox enzymes prior to the export by the Tat pathway. Such proofreading is operated by a series of cytosolic chaperones which are encoded in the same operon with related Tat substrates, and generally assist in the assembly of cofactor, correctly folding, and complex additions (Turner *et al.,* 2004; Jack *et al.,* 2004). To date, the specific chaperones are named as REMPs (redox enzyme maturation proteins) and most of them are important for transportation by the Tat pathway. The well-studied examples include TMAO reductase (TorA), DMSO reductase (DmsA) and Nitrate reductase (NapA) with their REMP TorD, DmsD and NapD which associate with the signal peptide to assist in the maturation of proteins (Ilbert *et al.*, 2004; Genest *et al.*, 2008; Chan *et al*., 2009; Potter and Cole, 1999).

In this chapter, TorA translocation was examined in the absence of TorD/DmsD in the *torD* and *dmsD* strains. The data showed that TorA export was blocked in *torD* deleted cells but remained in *dmsD* deleted cells implying TorD is the REMP of TorA. Similarly, NapD is confirmed to be the REMP of NapA because the export is inhibited when *napD* is deleted. Thus, the REMPs indeed play an important role in the Tat pathway and function in the proofreading process.

Furthermore, DmsD/TorD have been proposed to be REMPs of YedY on the basis of sequence alignment and indirect protein-protein interaction studies (Chan *et al.*, 2008). Since the mutants of YedY have been shown to be rejected in export and degraded by the proteolysis, TorD, DmsD and NapD were investigated to test whether they are indeed involved in proofreading of YedY. Perhaps surprisingly, the absence of *torD, dmsD* and *napD* did not affect the export of YedY by the Tat pathway, or affect the turnover of Arg194 mutant. On the basis of these results, it is proposed TorD, DmsD and NapD should not be the real and specific REMPs of YedY as predicted.

To understand whether other specific chaperones may be essential for YedY proofreading before or during the translocation, interactions between the YedY and cytoplasmic extract was performed using a variety of protein interaction approaches.

The affinity chromatography, cross-linking and label transfer studies along with the identification of protein by mass spectrometry revealed potential interactors. The results did not provide a conclusive result for the interaction, but the examination of YedY interactors reveals a few interesting proteins that may be relevant.

GroEL belongs to a molecular chaperone family and is found in a large number of bacteria (Zeilstra-Ryalls *et al*., 1991). It mediates proper protein folding processes through a workflow of binding, encapsulation, and release of substrate protein (Horwich *et al*., 2007). Evidence suggests that GroEL is involved in the insertion of the molybdenum–iron cofactor into the nitrogenase enzyme in *Azobacter vinelandii* (Ribbe and Burgess, 2001). Importantly, GroEL is shown to interact with DmsD and NapD in *E. coli* which function as REMP in the maturation of molybdopterin enzymes DmsA and NapA containing both molybdenum cofactors as well as iron–sulfur clusters (Li *et al*., 2010; Butland *et al*., 2005). Here, my study shows a direct interaction between GroEL and YedY which suggests that GroEL could be more generally involved in metalloenzyme maturation even in the absence of REMPs. In addition, GroEL was found to cooperate with DnaK which associates with RR signal peptides to stabilize the substrates (Perez-Rodriguez *et al*., 2007; Graubner *et al*., 2007).

In the cross-linking and label transfer studies, other potential weakly and/or transiently YedY-interacting proteins were identified, such as tufA (translation elongation factor), GAPDH (Glyceraldehyde 3 phosphate dehydrogenase), wrbA (Trp repressor-binding protein) and Tryptophanase tnaA.

Firstly, the interaction with translation elongation factor tufA suggests tufA binds to the YedY signal peptide before the termination of synthesis of the polypeptide chain. tufA was also showed to associate with DmsD, which implies that DmsD might similarly approach to the newly-synthesized DmsA leader to facilitate the folding-maturation process (Li *et al*., 2010).

Next, wrbA and Tryptophanase tnaA are unexpected interactors of YedY and it is difficult to understand the interaction. According to the literature, the flavoprotein wrbA can enhance the formation and/or stability of noncovalent complexes between

the trp repressor protein and operator-bearing DNA. Also, WrbA seems to block trpR-specific transcriptional processes that may lead to physiologically disadvantageous in the stationary phase of the bacterial life cycle (Yang *et al*., 1993; Grandori *et al*., 1998). On the other hand, Tryptophanase tnaA is suggested to participate in tryptophan metabolism and nitrogen metabolism (Burns and Demoss, 1962). Neither of these roles appear to be related to the Tat export systems.

Another interesting interactor is GAPDH. The metabolic function of the enzyme is required to catalyse the conversion of glyceraldehyde 3-phosphate which takes part in the pathway of energy and carbon molecule supply.

In light of the studies, the above data suggest that YedY may have no specific REMP to assist translocation and proofreading in Tat pathway, but it may rely on chaperones or small proteins which normally bind to REMPs to assist in synthesis and maturation.

Chapter 6

In Escherichia coli, the signal peptide is essential for export and maturation of proteins by the Tat pathway

6.1 Introduction

In bacteria, the Sec or Tat protein translocation pathways are normally used to transport extracytoplasmic proteins across the plasma membrane (Lee *et al*., 2006). The proteins are synthesized with an amino-terminal extension sequence named the signal peptide which is generally removed by a specific processing peptidase (von Heijne, 1990). The characteristics of typical signal peptides were identified during the extensive studies of Sec–dependent proteins which indicated it is composed of three domains: a positive N-terminal region, a hydrophobic central core and a polar C-terminal region ending with the short-chain amino acids at the -3 and -1 positions, relative to the cleavage site (von Heijne, 1990; Izard and Kendall, 1996). The maturation of exported proteins is performed by the leader (or signal) peptidase which recognizes the '(-3, -1) cleavage site'. In *E. coli*, the evidence demonstrated that the primary signal peptidase, LepB, is responible for the maturation of both Sec and Tat substrates (Luke *et al*., 2009).

The Tat pathway has the unique ability to export pre-folded proteins across the tightly sealed bacterial plasma membrane and chloroplast thylakoid membrane (Robinson *et al*., 2011). The signal peptides of Tat substrates contain the invariable 'twin-arginine' in conserved consensus *S–R–R–x–F–L–K* motif between the N- and H- domains, and end with typically Ala-Xaa-Ala (Stanley *et al*., 2000).

In *E. coli*, TatA, TatB and TatC appear to form the minimally functional translocase. Although the precise details of the translocation event are poorly understood, the currently proposed mechanism of the Tat translocation pathway suggest that the signal peptide of correctly folded proteins is firstly recognized and associated with the TatBC receptor complex, and the substrate-bound TatBC complex subsequently triggers the generation of the transport channel which is composed of TatA protomers for the export of proteins. Finally, the precursor is processed to mature form in the periplasm by removal of the signal peptide (Muller and Klosgen, 2005; Robinson *et al*., 2011).

Although a few native Tat substrates have been identified to lack a RR-motif (Hinsley *et al.*, 2001; Ignatova *et al.*, 2002), studies have underlined the importance

of RR-motif in the signal peptide function which mediates the export of proteins via the Tat pathway. Export of the cofactor-less Tat substrate SufI was completely blocked when both arginines were replaced with lysine, however, single substitutions of arginine with lysine can be tolerated (Stanley *et al*., 2000; Yahr and Wickner, 2001). Similarly, substitution of RR with KK in the signal peptide of trimethylamine N-oxide reductase (TorA) completely stopped Tat translocation of a reporter protein (GFP) (Cristobal *et al*., 1999). Nevertheless, previous studies also showed that the single arginine replacement of some Tat substrates such as *E. coli* CueO and *Pseudomonas stutzeri* nitrous oxide reductase can abolish the export of passenger proteins (Dreusch *et al*., 1997; Halbig *et al*., 1999; Stanley *et al*., 2000).

Studies on the Sec system have shown that the removal of the signal peptide is not essential for translocation of the protein, but it leads to the accumulation of precursor proteins in the membrane (Dalbey & Wickner, 1985). The positions -1 and -3 in the prokaryotic signal peptides are particularly critical for defining the cleavage site and Ala-X-Ala is the most frequently observed sequence preceding the cleavage site (von Heijne, 1983, 1984, 1986; Perlman and Halvorson, 1983). Studies performed by Fikes *et al* have indicated that mutations in the precursor of *E. coli* Maltose-binding Protein (MBP) at the -1 and -3 positions resulted in varying processing efficiencies (Fikes *et al*., 1990). In addition, the substitutions at the -1 Ala by Gly, Ser, Thr, Leu, Lys or Glu led to either a substantial or complete block in the processing of proteins (Fikes *et al*., 1990; Shackleton and Robinson, 1991).

In this Chapter, I firstly examined the critical importance of the signal peptide, especially the twin arginine of TorA and YedY, for the Tat pathway and then identified the relationship between Tat translocation and final maturation processing of proteins.

6.2 Results

6.2.1 Translocation of a Tat substrate TorA, containing mutations in the signal peptide

TMAO reductase (TorA) is one of the largest Tat substrates identified in *E. coli* at ~90 kDa. It contains a molybdopterin cofactor that it aquires in the cytoplasm before translocation by the Tat pathway to the periplasm (Silvestro *et al*., 1989; Czjzek *et al*., 1998). A cytoplasmic chaperone, TorD, has been identified that binds to the signal peptide of TorA to prevent premature interaction with the Tat machinery before cofactor insertion and correct protein folding has occurred (Jack *et al*., 2004; Genest *et al*., 2006). TorA acts as a final electron acceptor of the respiratory chain during anaerobic growth of *E. coli* on TMAO medium.

In order to identify the importance of the signal peptide which targets synthesized proteins to the translocation pathway, especially the twin-arginine motif in the Tat machinary, the translocation ability of TorA and mutations of the signal peptide were tested. TorA was amplified with a c-terminal *strep*-tag into the IPTG inducible plasmid pEXT22. Thereafter, the mutant TorA_KK was made by the substitution of twin arginines with lysine, and TorA_M was produced by the deletion of the whole signal peptide of TorA. TorA, TorA_KK and TorA_M were expressed for 3 h in *E. coli* wild-type MC4100 and Δ*tatABCDE* cells. Cells were harvested and separated into the cytoplasm (C), the membrane (M) and the periplasm (P) compartments. Total cell extract (TC) and each fraction (C, M, P) were measured to the same amount to analyse by SDS-PAGE and subsequent immunoblotting with antibody against *strep*. Figure 6.2.1 shows TorA is detected in the cytoplasmic fraction when TorA, TorA_KK and TorA_M are expressed in both wild-type and Δ*tatABCDE* cells, and some TorA is also bound to the membrane. It is worth pointing out that TorA is only observed in the cytoplasmic fraction when TorA_M is expressed, possibly indicating that the precursor form has an affinity with the membrane. Furthermore, pre-TorA and the dimer of pre-TorA are presented in the membrane fraction when TorA and TorA_KK are expressed in both wild-type and Δ*tatABCDE* cells. Importantly, mature TorA is only present in the periplasm when TorA is expressed in wild-type MC4100 cells. Hence, the result demonstrates that the signal peptide is

required to mediate the export of proteins, and the twin arginines of TorA are essential for translocation by the Tat pathway.

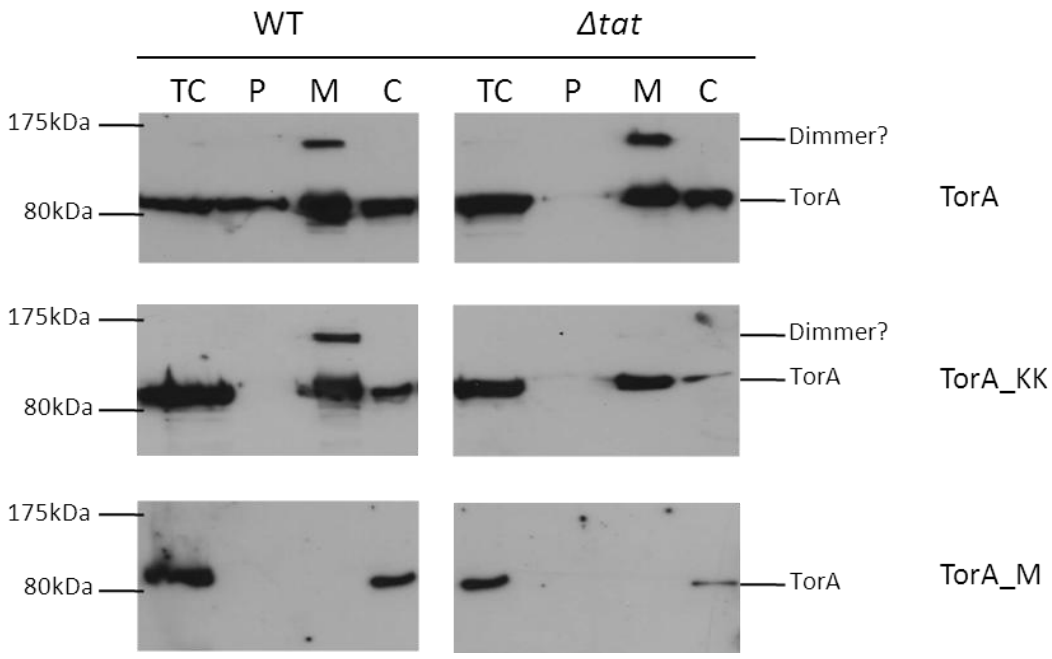


Figure 6.2.1 Translocation of Tat subtrate TorA, RR motif mutant TorA_KK and the deletion of signal peptide TorA_M TorA, TorA_KK and TorA_M were amplified with *strep-tag* into pEXT22 expression system and expressed in wild-type *E. coli* MC4100 cells (WT) and Δ*tatABCDE* (Δ*tat*) cells. Samples are taken after 3 h incubation, and total cell contents (TC) together with cytoplasm, membrane and periplasm samples (C, M, P) are analysed by immunoblot using anti-strep antibody. TorA and dimmer of TorA are indicated.

6.2.2 Translocation of mutated YedY and YedY_Arg by the Tat pathway

In the Chapter 3, it was shown that the Tat substrate YedY is exported to the periplasm, while the cofactor-less mutant YedY_Arg is rapidly degraded in wild-type MC4100 cells. To test the importance of the RR motif in the Tat translocation pathway, I analysed substitutions in the RR motif of YedY. The part of the amino acid sequence of *E. coli* pre-YedY is shown in Fig.6.2.2 which the signal peptide is outlined with RR motif bolded.

Figure 6.2.2 The part of primary structure of *E. coli* pre-YedY The picture underlines the YedY signal peptide with the bold RR motif. The predicted SP cleavage site is denoted by a dashed line and the amino acid in -3 to -1 are typed in Italic. Residue substituted in mutant YedY_A44L is indicated by asterisks.

Although previous studies have indicated that single substitutions with lysine could usually be tolerated by the Tat pathway (but not the loss of both arginines of the RR motif), I decided to analyse the importance of RR motif in YedY for export by the Tat pathway. Single substitutions of arginine with lysine were introduced in the signal sequence of YedY, with two mutants, YedY_RK and YedY_KR, obtained. The substitutions were made by the mutagenesis of YedY which was amplified with C-terminal *his*-tag into the arabinose inducible plasmid pBAD24. The YedY and mutants YedY_RK/YedY_KR were expressed in wild-type and Δ*tatABCDE* cells over a 5-h period after addition of arabinose to induce synthesis (Figure 6.2.3 A). When YedY was expressed in wild-type cells, pre-YedY was first detected after 2 h and mature YedY was detected (at low levels) after 2 h, thereafter, the mature protein gradually becomes more prominent. In contrast, pre-YedY was detected after 2 h but no mature YedY was present over 5 h when either YedY_RK or YedY_KR were expressed in wild-type cells. The result in Figure 6.2.3 B displays the fractionation of cells expressing YedY_RK and YedY_KR for 2.5 h, and the data show that export was blocked and pre-YedY was detected in the membrane fraction. Thus, it is demonstrated that single substitutions of RR motif are enough to completely inhibit the translocation of YedY by the Tat pathway.

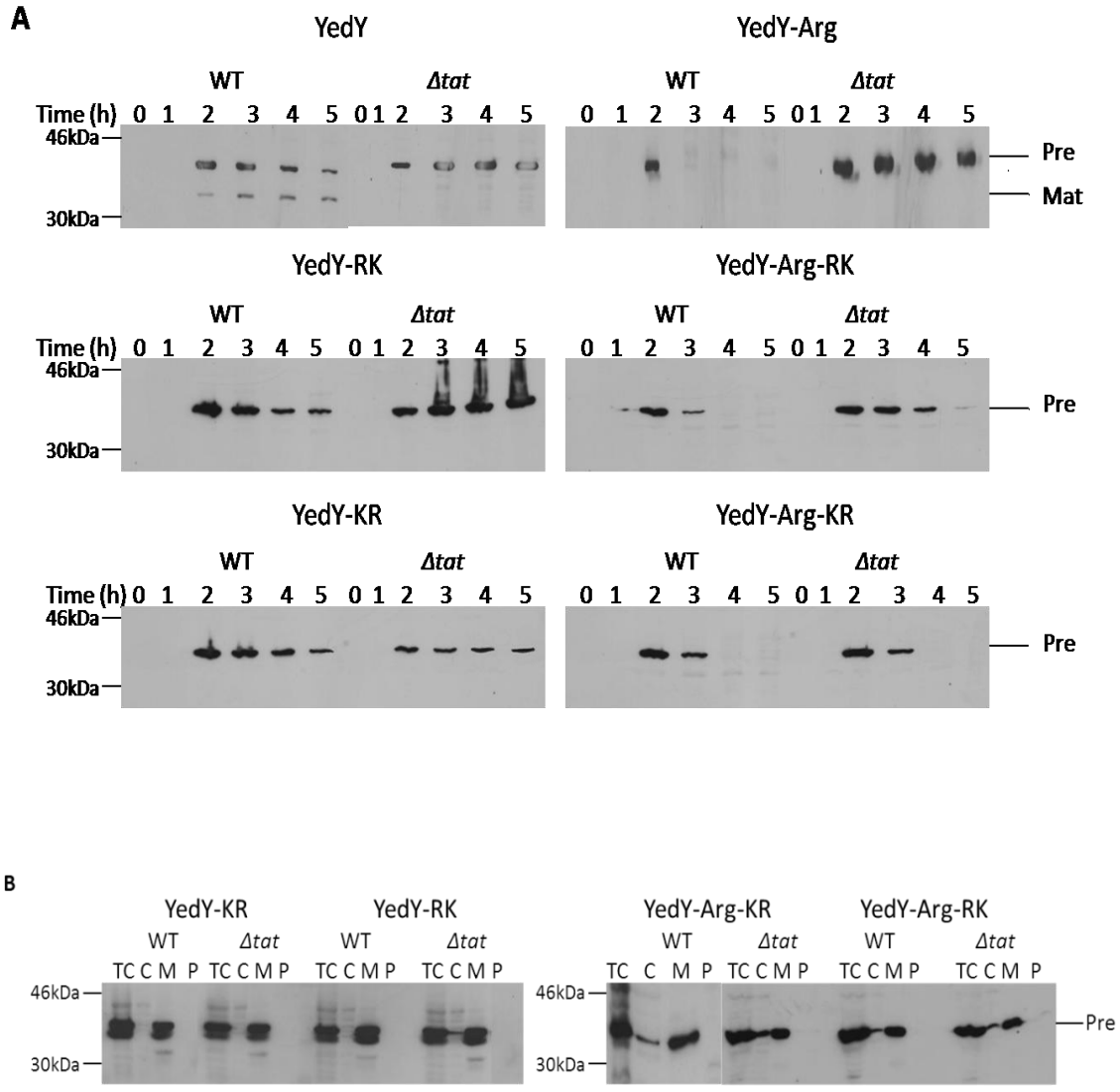


Figure 6.2.3 Translocation and fractionation of YedY-RK, YedY-KR and YedY-Arg-RK, YedY-Arg-KR in *E. coli* The single substitution of RR motif to RK/KR in YedY and YedY_Arg were amplified with C-terminal *his*-tag into pBAD24 vector and expressed in *E. coli* wild-type MC4100 (WT) and Δ*tatABCDE* (Δ*tat*) cells. (A) Samples were taken at the indicated times (in hours) after induction of synthesis with arabinose and immunoblotted by anti-his antibody. (B) Cells were taken after 2.5 h induction, and total cell contents (TC) together with cytoplasm, membrane and periplasm (C, M, P) are detected by immunoblot using anti-his antibody.

In addition, in order to test whether the presence/absence of the RR motif affects the degradation of YedY_Arg, I replaced the arginine of RR motif in the YedY_Arg mutant and analysed it in the same manner as YedY. Figure 6.2.3 shows that YedY_Arg can not be exported by the Tat pathway and was rapidly degraded in the wild-type cells, however, pre-YedY-Arg was stable in the Δ*tatABCDE* cells over a 5-h period. The single substitutions of arginine with lysine, YedY-Arg-RK and YedY-Arg-KR, resulted in a slower rate of degradation of YedY-Arg in the wild-type cells. Hence, the RR motif is not essential for the turnover of YedY-Arg.

In short, it was shown that the RR motif is essential for the transportation of YedY by the Tat pathway and even single substitutions of arginine can completely block export of YedY in wild-type cells. Also, the RR motif is not directly involved in the degradation of YedY-Arg implying proofreading and Tat-independent proteolysis.

6.2.3 Substitution at -1 of the signal peptide blocks the maturation of YedY

Previous studies on signal peptides have indicated that -1 and -3 residues near SP cleavage sites are essential for the efficient maturation of proteins (Perlman and Halvorson, 1983; von Heijne, 1990), and substitutions at -1 block the processing (Fikes *et al*., 1990; Shackleton and Robinson, 1991). In this study, YedY was used as a candidate Tat substrate to examine the relationship between the translocation by the Tat system and the terminal processing.

Here, the Alanine (Ala) at -1 position of the YedY signal peptide was substituted by Leucine (Leu) and the mutant protein YedY_A44L was expressed to analyse the effects. The AxA cleavage site highlighted in the signal sequence of *E. coli* pre-YedY is shown in Fig.6.2.2. YedY_A44L was amplified in the pBAD24 and pEXT22 vectors with a C-terminal *his*-tag appended to aid the identification of the proteins. The precursors were expressed in wild-type MC4100 *E. coli* cells after induction with arabinose and IPTG respectively. After 3 hours, cells were fractionated to give cytoplasm, membrane, periplasm samples (C, M, P) and analyzed by SDS-PAGE. Also, the non-mutated YedY expressed in both vectors (as in Chapter 3) are shown here for comparison. The result in Fig.6.2.4 shows that the YedY_A44L precursor in pBAD24 is not processed to any significant degree and the

protein is found almost exclusively in the membrane fraction as the precursor form. This demonstrates that the substitution of the -1 Ala completely blocks the maturation of YedY. Interestingly, somewhat different results are obtained when YedY_A44L is expressed in the pEXT22 plasmid. Here, the protein is found mostly in the periplasm, but as a degradation product that is smaller than the mature size protein. Apparently, maturation is again blocked and another, unidentified protease has cleaved the precursor protein. It is likely that this result reflects the different properties of the plasmids; YedY_A44L is expressed only for a short period from the pBAD24 plasmid whereas expression from pEXT22 is continuous throughout the bacterial growth period, including several hours before induction commences. It was reasoned that this would provide more time for proteases to cleave the unprocessed precursor.

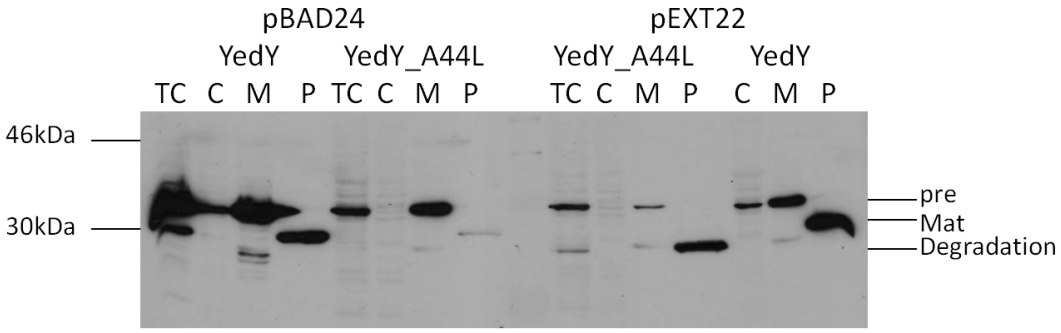


Figure 6.2.4 Translocation of YedY and mutant YedY_A44L in *E. coli*
Mutagenesis of the SP cleavage site blocks the export of pre-YedY cloned in pBAD24, but inhibits the correct process of pre-YedY cloned in pEXT22 by the Tat pathway. Pre-YedY and YedY_A44L are expressed in wild-type *E. coli* MC4100 cells. After induction with arabinose or IPTG for 3 h, cells are harvested and total cell contents (TC) are analyzed together with cytoplasm, membrane and periplasm samples (C, M, P) after osmotic shock and fractionation. Precursor and mature forms of YedY are indicated (pre, Mat), and the wrong processed degradation form is indicated (degradation).

In order to study further how the correct removal of the signal peptide after the translocation is important to the maturation of proteins, I analysed a chimeric protein comprising green fluorescent protein (GFP) linked to the signal peptide of YedY and its cleavage site mutant. The YedY-GFP was constructed from pJDT1 plasmid (Thomas *et al.*, 2001), which expresses TorA-GFP from the pBAD24 plasmid. The signal peptide of TorA in pJDT1 was replaced by the signal peptide of YedY to make the YedY-GFP, and the -1 position Ala was replaced with Leu to obtain YedY-A44L-GFP. The translocation of TorA-GFP and YedY-GFP to the *E. coli* periplasm was shown by confocal microscopy (Figure 6.2.5). It is seen that clear halos of GFP fluorescence are around the *E. coli* MC4100 cells over-expressing TorA-GFP and YedY-GFP because the GFP has been correctly folded before translocation to the periplasm.

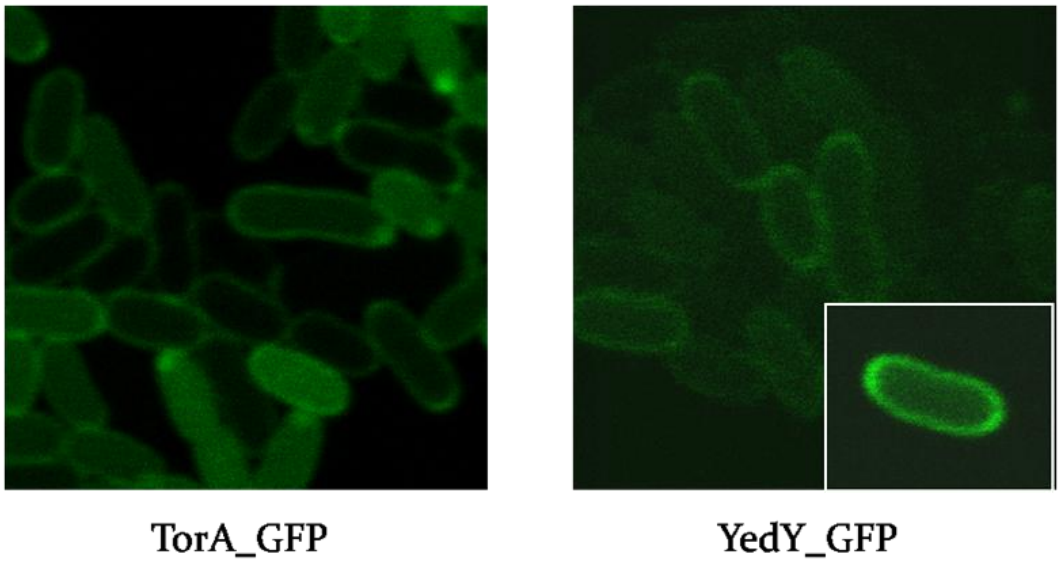


Figure 6.2.5 Translocation of active GFP by the Tat pathway Plasmids expressing TorA-GFP and YedY-GFP were expressed in the *E. coli* MC4100 cells. Cells were analysed by confocal microscopy using the 488nm laser line.

Next, YedY-GFP and YedY-A44L-GFP were expressed for 3 h after the induction of anabinose in wild-type MC4100 and Δ*tatABCED* cells. Cells were collected and fractionated into cytoplasm (C), membrane (M) and periplasm (P) compartments. YedY-GFP was export to the periplasm in the WT cells but denied in the Δ*tatABCDE* cells, wheras YedY-A44L-GFP was blocked to the transportation in both strains (Figure 6.2.6A). The results also displayed that the signal of fluorescence in YedY-GPF expressed in WT is higher than others (Figure 6.2.6B). Consequently, the mutation in the cleavage site retards the maturation of GFP.

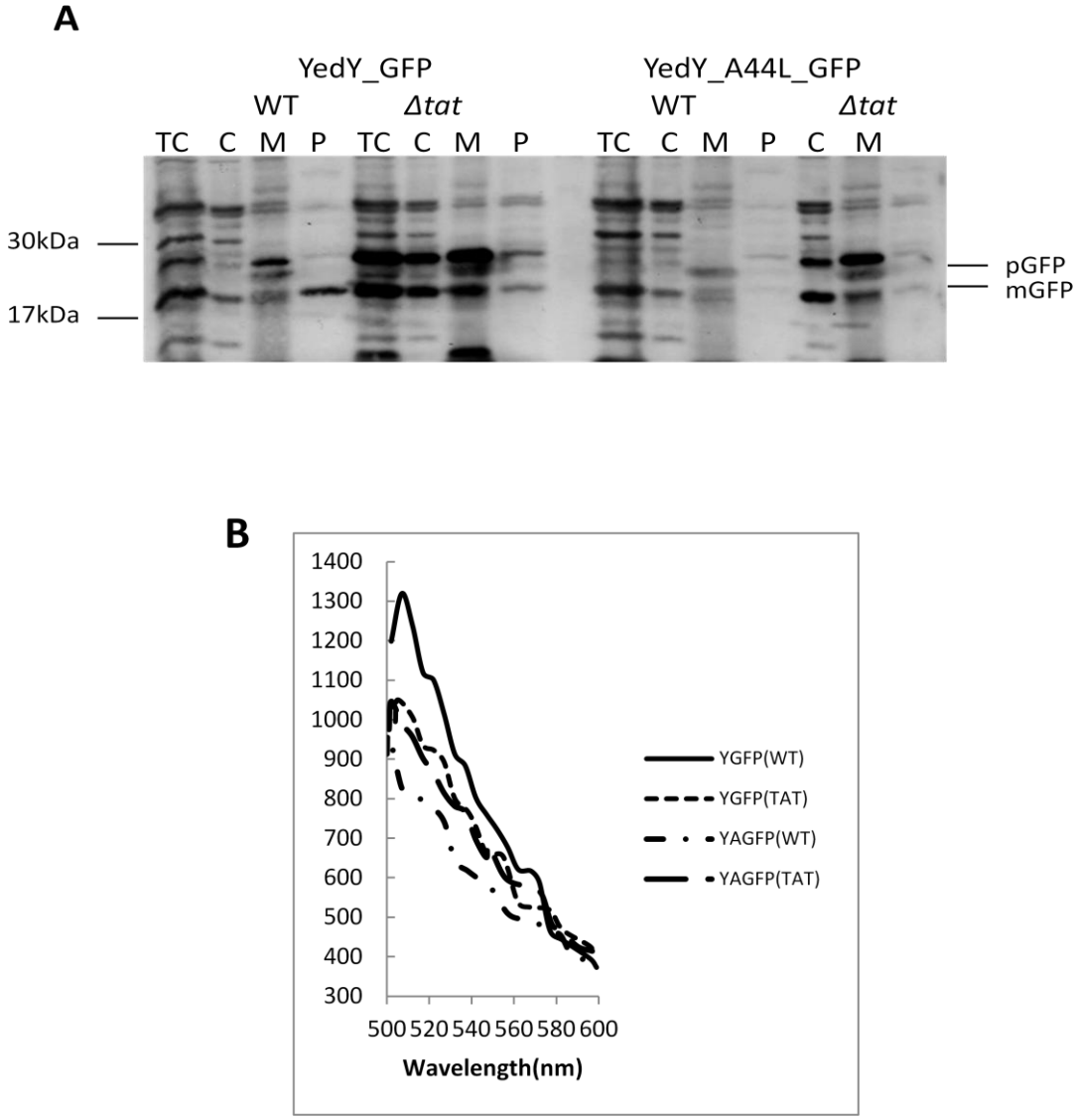


Figure 6.2.6 Translocation of YedY-GFP and YedY-A44L-GFP by the Tat pathway YedY signal peptide is fused with GFP to construct the YedY-GFP. Mutagenesis of the SP cleavage site of YedY-GFP makes the YedY_A44L-GFP. (A)YedY-GFP and YedY_A44L_GFP are expressed in *E. coli* wild-type MC4100 cells (WT) and Δ*tatABCDE* cells (Δ*tat*). After induction with arabinose for 3 h, cells are harvested and total cell contents (TC) are analyzed together with cytoplasm, membrane and periplasm samples (C, M, P) are detected by anti-GFP immunoblotting. Precursor and mature forms of GFP are indicated (pGFP, mGFP). (B) The graph shows the intensity of fluorescence of GFP when GFP fused with signal peptide of YedY and YedY_A44L expressed in WT cells, indicating as YGFP (WT) and YAGFP (WT). Also, YGFP (TAT) and YAGFP (TAT) are represented proteins expressed in Δ*tat* cells.

6.2.4 Unprocessed YedY_A44L is membrane-bound with the mature protein exposed to the periplasm

The presence of the YedY-A44L degradation product in the periplasmic fraction (when expressed from pEXT22) suggests that proteolysis of a membrane-bound precursor has resulted in the release of a soluble fragment into the periplasm. This would suggest that the mature domain has been efficiently transported to the periplasmic side of the membrane. To test this possibility more directly spheroplasts were prepared from cells expressing YedY-A44L from pBAD24 and subjected to proteolysis using proteinase K. Although the SP cleavage site of YedY is mutated to prevent its removal, the mature domain should presumably be translocated to the periplasmic side of the membrane where it would become protease-accessible. The spheroplasts from MC4100 cells expressing YedY-A44L were separated into three parts, one was analysed without further treatment (SP), one was incubated with buffer (K-) or proteinase K (K+), and the last one was treated with freezing-warming cycles following the proteinase K which break the spheroplasts and allow access of the protease to the cell interior. The data show that the YedY-A44L is fully degraded by proteinase K treatment of spheroplasts, with or without freeze-warming, strongly indicating that the mature protein has been successfully transported to the periplasmic side of the membrane. As a control, the Δ*tatABCDE* strain expressing non-mutated YedY was studied in exactly the same manner which showed pre-YedY is resistant to proteinase K treatment of the spheroplasts but is degraded when the spheroplasts were broken, confirming that the pre-YedY in Δ*tatABCDE* is on the cytoplasmic side of the membrane (Figure 6.2.7). Thus, YedY-A44L is exported to the periplasmic side of the membrane by the Tat pathway in wild-type cells, where it remains bound to the membrane. When expressed continuously using the pEXT22 plasmid, a large fragment is susceptible to proteolysis and released into the periplasm.

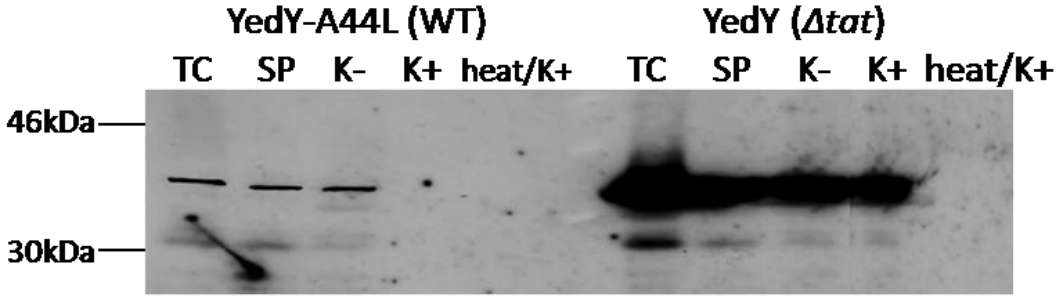


Figure 6.2.7 Location of YedY_A44L in *E. coli* wild-type cells YedY_A44L is amplified in the pBAD24 and expressed in *E. coli* wild-type MC4100 cells (WT) for 3 h after induction. Cells are harvested and total cell contents (TC) are analysed together with spheroplast untreated (K-) o, treated (K+) by proteinase K or treated (heat/k+) by 5×freeze-thaw cycles following proteinase K. (YedY is expressed in Δ*tat* as the positive control)

6.2.5 The signal peptide of YedY-A44L is fully embedded in the plasma membrane

The correct removal of signal peptides after translocation leads to the successful maturation of proteins. YedY-A44L fails to undergo maturation and is found almost entirely in the membrane fraction, which means that the uncleaved signal peptide is attached the membrane. The key question is how deeply/firmly the signal peptide is actually embedded in the membrane? To determine whether it is fully integrated into the membrane or only peripherally bound, methods were used that distinguish between integral and peripheral membrane proteins. The integral membrane proteins span the membrane can be displaced by detergents or some other apolar solvent, and the peripheral membrane proteins are defined as proteins that attached temporarily or more loosely to the lipid bilayer which can be removed by treatment with a polar reagent such as high pH solutions (Singer and Nicolson, 1972; Steck and Yu, 1973; Steck, 1974). Many studies indicated that alkaline carbonate procedure discriminates between peripheral and integral membrane proteins because pH 11.5 to pH 12 can effectively remove peripheral membrane proteins (Steck and Yu, 1973; Fujiki *et al*., 1982). Here, the spheroplasts of YedY-A44L were treated with sodium carbonate pH 11.5 or, as a control, Tris- HCl pH 7.5 (see methods). The single-span protein TatAd which is a component of the *Bacillus subtilis* Tat system also analyzed as a control test because it has been shown to be an integral membrane protein which is partly localised in the cytoplasm (Barnett *et al*., 2008).

The data (Fig 6.2.8) show that YedY-A44L is in fact very firmly integrated into the plasma membrane. The protein is not extracted by treatment with Na_2CO_3, and instead remains almost exclusively in the membrane fraction (M). In the control experiment, TatAd was expressed in *E. coli* cells and the data show that most of the TatAd remains in the membrane but a sizeable proportion is released by Na_2CO_3 and found in the supernatant (Sup). These data therefore demonstrate that the signal peptide of YedY-A44L is integrated within the membrane, rather than bound loosely to the periphery of the bilayer.

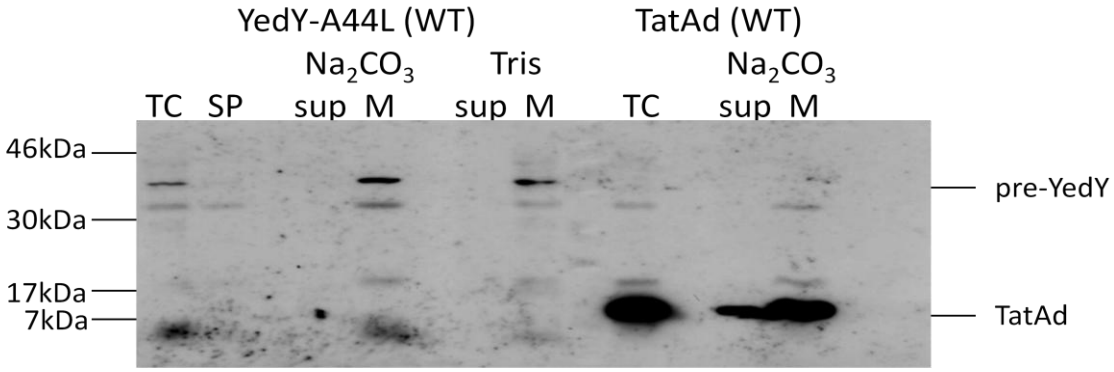


Figure 6.2.8 Location of the uncleaved signal peptide of YedY_A44L YedY_A44L is expressed in *E. coli* wild-type MC4100 cells (WT) for 3 h after induction. Spheroplast is extracted and treated under the alkali conditions (Na_2CO_3 pH11.5). Total cell contents (TC), spheroplast (SP), membrane (M) and supernatant (sup) after carbonate extraction are analyzed by immunoblotting using anti-his antibody. (YedY_A44L is treated by Tris as the negative control; the transmembrane protein TatAd is performed by the same manner as the positive control.)

6.2.6 Pre-YedY maturation is essential for efficient growth of *E. coli* cells

The data have shown that a Tat substrate can be translocated in the absence of processing, but additional studies were carried out to assess the overall importance of Tat substrate maturation. Clear effects were observed on cell growth rates when YedY_A44L was expressed in these experiments. Pre-YedY and YedY-A44L were expressed from pBAD24 in wild-type cells and I measured growth rates before and after induction with arabinose. The graph in Fig.6.2.9 shows that cells expressing non-mutated pre-YedY exhibited a typical growth curve with a characteristic lag phase (0-1 h), exponential phase (1-6 h), and stationary phase (6-10 h). However, expression of YedY-A44L yielded different growth curves in which exponential growth suddenly slows dramatically around the 2 h time point.

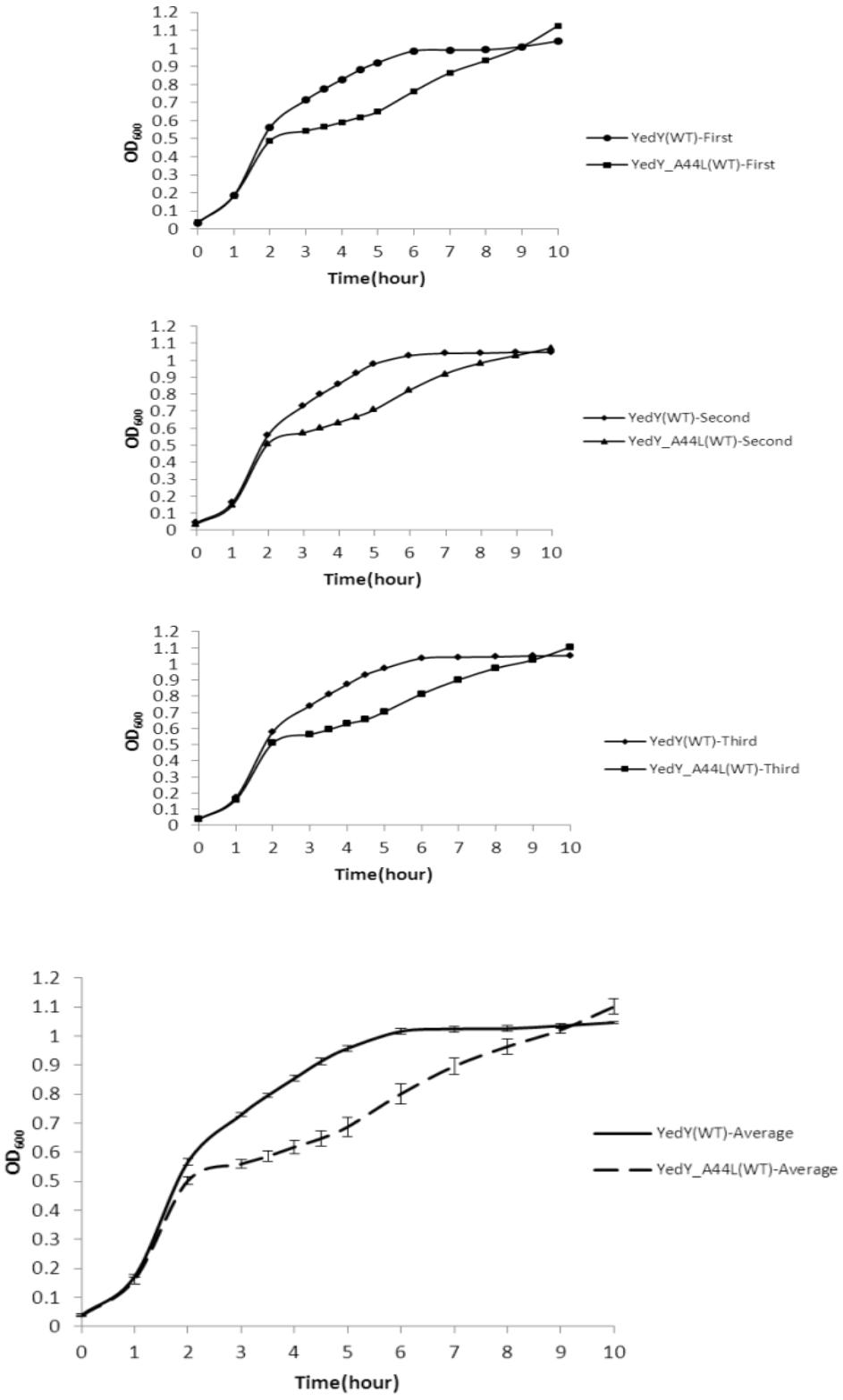


Figure 6.2.9 Growth curve of *E. coli* when expressed YedY_A44L and YedY
YedY and YedY_A44L are amplified in pBAD24 and expressed in *E. coli* wild-type MC4100 cells (WT). The growth rate of cells are measured after induction of synthesis with arabinose and analysed at times (in hours). The graph shows the growth curve of cells by OD_{600} in terms of time (h).

In addition, the expression of YedY_A44L in the period of 5 hours was analysed by immunoblotting (Fig. 6.2.10). The results show that the uncleavable YedY-A44L protein appears around the 2 h time point, after which it is subjected to degradation by the 4 h and 5 h time points, with a variety of degradation products apparent. This almost certainly coincides with the cell growth rate slowing when YedY_A44L was expressed, as shown in Fig. 6.2.9. Growth returns to normal by about the 5 h point, consistent with the disappearance of YedY_A44L. Clearly, the presence of the uncleavable precursor protein is detrimental to cell growth.

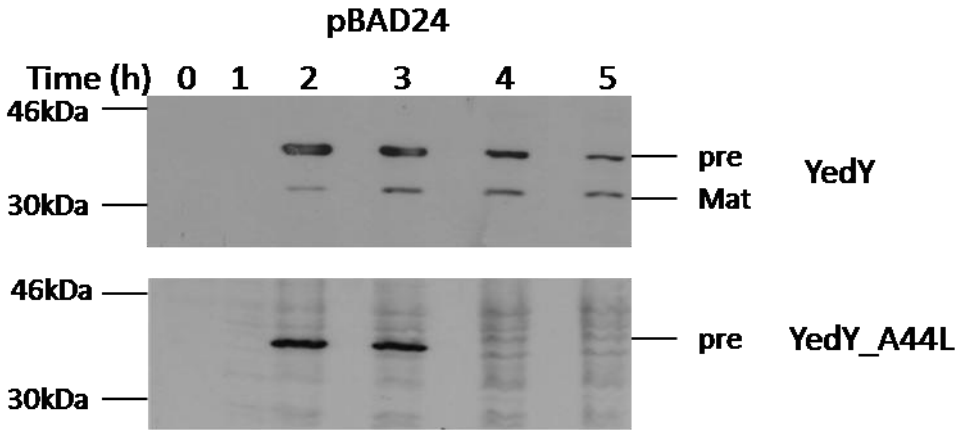


Figure 6.2.10 Translocation of YedY and YedY_A44L in *E. coli* MC4100 cells YedY_A44L cloned in pBAD24 is degraded in wild-type cells. YedY and YedY_A44L are amplified in pBAD24. Cells were analyzed at times (in hours) after induction and detected by the immunoblotting using anti-his antibody, with precursor and mature forms of the protein indicated (pre, Mat).

6.2.7 Membrane-integration of YedY-A44L occurs at a late stage of the overall translocation process, or after its completion

Studies have indicated YedY-A44L is firmly integrated into the membrane after translocation, and the expression of YedY-A44L affects the growth of cells until it is degraded, we considered it possible that the precursor protein had become stuck in the Tat translocase during the translocation process. To elucidate the relationship between translocation and maturation, YedY_A44L and other Tat substrates co-expressed to examine whether the former blocked export of the latter.

YedY_A44L was amplified with a c-terminal *his*-tag in both pBAD24 and pEXT22 vectors. Correspondingly, YedY was cloned in pEXT22 and TorA-GFP was cloned in pBAD24. YedY-A44L and another Tat substrate such as YedY or TorA-GFP were co-expressed in wild-type MC4100 cells. In the Figure 6.2.11A, YedY-A44L in pBAD24 was produced using arabinose for 3 hours, and then pre-YedY in pEXT22 was induced with IPTG for a further 2.5 hours. The longer expression time of YedY-A44L should ensure that YedY_A44L had indeed become stuck in the translocation channel if this could occur. The data show that some precursor protein is present in the membrane fraction (M) and this is probably mostly YedY-A44L, perhaps with some unprocessed pre-YedY that has not yet been exported. However, mature-size YedY is present in the periplasm, which suggests that the export process has not been badly affected by the prolonged synthesis of YedY-A44L.

Alternatively, YedY-A44L in pEXT22 was expressed for 3 hours and then expression of a chimeric Tat substrate, TorA-GFP, was induced using pBAD24. The TorA-GFP substrate comprises the signal peptide of TorA linked to GFP, and this protein has been shown to be exported by the Tat pathway (Thomas *et al*., 2001). The results in the Figure 6.2.11B resemble those shown above in Fig. 6.2.11A: the YedY-A44L precursor protein is found in the membrane and no mature protein can be detected, and mature-size GFP is present in the periplasmic fraction. This shows that extended prior expression of YedY-A44L does not block the export of TorA-GFP. In conclusion, the Tat translocon is not saturated with unprocessed YedY-A44L, and the protein has clearly completed interaction with the translocon and has moved to a different location.

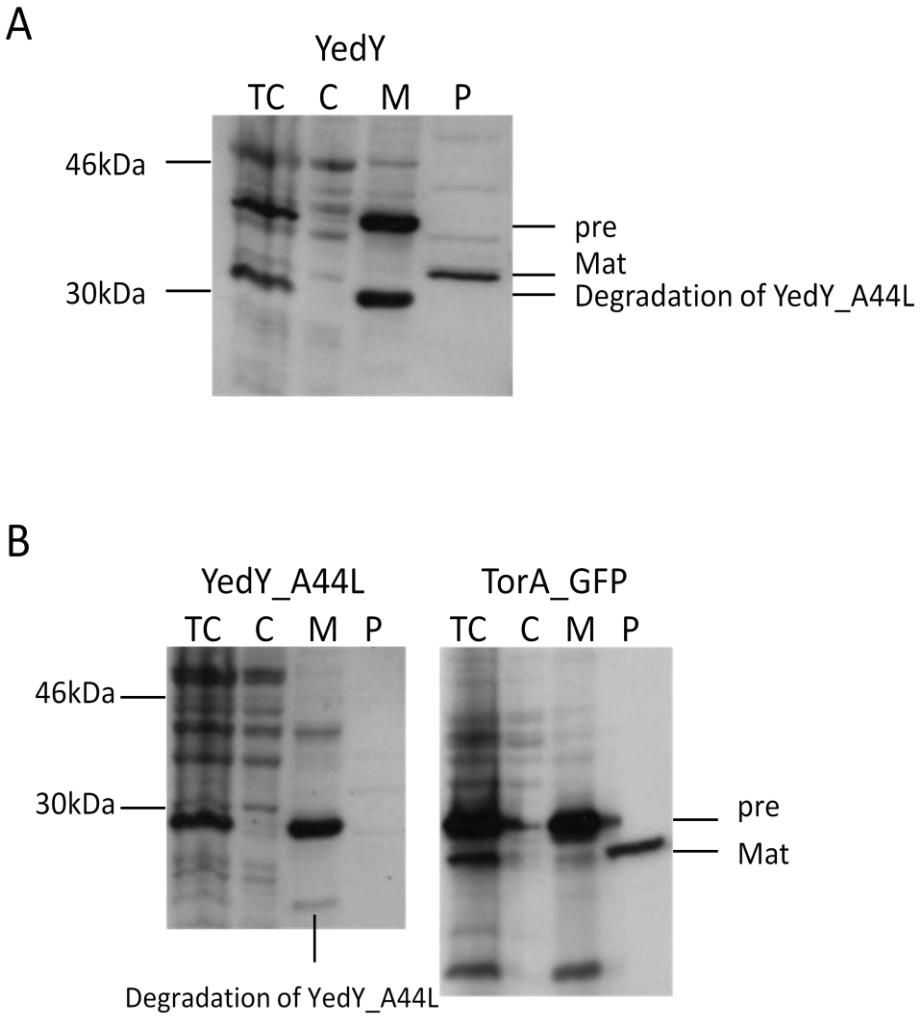


Figure 6.2.11 Translocation of YedY and TorA-GFP after the expression of YedY_A44L in *E. coli* MC4100 cells. A. YedY_A44L in pBAD24 was firstly expressed in *E. coli* wild-type MC4100 cells for 3 h, and followed by the expression of YedY in pEXT22 for 2.5 h. Total cell contents (TC) are analyzed together with cytoplasm, membrane and periplasm samples (C, M, P) and detected by immunoblotting using anti-his. Precursor, mature forms and degradation are indicated (pre, Mat, degradation). **B**. YedY_A44L in pEXT22 was firstly expressed in wild-type cells for 3 h, and then expressed TorA-GFP in pBAD24. Precursor of YedY_A44L and TorA_GFP are analyzed, both proteins are detected by immunoblotting using anti-his and anti-GFP respectively. Precursor and mature forms of TorA-GFP are indicated (pre, Mat).

6.3 Discussion

The signal peptides of most Tat substrates bear a highly conserved consensus motif SRRxFLK and the invariable twin-arginine residues gave rise to the name of the pathway (Berks *et al*., 2002; Sargent *et al*., 2002).

The study of Tat substrate SufI using cross-linking approach has revealed the possible mechanism of the specific recognition between substrate and translocase (Alami *et al*., 2003). Importantly, no interaction of signal peptide-TatBC complex is observed when both arginine residues are substituted by lysine, implying the importance of RR motif in substrate recognition. On the other hand, the replacement of the amino acid in the hydrophobic core of the signal peptide of TMAO reductase (TorA) has indicated that the Tat signal peptide is required to bind the Tat proofreading chaperone (Buchanan *et al*., 2008).

Various site-directed mutagenesis studies have shown the importance of the twin-arginine for Tat-dependent export. In *E. coli*, mutation of both arginines completely abolishes export by the Tat pathway but the conservative substitution of a single Arg sometimes can be tolerated. For example, the export of the cofactor-less Tat substrate SufI was completely blocked when the twin arginines were replaced with twin lysines, but the transpotation was just reduced when a single arginine was substituted with lysine (Stanley *et al*. 2000; Yahr and Wickner 2001). Similarly, the substitution of both arginines of TMAO reductase (TorA) with lysine completely inhibited Tat translocation. However, the second conserved arginine of TorA appears to be more significant for targeting by the Tat pathway, since the replacement of first arginine still retains periplasmic TMAO reductase activity whereas the export is effectively abolished for the second arginine (Delisa *et al*., 2002). In contrast, for some other Tat signal peptides such as *E. coli* CueO, *Zymomonas mobilis* glucose-fructose oxidoreductase (GFOR), *Pseudomonas stutzeri* nitrous oxide reductase, the Tat transloction is completely blocked by the single substitution of arginine (Dreusch *et al*., 1997; Halbig *et al*., 1999; Stanley *et al*., 2000).

In the first part of this chapter, evidence was shown that a single replacement of either arginine in the signal peptide of YedY completely abolished the export, but did

not affect the turnover of the Arg194 mutant. It was demonstrated that both of arginines are important for translocation, and the RR motif may not be involved in the proofreading process. According to the previous studies of Tat substrates, TorA and DmsA, the proofreading processes are performed by the association between their signal peptides or the core of mature domain and the specific chaperones. Thus, further studies are needed to elucidate whether other parts of the signal peptide play a role in the proofreading of YedY.

Proteins are targeted to the Tat translocation pathway by signal peptides that contain three distinct domains following an Ala–X–Ala cleavage site (Berks, 1996). The presence of short-chain residues at the -3 and -1 position which compose of the cleavage site are essential to be recognized by the unique signal peptidase resulting in the maturation of exported proteins. Numerous studies on Sec-type signal peptides and some mutagenesis studies of Tat substrates in plant thylakoids have confirmed the importance of the -3 and -1 position residues, and in most case the substitution of them completely blocks the processing at the correct site, and may lead to false cleavage at an alternative site (Dalbey and Von Heijne, 1992; Shackleton and Robinson, 1991., Fikes *et al*., 1990). In *E. coli*, it appears that a single specific enzyme, leader peptidase (LepB) is responsible for the maturation of most exported precursor proteins.

In my study, the -1 Ala residue of the YedY signal peptide was substituted by Leu and the data show that this leads to an essentially complete block in maturation. The result is an accumulation of membrane-bound precursor protein when YedY-A44L is synthesised using the pBAD24 plasmid, or the appearance of a degradation product in the periplasm when pEXT22 is used for expression studies. It is not clear why the mutated precursor is degraded to yield a specific degradation product when synthesised using pEXT22, yet not with pBAD24, but this presumably reflects slightly differing physiological states when the cells are induced using IPTG in one case or arabinose in the other.

The rapid induction kinetics of the pBAD24 system provide an opportunity to study the maturation of Tat precursors in the context of the overall translocation process, and several points emerge from this study.

Firstly, the data show that the mature domain of pre-YedY is exported across the bacterial plasma membrane, but the maturation is inhibited. This is supported by the appearance of a soluble degradation product after synthesis using the pEXT22 plasmid, and by the degradation of the membrane-bound precursor protein when proteinase K is added to spheroplasts after synthesis using the pBAD24 plasmid.

Secondly, there is evidence that the uncleaved precursor is in a transmembrane state. The study has exhibited that YedY-A44L protein is absent from the periplasm and fully resistant to extraction by alkaline carbonate. Therefore, the signal peptide is integrated into the membrane bilayer, rather than associating peripherally. While the signal peptide would be expected to have an affinity for membranes, there is no evidence that Tat-dependent precursor proteins normally interact with membranes with such a high affinity outside the translocation event. In contrast, they are normally soluble in both bacteria and chloroplasts, with only a slight affinity for membranes (Thomas *et al*., 2001; Alami *et al*., 2003; Mould and Robinson, 1991).

In addition, these data suggest a 'loop' type mechanism, whereby the N-terminus of signal peptide binds to the *cis*-side of the membrane and remains at this location while the passenger protein is translocated through the translocon channel. There is no evidence that the signal peptide leaves the membrane at any point. Loop mechanisms have been shown to operate during Sec-dependent translocation and there is evidence for a basically similar process in the transport of thylakoid Tat substrates (Shaw *et al*., 1988; Kuhn *et al*., 1994; Fincher *et al*., 1998). Here, my study provides the first clear data on this subject in bacteria. In this sense, the Tat translocation process appears to resemble the Sec-type mechanism, where LepB plays an important role in releasing the mature protein from a membrane-bound precursor (Dalbey and Wickner, 1985; Fikes and Bassford, 1987). However, it should be emphasized that, in other key respects, the Sec and Tat systems appear to use completely different translocation mechanisms.

This process of lateral exit from the transoclation channel has been examined in several studies on the thylakoid Tat system. The Rieske Fe/S protein, the cytochrome b_6/f complex in chloroplasts, is transported by the Tat translocation pathway, but the

signal peptide anchors the protein from the lumenal side to the thylakoid membrane instead of being cleaved (Molik *et al*., 2001; Karnauchov *et al.*, 1997). Also, the study of 16/23 precursor protein in chloroplasts suggested that the Tat substrate can be laterally released from membrane to the lipid bilayer prior to processing implying the independence of Tat transport and terminal maturation (Frielingsdorf and Klosgen, 2007).

Overall, the YedY_A44L protein contains only a single amino acid substitution, and it is therefore likely that the 'natural' translocation pathway for YedY has been observed. It remains possible that some Tat substrates may undergo complete transport to the periplasm without entering the bilayer laterally, just as a Sec substrate was found be located in the periplasm instead of the membrane, when the signal peptide was rendered less hydrophobic (Fikes and Bassford, 1987). However, the evidence suggests that at the very least, lateral exit from the Tat translocation channel can occur with some ease.

According to the data, it is possible to propose an overall Tat translocation process in Figure 6.3. After the posttranslation process, the mature domain of the newly synthesized precursors is folded. Then the signal peptide binds to the TatBC subunits of the TatABC complex, triggering recruitment of the separate TatA complex and initiation of translocation. When the passenger protein completes the translocation, the signal peptide is still bound at the *cis* face of the membrane in a loop configuration proposed above. It is presumed that the signal peptide never reaches the *trans*- side of the membrane and the precursor protein is released laterally into the lipid bilayer. Finally, the processing peptidase removes the signal peptide of the precursor protein to release the mature protein. If this is prevented, as is the case in this study, the outcome may involve proteolytic cleavage of the exposed mature domain.

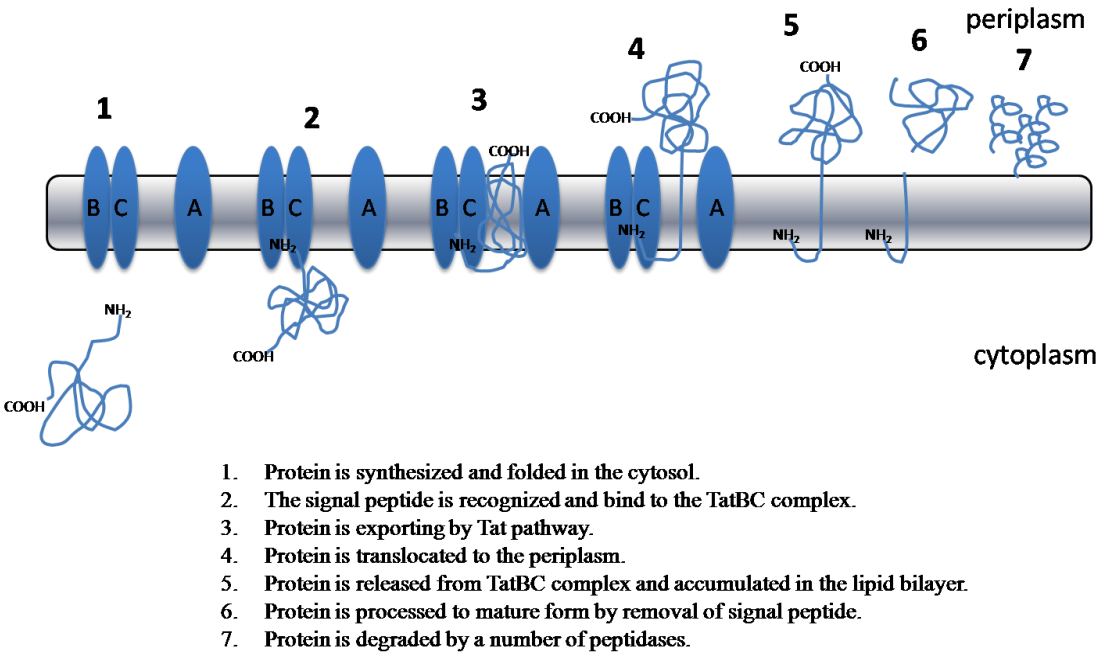


Figure 6.3 Model for the Tat-dependent export of YedY Pre-YedY is synthesized in the cytosol (step 1) and binds post-translationally to the TatBC subunits of the TatABC complex (step 2). This triggers recruitment of the separate TatA complex and translocation of the mature protein through the translocation channel ensues (step 3). In step 4, translocation of the mature domain has been completed and the signal peptide spans the membrane. Step 5 depicts the precursor protein exits laterally from the translocation channel into the bilayer, remaining in a transmembrane form. In the normal course of events, signal peptidase cleaves the signal peptide and releases the mature protein into the periplasm (step 6), whereas the exposed mature domain can be subjected to proteolysis if maturation is prevented (step 7).

Lastly, it is demonstrated that Tat substrate maturation is important for cell viability. The accumulation of unprocessed YedY-A44L has seriously adverse effects on growth rate and this clearly indicates that the presence of unprocessed precursors in the membrane causes problems. Previous studies suggested that the repression of the synthesis of the leader peptidase decreased the growth rate of cells and led to the accumulation of precursor proteins, however, the growth rate and the processing of precursors can be restored by the expression of the leader peptidase (Date, 1983; Dalbey and Wickner, 1985). Thus, our data serve to underline the importance of precursor processing in the overall Tat export pathway.

Chapter 7

Final Discussion

The twin-arginine translocation (Tat) pathway is uniquely able to export folded proteins across tightly sealed bacterial membranes, relying only on the membrane proton motive force (PMF) (Robinson and Bolhuis, 2004; Muller, 2005). Translocation is initiated by the interaction between the invariant twin-arginine signal peptide of substrates and the Tat translocase (Alami *et al*., 2003). A large number of Tat substrates which contain redox cofactors such as FeS, NiFe, copper and molybdopterin centers are required to assembly and correctly fold in the cytoplasm prior to export (Berks, 1996). Over the 20 years, Tat-related research has helped to clarify this transport system, moreover, intensive studies of medical and biotechnological applications of Tat system are ongoing (Sargent, 2007).

The Tat pathway in Gram-negative bacteria such as *E. coli*, is minimally comprised of three integral membrane proteins, TatA, TatB and TatC (Sargent *et al*., 1998; Sargent *et al.*, 1999). They organise into two distinct functional complexes: a 370 kDa Tat(A)BC complex which was identified as the twin-arginine signal peptide recognition site, whereas TatA complex is proposed to form a protein-conducting channel of varying size (Sargent, 2001). The exact working mechanism of Tat transport is still not unraveled fully but, based on different interaction studies, a current model for a 'Tat-transport cycle' model has been proposed.

Besides studies of the structure and mechanism of Tat systems, an important functional aspect, proofreading, is poorly understood. Comprehensive studies have shown that Tat substrates that either bind cofactors or oligomerise in the cytoplasm need to be prevented from interacting with the export machinery until their assembly has been completed. This essential process of "proofreading"—the term used here of Tat transport relates to the quality-control system, ensures that the *E. coli* Tat pathway would only undertake protein translocation when the substrate protein was in the folded state (Berks *et al.*, 2005). The Tat pathway has an inherent ability to preferentially transport folded proteins, but a more rigorous process of chaperone-mediated proofreading monitors the export for some substrates using additional safeguards (DeLisa *et al.*, 2003). Recently, the study of two FeS Tat substrates, NrfC and NapG, has indicated that substrates were completely inhibited in export in wild-type cells if FeS cluster ligands were mutated, and the mutants were rapidly degraded (Matos *et al*., 2008).

The work presented in this thesis sought to gain further understanding of the Tat pathway in *E. coli* using biochemical, bioimaging and biophysical techniques. The content of thesis includes two main aspects: the proofreading of molybdoproteins in chapter 3, 4 and 5; and the importance of the Tat signal peptide as shown in chapter 6.

7.1 Proofreading of the molybdopterin substrate YedY

A large proportion of redox enzymes exported by the Tat pathway belong to the molybdenum-containing group of proteins which contain molybdopterin centres as the cofactor. In general, molybdenum is coordinated to the dithioline group and made up of the most simple cofactor molybdopterin (MPT), and in most cases, the variant of MPT is modified by attachment of GMP to generate molybdopterin guanine dinucleotide (MGD) or *bis*-MGD (Berks *et al.*, 2003). YedY, a soluble periplasmic protein that binds MPT as a single cofactor, bears a twin-arginine signal peptide for export by the Tat pathway (Sargent *et al.*, 1998; Weiner *et al.*, 1998). YedY represents the first structural characterization of MPT in prokaryotes, and the overall architecture of YedY showed the construction of MPT which implicated some important ligands in the assembly of the cofactor (Loschi *et al.*, 2004). In order to investigate the proofreading of molybdoproteins in the Tat pathway, YedY was a novel candidate substrate to analyze.

The data showed that the substitution of ligands (Cys102, Glu48 and Arg194) coordinated with MPT was sufficient to completely block export of YedY, and MPT-deficient proteins were degraded at varying rates. The turnover of rejected proteins was observed in wild-type cells as well as *tatA/tatE/tatAE* deleted strains when the experiment was performed in the arabinose-inducible pBAD24 expression system. However, the degradation was hard to detect in the IPTG-inducible pEXT22 expression system (results were not shown in the thesis). Because expression in pBAD24 leads to different production, aggregation and misfolding levels (Lindenstrauss *et al*., 2010), the degradation of malfolded proteins was suggested to be by Tat-independent proteolysis. Although rejected proteins are degraded in this way, it is still possible that the turnover is triggered by the Tat pathway, since the overall pathway closely monitors folding of proteins before transport.

To understand how the assembly of cofactor affects the folding of a protein, and how the Tat pathway recognizes the incorrectly folded protein and inhibits export, YedY and three mutants were purified to measure the levels of molybdenum in the proteins and to analyze the change of secondary structure. CD spectroscopic analysis showed a typical α-helical structure in pre-YedY, but the observed ellipticity (degree) at wavelength of three mutants was shifted which indicated the proteins adopt different conformations. Pre-YedY comprises by 47% α-helix and 24% β-strand, as determined by the fitting program CDSSTR using the SP175 reference set. Compared with crystal structure of YedY at 2.5- Å shown in PBD (codes *1XDQ*), the higher α-helical content is speculated to reflect the presence of the N-terminal signal peptide of pre-YedY. In addition, the insertion of MPT was directly measured using the measure of molybdenum in proteins by ICP-MS. The cytoplasmic pre-YedY and periplasmic mature YedY contain detectable levels of molybdenum, and pre-YedY has less molybdenum than the mature form which implies a proportion of pre-YedY is in the slow cofactor insertion process to form the active molybdpterin apo-protein in the cytosol. In contrast, the three mutants totally lack molybdenum.

In short, the proofreading system in the Tat pathway can recognize the absence of cofactor in three mutants (Cys102, Glu48 and Arg194) or the overall secondary structural changes, and prevent the export process. Further structural studies would be useful to determine exactly how different the structures of the mutants are, for example using techniques that measure tertiary structure changes.

7.2 The REMP of YedY and chaperone-mediated proofreading in the Tat pathway

In chapter 3, the absence of cofactor resulted in the inhibition of export of YedY by the Tat pathway. The question is: does Tat translocase itself or other accessory molecules and complexes assist in sensing the unfolded or misfolded state of YedY and rejecting translocation? There is evidence for more complex and stringent quality control systems for some Tat substrates specific cytosolic chaperones (DeLisa *et al.*, 2003, Robinson *et al.*, 2011). These cytosolic chaperones were termed REMPs (redox enzyme maturation proteins), and most of them are either experimentally proved or strongly suggested to affect the Tat translocation of redox substrates

(Turner *et al.*, 2004; Jack *et al.*, 2004). Studies showed that the export of TMAO reductase (TorA) was blocked in the absence of REMP TorD but remained the efficiency if DmsD deleted (chapter 5, Palmer *et* al., 2004). Similarly, the REMP NapD was essential in the translocation of NapA.

In this study, the translocation and proofreading of YedY were examined in the *torD/dmsD/napD* deleted strains because DmsD and TorD are potential REMPs of YedY on the basis of sequence alignments and indirect protein-protein interaction studies (Chan *et al.*, 2008). Importantly, TorD, DmsD and NapD are not required for the export of YedY, nor do they play a role in the degradation of mutant forms.

Protein-protein interaction studies using affinity chromatography, cross-linking and label transfer did not detect any definite specific interactors with YedY that may be critical in the maturation process. Nevertheless, several accessory proteins were found to transient or weakly interact with YedY, such as the general folding chaperone GroEL and the translation elongation factor tufA, which was shown to associate with DmsD during the process of synthesis, folding and translocation of DmsA (Li *et al*., 2010).

Hence, this study suggests that TorD, DmsD and NapD are not the specific REMPs of YedY as predicted. Indeed, YedY may not have a specific REMP to assist translocation and proofreading in the Tat pathway, and those chaperones or small proteins which normally bind to REMPs might directly interact with YedY to asssit in the synthesis and maturation.

7.3 The importance of the Tat signal peptide of YedY in the translocation and maturation process

The Tat signal peptide bears a highly conserved consensus motif SRRxFLK and the invariant twin-arginine residues are important for Tat-dependent translocation (Berks *et al*., 2002; Sargent *et al*., 2002). In *E. coli* various site-directed mutagenesis studies have indicated that mutation of both arginines completely abolishes the export by Tat pathway but the conservative substitution of a single Arg sometimes can usually be tolerated (Stanley *et al*., 2000; Delisa *et al*., 2002; Halbig *et al*., 1999). In chapter 6,

YedY export was shown to be totally dependent on the twin-arginine motif in the signal peptide and the single replacement of arginine in the signal peptide of YedY completely abolished the translocation. However, the degradation of the Arg194 mutant was still detected if the single substitution within RR motif, implying the RR motif may not be essential for proofreading of YedY.

The cleavage of the signal peptide is a late and essential step in the Tat-dependent protein export pathway in *E. coli.* The signal peptide is generally removed during or just after translocation by a specific processing peptidase (von Heijne, 1990). It was shown that the removal of the signal peptide of YedY is not necessary for the translocation of the mature protein across the membrane, but it is essential for the maturation of exported pre-YedY and release into the periplasm. Moreover, the failure to cleave the signal peptide of YedY leads to the accumulation of precursor proteins and adversely affects the growth rate of cells. It is suggested that the accumulation affects the permeability and breaks the integrity of the membrane which might stimulate a stress response. Importantly, the translocation of uncleaved YedY by the Tat translocon does not jam the system, which implies that the signal peptide of the exported substrate had been released from the Tat translocase to the lipid bilayer. Overall, the data suggest an overall Tat translocation process including TatABC constituted translocation mechanism and a 'loop' type mechanism during the maturation process.

7.4 The proposed mechanism of translocation and proofreading by the Tat pathway in *E. coli*

The combination of the understanding of Tat system so far and all of results presented here, it can propose a comprehensive model for translocation and proofreading by the Tat pathway in *E. coli.* In Figure 7.1, substrates are synthesized with an N-terminal Tat signal peptide which is required to interact with the TatBC receptor to target the Tat tanslocase. Prior to export, substrates need to insert cofactors if necessary and fold in the correct conformation with the help of REMPs. The folded precursor proteins are transported by the Tat pathway but the signal peptides still associate with the membrane. The final maturation of substrates in the periplasm is completed by the cleavage of the signal peptides.

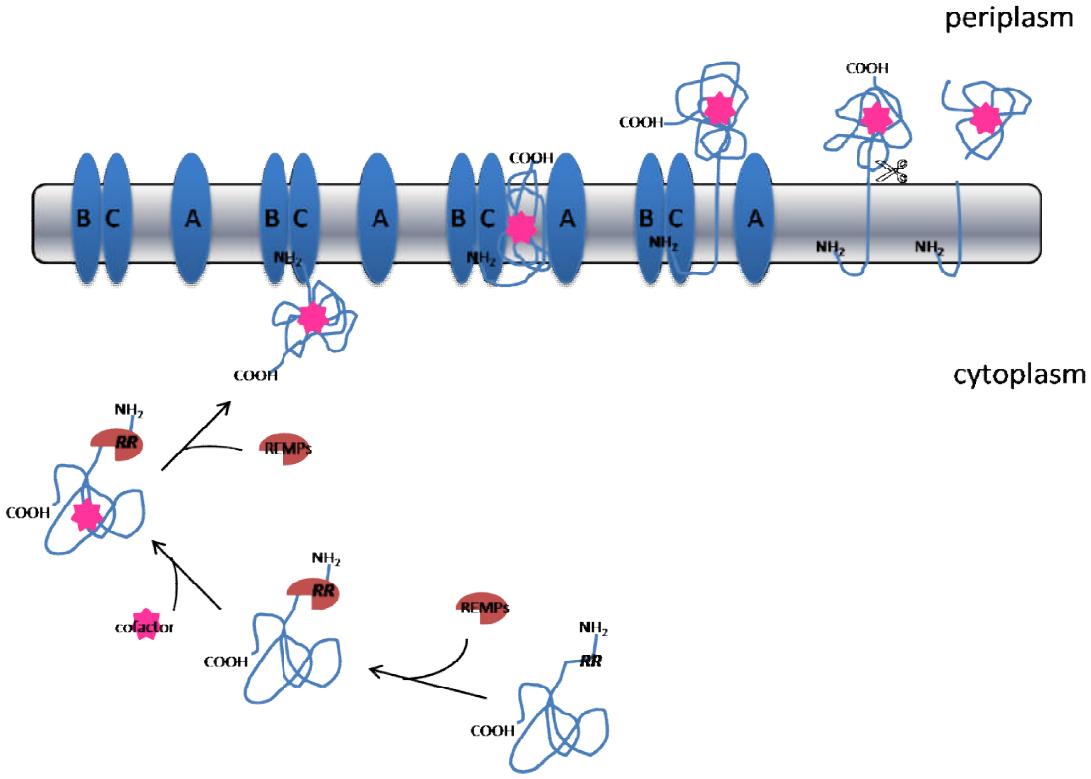


Figure 7.1 Proposed mechanism of translocation and proofreading by the Tat pathway in *E. coli* Protein is synthesized with N- signal peptide bearing RR motif. REMPs associate with SP to keep the unfolded state of the protein and competent of cofactor insertion. After the assembly of cofactor and folding of the protein, REMPs are released into cytosol. Then, the correct folded protein is targeted to the Tat pathway by the interaction between RR motif and TatBC compex. The TatA is composed of the Tat translocon to export the pre-protein. Thereafter, the pre-protein exits laterally from the translocation channel into the bilayer, remaining in a transmembrane form. At last, signal peptidase cleaves the signal peptide and releases the mature protein into the periplasm.

7.5 Future work

Despite the fast pace of research in the Tat pathway, there still remain many challenges to be resolved such as the translocation mechanism, energy utilisation and proofreading activity. Some studies have demonstrated that chaperone-mediated proofreading is essential for Tat substrates which associate with redox enzyme maturation proteins (REMPs), monitoring the assembly of cofactor and correct folding process.

The work presented in this thesis is focused on the proofreading of molybdopterin proteins in *E. coli*. It has indicated that the absence of cofactor seriously affects the secondary structure of YedY, and the structural changes lead to proofreading which inhibits the export of misfolded substrates by the Tat pathway. So far CD and ICP-MS have been used to compare the structures between the non-mutated substrate and three different mutants, and the results indicate a correlation of the proofreading event and structural changes. However, work is required to understand how the Tat system recognizes incorrectly folded/assembled proteins. For example, does it sense the lack of cofactor, or does it only detect aspects of the folded state of the proteins?

Another important future direction is the determination of the factors which play a critical role in the proofreading. Cytosol chaperone or REMPs have been identified as essential factors to help monitor the folding process of Tat substrates. Protein-protein interaction technologies are useful to examine if the signal peptides or any mature domains of proteins interact with other accessory molecules during the folding which may belong to the proofreading. However, the natures of these interactions, and how they control when the protein is recognized as ready for export, are not well understood.

Chapter 8

References

Ajm, D. (1992) Precusor protein translocation by the *Escherichia coli* translocase is directed by the protonmotive force. *EMBO J*, 847–853.

Akiyama, Y. and Ito, K. (1987) Topology analysis of the secy protein, an integral membrane-protein involved in protein export in *Escherichia-coli*. *EMBO J*. **6**, 3465-3470.

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

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

Alder, N. and Theg, S. (2003) Energetics of protein transport across biological membranes: A study of the thylakoid Delta pH-dependent/cpTat pathway. *Cell*. **112**, 231-242.

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

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

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

Bakker, E. and Randall, L. (1984) The requirement for energy during export of beta-lactamase in *Escherichia.coli* is fulfilled by the total protonmotive force. *EMBO J*. **3**, 895-900.

Barnett, J., Eijlander, R., Kuipers, O. and 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*. 2534-2542.

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

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

Basha, E., Lee, G., Breci, L., Hausrath, A., Buan, N., Giese, K. and Vierling, E. (2004) The identity of proteins associated with a small heat shock protein during heat stress in vivo indicates that these chaperones protect a wide range of cellular functions. *J. Biol. Chem.* **279,** 7566-7575.

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

Beloin, C., Valle, J., Latour-Lambert, P., Faure, P., Kzreminski, M., Balestrino, D., Haagensen, J., Molin, S., Prensier, G., Arbeille, B. and Ghigo, J. (2004) Global impact of mature biofilm lifestyle on *Escherichia coli* K-12 gene expression. *Mol. Microbiol*, **51**, 659-674.

Ben-Zvi, A. and Goloubinoff, P. (2001) Review: Mechanisms of disaggregation and refolding of stable protein aggregates by molecular chaperones. *J. Struc.Biol.* **135**, 84-93.

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

Berks, B., Palmer, T. and Sargent, F. (2003) The Tat protein translocation pathway and its role in microbial physiology. *Advances in Microbial Physiology.* **Vol 47,** 187-254.

Berks, B., Palmer, T. and Sargent, F. (2005) Protein targeting by the bacterial twin-arginine translocation (Tat) pathway. *Current Opinion in Microbiology*. 174-181.

Bernhardt, T. and de Boer, P. (2003) The *Escherichia coli* amidase AmiC is a periplasmic septal ring component exported via the twin-arginine transport pathway. *Mol. Microbiol*. **48**, 1171-1182.

Berthelmann, F. and Bruser, T. (2004) Localization of the Tat translocon components in *Escherichia coli*. *Febs Lett*. **569**, 82-88.

Blattner, F., Plunkett, G., Bloch, C., Perna, N., Burland, V., Riley, M., ColladoVides, J., Glasner, J., Rode, C., Mayhew, G., Gregor, J., Davis, N., Kirkpatrick, H., Goeden, M., Rose, D., Mau, B. and Shao, Y. (1997) The complete genome sequence of *Escherichia coli* K-12. *Science*. 1453-&.

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

Brundage, L., Hendrick, JP., Schiebel, E., Driessen, AJ. and Wickner, W. (1990) The purified *E.coli* integral membrane protein SecY/E is sufficient for reconstitution of SecA-dependent precursor protein translocation. *Cell*, 649–657.

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

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

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

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

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

Broeze, R., Solomon, C. and Pope, D. (1978) Effects of low temperature on in vivo and in vitro protein synthesis in *Escherichia coli* and *Pseudomonas* fluorescens. *J. Bacteriol*, 861–874.

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

Bruser, T. and Sanders, C. (2003) Hypothesis-review - An alternative model of the twin arginine translocation system. *Microbiol. Research*. **158**, 7-17.

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

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

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

Bukau, B. and Horwich, A. (1998) The Hsp70 and Hsp60 chaperone machines. *Cell*. **92**, 351-366.

Burns, R. O. and Demoss, R. D. (1962) Properties of tryptophanase from *Escherichia coli*. *Biochim.Biophys. Acta*. 233-244.

Butland, G., Peregrin-Alvarez, J., Li, J., Yang, W., Yang, X., Canadien, V., Starostine, A., Richards, D., Beattie, B., Krogan, N., Davey, M., Parkinson, J., Greenblatt, J. and Emili, A. (2005) Interaction network containing conserved and essential protein complexes in *Escherichia coli*. *Nature*. **433**, 531-537.

Cannon, K., Or, E., Clemons, W., Shibata, Y. and Rapoport, T. (2005) Disulfide bridge formation between SecY and a translocating polypeptide localizes the translocation pore to the center of SecY. *J.Cell.Biol*. **169**, 219-225.

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

Chan, C., Chang, L., Rommens, K. and Turner, R. (2009) Differential Interactions between Tat-Specific Redox Enzyme Peptides and Their Chaperones. *J.Bacteriol*. 2091-2101.

Chanal, A., Santini, C. and Wu, L. (1998) Potential receptor function of three homologous components, TatA, TatB and TatE, of the twin-arginine signal sequence-dependent metalloenzyme translocation pathway in *Escherichia coli*. *Mol.Microbiol*. ***30***, 674-676.

Cline, K., Ettinger, W. and Theg, S. (1992) Protein-specific energy-requirements for protein-transport across or into thylakoid membranes - 2 lumenal proteins are transported in the absence of ATP. *J.Biol.Chem*. **267**, 2688-2696.

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

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

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

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

Czjzek, M., Dos Santos, J., Pommier, J., Giordano, G., Mejean, V. and Haser, R. (1998) Crystal structure of oxidized trimethylamine N-oxide reductase from *Shewanella massilia* at 2.5 angstrom resolution. *J.Mol.Biol*. **284**, 435-447.

Dabney-Smith, C. and Cline, K. (2009) Clustering of C-Terminal Stromal Domains of Tha4 Homo-oligomers during Translocation by the Tat Protein Transport System. *Mol.Biol.Cell.* **20**, 2060-2069.

Dalbey, R. and Wickner, W. (1985) Leader peptidase catalyzes the release of exported proteins from the outer surface of the *Escherichia-coli* plasma-membrane. *J.Biol.Chem.* **260**, 5925-5931.

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

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

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

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

Di Cola, A., Klostermann, E. and Robinson, C. (2005) The complexity of pathways for protein import into thylakoids: it's not easy being green. *Biochemical Society Transactions.* **33,** 1024-1027.

Di Cola, A. and Robinson, C. (2005) Large-scale translocation reversal within the thylakoid Tat system in vivo. *J.Cell.Biol*. 281-289.

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

Dilks, M., Gimenez, M. and Pohlschroder, M. (2005) Genetic and biochemical analysis of the twin-arginine translocation pathway in *halophilic archaea*. *J.Bacteriol.* **187**, 8104-8113.

Ding, Z. and Christie, P. (2003) Agrobacterium tumefaciens twin-arginine-dependent translocation is important for virulence, flagellation, and chemotaxis but not type IV secretion. *J. Bacteriol.* **185,** 760-771.

Dionisi, H., Checa, S., Krapp, A., Arakaki, A., Ceccarelli, E., Carrillo, N. and Viale, A. (1998) Cooperation of the DnaK and GroE chaperone systems in the folding pathway of plant ferredoxin-NADP(+) reductase expressed in *Escherichia coli*. *Euro.J.Biochem.* **251**, 724-728.

Dreusch, A., Burgisser, D., Heizmann, C. and Zumft, W. (1997) Lack of copper insertion into unprocessed cytoplasmic nitrous oxide reductase generated by an R20D substitution in the arginine consensus motif of the signal peptide. *Biochim.Biophy.Acta.Bioenerget.* **1319**, 311-318.

Dubini, A. and Sargent, F. (2003) Assembly of Tat-dependent [NiFe] hydrogenases: identification of precursor-binding accessory proteins. *FEBS Lett.* **549**, 141-146.

Duffaud, G., Lehnhardt, S., March, P. and Inouye, M. (1985) Structure and function of the signal peptide. *Curr.Membr.Transport*. **24,** 65-104.

Duong, F. and Wickner, W. (1997) The SecDFyajC domain of preprotein translocase controls preprotein movement by regulating SecA membrane cycling. *EMBO J.* **16,** 4871-4879.

Fancy, D. (2000) Elucidation of protein-protein interactions using chemical cross-linking or label transfer techniques. *Curr.Chem.Biol* .**4**, 28-33.

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

Fekkes, P. and Driessen, A. (1999) Protein targeting to the bacterial cytoplasmic membrane. *Microbiol.Mol.Biol.Rev.* **63**, 161-+.
Fekkes, P., vanderDoes, C. and Driessen, A. (1997) The molecular chaperone SecB is released from the carboxy-terminus of SecA during initiation of precursor protein translocation. *EMBO J*. **16**, 6105-6113.

Feldherr, C., Akin, D., Littlewood, T. and Stewart, M. (2002) The molecular mechanism of translocation through the nuclear pore complex is highly conserved. *J.Cell.Sci.* **115**, 2997-3005.

Fikes, J., Barkocygallagher, G., Klapper, D. and Bassford, P. (1990) Maturation of *Escherichia-coli* maltose-binding protein by signal peptidase-I invivo - sequence requirements for efficient processing and demonstration of an alternate cleavage site. *J.Biol.Chem.* **265**, 3417-3423.

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

Fincher, V., McCaffery, M. and Cline, K. (1998) Evidence for a loop mechanism of protein transport by the thylakoid Delta pH pathway. *FEBS Lett.* **423**, 66-70.

Fisher, A. and DeLisa, M. (2004) A little help from my friends: Quality control of presecretory proteins in bacteria. *J.Bacteriol.* **186**, 7467-7473.

Frielingsdorf, S., Jakob, M. and Klosgen, R. (2008) A Stromal Pool of TatA Promotes Tat-dependent Protein Transport across the Thylakoid Membrane. *J.Biol.Chem.* **283**, 33838-33845.

Fujiki, Y., Fowler, S., Shio, H., Hubbard, A. and Lazarow, P. (1982) Polypeptide and phospholipid-composition of the membrane of rat-liver peroxisomes - comparison with endoplasmic-reticulum and mitochondrial-membranes. *J.Cell.Biol.* **93**, 103-110.

Genest, O., Neumann, M., Seduk, F., Stocklein, W., Mejean, V., Leimkuhler, S. and Iobbi-Nivol, C. (2008) Dedicated metallochaperone connects apoenzyme and molybdenum cofactor biosynthesis components. *J.Biol.Chem.* **283**, 21433-21440.

Genest, O., Seduk, F., Ilbert, M., Mejean, V. and Iobbi-Nivol, C. (2006) Signal peptide protection by specific chaperone. *Biochem.Biophy.Research.Commu*. **339**, 991-995.

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

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

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

Gouffi, K., Santini, C. and Wu, L. (2002) Topology determination and functional analysis of the *Escherichia coli* TatC protein. *FEBS Lett.* **525**, 65-70.

Grandori, R., Khalifah, P., Boice, J., Fairman, R., Giovanielli, K. and Carey, J. (1998) Biochemical characterization of WrbA, founding member of a new family of multimeric flavodoxin-like proteins. *J.Biol.Chem.* **273**, 20960-20966.

Graubner, W., Schierhorn, A. and Bruser, T. (2007) DnaK plays a pivotal role in Tat targeting of CueO and functions beside SlyD as a general Tat signal binding chaperone. *J.Biol.Chem.* **282**, 7116-7124.

Griwatz, C. and Junge, W. (1992) Cooperative transient trapping of protons by structurally distorted chloroplast atpase - evidence for the proton well. *Biochim.Biophy.Acta.* **101**, 244-248.

Guymer, D., Maillard, J., Agacan, M., Brearley, C. and Sargent, F. (2010) Intrinsic GTPase activity of a bacterial twin-arginine translocation proofreading chaperone induced by domain swapping. *FEBS J.* **277**, 511-525.

Halbig, D., Wiegert, T., Blaudeck, N., Freudl, R. and Sprenger, G. (1999) The efficient export of NADP-containing glucose-fructose oxidoreductase to the periplasm of Zymomonas mobilis depends both on an intact twin-arginine motif in the signal peptide and on the generation of a structural export signal induced by cofactor binding. *Euro.J.Biochem.* **263**, 543-551.

Hartl, F. (1996) Molecular chaperones in cellular protein folding. *Nature*. **381**, 571-580.

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

Heikkila, M. P., Honisch, U., Wunsch, P. and Zumft, W. G. (2001) Role of the Tat transport system in nitrous oxide reductase translocation and cytochrome cd1 biosynthesis in *Pseudomonas stutzeri*. *J. Bacteriol*. 1663-1671.

Herendeen, S., VanBogelen, R. and Neidhardt, F. (1979) Levels of major proteins of *Escherichia coli* during growth at different temperatures. *J Bacteriol*, 185–194.

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

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

Holland, I. (2004) Translocation of bacterial proteins - an overview. *Biochim.Biophy.Acta.Mol.Cell. Research*. **1694**, 5-16.

Horwich, A., Fenton, W., Chapman, E. and Farr, G. (2007) Two families of chaperonin: Physiology and mechanism. *Annu.Rev.Cell.Develop.Biol.* **23**, 115-145.

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

Howell, J., Winstone, T., Coorssen, J. and Turner, R. (2006) An evaluation of in vitro protein-protein interaction techniques: Assessing contaminating background proteins. *Proteomics* . **6**, 2050-2069.

Hu, Y., Zhao, E., Li, H., Xia, B. and Jin, C. (2010) Solution NMR Structure of the TatA Component of the Twin-Arginine Protein Transport System from Gram-Positive Bacterium *Bacillus subtilis*. *J.American.Chem.Soci*. **132**, 15942-15944.

Huang, S., Ratliff, K. and Matouschek, A. (2002) Protein unfolding by the mitochondrial membrane potential. *Nature Structural Biology*. **9**, 301-307.

Hwang, Y. and Miller, D. (1987) A mutation that alters the nucleotide specificity of elongation factor-tu, a GTP regulatory protein. *J.Biol.Chem*. **262**, 13081-13085.

Ignatova, Z., Hornle, C., Nurk, A. and Kasche, V. (2002) Unusual signal peptide directs penicillin amidase from *Escherichia coli* to the tat translocation machinery. *Biochem.Biophy.Research.Commu.* **291**, 146-149.

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

Inoue, K. (2003) Protein translocation across biological membranes: cost what it may., *Trends Plant. Sci*. **8**, 360-363.

Ishijima, A., Kojima, H., Funatsu, T., Tokunaga, M., Higuchi, H., Tanaka, H. and Yanagida, T. (1998) Simultaneous observation of individual ATPase and mechanical events by a single myosin molecule during interaction with actin. *Cell*. **92**, 161-171.

Izard, J. and Kendall, D. (1994) Signal peptides - exquisitely designed transport promoters. *Mol.Microbio.* **13**, 765-773.

Izard, J., Rusch, S. and Kendall, D. (1996) The amino-terminal charge and core region hydrophobicity interdependently contribute to the function of signal sequences. *J. Biol. Chem.* **271**, 21579-21582.

Ize, B., Gerard, F., Zhang, M., Chanal, A., Voulhoux, R., Palmer, T., Filloux, A. and Wu, L. (2002) In vivo dissection of the tat translocation pathway in *Escherichia coli*. *J. Mol. Biol.* **317**, 327-335.

Ize, B., Stanley, N., Buchanan, G. and Palmer, T. (2003) Role of the *Escherichia coli* Tat pathway in outer membrane integrity. *Mol. Microbiol.* **48**, 1183-1193.

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

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

Jacquet, E. and Parmeggiani, A. (1989) Substitution of val20 by gly elongation-factor tu - effects on the interaction with elongation-factors Ts, aminoacyl-transfer RNA and ribosomes. *Euro.J.Biochem.* **185**, 341-346.

Jarvis, P. and Robinson, C. (2004) Mechanisms of protein import and routing in chloroplasts. *Curr.Biol.* **14**, R1064-R1077.

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

Jepson, B., Mohan, S., Clarke, T., Gates, A., Cole, J., Butler, C., Butt, J., Hemmings, A. and Richardson, D. (2007) Spectropotentiometric and structural analysis of the periplasmic nitrate reductase from *Escherichia coli*. *J. Biol. Chem.* **282**, 6425-6437.

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

Joly, J. and Wickner, W. (1993) The secA and secY subunits of translocase are the nearest neighbors of a translocating preprotein, shielding it from phospholipids. *EMBO J.* **12**, 255-263.

Jongbloed, J., Grieger, U., Antelmann, H., Hecker, M., Nijland, R., Bron, S. and van Dijl, J. (2004) Two minimal Tat translocases in *Bacillus*. *Mol. Microbio.* **54**, 1319-1325.

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

Keenan, R., Freymann, D., Stroud, R. and Walter, P. (2001) The signal recognition particle. *Annu. Rev. Biochem.* **70**, 755-775.

Kihara, A., Akiyama, Y., and Ito, K. (1995) Ftsh is required for proteolytic elimination of uncomplexed forms of secY, an essential protein translocase subunit. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 4532-4536.

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

Kisker, C., Schindelin, H. and Rees, D. (1997) Molybdenum-cofactor-containing enzymes: Structure and mechanism. *Annu. Rev. Biochem.* **66**, 233-267.

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

Kuhn, A., Kiefer, D., Kohne, C., Zhu, H., Tschantz, W. and Dalbey, R. (1994) Evidence for a loop-like insertion mechanism of pro-omp-a into the inner membrane of *Escherichia-coli*. *Euro. J. Biochem.* **226**, 891-897.

Kumamoto, C. (1991) Molecular chaperones and protein translocation across the *Escherichia-coli* inner membrane. *Mol. Microbiol.* **5**, 19-22.

Kumamoto, C. and Francetic, O. (1993) Highly selective binding of nascent polypeptides by an *Escherichia-coli* chaperone protein in vivo. *J. Bacteriol.* **175**, 2184-2188.

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

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

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

Lee, P., Tullman-Ercek, D. and Georgiou, G. (2006) The bacterial twin-arginine translocation pathway. *Annu. Rev. Microbiol.* **60**, 373-395.

Li, H., Chang, L., Howell, J. and Turner, R. (2010) DmsD, a Tat system specific chaperone, interacts with other general chaperones and proteins involved in the molybdenum cofactor biosynthesis. *Biochim. Biophy. Acta. Proteins and Proteomics.* **1804**, 1301-1309.

Li, H., Jacques, P., Ghinet, M., Brzezinski, R. and Morosoli, R. (2005) Determining the functionality of putative Tat-dependent signal peptides in *Streptomyces coelicolor* A3(2) by using two different reporter proteins. *Microbiol. SGM*. **151**, 2189-2198.

Lindenstrauss, U. and Bruser, T. (2006) Conservation and variation between *Rhodobacter capsulatus* and *Escherichia coli* Tat systems. *J. Bacteriol.* **188**, 7807-7814.

Loschi, L., Brokx, S., Hills, T., Zhang, G., Bertero, M., Lovering, A., Weiner, J. and Strynadka, N. (2004) Structural and biochemical identification of a novel bacterial oxidoreductase. *J. Biol. Chem.* 50391-50400.

Luke, I., Handford, J., Palmer, T. and Sargent, F. (2009) Proteolytic processing of *Escherichia coli* twin-arginine signal peptides by LepB. *Archiv. Microbiol.* **191**, 919-925.

Ma, X. and Cline, K. (2010) Multiple precursor proteins bind individual Tat receptor complexes and are collectively transported. *EMBO J.* **29**, 1477-1488.

Madueno, F., Napier, J. and Gray, J. (1993) Newly imported rieske iron-sulfur protein associates with both cpn60 and hsp70 in the chloroplast stroma. *Plant Cell*. **5**, 1865-1876.

Maillard, J., Spronk, C., Buchanan, G., Lyall, V., Richardson, D., Palmer, T., Vuister, G. and Sargent, F. (2007) Structural diversity in twin-arginine signal peptide-binding proteins. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 15641-15646.

Maldonado, B., Kneuper, H., Buchanan, G., Hatzixanthis, K., Sargent, F., Berks, B. and Palmer, T. (2011) Characterisation of the membrane-extrinsic domain of the TatB component of the twin arginine protein translocase. *FEBS Lett*. **585**, 478-484.

Manting, E. and Driessen, A. (2000) *Escherichia coli* translocase: the unravelling of a molecular machine. *Mol.Microbiol*. **37**, 226-238.

Marr, A. G., Ingraham, J. L. and Squires, C. L. (1964) Effect of the temperature of growth of *Escherichia coli* on the formation of β-galactosidase. *J. Bacteriol*. 356–362.

Matos, C., Di Cola, A. and 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.* 474-479.

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

Maurer, C., Panahandeh, S., Jungkamp, A., Moser, M. and Muller, M. (2010) TatB Functions as an Oligomeric Binding Site for Folded Tat Precursor Proteins. *Mol. Biol. Cell.* **21**, 4151-4161.

Mcgeoch, D. (1985) On the predictive recognition of signal peptide sequences. *Virus Research.* **3**, 271-286.

Mejean, V., Iobbinivol, C., Lepelletier, M., Giordano, G., Chippaux, M. and Pascal, M. (1994) TMAO anaerobic respiration in *Escherichia coli* involvement of the tor operon. *Mol. Microbiol.* **11**, 1169-1179.

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

Mitchell, P. (1967) Proton current flow in mitochondrial systems. *Nature.* **214**, 1327-1328.

Mitra, K., Frank, J. and Driessen, A. (2006) Co- and post-translational translocation through the protein-conducting channel: analogous mechanisms at work? *Nature Structural & Molecular Biology.* **13**, 957-964.

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

Molik, S., Karnauchov, I., Weidlich, C., Herrmann, R. and Klosgen, R. (2001) The Rieske Fe/S protein of the cytochrome b(6)/f complex in chloroplasts - Missing link in the evolution of protein transport pathways in chloroplasts? *J.Biol.Chem.* **276**, 42761-42766.

Mori, H. and Cline, K. (2001) Post-translational protein translocation into thylakoids by the Sec and Delta pH-dependent pathways. *Biochim. Biophy. Acta. Mol. Cell Research.* **1541**, 80-90.

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

Mori, H. and Cline, K. (2002) Assembly and disassembly of Tha4 and CpTatC-Hcf106 complexes in the Delta pH/TAT-dependent pathway. *Plant. Cell. Physiol.* **43**, S222-S222.

Mori, H. and Ito, K. (2003) Biochemical characterization of a mutationallv altered protein translocase: Proton motive force stimulation of the initiation phase of translocation. *J.Bacteriol.* **185**, 405-412.

Mould, R. and Robinson, C. (1991) A proton gradient is required for the transport of 2 lumenal oxygen-evolving proteins across the thylakoid membrane. *J. Biol. Chem.* **266**, 12189-12193.

Muller, M. and Klosgen, R. (2005) The Tat pathway in bacteria and chloroplasts (Review). *Mol. Membr. Biol.* 113-121.

Nishiyama, K., Fukuda, A., Morita, K. and Tokuda, H. (1999) Membrane deinsertion of SecA underlying proton motive force-dependent stimulation of protein translocation. *EMBO J.* **18**, 1049-1058.

Nishiyama, K., Mizushima, S. and Tokuda, H. (1993) A novel membrane-protein involved in protein translocation across the cytoplasmic membrane of *Escherichia-coli*. *EMBO J.* **12**, 3409-3415.

Nolandt, O., Walther, T., Roth, S., Burck, J. and Ulrich, A. (2009) Structure analysis of the membrane protein TatC(d) from the Tat system of B. subtilis by circular dichroism. *Biochim. Biophy. Acta. Biomembr.* **1788**, 2238-2244.

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

Oreb, M., Tews, I. and Schleiff, E. (2008) Policing Tic 'n' Toc, the doorway to chloroplasts. *Trends. Cell. Biol.* **18**, 19-27.

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

Orriss, G., Tarry, M., Ize, B., Sargent, F., Lea, S., Palmer, T. and Berks, B. (2007) TatBC, TatB, and TatC form structurally autonomous units within the twin arginine protein transport system of *Escherichia coli*. *FEBS Lett.* **581**, 4091-4097.

Osborne, A., Rapoport, T. and van den Berg, B. (2005) Protein translocation by the Sec61/SecY channel. *Annu. Rev. Cell. Develop. Biol.* **21**, 529-550.

Palleros, D., Shi, L., Reid, K. and Fink, A. (1993) 3-state denaturation of Dnak induced by guanidine-hydrochloride - evidence for an expandable intermediate. *Biochemistry.* **32**, 4314-4321.

Palleros, D. R., Reid, K. L., Shi, L., Welch, W. J. and Fink, A. L. (1993) ATP induced protein-Hsp7O complex dissociation requires K+ but not ATP hydrolysis. *Nature.* **365**, 664-666.

Palmer, T., Sargent, F. and Berks, B. (2005) Export of complex cofactor-containing proteins by the bacterial Tat pathway. *Trends. Microbiol.* 175-180.

Panahandeh, S., Maurer, C., Moser, M., DeLisa, M. and Muller, M. (2008) Following the Path of a Twin-arginine Precursor along the TatABC Translocase of *Escherichia coli*. *J.Biol.Chem.* **283**, 33267-33275.

Papish, A., Ladner, C. and Turner, R. (2003) The twin-arginine leader-binding protein, DmsD, interacts with the TatB and TatC subunits of the *Escherichia coli* twin-arginine translocase. *J. Biol. Chem.* **278**, 32501-32506.

Patterso.D, and Gillespi.D. (1972) Effect of elevated-temperatures on protein-synthesis in *Escherichia coli*. *J. Bacteriol.* **112**, 1177-&.

Perez-Rodriguez, R., Fisher, A., Perlmutter, J., Hicks, M., Chanal, A., Santini, C., Wu, L., Palmer, T. and DeLisa, M. (2007) An essential role for the DnaK molecular chaperone in stabilizing over-expressed substrate proteins of the bacterial twin-arginine translocation pathway. *J. Mol. Biol.* **367**, 715-730.

Perlman, D. and Halvorson, H. (1983) A putative signal peptidase recognition site and sequence in eukaryotic and prokaryotic signal peptides. *J. Mol. Biol.* **167**, 391-409.

Phizicky, E. and Fields, S. (1995) Protein-protein interactions - methods for detection and analysis. *Microbiol. Rev.* **59**, 94-123.

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

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

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

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

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

Ramasamy, S. and Clemons, W. (2009) Structure of the twin-arginine signal-binding protein DmsD from *Escherichia coli*. *Acta Crystallographica Section F-Structural Biology and Crystallization Communications*. **65**, 746-750.

Randall, L., Topping, T. and Hardy, S. (1990) No specific recognition of leader peptide by secB, a chaperone involved in protein export. *Science.* **248**, 860-863.

Rapoport, T. (2008) Protein transport across the endoplasmic reticulum membrane. *FEBS J.* **275**, 4471-4478.

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

Ray, N., Oates, J., Turner, R. and 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**, 156-160.

Ribbe, M. and Burgess, B. (2001) The chaperone GroEL is required for the final assembly of the molybdenum-iron protein of nitrogenase. *Proc. Natl. Acad. Sci.U.S.A.* **98**, 5521-5525.

Richter, S. and Bruser, T. (2005) Targeting of unfolded PhoA to the TAT translocon of *Escherichia coli*. *J.Biol.Chem.* **280**, 42723-42730.

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

Robinson, C, and Bolhuis, A. (2004) Tat-dependent protein targeting in prokaryotes and chloroplasts. *Biochim.Biophy.Acta.Mol.Cell.Research*, 135-147.

Robinson, C., Matos, C., Beck, D., Ren, C., Lawrence, J., Vasisht, N. and Mendel, S. (2011) Transport and proofreading of proteins by the twin-arginine translocation (Tat) system in bacteria. *Biochim.Biophy.Acta.Biomembr.***1808**, 876-884.

Robinson, C., Thompson, S. and Woolhead, C. (2001) Multiple pathways used for the targeting of thylakoid proteins in chloroplasts. *Traffic.* **2**, 245-251.

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

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

Rudiger, S., Germeroth, L., SchneiderMergener, J. and Bukau, B. (1997) Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries. *EMBO J.* **16**, 1501-1507.

Saier, M. (2006) Protein secretion and membrane insertion systems in gram-negative bacteria. *J.Membr.Biol.* **214**, 75-90.

Sanders, C., Wethkamp, N. and Lill, H. (2001) Transport of cytochrome c derivatives by the bacterial Tat protein translocation system. *Mol.Microbiol.* 241-246.

Santini, C., Bernadac, A., Zhang, M., Chanal, A., Ize, B., Blanco, C. and Wu, L. (2001) Translocation of jellyfish green fluorescent protein via the Tat system of *Escherichia coli* and change of its periplasmic localization in response to osmotic up-shock. *J.Biol.Chem.* 8159-8164.

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

Sargent, F. (2007) The twin-arginine transport system: moving folded proteins across membranes. *Bio.Soci.Transactions*. 835-847.

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

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

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

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

Schiebel, E., Driessen, A., Hartl, F. and Wickner, W. (1991) Delta-mu-h+ and ATP function at different steps of the catalytic cycle of preprotein translocase. *Cell.* **64**, 927-939.

Schleiff, E., Soll, J., Kuchler, M., Kuhlbrandt, W. and Harrer, R. (2003) Characterization of the translocon of the outer envelope of chloroplasts. *J.Cell.Biol.* **160**, 541-551.

Schmid, D., Baici, A., Gehring, H. and Christen, P. (1994) Kinetics of molecular chaperone action. *Science.* **263**, 971-973.

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

Shackleton, J. and Robinson, C. (1991) Transport of proteins into chloroplasts - the thylakoidal processing peptidase is a signal-type peptidase with stringent substrate requirements at the -3-position and -1-position. *J.Biol.Chem.* **266**, 12152-12156.

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

Shaw, A., Rottier, P. and Rose, J. (1988) Evidence for the loop model of signal-sequence insertion into the endoplasmic-reticulum. *Proc. Natl. Acad. Sci. U.S.A.* **85**, 7592-7596.

Silvestro, A., Pommier, J., Pascal, M. and Giordano, G. (1989) The inducible trimethylamine n-oxide reductase of *Escherichia-coli*-k12 - its localization and inducers. *Biochim.Biophy.Acta.* **999**, 208-216.

SimalaGrant, J. and Weiner, J. (1996) Kinetic analysis and substrate specificity of *Escherichia coli* dimethyl sulfoxide reductase. *Microbiology-Uk.* **142,** 3231-3239.

Singer, S. and Nicolson, G. (1972) Fluid mosaic model of structure of cell-membranes. *Science.* **175**, 720-&.

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

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

Steck, T. (1974) Organization of proteins in human red blood-cell membrane - review. *J. Cell. Biol.* **62**, 1-19.

Steck,TL. and Yu,J. (1973) Selective solubilization of proteins from red blood cell membranes by protein perturbants. *J Supramol Struct*, 220-232.

Takyar, S., Hickerson, R. and Noller, H. (2005) MRNA helicase activity of the ribosome. *Cell.* **120**, 49-58.

Tarry, M., Schafer, E., Chen, S., Buchanan, G., Greene, N., Lea, S., Palmer, T., Saibil, H. and Berks, B. (2009) Structural analysis of substrate binding by the TatBC component of the twin-arginine protein transport system. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 13284-13289.

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

Tullman-Ercek, D., DeLisa, M., Kawarasaki, Y., Iranpour, P., Ribnicky, B., Palmer, T. and Georgiou, G. (2007) Export pathway selectivity of *Escherichia coli* twin arginine translocation signal peptides. *J.Biol.Chem.* 8309-8316.

Turner, R., Papish, A. and Sargent, F. (2004) Sequence analysis of bacterial redox enzyme maturation proteins (REMPs). *Canadian. J.Microbiol.* 225-238.

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

van der Slurs, E., Nouwen, N., Koch, J., de Keyzer, J., van der Does, C., Tampe, R. and Driessen, A. (2006) Identification of two interaction sites in SecY that are important for the functional interaction with SecA. *J.Mol.Biol.* **361**, 839-849.

van der Wolk, J., de Wit, J. and Driessen, A. (1997) The catalytic cycle of the *Escherichia coli* SecA ATPase comprises two distinct preprotein translocation events. *EMBO J.* **16**, 7297-7304.

Voelker, R. and Barkan, A. (1995) 2 nuclear mutations disrupt distinct pathways for targeting proteins to the chloroplast thylakoid. *EMBO J.* **14**, 3905-3914.

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

von Heijne, G. (1995) Protein sorting signals: simple peptides with complex functions. *EXS*. 67-76.

Vonheijne, G. and Abrahmsen, L. (1989) Species-specific variation in signal peptide design - implications for protein secretion in foreign hosts. *FEBS Lett.* **244**, 439-446.

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

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

Weiner, J., Rothery, R., Sambasivarao, D. and Trieber, C. (1992) Molecular analysis of dimethylsulfoxide reductase - a complex iron-sulfur molybdoenzyme of *Escherichia-coli*. *Biochim.Biophy.Acta.* **1102**, 1-18.

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

White, G., Schermann, S., Bradley, J., Roberts, A., Greene, N., Berks, B. and Thomson, A. (2010) Subunit Organization in the TatA Complex of the Twin Arginine Protein Translocase a site-directed epr spin labeling study. *J.Biol.Chem.* **285**, 2294-2301.

Widdick, D., Dilks, K., Chandra, G., Bottrill, A., Naldrett, M., Pohlschroder, M. and Palmer, T. (2006) The twin-arginine translocation pathway is a major route of protein export in Streptomyces coelicolor. *Proc. Natl. Acad.Sci. U.S.A.* **103**, 17927-17932.

Widdick, D., Eijlander, R., van Dijl, J., Kuipers, O. and Palmer, T. (2008) A facile reporter system for the experimental identification of twin-arginine translocation (Tat) signal peptides from all kingdoms of life. *J.Mol.Biol.* **375**, 595-603.

Wild, J., Rossmeissl, P., Walter, W. and Gross, C. (1996) Involvement of the DnaK-DnaJ-GrpE chaperone team in protein secretion in *Escherichia coli*. *J.Bacteriol.* **178**, 3608-3613.

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

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

Yang, W., Ni, L. and Somerville, R. (1993) A stationary-phase protein of *Escherichia-coli* that affects the mode of association between the Trp repressor protein and operator-bearing DNA. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 5796-5800.

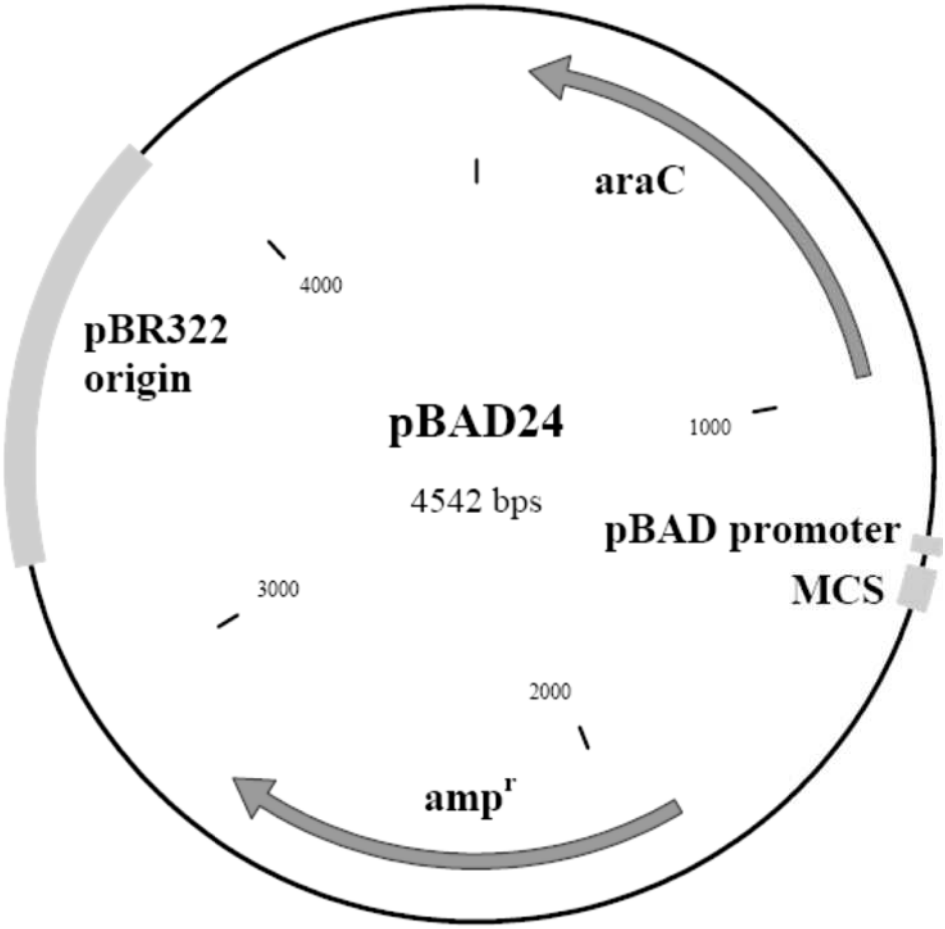
Yen, M., Tseng, Y., Nguyen, E., Wu, L. and Saier, M. (2002) Sequence and phylogenetic analyses of the twin-arginine targeting (Tat) protein export system. *Archi.Microbiol.* **177**, 441-450.

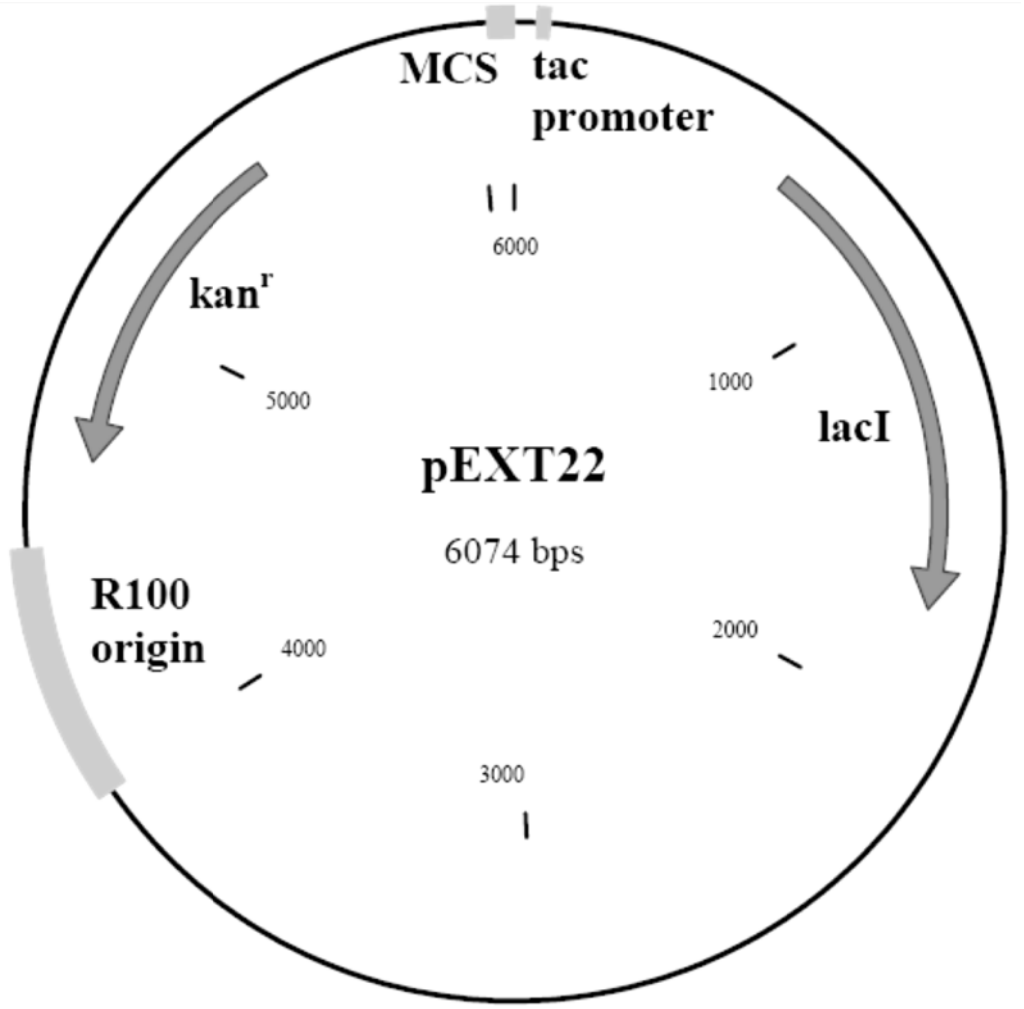
Young, C. and Bernlohr, R. (1991) Elongation factor-Tu is methylated in response to nutrient deprivation in *Escherichia-coli*. *J.Bacteriol.* **173**, 3096-3100.

Zeilstraryalls, J., Fayet, O. and Georgopoulos, C. (1991) The universally conserved GroE (Hsp60) chaperonins. *Annu.Rev.Microbiol.* **45**, 301-325.

Appendix

Vector maps of pBAD24 and pEXT22




MCS
tac
promoter
6000
kan^r
5000
1000
lacI
pEXT22
6074 bps
R100
origin
4000
2000
3000