DECIPHERING THE MECHANISM OF *E. COLI* TAT PROTEIN TRANSPORT: KINETIC SUBSTEPS AND CARGO PROPERTIES

A Dissertation

by

NEAL WILLIAM WHITAKER

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Co-Chairs of Committee,	Siegfried M. Musser
	J. Martin Scholtz
Committee Members,	Michael Benedik
	Jean-Philippe Pellois
Head of Department,	Gregory D. Reinhart

December 2012

Major Subject: Biochemistry

ABSTRACT

The *Escherichia coli* twin-arginine translocation (Tat) system transports fully folded and assembled proteins across the inner membrane into the periplasmic space. The *E. coli* Tat machinery minimally consists of three integral membrane proteins: TatA, TatB and TatC. A popular model of Tat translocation is that cargo first interacts with a substrate binding complex composed of TatB and TatC and then is transported across the inner membrane through a channel comprised primarily of TatA.

The most common method for observing the kinetics of Tat transport, a protease protection assay, lacks the ability to distinguish between individual transport sub-steps and is limited by the inability to observe translocation in real-time. Therefore, a real-time FRET based assay was developed to observe interactions between the cargo protein pre-SufI, and its initial binding site, the TatBC complex. The cargo was found to first associate with the TatBC complex, and then, in the presence of a membrane potential ($\Delta\psi$), migrate away from the initial binding site after a 20-45 second delay. Since cargo migration away from the TatBC complex was not directly promoted by the presence of a $\Delta\psi$, the delay likely represents some preparatory step that results in a transport competent translocon.

In addition, the Tat system has long been identified as a potential biotechnological tool for protein production. However, much is still unknown

about which cargos are suitable for transport by the Tat system. To probe the Tat system's ability to transport substrates of different sizes and shapes, 18 different cargos were generated using the natural Tat substrate pre-SufI as a base. Transport efficiencies for these cargos indicate that not only is the Tat machinery's ability to transport substrates determined by the protein's molecular weight, as well as by its dimensions.

In total, these results suggest a dynamic translocon that undergoes functionally significant, $\Delta\psi$–dependent changes during translocation. Moreover, not every protein cargo can be directed through the Tat translocon by a Tat signal peptide, and this selectivity is not only related to the overall size of the protein, but also dependent on shape.

DEDICATION

To my wife and family.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Siegfried Musser, for his guidance, advice and patience. His continued interest in my projects, as well as his overall enthusiasm for science, was invaluable during my graduate studies. Without his help, the research described in this dissertation would not have been possible. I would also like to thank my current committee members: Dr. Scholtz, Dr. Benedik and Dr. Pellois. Additionally, I'd like to thank Dr. Johnson and Dr. LiWang for their advice during their time on my committee.

I would like to thank all the current and former members of the Musser Lab with whom I worked on a daily basis. Special thanks must be given to Dr. Umesh Bageshwar who developed, and trained me on, the *in vitro* Tat transport assay that I used continuously throughout my research.

Finally, I would like to thank B. Niederhauser for the rhizavidin expression plasmid, T.L. Yahr for pTatABC and the TatA, TatB, and SufI antibodies, and T. Palmer for MC4100 and MC4100ΔTatABCDE *E. coli* strains.

NOMENCLATURE

$\Delta\Psi$:	membrane potential
ΔpH:	pH gradient
BSA:	bovine serum albumin
FRET:	fluorescence resonance energy transfer
GFP:	green fluorescent protein
IMVs:	inner-membrane vesicles
NADH:	β-Nicotinamide adenine dinucleotide, reduced
PMF:	proton motive force

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

SUMMARY

All proteins in *Escherichia coli* are synthesized in the cytoplasm. However, many important functions are performed by proteins outside the cell, inside the periplasm, or within the cell membranes. The movement of proteins across the inner membrane is facilitated by two general transport systems; the Secretory pathway (Sec) and the Twin arginine translocase (Tat). In contrast to the Sec machinery, which transports unfolded proteins, Tat cargos are translocated in a fully folded manner (Clark and Theg, 1997; Hynds et al., 1998). The Tat machinery derives its name from a conserved double arginine motif present in the signal peptide of its substrates (Chaddock et al., 1995). It has been identified in archea, bacteria, plants and at least one animal (albeit a sea sponge) (Frobel et al., 2012; Wang and Lavrov, 2007). Although it transports fewer substrates than the Sec machinery in *E. coli*, the cargos translocated by the Tat translocase are nevertheless responsible for a wide range of important functions such as anaerobic respiration, biogenesis, cell division and virulence (Lee et al., 2006; Widdick et al., 2008). The Tat pathway in *E. coli* minimally consists of three transmembrane proteins: TatA (or its homolog TatE), TatB and TatC (Bogsch et al., 1998; Sargent et al., 1999; Weiner et al., 1998). Despite

the relative scarcity of high resolution structures of the translocon and its constituents (the only one being a NMR structure of a truncated TatA monomer), numerous models of the transport cycle have been proposed based on biochemical studies. The most popular of these models describes the active translocation site as being an oligomer primarily composed of TatA, and the initial binding site consisting of a complex of TatB and TatC (De Leeuw et al., 2001; Gohlke et al., 2005). The stoichiometry of the translocon as well as a precise mechanistic model remains a subject of debate. This chapter summarizes the current understanding of the Tat machinery derived from genetic, biochemical and structural studies performed in the past two decades.

DIVERSITY AND SIGNIFICANCE

Discovery of the Translocon

The earliest evidence of an alternate pathway to Sec came from studies on hydrogenases in the gram-negative bacteria *Desulfovibrio vulgari* (Prickril et al., 1986; van Dongen et al., 1988). These described the export of a heteroligomeric protein that contains a signal peptide on only one of its subunits, which suggested that it was being transported in a folded and assembled state. The discovery of a different kind of signal peptide, one that contained a strictly conserved twin-arginine motif, further strengthened the idea that these substrates were being conveyed across the inner membrane by a new, Sec-

independent, pathway (Niviere et al., 1992). At around the same time, a new, ATP-independent pathway was being described in thylakoids (Cline et al., 1992; Klosgen et al., 1992; Robinson and Klosgen, 1994). Studies had indicated that there was an important double-arginine in the signal sequences for this pathway, and that the pathway could function without the presence of any soluble factor (Chaddock et al., 1995; Cline et al., 1992; Robinson et al., 1993). It was eventually suggested that these two pathways may be related, and that they might share functional and mechanistic characteristics (Berks, 1996). This was confirmed by sequence homology between the thylakoid protein Hcf106 and an *E. coli* protein that would eventually be called TatB (Settles et al., 1997).

Since its discovery, the vast majority of mechanistic studies have been performed on the plant and *E. coli* versions of the Tat machinery, and genomic studies have revealed the presence of Tat homologs in a wide array of organisms (Dilks et al., 2003; Yen et al., 2002). The machinery appears to have a particularly important role in some archea, namely halophiles like *Haloferax volcanii,* one of the few organisms in which Tat has been shown to be essential for viability (Dilks et al., 2005). This dependence on Tat as the organism's primary export machinery has been suggested to be due to rapid folding of protein cargos under high-salt conditions (Bolhuis, 2002; Rose et al., 2002).

Potential Drug Target

Despite Tat not being essential in most organisms, the system has nevertheless been identified as a potential drug target in some pathogens, in part due to the absence of any Tat homologs in humans (De Buck et al., 2008b). Interest in this has led to some studies in organisms like *Mycobacterium smegmatis*, the non-pathogenic mycobacterium frequently used as a model for *Mycobacterium tuberculosis*. *M. smegmatis tat* mutants failed to export active β-lactamase and exhibited slower growth rates (McDonough et al., 2005; Posey et al., 2006). Surprisingly, *M. tuberculosis* requires the Tat machinery for cell growth (at least under laboratory conditions) (Saint-Joanis et al., 2006). In *Legionella pneumophila*, the cause of Legionnaire's disease, the absence of a functional Tat machinery results in the failure to export 3',5'-cyclic nucleotide phosphodiesterase and the iron–sulfur subunit of the ubiquinol-cytochrome c reductase, resulting in abnormal cell growth and motility (De Buck et al., 2008a; De Buck et al., 2007). The Tat machinery is important for virulence of the pathogenic *E. coli* strain O157:H7, since a *tatABC* deletion shows a five-fold decrease in the secretion of Shiga toxin 1 (Pradel et al., 2003).

In addition to the studies described above, the deletion or mutation of Tat genes in various plant and animal pathogens result in phenotypes that affect iron acquisition, cell growth, motility, infection and biofilm formation (Bronstein et al., 2005; Caldelari et al., 2006; De Buck et al., 2005; Ding and Christie, 2003; Ize et al., 2004; Lavander et al., 2006; Ochsner et al., 2002; Reynolds et al.,

2011; Rossier and Cianciotto, 2005). It is apparent that, despite being often thought of as the “alternative” pathway for translocation, the Tat system is important for a wide array of functions in a variety of organisms. Extensive investigations to elucidate its precise role in many economically and medically important pathogens is needed.

Biotechnological and Protein Engineering Applications

The value of secreting recombinant proteins into the periplasm or outside of the bacterial cell has been known for many years (Mergulhao et al., 2005). Exporting proteins from the cytoplasm allows easier and less costly purification downstream. While the Sec pathway can transport most proteins faster and with a higher efficiency, the Tat machinery does have several advantages. The first and most important advantage is that the Tat system typically only transports correctly folded cargos (Clark and Theg, 1997; Hynds et al., 1998). Misfolded cargos would therefore be excluded from the periplasmic fraction, allowing for the isolation of only correctly folded species. Additionally, the pathway can accommodate proteins that the Sec system cannot, specifically oligomeric, fast folding, or cofactor-containing substrates. For these reasons, the Tat system has been identified as a potential tool for the production of protein products (Bruser, 2007; Yoon et al., 2010).

The Tat machinery has also been applied to protein engineering. It has been used to select solubility enhanced proteins (Fisher et al., 2011b), to direct

the evolution of single-chain antibodies for both solubility (Fisher and DeLisa, 2009) and for faster folding (Ribnicky et al., 2007), to develop a bacterial two-hybrid assay (Strauch and Georgiou, 2007a), and as a tool for glycoengineering in *E. coli* (Fisher et al., 2011a).

TAT CARGOS

The physiological effects of the Tat knockouts and mutations discussed in the previous section are not caused by the absence of Tat directly, but by the absence of its properly translocated substrates. In this section the characteristics of these cargos are discussed.

The Signal Peptide

For Sec and Tat cargos, short, N-terminally attached signal peptides target precursors to the translocons. Tat and Sec signals share several characteristics, most notably their tripartite structure (Figure 1.1). This consists of a positively charged amino-terminal sequence (n-region), a hydrophobic central sequence (h-region) and a carboxy-terminal polar sequence (c-region) (Natale et al., 2008). Tat signal peptides, however, are typically longer than Sec peptides and nearly always contain two consecutive arginines in their n-region. It has been noted that these arginines are part of a larger S-R-R-x-F-L-K consensus motif (in *E. coli* systems), and it is present in about 50% of bacterial cargos (Berks, 1996). A more generalized view of this motif which

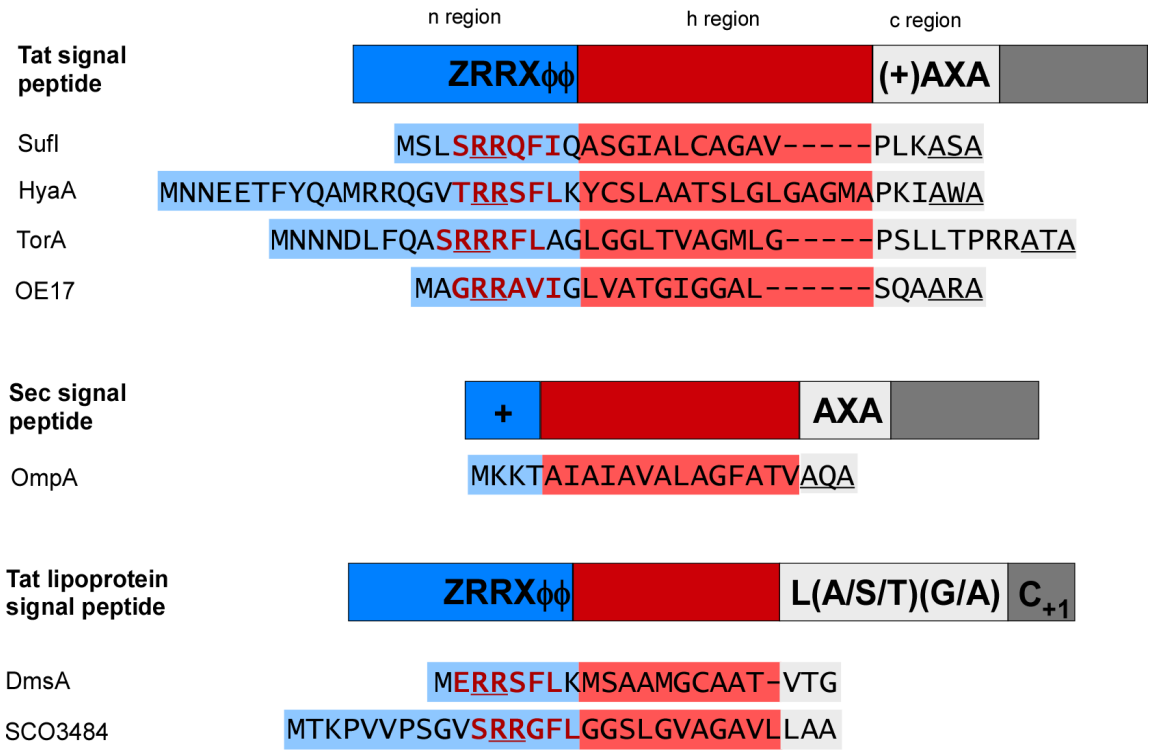


Figure 1.1. Tat signal peptides. Peptides for *E. coli* substrates SufI, HyaA, TorA and DmsA are shown. OE17 is a thylakoid Tat cargo and SCO3484 is a *Streptomyces coelicolor* cargo. The Sec signal peptide of *E. coil* OmpA is shown for comparison. Blue indicates the polar n regions, red indicates the hydrophobic core region and white indicates the c region. The Tat consensus motif is shown in bold with Z as any polar residue, Φ is a hydrophobic residue and + indicates any basic residue. The peptide cleavage signal is shown in bold in the c regions. Figure adapted with modifications from (Palmer and Berks, 2012).

encompasses thylakoid Tat substrates has been suggested: Z-R-R-x-Φ-Φ (Robinson and Bolhuis, 2004), where Z is any polar residue and Φ is a hydrophobic residue. The twin-arginines are highly conserved, with only a few naturally occurring variations known (typically single lysine, asparagine, or glutamine residue) (Hinsley et al., 2001; Ignatova et al., 2002; Molik et al., 2001; Widdick et al., 2008). Introducing these mutations into other Tat cargos appears to allow proper targeting and transport, albeit to a lesser extent (DeLisa et al.,

2002; Halbig et al., 1999; Ize et al., 2002b; Stanley et al., 2000). However, a "conservative" double mutation to twin lysines abolishes transport (Chaddock et al., 1995; Stanley et al., 2000). Clearly, the twin-arginine motif is important for proper recognition by the Tat machinery. However, suppressor mutations have been generated in the cargo binding protein TatC that allow for limited transport of the double lysine mutant (Strauch and Georgiou, 2007b).

Overall, Sec signal peptides are more hydrophobic than Tat peptides, and mutations to either the n-region or the h-region that increase their hydrophobicity have been shown to cause Tat cargos to be directed to the Sec machinery instead (Cristobal et al., 1999; Ize et al., 2002a). The hydrophobicity of the middle of the h-region (amino acid position 17) in at least one substrate seems to be of particular importance, with more hydrophobic residues leading to greater Tat dependent binding and transport (Bageshwar et al., 2009).

As in the Sec system, the c-region of Tat signal peptides typically contains a number of polar and charged residues. The number of these residues varies, but is usually higher in Tat peptides (Ize et al., 2002a). One of these charged residues was reported to function as a "Sec-avoidance signal" (Blaudeck et al., 2003; Bogsch et al., 1997). The c-region of Sec and Tat precursors also contains a recognition sequence for signal peptidase I (LepB) (Luke et al., 2009), which is required for post-transport processing of the precursor into its mature form.

Protein	Physiological role	Cofactors	Co-exported partner	Signal chaperone
HyaA	Hydrogen oxidation	3×Fe-S clusters	HyaB	HyaE
HybO	Hydrogen oxidation	3×Fe-S clusters	HybC	HybE
HybA	Hydrogen oxidation	4×Fe-S clusters	Unknown	Unknown
NapG	Nitrate reduction	4×Fe-S clusters	Unknown	Unknown
NrfC	Nitrite reduction	4×Fe-S clusters	Unknown	Unknown
YagT	Unknown	2×Fe-S clusters	YagR, YagS	YagQ
YdhX	Unknown	4×Fe-S clusters	Unknown	Unknown
TorA	TMAO reduction	MGD	None	TorD
TorZ	TMAO reduction	MGD	None	YcdY
NapA	Nitrate reduction	MGD, 1×Fe-S cluster	None	NapD
DmsA	DMSO reduction	MGD, 1×Fe-S cluster	DmsB	DmsD
YnfE	DMSO reduction	MGD, 1×Fe-S cluster	YnfG	DmsD
YnfF	DMSO reduction	MGD, 1×Fe-S cluster	YnfG	DmsD
FdnG	Formate oxidation	MGD, 1×Fe-S cluster	FdnH	FdhD, FdhE
FdoG	Formate oxidation	MGD, 1×Fe-S cluster	FdoH	FdhD, FdhE
YedY	TMAO/DMSO reduction	MPT	None	Unknown
CueO	Copper homeostasis	4×Cu ions	None	Unknown
SufI	Cell division	None	Unknown	None
YahJ	Unknown	1×Fe ion	Unknown	Unknown
WcaM	Colanic acid biosynthesis	Unknown	Unknown	Unknown
MdoD	Glucans biosynthesis	Unknown	Unknown	Unknown
YcdB	Unknown	Unknown	Unknown	Unknown
YcdO	Unknown	Unknown	Unknown	Unknown
YaeI	Possible phosphodiesterase	Unknown	Unknown	Unknown
AmiA	Cell wall amidase	None	Unknown	Unknown
AmiC	Cell wall amidase	None	Unknown	Unknown
FhuD	Ferrichrome binding	None	Unknown	Unknown
YcbK	Unknown	Unknown	Unknown	Unknown

Table 1.1. List of known and predicted *E. coli* Tat substrates. Adapted from (Palmer et al., 2005).

Cargo Function, Folding and "Proofreading"

As described previously, the degree to which the Tat machinery is utilized varies widely from species to species. In *E. coli,* there are 28 known or

predicted substrates (Table 1.1). Most of these proteins perform redox functions, and contain metallocofactors and/or a co-exported partner (Palmer et al., 2005). Proper assembly of these cofactor containing proteins necessitates that they are folded in the cytoplasm. This was demonstrated directly by the observation that TMAO reductase (TorA) required the insertion of its molybdocofactor in the cytoplasm before translocation (Santini et al., 1998). Proper insertion of cofactors in artificial Tat substrates was also seen to be necessary prior to transport (Hynds et al., 1998; Sanders et al., 2001). Similarly, hetero-oligomeric cargos with subunits that lack a signal peptide must obtain their quaternary structure before translocation. This transport of a signal peptide lacking subunit has been termed the "hitch-hiker mechanism," and was shown experimentally with dimethyl sulphoxide reductase's catalytic dimer (DmsAB) and hydrogenases in *E. coli* (Rodrigue et al., 1999; Sambasivarao et al., 2000).

The hitch-hiker mechanism is a salient feature of the Tat machinery, but natural Tat cargos that were tightly bound to an avidin tetramer failed to be translocated under *in vitro* conditions (Bageshwar et al., 2009; Musser and Theg, 2000). These substrates still maintained the ability to bind to the translocon, and are not larger than the largest known Tat cargos. This suggests a possible dependence of "transportability" on the shape of the precursor, as opposed to its overall size. This relationship is explored in detail in Chapter III of this dissertation.

Tat substrates also include lipoproteins (Gralnick et al., 2006; Shruthi et al., 2010b). These substrates contain modified signal peptides (termed type II), whose cleavage site is next to a lipidated cysteine residue that is eventually cleaved by the lipoprotein signal peptidase (Figure 1.1) (Yamada et al., 1984). In *E. coli*, the Tat machinery was found to be both capable of translocating a fast folding lipoprotein and essential for its lipid modification (Shruthi et al., 2010a).

The function of some Tat substrates (and for what reasons, if any, they are folded before export) are still unclear. SufI's function, for instance, has yet to be elucidated despite its frequent use as a model cargo. However, there is significant evidence that it is involved in cell division (Tarry et al., 2009a). Also, despite the high sequence identity (~80%) between the *E. coli* Tat proteins and their counterparts in *Salmonella enterica,* approximately 40% of the predicted substrates differ between the two closely related organisms (Reynolds et al., 2011). Although the Tat system is sometimes thought of as a back-up translocation machinery to Sec, to be used only when necessary (since transporting a fully folded protein across an ion tight membrane is a difficult task (Palmer and Berks, 2012)), it may simply be an alternative method of export.

There have been numerous studies linking specific cytosolic chaperones to the transport of cargos through the Tat translocon. These have been termed redox enzyme maturation proteins (REMPs) (Turner et al., 2004). The first of such proteins described was DmsD, a specific chaperone for DmsA (Oresnik et al., 2001; Weiner et al., 1992). Surprisingly, DmsD does not recognize the

signal peptide of DmsA when it is fused to a different protein, indicating that DmsD recognizes the mature domain of the cargo, at least in part (Ray et al., 2003). DmsD appears to also interact with several general chaperones (GroEL, DnaK, and others) leading some to suggest that DmsD directs its substrate into a cascade of chaperones to assist in the protein's proper folding and assembly (Li et al., 2010). Other highly studied Tat specific chaperones are TorD, NapD, HyaE, and HybE, which bind to TMAO reductase, the NapA subunit of nitrate reductase, the HyaA subunit of hydrogenase-1, and the HybO subunit of hydrogenase-2 respectively (Dubini and Sargent, 2003; Genest et al., 2006; Jack et al., 2004; Tranier et al., 2002). Most of these chaperones appear to dimerize (Turner et al., 2004). It has been proposed that these substrate-specific chaperones act as proofreading checkpoints, allowing the cargo to mature before transport (Sargent, 2007a).

In addition to these apparently dedicated chaperones, a more generalized proofreading system has been proposed in which the Tat translocon directs malfolded proteins to degradation pathways. Two studies seemed to indicate the presence of such a mechanism, with TatD playing a central role (Matos et al., 2009; Matos et al., 2008). However, it was recently determined that these results were caused by an overexpression artifact, and that malfolded Tat substrates are degraded by Tat independent cytosolic systems (Lindenstrauss et al., 2010). With these findings, the biological role of TatD remains unclear (Wexler et al., 2000). A general proofreading system remains a possibility, as it

explains several results which demonstrates the machinery's ability to apparently "reject" certain untransportable substrates (Bageshwar et al., 2009; Musser and Theg, 2000). It is unclear if this ability is linked to a specific action of the translocon, or if it is simply a result of the machinery not being able to accommodate certain substrates, resulting in their disassociation. A recent study has determined that this quality control system can be disabled through random mutagenesis, leading the authors to suggest that a specific region of the Tat translocon performs this function (Rocco et al., 2012). It is apparent from this data that the proof-reading abilities of the Tat translocase are embedded in the Tat proteins themselves, since mutations solely in TatA, TatB and TatC were able to give rise to the machinery's ability to transport previously untransportable cargo. However, the claim that a region of the translocon is specifically responsible for this function cannot fully be supported by any current data, and would likely require high resolution structures of active Tat machinery to validate.

THE TRANSLOCON

As stated previously, genetic studies have indicated that the *E. coli* Tat translocon minimally requires three components (TatA, TatB, and TatC) for proper function (Figure 1.2). This was confirmed with an *in vitro* assay utilizing a purified Tat cargo and isolated bacterial inner membrane vesicles (Yahr and Wickner, 2001). Despite many studies by numerous different laboratories, the

precise manner in which these three components interact to transport cargo is not fully understood. However, the most well-accepted model proposes that TatB and TatC act together to form the initial binding site for precursor proteins, and a structure composed primarily, or entirely of TatA (typically thought to be a pore) comprises the protein conducting channel. This section discusses what is known about the translocon, the individual roles of the components, their stoichiometry and the kinetics of the actual translocation event.

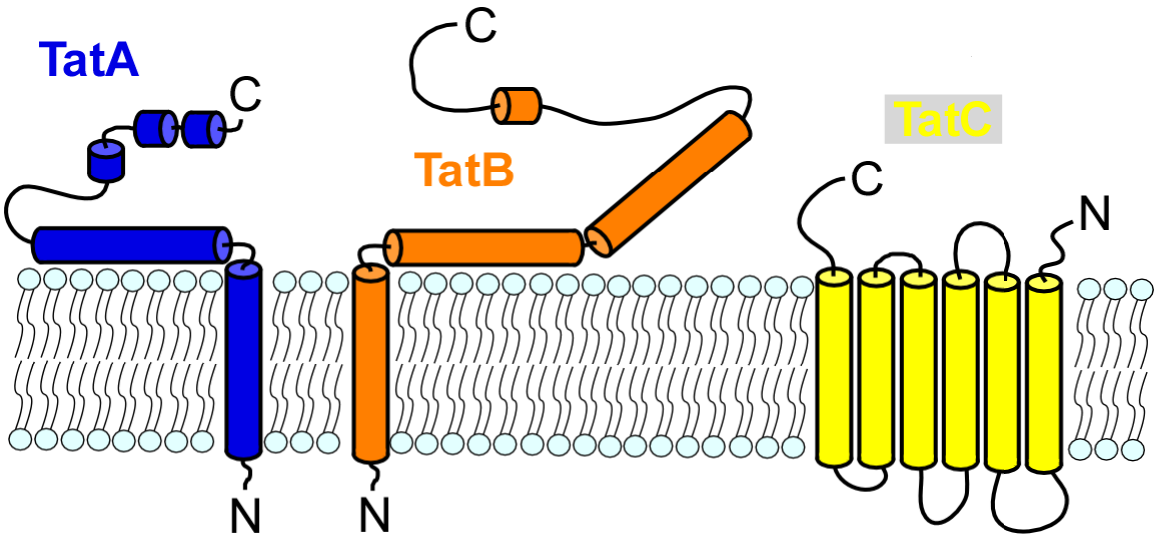


Figure 1.2. Components of the *E. coli* twin arginine translocase. TatA (blue) contains a predicted single transmembrane spanning helix, an amphipathic helix and an unstructured C-terminal region. TatB (orange) contains two helices with significant homology to those in TatA. In addition, TatB also is predicted to contain a soluble helix near its C-terminus. TatC (yellow) is composed of six transmembrane helices.

The TatBC Complex

TatB and TatC appear to work congruently in a stable complex (De Leeuw et al., 2001). In *E. coli*, TatB has 20% sequence identity with TatA (Hicks

et al., 2003). In some organisms (like *Mycobacterium smegmatis*), TatB is not present, with TatA presumably assuming its functional role (McDonough et al., 2005). For this reason, Tat proteins have been broadly described as pertaining to one of two types: TatA and TatC (Palmer and Berks, 2012). However, in *E. coli,* TatB's functional diversity with the similar TatA molecule was discovered fairly early on (Sargent et al., 1999). TatB is a 171 amino acid protein with a molecular mass of 18.5 kDa. Structurally, TatB is predicted to have a single membrane spanning helix, an amphipathic helix, and a short soluble region on the C-terminal, cytoplasmic side (Figure 1.2) (Barrett et al., 2003). The highest homology between TatB and TatA occurs in the amphipathic and transmembrane regions, with the C-terminus being mostly absent in TatA. Truncation analyses revealed that the C-terminal domain is unnecessary for proper TatB function (Lee et al., 2002). TatB's role in cargo binding has been examined in crosslinking studies. These suggest a model of precursor binding in which cargo transiently associates with the transmembrane and amphipathic helices of TatB before transferring to TatC (Alami et al., 2003; Maurer et al., 2010).

TatC is the largest of the *E. coli* Tat proteins with a mass of 28.9 kDa. It has six predicted transmembrane spanning helices (Figure 1.2), which were confirmed by cysteine-labeling and reporter fusions (Behrendt et al., 2004; Ki et al., 2004; Punginelli et al., 2007). Site specific crosslinking studies in both active and resting translocons revealed that the most prominent contacts of TatC to

cargo are between the signal peptide of the precursor protein and the first cytosolic loop of TatC (Zoufaly et al., 2012). This is in contrast to the aforementioned experiments that revealed that TatB makes contacts primarily with the mature domain of the translocating protein (Alami et al., 2003; Maurer et al., 2010). Thus, TatB and TatC together bind the precursor protein but interact with distinct regions (Maurer et al., 2010).

Several studies have indicated that TatB and TatC are found in stable, high molecular weight complexes in an equimolar ratio. In resting (no pmf) membranes these complexes range in size, typically between 350 and 600 kDa (Bolhuis, 2002; De Leeuw et al., 2001). Structures generated by low resolution single-particle electron microscopy indicate that TatBC complexes are irregularly shaped with varying stoichiometries (Oates et al., 2003). However, a more recent, higher resolution (15 Å) structural study observed complexes with what was first described to be 7-8 fold symmetry (Tarry et al., 2009b). Later studies with fused TatC dimers favor an octomer (Palmer and Berks, 2012). Interestingly, for structures determined in the presence of a cargo protein, it was found that at most two precursors were found bound to a single TatBC octomer. In contrast, recent experiments performed with the thylakoid Tat system indicated that a single cpTatC-Hcf106 complex (the chloroplast equivalent to the TatBC complex in *E. coli*) can bind up to eight cargos (Celedon and Cline, 2012). This apparent discrepancy could be due to inherent differences between the plant and bacterial systems. However, it must also be noted that the

structural investigations were performed with purified *E. coli* proteins in the absence of the membrane, and therefore may not accurately reflect the TatBC complex in its active state.

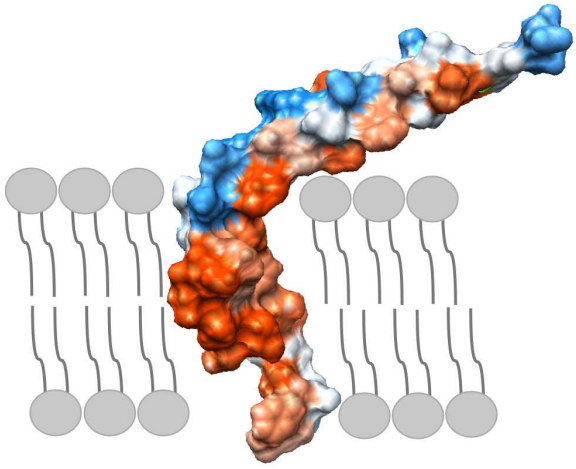


Figure 1.3. Structure of a Tat monomer from *Bacillus subtilus*. This structure was determined by NMR in dodecyl-phosphocholine (DPC) micelles (Hu et al., 2010). The position of TatA is in accordance with the membrane alignment of (Walther et al., 2010) and the PEGylation analysis of (Koch et al., 2012; Tarry et al., 2009a). The amphipathic helix has a slanted alignment with respect to the lipid bilayer, with its N-terminal residues buried. Red color indicates regions with high hydrophobicity and blue indicate regions with low hydrophobicity. Only residues 2-45 are shown. Structure drawn in UCSF Chimera.

The TatA Channel

TatA is a 9.6 kDa protein comprised of one predicted transmembrane helix, a cytosolic amphipathic helix and a short hinge region linking the two. These two helices were confirmed by two separate solution NMR analyses (Chan et al., 2011; Hu et al., 2010). Solid-state NMR performed with TatA

reconstituted into planar bicelles revealed that the hinge region and the end of the amphipathic helix were buried into the lipid bilayer (Figure 1.3). Additionally, the transmembrane helix was found to be tilted 13º with respect to the bicelle (Walther et al., 2010). This slanted conformation was confirmed by a recent PEGylation accessibility study in which the residues in the hinge region were protected, indicating that they are indeed buried in the membrane (Koch et al., 2012).

The topology of TatA during a translocation event, particularly that of the transmembrane helix, has been the subject of some debate. Some studies have indicated that the N-terminus faces the cytoplasmic side of the membrane at least some of the time (Chan et al., 2007), leading some to suggest that it possesses a dual topology (Gouffi et al., 2004). This configuration of TatA would be at odds with the "positive inside rule," which states that integral membrane proteins favor a topology that places positive charges inside the cytoplasm (von Heijne and Gavel, 1988). Some have argued that the opposite is true, with data from protease and oxidant accessibility experiments indicating that the N-terminus faces the periplasm (Greene et al., 2007; Porcelli et al., 2002). A recent study may have finally settled the issue by probing TatA via PEGylation in intact cells, arguing that previous "N-in" results have arisen from damaged or inverted cytoplasmic membranes (Koch et al., 2012).

There have been a number of studies that suggest that a soluble (i.e., cytoplasmic or stromal) form of TatA exists. These were performed in both the

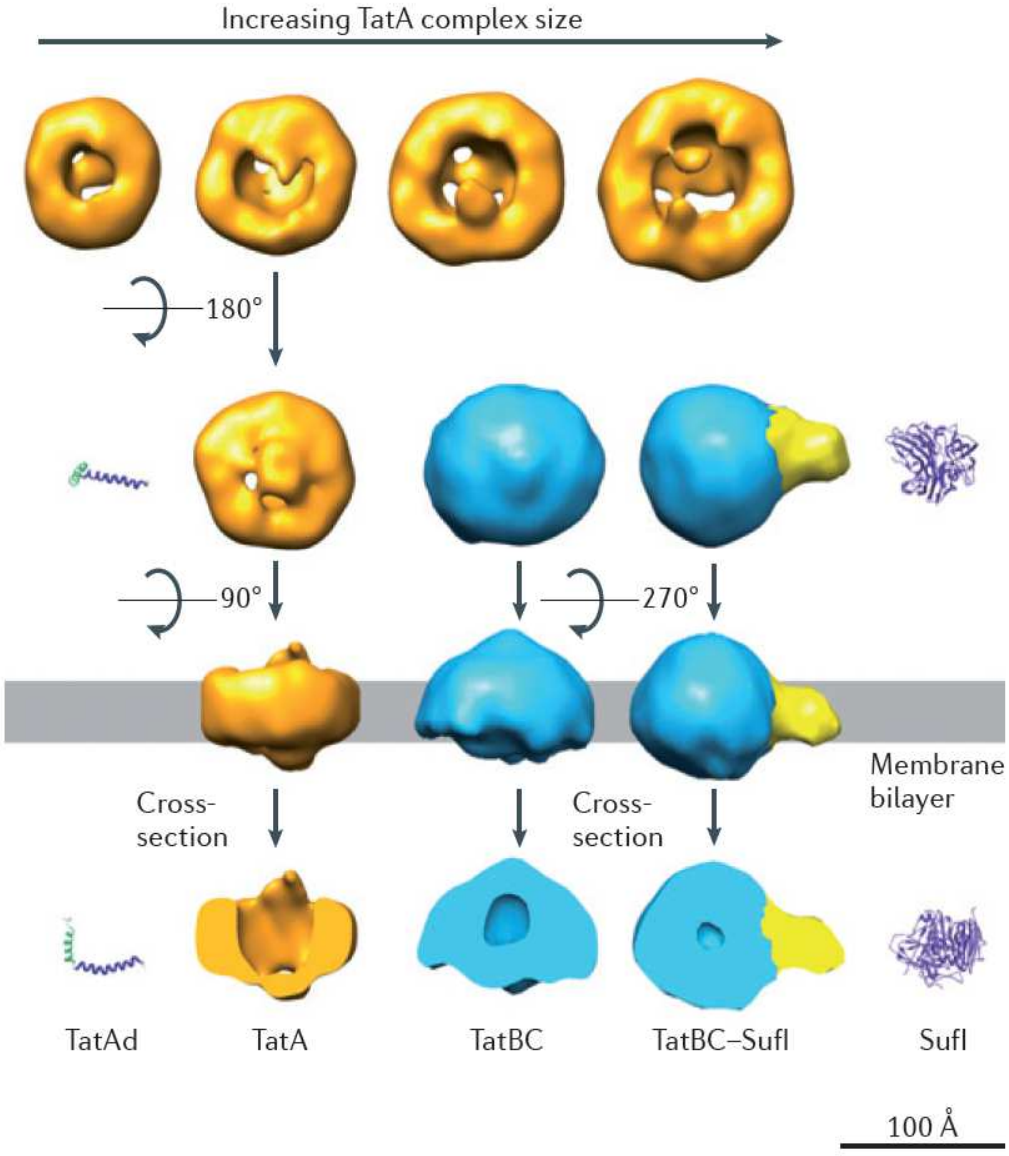


Figure 1.4. 25 Å and 15 Å resolution electron tomography structures of the TatA and TatBC complexes. Structures obtained using single-particle negative-stain electron microscopy and detergent solubilized *E. coli* proteins (Gohlke et al., 2005; Tarry et al., 2009b). The top row shows four different size classes. The third row shows the assumed position of the structures within the inner membrane. High resolution structures of *Bacillus subtilus* TatA (TatAd) and the *E. coli* Tat cargo SufI was included for comparison to their low resolution counterparts. Reprinted with permission from "The twin-arginine translocation (Tat) protein export pathway" by T. Palmer and B.C. Berks, 2012. *Nat Rev Microbiol*. (10) 483-96. Copyright 2012 Nature Publishing Group.

chloroplasts of plants and in gram-positive bacteria (Barnett et al., 2009; De Keersmaeker et al., 2005; Frielingsdorf et al., 2008; Pop et al., 2003; Schreiber et al., 2006; Westermann et al., 2006). A possibly related finding describes TatC-dependent TatA tube-like structures observed in the cytoplasm of *E. coli* (Berthelmann et al., 2008). Whether this is part of the biogenesis of TatA, or evidence of some unknown step in Tat mediated protein transport is still unclear.

TatA has been suggested to be the primary component of the protein conducting channel in Tat dependent translocation for many years due to its observed ability to form many different sized oligomers (Behrendt et al., 2007; De Leeuw et al., 2001; Oates et al., 2005). These homo-oligomers have been observed using detergent solubilized TatA molecules and single-particle electron microscopy (Figure 1.4, top row) (Gohlke et al., 2005). Ring-like structures of variable diameter were observed with a “lid” on what the authors suggest is the cytoplasmic side. It was proposed that the variable diameters may be important to the machinery’s ability to accommodate substrates of different sizes and that the lid could be part of a gating mechanism. Importantly, smaller, but similar, structures were found when identical experiments were performed with the TatA functional paralog TatE (Baglieri et al., 2012). The authors of this study suggest that in the case of small substrates, channels comprised of TatE could be used.

A comparison of TatA proteins from different species reveals that the residues found within the amphipathic and transmembrane helices are the most conserved (Frobel et al., 2012). Alanine substitution (Hicks et al., 2003),

cysteine scanning (Greene et al., 2007), and loss of function mutagenesis (Barrett et al., 2003; Hicks et al., 2005) studies indicated that, predictably, these regions are important for TatA function. Mutations that structurally affect the hinge region also cause significant loss of function (Hicks et al., 2003; Hicks et al., 2005). Because the helices and hinge region are within the first 42 amino acids, it is not surprising then that TatA retains its functionality even when 40 amino acids are removed from the C-terminus in truncation analyses (Lee et al., 2002).

The Mechanism of Tat Translocation

Despite substantial progress in understanding the Tat machinery, there is yet no consensus on the events of the translocation cycle. The translocon itself appears to be transient (Cline and McCaffery, 2007), and no stable translocation intermediates have been generated that would allow researchers to characterize sub-steps of the mechanism. *In vitro* studies utilizing isolated inverted inner-membrane vesicles and thylakoids have proven the most valuable in terms of developing a working model. When comparing these studies, it is convenient to assume that data derived from experiments in the two systems (plant and bacterial Tat) can be compared directly. It is important to remember, however, that there could easily be substantial differences, especially considering that it has been suggested that the evolutionary split between TatA and TatB (Tha4

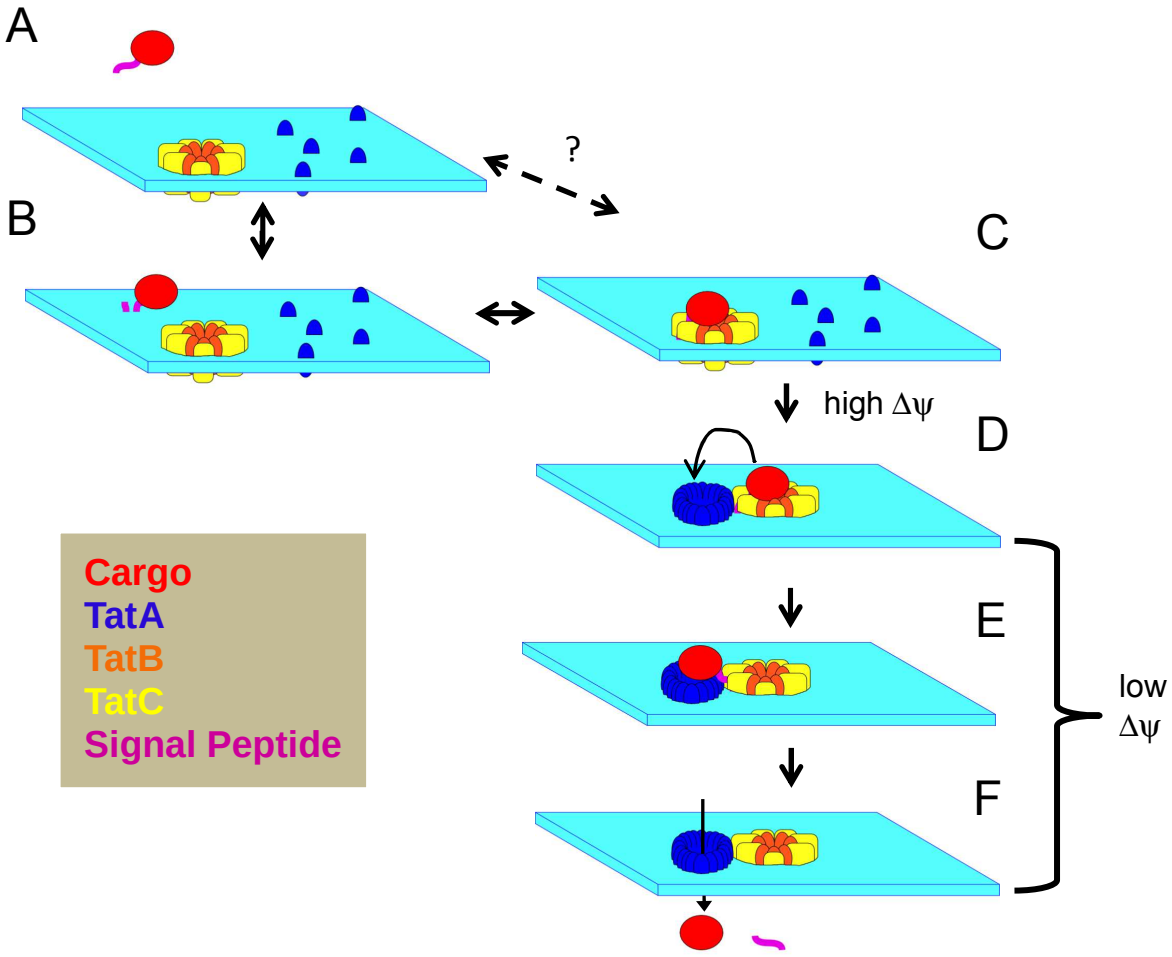


Figure 1.5. A TatA pore model of Tat transport. (A) TatB (orange) and TatC (yellow) form a complex in the membrane with a 1:1 stoichiometry. TatA (blue) is shown here as membrane integrated monomers. (B) Tat cargos can associate with the lipid bilayer in a signal peptide dependent fashion. (C) The precursor can then bind to the TatBC receptor complex (Bageshwar et al., 2009). Formally, it is possible that the cargo can forgo the membrane binding step and associate directly with TatBC (A → C), but this has yet to be directly observed. (C → F) From the TatBC-bound state, the cargo is translocated across the membrane in a process that includes a high Δψ-dependent step (C → D) and a low Δψ-dependent step (D → F) (Bageshwar and Musser, 2007). After translocation, the signal peptide is cleaved from the cargo to generate a mature protein (Luke et al., 2009).

and Hcf106 in plants) appears to have occurred independently within the two system's lineages (Greene et al., 2007).

One such difference may be how the plant and bacterial Tat systems utilize the proton motive force to drive the translocation process. The pmf was first identified as the sole source of energy for the thylakoid translocon (Mould and Robinson, 1991) and was later confirmed in *E. coli* (Yahr and Wickner, 2001). Early results on the plant system seemed to indicate that the ΔpH component of the pmf alone drove cpTat specific transport (Cline et al., 1992), with no detectable contribution of the electrical potential ($\Delta\Psi$). This result was apparently confirmed in a study that demonstrated a substrate specific minimum ΔpH is required transport, suggesting that larger cargos require a larger ΔpH for translocation (Alder and Theg, 2003). This is in stark contrast to a more recent study in the bacterial Tat system that indicates that the exact opposite is true, i.e., that the ΔpH appears to not be necessary for transport, and that the $\Delta\Psi$ is absolutely required (Bageshwar and Musser, 2007). More recent evidence indicates that the $\Delta\Psi$ is utilized in the thylakoids of other organisms, such as the green algea *Chlamydomonas reinhardtii* and in the leaves of barley plants (Finazzi et al., 2003). It is somewhat surprising that such a fundamental part of the transport machinery would differ so much from species to species, leading some to suggest that an unknown component is absent from these *in vitro* assays (Theg et al., 2005), that the ionophore used to collapse the $\Delta\Psi$ (valinomycin) does not completely remove the gradient (Bageshwar and Musser,

2007) or that both the $\Delta\Psi$ and the ΔpH can support transport under certain conditions. The latter case is supported by another study that suggested that, contrary to previous results, the $\Delta\Psi$ can contribute, at least in part, to *in vitro* translocation in thylakoids (Braun et al., 2007; Braun and Theg, 2008).

How the pmf is specifically utilized to drive the transport of cargo is also a subject of some debate. Part of the discussion has centered on whether the gradient "drives" the process at all. Models have been suggested in which the pmf is instead coupled to a gating reaction, not unlike those which have been demonstrated to occur in voltage-gated ion channels (Tombola et al., 2005). In this scenario, the cargo would be recruited to the translocon in a pmf independent manner, then migrate across the inner membrane after a "gate" is opened. The translocation event is energetically favorable because the concentration of precursor proteins is higher in the cytoplasm as compared to the periplasm, where they are processed into their mature forms by the signal peptidase LepB (Luke et al., 2009), creating a concentration gradient. The alternative model is that the ion movement is coupled to the protein transport event itself. This would mean that a certain amount of the pmf would be "consumed" in every transport reaction. This appears to be supported by data that suggests that an estimated 10^5 protons are transferred per protein transported (Alder and Theg, 2003). The amount of potential energy that would be used in this transfer would be similar in magnitude to the energy that is estimated to be consumed by the ATP-driven Sec dependent translocon. It is

possible, however, that these protons are not providing energy for translocation, but are simply moving across the membrane when the putative “gate” is opened (a leak).

The channel, regardless of the existence of a gating mechanism or not, must first be assembled in order to function. This issue has been addressed with some success in some of the crosslinking studies that have been discussed in the previous section. The TatBC complex exists in stable high-molecular weight structures in the absence of both the pmf and cargo proteins (Bolhuis et al., 2001; De Leeuw et al., 2001). It has been observed that cargo can bind directly to the lipid bilayer before associating with the TatBC binding complex (Figure 1.5 A→B) (Bageshwar et al., 2009; Brehmer et al., 2012; Hou et al., 2006). This process can occur in the absence of any detectable pmf, consistent with results from the thylakoid system that suggest that the pmf is not necessary for the earliest steps in transport (Di Cola et al., 2005). In the presence of a pmf, however, small amounts of TatA are recruited to the TatBC receptor complex during the early stages of the Tat cycle (Zoufaly et al., 2012). In the later stages of transport, the substrate primarily makes contacts with TatA (Maurer et al., 2010). It has been suggested that the initial small amount of TatA may serve as a nucleation site for assembly of higher order TatA oligomers (Frobel et al., 2011). Several studies in the thylakoid Tat system have suggested that this assembly of Tha4 (TatA) occurs only in the presence of a substrate (Dabney-Smith and Cline, 2009; Dabney-Smith et al., 2006). Support for this comes from

real-time fluorescence experiments (using TatA fused to the fluorescent protein YFP) that found that in the absence of cargo, TatA formed lower-order oligomers than it did when substrates were available (Leake et al., 2008). It should be noted, however, that this TatA-YFP fusion transported cargo with a much lower efficiency compared to wild-type, possibly a consequence of an oligomerization effect.

Assembly of TatA pores in response to a bound substrate is a popular hypothesis because it addresses one of the major challenges in describing a model of Tat translocation: the necessity of maintaining an ion-tight membrane. The translocon appears to be quite versatile in its ability to accommodate a wide range of differently sized substrates, and any model that seeks to explain it must allow for this flexibility without uncoupling the membrane potential. The Sec machinery, in contrast, needs only a single sized channel capable of accommodating an unfolded polypeptide, highlighting the uniqueness of the bacterial Tat system. It has been proposed that by having TatA assemble from monomers or low-order oligomers in response to a bound cargo, the machinery can create a pore that is large enough to accommodate the precursor protein without allowing the unnecessary passage of ions (Palmer and Berks, 2012). Despite the popularity of this “channel model” where TatA (or TatE) are part of a pore, an alternative model has been proposed in which cargos are conducted across the membrane via a destabilized region of lipids (Cline and McCaffery, 2007). This model suggests that the lipids are allowed to assemble around the

translocating cargo, preventing any leakage of ions. Other models have suggested a more active role of the amphipathic helix of TatA. In one such model, the amphipathic helix reorientates itself into the lipid bilayer in response to a force on the translocating substrate. In this scheme, the helices would shield the cargo from the hydrophobic bilayer's interior by having the hydrophilic side face the substrate and the hydrophobic side face the lipids (Dabney-Smith et al., 2006). A similar model has also been suggested in which the membrane bends upon insertion of the amphipathic helices into the cytoplasmic leaflet, allowing the opening of an aqueous channel (Greene et al., 2007). Once this channel is open, the cargo would then diffuse across the membrane spontaneously in accordance with its concentration gradient. These two models require that the amphipathic helix is able to reorientate itself into the lipid bilayer, in agreement with some publications that suggested a dual TatA topology (Chan et al., 2007; Gouffi et al., 2004). However, as previously mentioned, a recent study seems to refute this (Koch et al., 2012) and the limited structural data on TatA suggests that the amphipathic helix is held in a rigid ~90º angle to the transmembrane helix due to interactions of side chains at the helical joint, which also suggests that any major topological reorientation is unlikely (Chan et al., 2011; Hu et al., 2010; Walther et al., 2010).

Another unanswered question with regards to a Tat transport model is the fate of the signal peptide. It has been shown through crosslinking studies that the signal peptide alone can be recruited to the Tat machinery in a pmf

dependent manner (Gerard and Cline, 2007). Furthermore, it seems that in some cases the signal peptide can become cleaved even if the mature domain is not translocated, indicating that at some point the C-terminal end of the peptide must cross the membrane since the LepB peptidase is only active in the periplasm (Cline and McCaffery, 2007; Di Cola and Robinson, 2005). It is unknown, however, if the signal peptide then remains in the bilayer, enters the periplasm (as shown in Figure 1.5F), or the cytoplasm. More fundamentally, the way the signal peptide interacts with the lipid bilayer in the absence of the translocon is still not fully known. At least one study, albeit with an artificial substrate, has suggested that it may form a hairpin loop with two transmembrane spans (Schlesier and Klosgen, 2010). It should be noted that in at least some substrates (such as SufI), a portion of the mature domain would need to unfold to have a long enough polypeptide to form a hairpin structure like this. Alternative hypotheses include a scenario in which the signal peptide crosses the membrane only once, implying that the mature domain must be translocated and that the signal peptide must flip in this process in order for LepB to have access to the cleavage site. Another scenario is that the signal peptide forms a hairpin that does not fully transverse the bilayer, or that it merely is associated with the lipid headgroups and does not enter the bilayer at all. If either one of these conditions is the actual state of the cargo before it interacts with the translocon, significant movement of the signal peptide during transport must take place to allow LepB to process the precursor to its mature form.

Although much progress in obtaining an accurate and complete understanding of the Tat machinery has been made, our overall picture of the process remains unclear. Part of the reason for this is the lack of approaches that enable probing of Tat-dependent transport events with high time resolution. Furthermore, it has proven difficult to obtain data elucidating sub-steps of the Tat transport cycle. Most *in vitro* approaches in the field can either only observe the translocation endpoint (i.e, when the cargo has been successfully transported into the lumen of a vesicle or thylakoid) or are performed on detergent solubilized Tat assemblies that are therefore incapable of pmf-dependent transport. Chapter II of this dissertation addresses both these issues by describing a novel real-time assay that utilizes fluorescence resonance energy transfer (FRET) to observe the kinetics of interactions between Tat cargos and the translocon. Another, very different, but important issue concerning the Tat system is its potential to be utilized in the export of proteins with biotechnological applications. Several studies have provided some data on this subject by examining the Tat system's ability to transport substrates of different molecular weights or lengths (Cline and McCaffery, 2007; Musser and Theg, 2000). Chapter III uses a different approach and seeks to probe the Tat machinery by investigating its ability to transport cargos that not only vary in size, but in shape.

CHAPTER II

KINETICS OF PRECURSOR INTERACTIONS WITH THE BACTERIAL TAT TRANSLOCASE DETECTED BY REAL-TIME FRET*

SUMMARY

The *Escherichia coli* twin-arginine translocation (Tat) system transports fully folded and assembled proteins across the inner membrane into the periplasmic space. Traditionally, in vitro protein translocation studies have been performed using gel-based transport assays. This technique suffers from low time resolution, and often, an inability to distinguish between different steps in a continuously occurring translocation process. To address these limitations, we have developed an in vitro FRET-based assay that reports on an early step in the Tat translocation process in real-time. The natural Tat substrate pre-SufI was labeled with Alexa532 (donor) and the fluorescent protein mCherry (acceptor) was fused to the C-terminus of TatB or TatC. The colored Tat proteins were easily visible during purification, enabling identification of a highly active inverted membrane vesicle (IMV) fraction yielding transport rates with NADH almost an order of magnitude faster than previously reported. When pre-SufI was bound to the translocon, FRET was observed for both Tat proteins. FRET was diminished upon addition of nonfluorescent pre-SufI, indicating that the initial binding step is reversible. When the membranes were energized with

* This research was originally published in The Journal of Biological Chemistry. Whitaker N, Bageshwar UK, and Musser SM. "Kinetics of precursor interactions with the bacterial Tat translocase detected by real-time FRET" *J Biol Chem. 287*(14):11252-60.

NADH, the FRET signal was lost after a short delay. These data suggest a model in which a Tat cargo initially associates with the TatBC complex, and an electric field gradient is required for the cargo to proceed to the next stage of transport. This cargo migration away from the TatBC complex requires a significant fraction of the total transport time.

INTRODUCTION

The bacterial Tat machinery transports protein cargos from the cytoplasm to the periplasm (Santini et al., 1998; Thomas et al., 2001). A functional *E. coli* Tat machinery minimally consists of three membrane proteins, TatA (or TatE), TatB and TatC (Natale et al., 2008; Sargent, 2007b; Sargent et al., 1998; Sargent et al., 1999; Weiner et al., 1998). TatB and TatC are expressed at approximately equal levels (Bolhuis et al., 2001), and together act as the receptor for precursor proteins (Alami et al., 2003; Jack et al., 2001). TatA is expressed at higher levels than TatB and TatC (Bolhuis et al., 2001), and can form ring-like structures *in vitro* (Gohlke et al., 2005). A popular model is that cargos are first recruited by TatBC complexes, and are then conveyed across the membrane by a protein conducting channel comprised of TatA oligomers (De Leeuw et al., 2001; Gohlke et al., 2005). After transport, Tat signal peptides are cleaved from precursor proteins by the LepB peptidase (Luke et al., 2009; Yahr and Wickner, 2001). Recent studies have suggested that precursor binding to the cytoplasmic face of the inner membrane is a functional

intermediate in the transport process (Bageshwar et al., 2009; Hou et al., 2006). A proton motive force (PMF) is essential for translocon assembly and cargo transport (Braun et al., 2007; Cline et al., 1992; Mould and Robinson, 1991). In *E. coli,* only the membrane potential ($\Delta\psi$) component of the PMF is required for transport (Bageshwar and Musser, 2007).

Both TatB and TatC participate in binding and recognition of signal peptides. Bacterial Tat signal peptides contain an (S/T)RRxFLK consensus motif, a hydrophobic domain, and a short polar domain which precedes the signal sequence cleavage site (Bendtsen et al., 2005; Berks et al., 2003; Shanmugham et al., 2006). Based on mutant suppression, alanine substitutions and crosslinking experiments, the N-terminal half of TatC interacts with the twin-arginine portion of the signal peptide (Gerard and Cline, 2006; Holzapfel et al., 2007; Kreutzenbeck et al., 2007; Strauch and Georgiou, 2007b). Crosslinking indicates that TatB interacts with the C-terminal end of the signal peptide (Gerard and Cline, 2006; Holzapfel et al., 2007). TatB also makes extensive contacts with the cargo mature domain, presumably through its cytoplasmic domain (Maurer et al.).

It is unclear how the cargo proceeds across the membrane after recognition by the TatBC complex. A severe constraint on possible models is the finding that the mature domain can be efficiently translocated when the signal sequence is crosslinked to TatC near the twin-arginine motif (Gerard and Cline, 2006). One possibility is that TatA somehow assists with flipping the

mature domain from one side of the membrane to the other while the signal sequence remains tethered (Gerard and Cline, 2006). In this picture, the signal peptide may initially only partially occupy a transmembrane binding pocket before transport, and full binding only occurs during or after transport, consistent with the identification of deep insertion of the signal peptide and distinct translocation intermediates (Frielingsdorf and Klosgen, 2007; Gerard and Cline, 2007; Schlesier and Klosgen). Electron microscopic images of purified TatBC complexes revealed cargo bound within the membrane plane, possibly reflecting cargo in transit (Tarry et al., 2009b).

Fluorescence resonance energy transfer (FRET) can provide time-dependent distance information at low protein concentrations. Here we report the use of FRET to probe the binding of a cargo to the TatBC complex, and its migration away from this complex, either by dissociation or by movement along the transport pathway. We found that the TatBC complex has a nanomolar affinity for pre-SufI, and that an electric field gradient is required for migration beyond the initial binding step.

RESULTS

Experimental Design

Our goal was to examine the interaction between the TatBC receptor complex and a cargo under real-time transport conditions. To this end, we used FRET, with a donor fluorophore (Alexa532) on the cargo and an acceptor

fluorophore (mCherry) on one of the Tat proteins. We assumed that binding would result in a decrease in donor fluorescence due to FRET. The fluorescent cargo selected was the natural Tat substrate pre-SufI. Eight different single cysteine mutants of pre-SufI were generated. The different cysteine mutations were spread over the surface of the protein (Figure 2.1A). These mutants were labeled with Alexa532 maleimide. The C-terminal cysteine mutant, pre-SufI(479C), serves as our wildtype (wt) reference because previous experiments indicated that a C-terminal dye had no effect on transport efficiency (Bageshwar et al., 2009). The T96C and 479C mutants yielded the highest transport efficiencies (Figure 2.1B).

In order to exclude any endogenous (and therefore unlabeled) Tat proteins from IMV preparations, $TatB^{cherry}$ and $TatC^{cherry}$ were expressed in the Tat deletion strain MC4100ΔTatABCDE. The mCherry domain did not significantly affect transport efficiency (Figure 2.1C) or transport kinetics (Figure 2.1D). The fluorescent TatB and TatC proteins allowed us to visually monitor the IMV purification process. Consequently, we identified a minor membrane-containing band within the 0.5 M sucrose region of the sucrose gradient that had the majority of the mCherry protein (Figure A11). The MC4100 membranes recovered from this band catalyzed pre-SufI transport about an order of magnitude faster (Figure 2.1D) than reported previously for JM109 IMVs (Bageshwar and Musser, 2007; Bageshwar et al., 2009). Further, IMVs prepared with our current protocol were more consistently active. Previously,

we were not able to consistently obtain active membranes from MC4100 (Bageshwar and Musser, 2007).

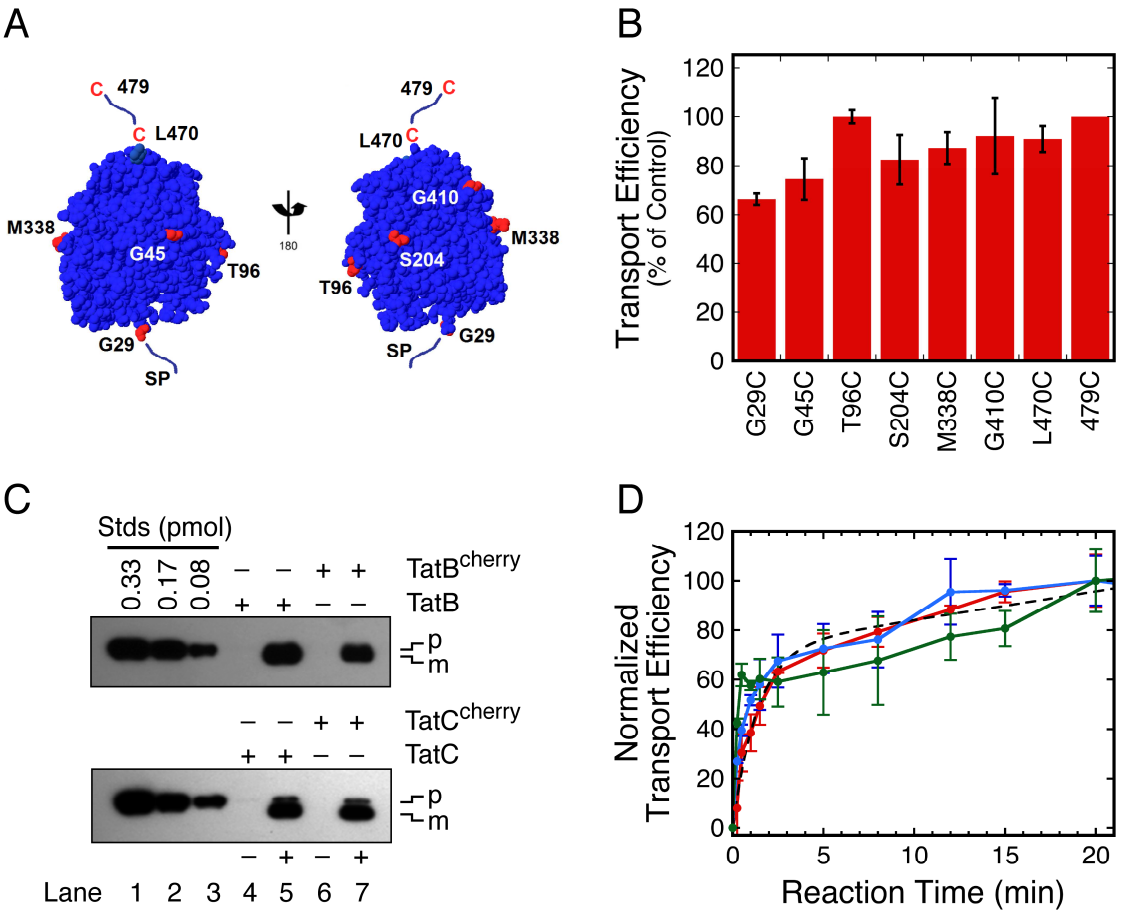


Figure 2.1. Pre-SufI cysteine mutants and influence of TatB^{cherry} and TatC^{cherry} on Tat transport. (*A*) Location of single cysteine mutations. Residues mutated to cysteine are indicated in *red* (PDB accession number: 2UXT). Cysteines were labeled with Alexa532 maleimide for transport and FRET experiments. (*B*) Transport efficiencies of Alexa532 labeled pre-SufI mutants. Transport reactions (30 min) were performed with TatABC IMVs (A_{280} = 5), 4 mM NADH and 90 nM pre-SufI. Transport efficiencies were normalized to the 479C mutant (N = 3). (*C*) Transport of pre-SufI(479C)^{Alexa532} into IMVs containing TatAB^{cherry}C, TatABC^{cherry} or wt TatABC. Transport requires NADH to generate the necessary $\Delta\psi$. The mCherry domain has no apparent effect on transport efficiency. Lanes 1-3 are pre-SufI concentration standards. The location of precursor (p) and mature (m) molecular weight bands are identified. Conditions are the same as in (*B*). (*D*) Kinetics of pre-SufI(479C)^{Alexa532} (20 nM) transport into IMVs (A_{280} = 2) containing wt TatABC (*blue*), TatAB^{cherry}C (*green*), and TatABC^{cherry} (*red*). Transport efficiencies were normalized based on the 20 min time point. The black dashed curve is a single exponential plus linear baseline fit to the TatABC^{cherry} data (τ = 81 s).

The FRET Signal

IMVs isolated from *E. coli* expressing TatAB^{cherry}C or TatABC^{cherry} exhibited two fluorescence peaks at ~610 nm and ~641 nm (Figures A12A and A12B). These peaks both arise from mCherry (Figure A12C). Addition of pre-SufI^{Alexa532} resulted in a third peak at ~550 nm (Figure 2.2A). The decrease in donor fluorescence signal due to FRET was most easily verified by addition of a competitor protein (non-fluorescent cargo), which resulted in an increase in the 550 nm peak, presumably due to replacement of the bound fluorescent protein with the non-fluorescent competitor (Figure 2.2A). No increase in 550 nm emission was observed upon competitor addition if mCherry was not attached to the Tat proteins (Figure A12D), indicating that the decrease in 550 nm emission when the fluorescent cargo was bound to the Tat translocon (Figure 2.2A) was indeed due to FRET. The strong mCherry fluorescence emission far from the excitation wavelength of 500 nm indicates a high concentration relative to the cargo. The mCherry concentration was estimated as 322±88 nM (N = 5) under typical assay conditions, over an order of magnitude higher than the pre-SufI concentration of 20 nM (Figure 2.2A). These values are consistent with the amounts of pre-SufI^{Alexa532} and mCherry needed to approximately reproduce the emission spectra in Figure 2.2A in an IMV-free mixture (Figure A12D). Any increase in mCherry emission due to FRET was weak and not reliably detected. This is expected if the mCherry acceptor molecules self-quench due to their proximity in a TatBC oligomer (Tarry et al., 2009b).

Maximizing the FRET Efficiency

Having determined that FRET indeed occurred upon binding of the cargo to the Tat translocon, we next sought to maximize the FRET signal. This was accomplished by individually attaching the donor dye to the 8 locations identified in Figure 2.1A, which cover the surface of pre-SufI. In addition, the mCherry acceptor was attached to either TatB or TatC. FRET signal intensity was determined from kinetic experiments in which non-fluorescent competitor cargo was added to IMVs with prebound fluorescent cargo (Figure 2.2B). For pre-SufI(T96C), the competitor released the bound cargo with τ = 24±2 s (Figure 2.2B). Control kinetic experiments in the absence of a donor or acceptor fluorophore confirmed the transient FRET signal (Figure A13A). We found that $TatC^{cherry}$ yielded stronger FRET signals than $TatB^{cherry}$ (Figure 2.2C). Two pre-SufI mutants, G29C and T96C, yielded similarly strong FRET signals when the mCherry domain was attached to TatC (Figure 2.2C). However, labeling at G29C reduced transport efficiency (Figure 2.1B). To convert the relative FRET signals into reliable distance and/or orientation information, the data in Figure 2.2C need to be corrected based on binding affinity. We attempted to do this using our previous binding assay (Bageshwar et al., 2009), but the errors of these binding measurements and the errors of the FRET measurements were too high to provide meaningful information. Instead, we settled on using $TatC^{cherry}$ and the T96C mutant of pre-SufI for all subsequent experiments since

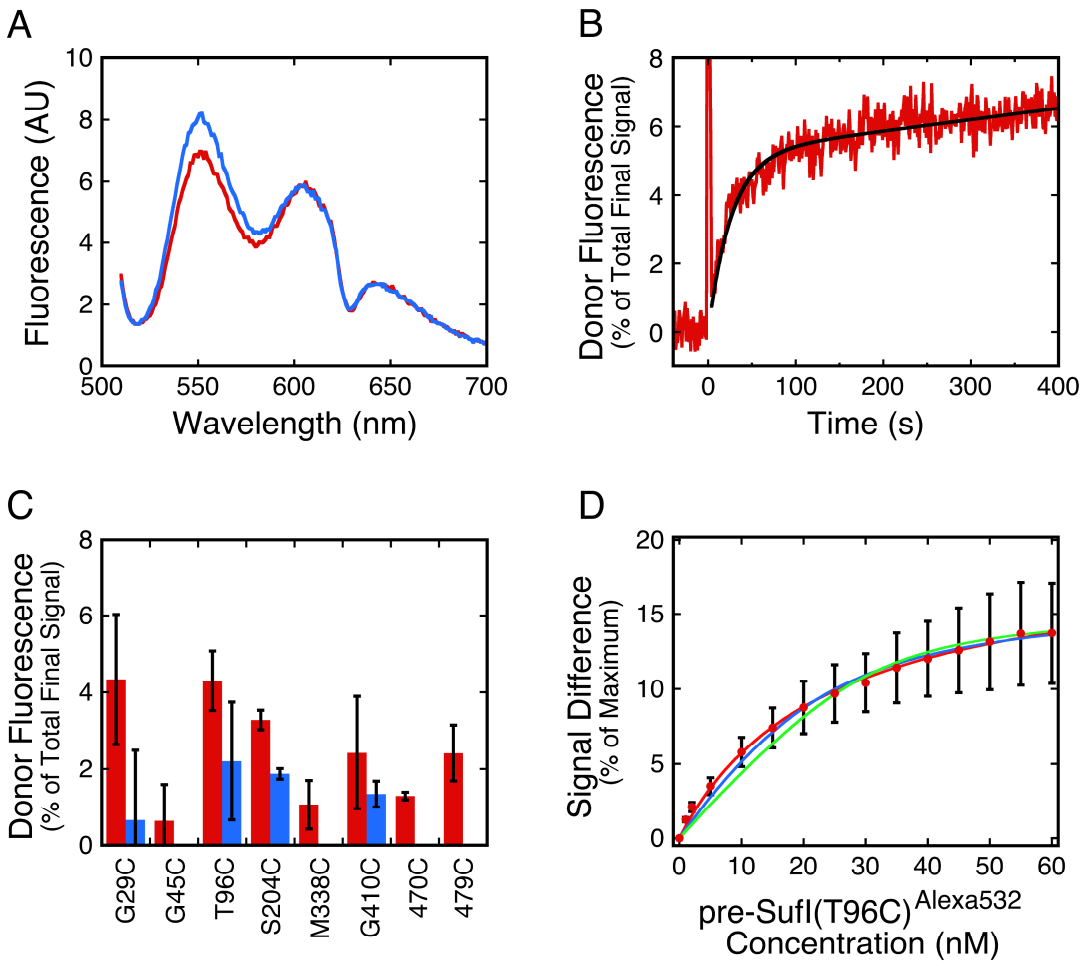


Figure 2.2. FRET between the pre-SufI cargo and the Tat translocon. (*A*) Emission scan (EX = 500 nm) of pre-SufI(T96C)^{Alexa532} (20 nM) and TatABC^{cherry} IMVs (A_{280} = 2) before (*red*) and after (*blue*) addition of 200 nM unlabeled pre-SufI(T96C). The increase in signal at 550 nm indicates loss of FRET. The overall signal change is not a good measure of FRET efficiency since a high fraction of the cargo is not bound to Tat translocons (i.e., free in solution or bound to the lipids). (*B*) Time trace of the fluorescence emission at 550 nm (donor) for the experiment in (*A*). Unlabeled pre-SufI(T96C) was added at τ = 0 s. The competitor-induced loss of FRET occurred with a time constant of 24±2 s (N = 3). (*C*) Total FRET signal observed for various pre-SufI mutants, determined as in (*B*). IMVs contained TatABC^{cherry} (*red*) or TatAB^{cherry}C (*blue*) (N = 3). The absence of blue bars for some mutants indicates no FRET to TatB^{cherry}. (*D*) Precursor binding affinity and receptor concentration estimated from the concentration dependence of the FRET signal. TatABC^{cherry} IMVs were titrated with pre-SufI(T96C)^{Alexa532} in the presence and absence of 300 nM unlabeled pre-SufI(T96C). Shown here is the average difference (N = 5) between two titration curves, such as those shown in Figure A14. Three fits are shown in which the receptor concentration (T_0) was fixed and the K_D and maximum signal were fitting parameters: (*red*) T_0 = 0.1 nM, K_D = 23 nM; (*blue*) T_0 = 20 nM, K_D = 7.5 nM; (*green*) T_0 = 30 nM, K_D = 3.6 nM. These fits indicate that T_0 ≈ 0-20 nM and K_D ≈ 7-23 nM. Details of the analysis are described in the Supplementary Material.

these proteins yielded the strongest FRET signal while retaining transport efficiency (Figures 2.1B and 2.2C).

Binding Affinity

The FRET signal was used to estimate the cargo-TatBC binding affinity. Two titrations were done. The first titration involved adding increasing amounts of fluorescent pre-SufI(T96C) to IMVs. The second titration was identical to the first, except that a large excess of non-fluorescent cargo was preincubated with the IMVs (Figure A14). The difference in donor fluorescence intensity between the two titrations (FRET signal) reflects the amount of fluorescent cargo bound to the translocon (Figure 2.2D). Based on the mCherry concentration, the Tat receptor concentration is high relative to the apparent K_D. Thus, the data in Figure 2.2D were fit with equation 2 (Experimental Procedures and Supplementary Material), which explicitly includes the receptor concentration as a fit parameter. Unfortunately, the receptor concentration and K_D are not uniquely determined from the data. However, the data indicate that the functional receptor concentration is ≤ ~20 nM and the K_D ≈ 7-23 nM (see Figure 2.2D and Supplementary Material for a detailed explanation).

FRET Decreases Upon Membrane Energization

When fluorescent pre-SufI was prebound to IMVs containing Tat-mCherry fusions and the membranes were subsequently energized by the addition of

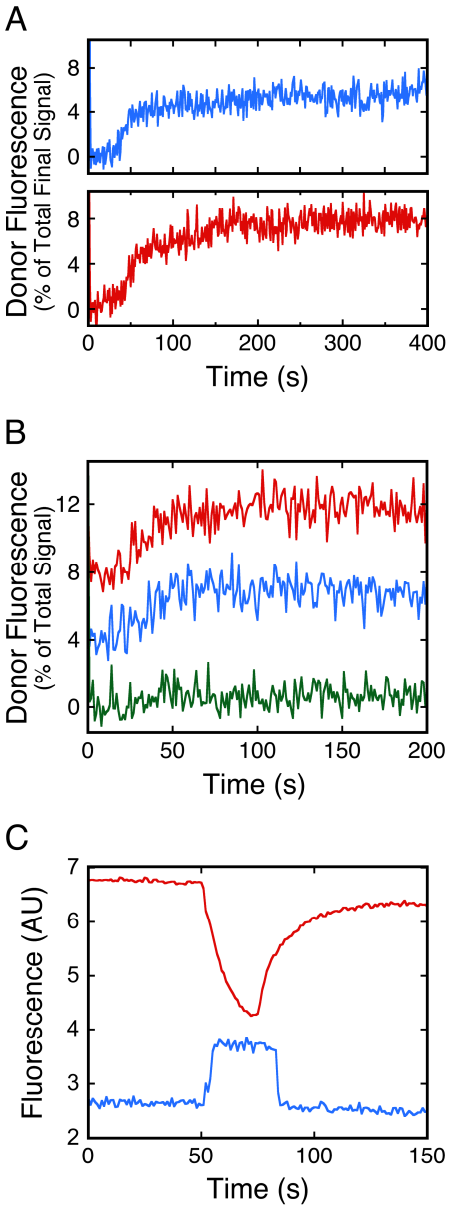


Figure 2.3. Effect of PMF components on the FRET signal. (*A*) Time trace of the fluorescence emission at 550 nm (donor) upon membrane energization. Reactions contained pre-SufI(T96C)^{Alexa532} (20 nM) and IMVs (A_{280} = 2) with TatAB^{cherry}C (*blue*) or TatABC^{cherry} (*red*). NADH (4 mM) was added at t = 0 s. (*B*) Sensitivity of NADH-dependent fluorescence changes to components of the PMF. The ΔpH and Δψ were reduced with nigericin (5 μM, *blue*) and valinomycin (5 μM, *green*), respectively. The control trace (*red*) contains no ionophores. Conditions as in (*A*) with TatABC^{cherry} IMVs (N = 2). (*C*) Gradients across membranes of TatABC^{cherry} IMVs (A_{280} = 2). The presence of Δψ (*blue*) and ΔpH (*red*) gradients were determined using 100 nM oxonol VI (EX = 610 nm, EM = 645 nm) and 2.5 μM quinacrine (EX = 420 nm, EM = 510 nm), respectively, as described previously (Bageshwar et al., 2009). NADH (4 mM) was added at t = 50 s.

NADH, the donor fluorescence signal increased (a decrease in FRET). This was observed for membranes containing either $TatB^{cherry}$ or $TatC^{cherry}$ (Figure 2.3A). Control experiments in the absence of donor or acceptor fluorophore confirmed the NADH-dependent loss of FRET (Figure A13B). Since NADH addition initiates transport across the membrane, these data are consistent with migration of the cargo away from its initial binding site on the TatBC complex as part of the transport process. It is unclear if the observed kinetics reflect migration elsewhere within the translocon (e.g., to the TatA pore) or immediate movement across the membrane. NADH generates a PMF, which is necessary for Tat transport. It was shown earlier that the $\Delta\psi$ and not the ΔpH component of the PMF is essential for Tat transport (Bageshwar and Musser, 2007). We therefore tested whether the observed FRET signal is sensitive to these two PMF components. We found that the decrease in FRET upon NADH addition requires a $\Delta\psi$ and not a ΔpH (Figure 2.3B), consistent with the hypothesis that the observed changes in the FRET signal report a transport substep.

A $\Delta\psi$-dependent Step Precedes Cargo Migration Away from the TatBC Complex

Exponential NADH-dependent FRET changes (τ = ~30 s) were preceded by a lag phase (Figure 2.3A). The duration of the delay was dependent on the batch of IMVs (e.g., compare Figures 2.3A and 2.3B), and ranged from ~20 to ~45 s (*N* = 9). This delay did not arise from slow formation of the $\Delta\psi$, as the $\Delta\psi$

forms within seconds (Figure 2.3C). While it takes ~25-30 s for the ΔpH to become fully established (Figure 2.3C), it was shown earlier (Bageshwar and Musser, 2007) and here (Figure 2.3B) that the ΔpH is not necessary for transport. Therefore, the slow formation of the ΔpH does not explain the lag phase. Consequently, we reasoned that the lag phase could be explained by a Δψ-dependent conformational change or oligomerization process that is required for the cargo to migrate away from the TatBC binding site. We tested this hypothesis by preincubating the IMVs with NADH and then adding the cargo after a short delay (Figure 2.4). The lag phase largely disappeared when the cargo was added 200 s after NADH (Figure 2.4C). Under these conditions, slower kinetics (τ = ~70-90 s) were observed.

Pre-SufI is Released From TatBC Complexes when Membranes are Energized in the Absence of TatA

As a control experiment for the NADH-dependent loss of FRET (Figure 2.4A), we repeated these experiments in the absence of TatA. Similar levels of TatB and $TatC^{cherry}$ were recovered in IMVs in the presence and absence of TatA (Figure A15). Since TatA is required for transport, our expectation was that the cargo would remain bound to the TatBC complex, and hence, the FRET signal would remain constant. Experiments similar to those in Figures 2.3A and 4A were performed, where pre-SufI was allowed to bind to IMVs, and NADH

was added at time zero. FRET immediately decreased (τ = 71 s) when the membranes were energized in the absence of TatA (Figure 2.5A). This differed

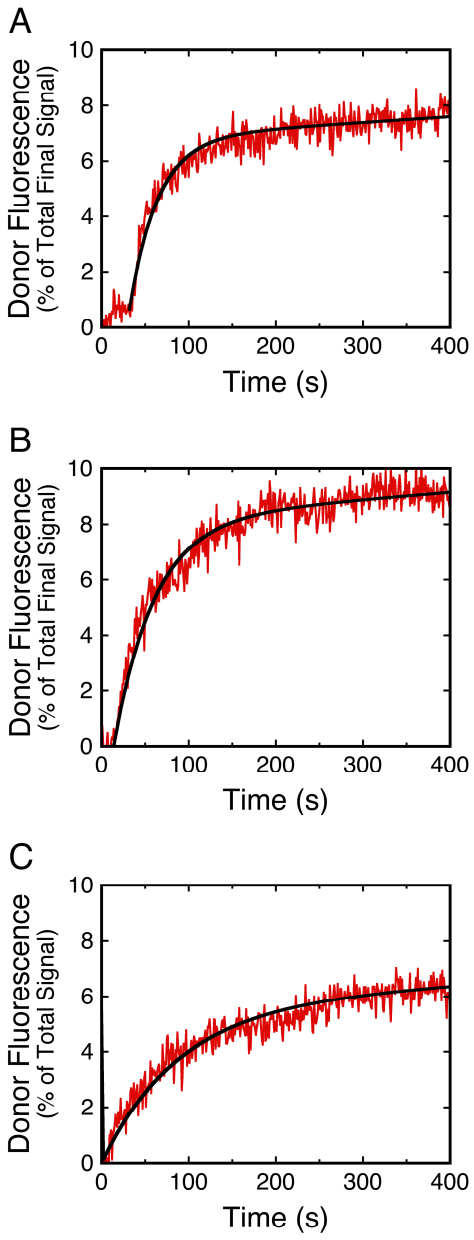


Figure 2.4. Cargo migration kinetics depend on the interval between energization and cargo addition times. The time-dependent donor fluorescence emission (550 nm) was measured for transport reactions containing TatABC^{cherry} IMVs (A_{280} = 2) and pre-SufI(T96C)^{Alexa532} (20 nM). Membranes were energized with NADH (4 mM) and the cargo was added 0 s (*A*), 10 s (*B*) or 200 s (*C*) later. Cargo was added at τ = 0 s. The kinetics after the lag phase in (*A*) and (*B*) were fit with a single exponential plus a linear baseline drift (Experimental Procedures), yielding τ = 33 s and 46 s, respectively. The entire trace in (*C*) was fit with the same equation, yielding t = 93 s. The linear baseline drift accounts for ~11% of the total fit fluorescence change for all three panels. Each trace represents an average of 3 individual runs.

from results obtained with IMVs containing TatA (Figures 2.3A and 2.4A), which exhibited a lag-phase and rapid loss of FRET after the lag phase (τ = ~30 s). Since the Tat machinery is incapable of transport without TatA (or its paralog, TatE) (Sargent et al., 1998), the observed loss of FRET is not due to cargo translocation across the membrane, as confirmed in Figure 2.5B. Rather, likely possibilities are that the loss of FRET arises either from a conformational change that results in the mature domain of the bound precursor moving away from the mCherry fluorophore, or from the precursor dissociating from the TatBC complex entirely. The latter appears to be the case since TatBC IMVs pelleted after energization retain ~53% less bound pre-SufI than unenergized IMVs (Figure 2.5B). Multiple groups have suggested the possibility that the signal peptide can penetrate fairly deeply into the membrane (Gerard and Cline, 2007; Panahandeh et al., 2008; Schlesier and Klosgen, 2010; Strauch and Georgiou, 2007b). The energy-dependent dissociation of pre-SufI from the receptor complex in the absence of TatA suggests the possibility that the signal sequence may not have penetrated as deeply under these conditions.

DISCUSSION

Tat transport requires both binding of a pre-protein to the Tat proteins as well as movement of the cargo across the membrane. Presumably, these two seemingly conflicting processes are coupled, i.e., the binding reaction gates the transport event. Critical to understanding this process is the development and

analysis of assays that can identify and discern the conversion between substeps of transport. To this end, we developed a FRET assay that reports the initial cargo binding step. Our major conclusions are: 1) transport occurs on the minute timescale with a highly purified membrane fraction (Figure 2.1D); 2.2) the cargo binding interaction occurs with an apparent K_D ≈ 7-23 nM (Figure 2.2D) and a competitor-induced k_{off} ≈ -0.042 s^{-1} (= 1/24 s) (Figure 2.2B); 3) the cargo mature domain appears to be nearer to the C-terminus of TatC than the C-terminus of TatB (Figure 2.2C); 4) migration of the cargo from its initial binding site requires a $\Delta\psi$ (Figure 2.3B); 5) a delay in cargo migration from the initial binding site occurs after membrane energization (Figures 2.3 and 2.4), indicating that the $\Delta\psi$ does not *directly* promote cargo migration; and 6) TatA increases the affinity of the TatBC receptor complex for the precursor in the presence of a membrane potential (Figure 2.5). The implications of these results are now discussed.

As far as we are aware, the binding experiments reported here provide the first estimate of the binding affinity of a signal peptide for functional *E. coli* Tat translocons. The estimated K_D of 7-23 nM indicates a fairly strong, highly specific interaction, as is reasonably expected for a selective process. From this measured K_D and the competitor-induced k_{off} ≈ 0.042 s^{-1} (Figure 2.2B), the apparent k_{on} is 10^6-10^7 M^{-1} s^{-1}. This second order rate constant is of the magnitude expected for enzymatic interactions, where orientation of the reactants is clearly important, and is complicated by the fact that it likely reflects

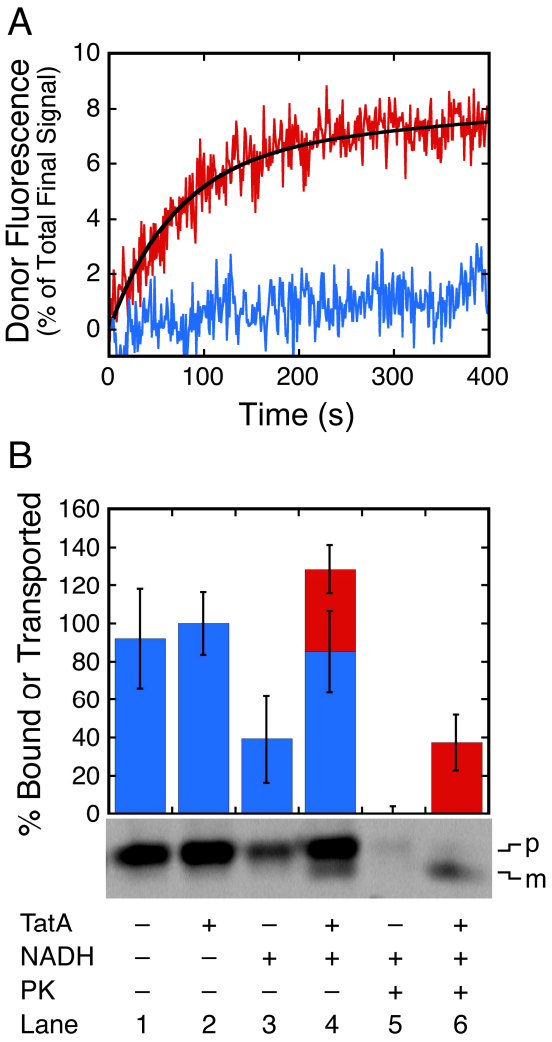


Figure 2.5. Transport and binding to energized membranes in the absence of TatA. (*A*) Time traces of donor emission (at 550 nm) in the presence or absence of mCherry after energization with 4 mM NADH at $\tau = 0$. Reactions contained pre-SufI(T96C)^{Alexa532} (20 nM) and IMVs with TatBC^{cherry} (*red*) or TatBC (*blue*) (A_{280} = 2). The TatBC^{cherry} data were fit with a single exponential plus a linear baseline drift, yielding τ = 71 s (the baseline drift accounts for 13% of the total fluorescent change). (*B*) Membrane binding and cargo transport in the presence and absence of TatA. Pre-SufI(T96C)^{Alexa532} (20 nM) was allowed to bind to membranes (TatBC or TatABC IMVs, A_{280} = 2) for 10 minutes at 37°C, respectively. After a 10 min binding period, control samples were then centrifuged (16,000 g, 45 min, room temperature) and the pellet fractions were analyzed by SDS-PAGE (lanes 1-2). Resuspended membranes were energized with 4 mM NADH for 20 minutes at 37°C (lanes 3-6). Proteinase K (PK, 0.7 mg/mL) was added to half the reactions and all reactions were immediately centrifuged (16,000 g, 45 min, room temperature). The resultant membrane pellets were resuspended in translocation buffer containing 68 mM PMSF and resolved via SDS-PAGE. The amounts of recovered precursor (p, *blue*) and mature (m, *red*) proteins are shown above the sample gel, as indicated (N = 3). Less pre-SufI was recovered with membranes lacking TatA (compare lanes 3 and 4), indicating that the TatBC affinity for pre-SufI is lower without TatA in the presence of a PMF. Matured protein was protease protected (compare lanes 4 and 6), indicating transport. The average amount of protein initially recovered with TatABC membranes (lane 2) was set to 100%.

both binding to the membrane lipids (Bageshwar et al., 2009) as well as lateral diffusion to the Tat translocon. Whereas the precursor can apparently exchange between lipid- and translocon-bound forms on the tens of seconds timescale, release from the membrane surface is significantly slower (Bageshwar et al., 2009).

It is unclear whether the signal peptide binding interaction was broken at the point at which the FRET interaction was lost under transport conditions, or whether there is sufficient flexibility between the mature domain and signal peptide such that large movements of the mature domain away from the C-termini of TatB and TatC can occur without dissociation of the signal peptide. The latter is consistent with the finding that the thylakoid Tat machinery can transport cargos covalently linked to TatC via their signal peptide (Gerard and Cline, 2006). In contrast, the NADH-dependent release of the cargo from membranes containing TatBC complexes but no TatA (Figure 2.5) indicates a weaker membrane binding affinity in the presence of a $\Delta\psi$ under these conditions, and thus clear dissociation of the signal peptide from the TatBC receptor complex under these conditions.

The analysis of the data in Figure 2.2D indicates that the functional TatBC receptor concentration is ≤ ~20 nM (Supplementary Material). Thus, the functional binding site concentration is significantly lower than the number of TatC molecules (322±88 nM) calculated based on the mCherry concentration. Note that TatABC^{cherry} IMVs contained only full-length TatC^{cherry} (Figure A15),

indicating that the mCherry fluorescence signal was not contaminated by degradation products of the fusion protein. One possibility is that most TatC molecules are inaccessible or non-functional, possibly a consequence of overexpression. Tarry and coworkers (Tarry et al., 2009b) found that cargos bind to only a few available binding sites in TatBC oligomers, which they estimated to be heptamers. According to our data, active heptamers could account for ~140 nM (= 7 x 20 nM) of the TatC molecules present. The remaining ~180 nM of the TatC molecules present may then be inactive or inaccessible. However, a recent report with covalent TatC dimers suggests that TatBC complexes contain an even number of TatC molecules (Maldonado et al., 2011). Thus, the oligomerization state of active TatBC complexes remains uncertain.

For all the pre-SufI single cysteine mutants tested, FRET signals were higher with $TatC^{cherry}$ than with $TatB^{cherry}$. This is seemingly at odds with a recent crosslinking study, which suggested that pre-SufI binds more closely to TatB than to TatC (Maurer et al., 2010). However, the observed FRET signal arises from the proximity of pre-SufI's mature domain to the mCherry protein, which was at the C-termini of the Tat proteins. Our data do not exclude the possibility that the dye molecules on the pre-SufI mutants are closer to the amphipathic region of TatB than to the TatB C-terminus. Consequently, the models of Maurer et al. (Maurer et al., 2010) are entirely consistent with our FRET data.

A model that has emerged for the *E. coli* Tat system and the pre-SufI cargo is that the cargo first binds to the membrane lipids. From there, the cargo laterally diffuses to the TatBC complex. Translocation across the membrane, presumably with the assistance of TatA oligomers, requires a $\Delta\psi$ (Bageshwar et al., 2009). From studies on the thylakoid Tat system, the oligomerization of a TatBC complex with a TatA complex requires a PMF (Mori and Cline, 2002). It seems safe to deduce, then, that such an oligomerization process is responsible for the lag phase observed in Figure 2.3A. Loss of FRET is fairly rapid after this lag phase, possibly indicating that migration of the cargo mature domain from TatBC is fairly quick after TatABC oligomerization is complete. Figure 2.4C indicates that migration after binding occurs essentially immediately if the cargo is added ~200 s after $\Delta\psi$ generation has been initiated, possibly indicating that TatABC oligomerization occurred before the cargo bound to TatBC. The cargo migration away from TatBC is ~2.6-fold slower under these conditions, likely due to the fact that the detectable $\Delta\psi$ had already collapsed (Figure 2.3C; see also (Bageshwar et al., 2009)) when the cargo was added. Receptor binding and migration of the mature domain away from the TatBC complex occurs concurrently in Figure 2.4C. The absence of a detectable increase in FRET at the beginning part of the kinetics in Figure 2.4C indicates a rapid binding step ($\tau \leq$ ~3 s), which includes both lipid binding and diffusion to the Tat translocon. Our data are therefore consistent with the hypothesis that TatBC and TatA oligomerization occurs in the presence of a $\Delta\psi$, with or without the cargo.

Further, the $\Delta\psi$ makes the translocation system receptive to cargo movement from the receptor binding site (e.g., by a conformational change, or the oligomerization itself), but does not appear to directly drive it energetically. We expect that under in vivo conditions where a large $\Delta\psi$ is consistently maintained, all three Tat components can and do oligomerize and cargo transport occurs on the sub-minute timescale.

According to the oligomerization model discussed in the previous paragraph, TatA and TatBC form separate oligomers in the absence of a cargo and a PMF (Bolhuis et al., 2001; de Leeuw et al., 2002; Gohlke et al., 2005; Maldonado et al., 2011). Thus, our expectation was that the stability of the TatBC-precursor complex should be unaffected by the absence of TatA in the presence or absence of a $\Delta\psi$. We were therefore surprised that pre-SufI was released from IMVs upon addition of NADH in the absence of TatA (Figure 2.5B). We interpret these data to imply that the receptor complex has a weaker affinity for pre-SufI in the absence of TatA when a $\Delta\psi$ is present. This appears inconsistent with an oligomerization model in which TatA is not part of the receptor complex. However, an alternative, revised model is that at least some TatA is part of the receptor complex, consistent with earlier results (Bolhuis et al., 2001; de Leeuw et al., 2002). This picture is also consistent with a recent study, which reports crosslinking of pre-SufI to TatA in the absence of a PMF (Frobel et al., 2011). In light of these data suggesting that TatA is part of the TatBC receptor complex, it is not at all surprising that the absence of TatA

affects the interaction between TatBC and pre-SufI. According to our FRET assay, a weaker receptor complex binding affinity in the absence of TatA is only apparent in the presence of a PMF. TatA is necessary to keep pre-SufI bound to the TatBC complex in the presence of a PMF, eventually leading to transport (compare Figures 2.4A and 2.5A).

In summary, we have isolated membranes exhibiting rapid Tat translocation activity. For TatABC^{cherry} IMVs, the overall translocation process is characterized by τ = ~80 s (Figure 2.1D). When a $\Delta\psi$ is generated across these IMV membranes with bound precursor protein, migration away from the TatBC binding site occurs with τ = ~30 s after a significant (20-45 s) lag phase (Figure 2.4A). Migration away from the TatBC binding site occurs more slowly, but with no lag phase, if the precursor is added to previously energized membranes (τ = ~90 s, Figure 2.4C), suggesting that the effect of the $\Delta\psi$ slowly wears off or that transport in the presence of a weak, undetectable $\Delta\psi$ (Bageshwar et al., 2009) occurs more slowly. Binding to the membrane surface and diffusion to the TatBC complex occurs rapidly ($\tau \leq$ ~3 s). Therefore, the migration process in which the cargo moves away from the TatBC complex requires a significant fraction (~40%) of the total transport time. It is unclear whether the migration step itself is rate-limiting or whether an upstream kinetic step limits the migration rate. The gel-based transport assay has poor time resolution, and, since it is not a real-time assay, it is not clear whether any transport substeps can occur after quenching the reaction. Consequently, it remains unclear whether the cargo

proceeds directly across the membrane bilayer upon leaving the TatBC complex or whether there is an additional kinetically significant intermediate, e.g., cargo within the TatA pore. Further experiments are needed to clarify these downstream events.

METHODS

Bacterial Strains, Plasmids, and Growth Conditions

E. coli strains MC4100ΔTatABCDE, JM109, and BL21(λDE3) were described previously (Casadaban and Cohen, 1979; Studier et al., 1990; Wexler et al., 2000; Yanisch-Perron et al., 1985). Cultures for IMV preparations were grown at 37°C as described (Bageshwar and Musser, 2007). Overexpression cultures for pre-SufI and spTorA-mCherry-MA were grown in Luria-Bertani (LB) medium at 37°C, and shifted to 15°C upon induction with 1 mM IPTG or 0.6% arabinose, respectively. All cultures were supplemented with ampicillin (50 µg/mL). Tat proteins were induced for 4 hours with 0.7% arabinose, as previously described (Bageshwar and Musser, 2007).

The plasmid pIntein-preSufI-MA was generated by inserting a cysteine-free form of pre-SufI with a 6xHis tag into the vector pTYB11 (New England Biochem). Plasmid pIntein-preSufI-MA was used as the template for all pre-SufI mutants, as described in Figure A5. Plasmids pTatAB^{cherry}C and pTatABC^{cherry} were generated by insertion of the mCherry sequence (plasmid pmCherry, Clontech) into pTatABC (Yahr and Wickner, 2001) immediately after the coding

regions of TatB and TatC, respectively (Figures A6 and A7). pTatBC and pTatBC^{cherry} were generated by excising the TatA sequence from pTatABC and pTatABC^{cherry}, respectively (Figure A8). All mutations were generated via the Quikchange protocol (Stratagene), and all plasmid coding sequences were verified by DNA sequencing. Plasmid pTorA-mCherry-H6 was generated by inserting mCherry along with the TorA signal peptide and 4 amino acids (AQAA) of the TorA mature domain into the pBAD24 vector (Figure A9).

Protein Purification and Labeling

All pre-SufI proteins were purified from expression cultures via chitin chromotography as described previously (Chong et al., 1998) with an on-column cleavage (50 mM dithiothreitol) at 4°C for 20 hours (Figure A10). Excess dithiothreitol was removed from the eluate by dialysis (10,000 MW cutoff) against phosphate buffered saline (1X PBS: 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, and 2 mM KH_2PO_4, pH 7.5) for 4 hours. Ni-NTA chromatography was used for further purification (Bageshwar and Musser, 2007). Alexa532 labeling of pre-SufI proteins was performed as described previously (Bageshwar et al., 2009). After a 4 hour dialysis against 1X PBS buffer, excess dye was removed by an additional Ni-NTA purification step (Bageshwar et al., 2009). The final storage buffer was 20 mM Tris-HCl, 50 mM NaCl, 250 mM imidazole, 50% glycerol, pH 8.0. Purification of spTorA-mCherry-H6 was performed by Ni-NTA chromatography, by the procedure used for pre-SufI (Bageshwar and Musser,

2007). The mCherry protein was obtained by purifying the mature protein from a spTorA-mCherry-H6 preparation via Resource Q (GE Healthcare) chromatography.

The protein concentrations of all pre-SufI proteins were quantified by SDS-PAGE using bovine serum albumin as the standard. Spot intensities after Coomassie Blue staining were determined with a phosphorimager (model FX PhosphoImager, Bio-Rad Laboratories). Alexa532 concentrations were determined by absorbance spectroscopy at 532 nm (e = 81,000 $cm^{-1}M^{-1}$). Typical dye-to-protein ratios after labeling indicated that 80-90% of the cysteines were successfully tagged. mCherry concentrations were estimated in 2% SDS based on the mCherry absorbance at 587 nm in and its extinction coefficient at this wavelength (Shaner et al., 2004).

Isolation of IMVs

IMVs were isolated from MC4100ΔTatABCDE as described (Bageshwar and Musser, 2007), with the following modifications. Cell lysis by French press was performed at ~16,000 psi instead of ~6,000 psi. In addition, the 2.3 M sucrose cushion was replaced with a 3-step (0.5 M, 1.5 M, and 2.3 M) sucrose gradient, which enabled enrichment of a highly active inner membrane fraction. A band in the 0.5 M region was faintly pink when mCherry Tat fusions were expressed (Figure A11), and translucent dark brown when wild type TatABC

was expressed. IMV concentrations were determined as the A_{280} in 2% SDS (Bageshwar and Musser, 2007).

Transport and Membrane Binding Assays

In vitro transport assays and precursor-membrane binding assays were performed as described (Bageshwar and Musser, 2007; Bageshwar et al., 2009), unless otherwise indicated. Gel based transport kinetics were performed in a heated cuvette to mimic conditions in which the real-time FRET assays were performed. Aliquots (35 µL) of the reaction mixture were removed from the cuvette at the indicated time points and quenched with a mixture of 5 µM nigericin and 5 µM valinomycin on ice. Visualization of protein bands was performed by direct in-gel fluorescence imaging (model FX PhosphorImager, Bio-Rad Laboratories).

Real-time FRET Assay

Translocation buffer (5 mM $MgCl_2$, 50 mM KCl, 200 mM sucrose, 57 µg ml^{-1} BSA, 25 mM MOPS, 25 mM MES, pH 8.0) (Bageshwar and Musser, 2007) and IMVs (A_{280} = 2) were added to a heated cuvette (final volume: 800 µL) in an SLM-Aminco fluorometer and allowed to equilibrate to 37°C for 5 min. Precursor protein was then added and equilibration was continued for at least 200 s. Reactions were initiated by addition of competitor precursor or NADH.

Excitation and emission wavelengths were 500 nm (4 nm slits) and 550 nm (8 nm slits), respectively.

Analysis

Kinetic data were fit with a single exponential plus a linear baseline drift:

$$y = ax + b(1 - e^{-kx}) \qquad \text{(equation 1)}$$

where *a*, *b*, and *k* (= 1/τ) are fit parameters. The data in Figure 2.2D were fit to the equation for single site binding when the K_D and receptor concentration are of similar magnitude (derived in the Supplementary Material)(equation 2)

where *C*, T_0, and K_D are fit parameters.

$$y = C * \frac{(x + T_0 + K_D) - \sqrt{(x + T_0 + K_D)^2 + 4(x)T_0}}{2T_0}$$

(equation 2)

CHAPTER III

EFFECT OF CARGO SHAPE ON TRANSPORT EFFICIENCY OF THE BACTERIAL TAT TRANSLOCASE

SUMMARY

The Tat machinery's ability to translocate fully folded and, in some cases, oligomeric substrates have caused it to attract attention as a potential biotechnological tool. Previous studies have implied that a cargo's molecular weight and possibly its length can adversely affect its transport efficiency. It was hypothesized that transport efficiency may not only be dependent on size, but shape. To this end a series of cargos of different sizes and shapes were generated using the Tat substrate pre-SufI as a base. Of the eighteen cargos that were tested, four were found to transport with significant (>15% of wild–type) efficiencies. The lower efficiencies were not due to slower transport kinetics. Despite identical molecular weights, the same theoretical isoelectric point, and the same signal peptide, many of the other cargos failed to transport. These results indicate that the shape of the cargo is indeed important for its transportability.

INTRODUCTION

As described in previous chapters, the twin arginine translocase transports fully folded and assembled cargos from the *Escherichia coli*

cytoplasm to the periplasm (Clark and Theg, 1997). The system is named for two highly conserved arginines found within the consensus motif in the signal peptides of Tat directed precursor proteins (Chaddock et al., 1995). The *E. coli* machinery is composed of three minimal components: TatA, TatB and TatC (Bogsch et al., 1998; Sargent et al., 1999; Weiner et al., 1998). The working model of Tat dependent transport describes a scenario in which cargos first interact with a complex of TatB and TatC, and then are conveyed across the bilayer though a channel comprised primarily of TatA (De Leeuw et al., 2001; Gohlke et al., 2005). After the precursor has been transported, the signal peptide is removed by a periplasmically active signal peptidase resulting in the mature form of the protein (Luke et al., 2009).

The Tat system's ability to transport fully folded, and in some cases oligomeric, proteins in a vesicle-independent fashion is perhaps its most interesting feature. In fact, there are only two other known systems that have this capability: the nuclear pore complex and the peroxisomal import system, both of which are absent in bacteria (Strambio-De-Castillia et al., 2010; Walton et al., 1995). Due to its uniqueness in bacterial systems, the Tat machinery has attracted attention because of its potential for use in the production of protein products (Bruser, 2007; Yoon et al., 2010). The bacterial Sec system has long been used to transport proteins of interest into the periplasm or outside of the cell for increased ease and rapidity of purification (Mergulhao et al., 2005). The Sec system, however, cannot transport some proteins such as those that are

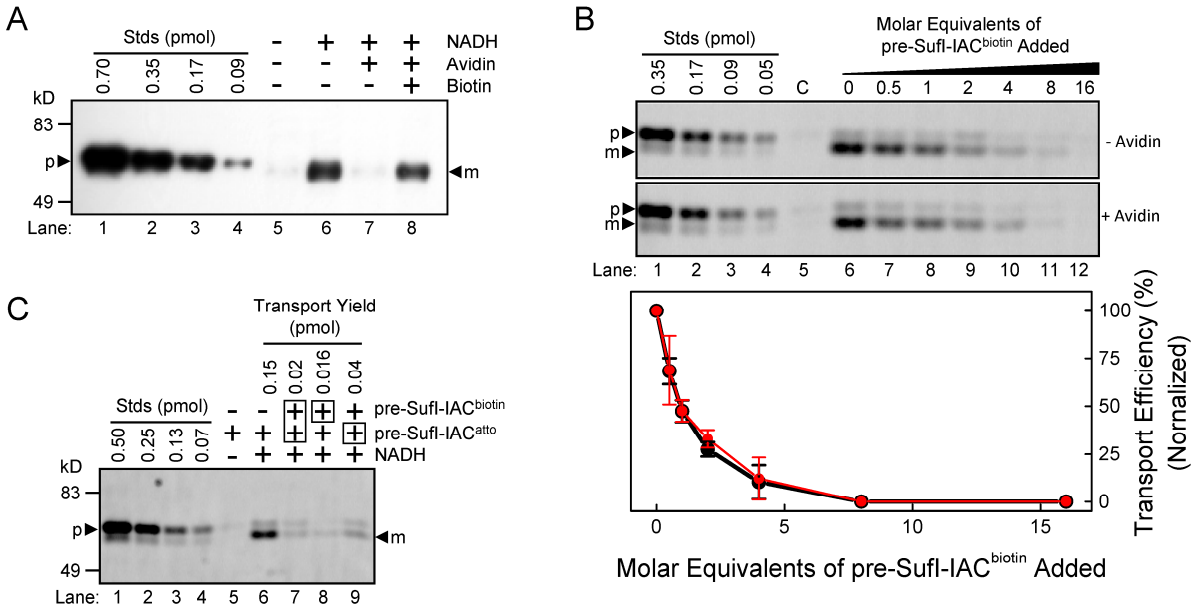


Figure 3.1. Pre-SufI-IAC fails to transport when bound to avidin. (A) Effect of avidin on the transport of pre-SufI-IAC^{biotin}. Proteins were detected by Western blotting using avidin-HRP. When present, biotin (16 μM) was in large excess over neutravidin (1.6 μM), and neutravidin was in large excess over the precursor protein (40 nM, 1.4 pmol). Thus, the binding interactions were saturated. [IMV] = 5 (A_{280}) (B) Competition between pre-SufI-IAC^{biotin} and pre-SufI-IAC^{atto} in the presence of avidin. The transport efficiency of pre-SufI-IAC^{atto} (40 nM, 1.4 pmol) was determined in the absence (black) and presence (red) of 25 μM neutravidin and the indicated molar equivalents of pre-SufI-IAC^{biotin} (n = 3). Gels were visualized by the Atto565 fluorescence emission. [NADH] = 4 mM; [IMV] = 5 (A_{280}). (C) Effect of precursor addition order on the transport efficiency of pre-SufI-IAC^{atto}. For lanes 5–9, the reactions contained pre-SufI-IAC^{atto}, 25 μM neutravidin, and IMVs containing over-expressed Tat (A_{280} = 5). For lanes 7–9, the reactions contained pre-SufI-IAC^{biotin}. The boxes indicate which precursor (90 nM, 3.1 pmol each) was added first to the reaction solution. In lane 7, both precursors were added simultaneously, and 4 mM NADH was added 5 min later. In lane 8, pre-SufI-IAC^{biotin} was added 5 min prior to the simultaneous addition of pre-SufI-IAC^{atto} and NADH. Lane 9 was obtained similarly to lane 8, but the precursors were reversed. This gel was visualized by the Atto565 fluorescence emission. Adapted from (!!! INVALID CITATION !!!).

oligomeric, fast folding, or those that contain cofactors. The Tat translocase offers an attractive alternative pathway that, in principle, could accommodate these problematic substrates. Additionally, the Tat translocase possesses a

poorly understood quality control mechanism that appears to be able to preclude incorrectly folded cargos, which would presumably allow for isolation of only correctly folded proteins (Rocco et al., 2012; Sargent, 2007a). In order to make use of the system in this way, it is important to understand the limitations of the machinery, especially in terms of its capability to translocate unnatural substrates.

The Tat system's ability to transport hetero-oligomeric proteins with only a single signal peptide has been termed the "hitch-hiker" mechanism. In *E. coli*, the small (HybO) and large (HybC) subunits of hydrogenase 2 (Hyd2) are a good example of such a situation, with the small subunit responsible for the co-transport of the large subunit (Rodrigue et al., 1999). Dimethyl sulfoxide (DMSO) reductase's catalytic dimer (DmsAB) also only has a single signal peptide, which is found on DmsA (Sambasivarao et al., 2000). Despite the presence of these naturally occurring hitch-hiker complexes, efforts to transport "unnatural" non-covalent hetero-dimers have been unsuccessful. The only two studies that address this directly have utilized biotinylated natural cargo bound to avidin tetramers. In both instances (OE17 and SufI avidin complexes), the precursors failed to transport under *in vitro* conditions despite their apparent ability to bind to the Tat translocon (Figure 3.1AB and (Musser and Theg, 2000)). These bound, non-transportable cargos were not, however, dead-end intermediates (i.e. they were not "stuck" in the translocon) because they could

be exchanged out with different cargos that were subsequently transported (Figure 3.1C).

Efforts have been made to probe the size limitations of the Tat machinery using fusion proteins consisting of a natural Tat substrate (or corresponding signal peptide) fused to a separate protein moiety. One study utilized the natural chloroplast Tat substrate OE17 fused to a Protein A moiety, separated by varying lengths of unstructured linkers (Cline and McCaffery, 2007). Like the avidin bound precursor proteins, these Protein A fused cargos failed to fully transport. Surprisingly, the fusions with the longest linkers formed membrane spanning intermediates that were not associated with the translocon. In light of these intermediates, it is possible that some attribute of Protein A itself prevents it from becoming transported, especially because it is not a natural substrate for the Tat machinery. Another unnatural Tat cargo has, on the other hand, been successfully transported. Green fluorescent protein (GFP) is translocated into IMVs containing the *E. coli* Tat system when it was fused to the signal peptide from the substrate TorA (Bageshwar and Musser, 2007). However, though pre-SufI is transported with high efficiency *in vitro*, when SufI's signal peptide is fused to GFP, it does not transport or even bind to the translocon (Bageshwar and Musser, 2007).

Based on these studies, experiments were performed that sought to probe the translocon by modifying the natural cargo pre-SufI. GFP fusions and noncovalent pre-SufI-avidin hetero-dimers formed from biotinylated versions of

the single cysteine mutants described in Chapter II were used. Additionally, the naturally occurring, high affinity, dimeric avidin homolog, rhizavidin, was used in place of the tetrameric avidin (Helppolainen et al., 2007). These results indicate that transport efficiency of a given cargo is not only dependent on its molecular weight, but also on its shape.

RESULTS

Generation of Pre-SufI Cargos

The natural Tat substrate, pre-SufI, was selected as a base for the series of cargos used in this study based on its relatively high transport affinity demonstrated in previous studies (Bageshwar and Musser, 2007; Bageshwar et al., 2009). For a preliminary set of experiments, GFP and a tandem GFP dimer were fused to pre-SufI at the C-terminus, generating new substrates with molecular weights of 86 and 110 kDa, respectively. As a Tat substrate, these constructs are similar to several that have been previously studied, i.e, a natural Tat cargos with a C-terminal addition (Cline and McCaffery, 2007; Lindenstrauss and Bruser, 2009; Musser and Theg, 2000) (Figure 3.1). The goal of this study, however, is to probe the Tat machinery with a range of cargos that differ not only in size, but in shape. To this end, the series of single cysteine mutations from Chapter II was utilized. Each mutant was labeled with N-(3-maleimidyl propionyl)biocytin, yielding a series of pre-SufI proteins with a single biotin at 8 different locations on the surface of the molecule (see Figure 2.1A). When

mixed with saturating levels of avidin, the biotinylated pre-SufI form cargos with a MW of 113 kDa, but with 8 distinct morphologies. The high affinity for biotin and slow off-time of the avidin tetramer (K_D ~ 10^{-15} M) ensures that the heterodimeric structure will be maintained during *in vitro* Tat dependent transport reactions, which take place on the order of a few minutes (see Figure 2.1D).

Recently, a naturally occurring, dimeric homolog of avidin was characterized, termed rhizavidin. Isolated from the nitrogen-fixing bacterium *Rhizobium etli,* rhizavidin does not have the 1-2 and 1-3 interfaces that tetrameric interfaces that avidin has, yet it still maintains a high affinity for biotin (Helppolainen et al., 2007; Meir et al., 2009). When the biotinylated pre-SufI variants are mixed with rhizavidin at saturating concentrations, a series of heterodimeric precursor proteins with a MW of 83 kDa are formed.

Transport Efficiency of Modified Pre-SufI Proteins

Transport efficiencies of all substrates were measured using the protease protection based *in vitro* transport assay described previously (Bageshwar and Musser, 2007). In short, cargo protein was mixed with IMVs containing over-expressed TatABC. Membranes were energized by addition of NADH which promotes translocation of the cargo into the lumen of the vesicles. Reactions were then treated with Proteinase K, which digests any untranslocated

substrate. The resultant mixtures were resolved via SDS-PAGE and visualized by Western blotting.

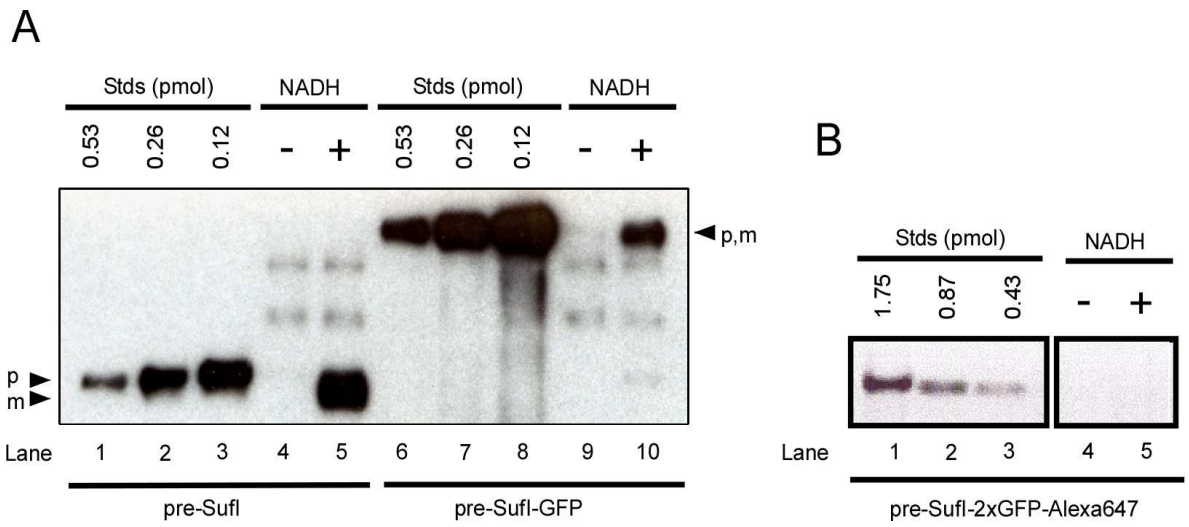
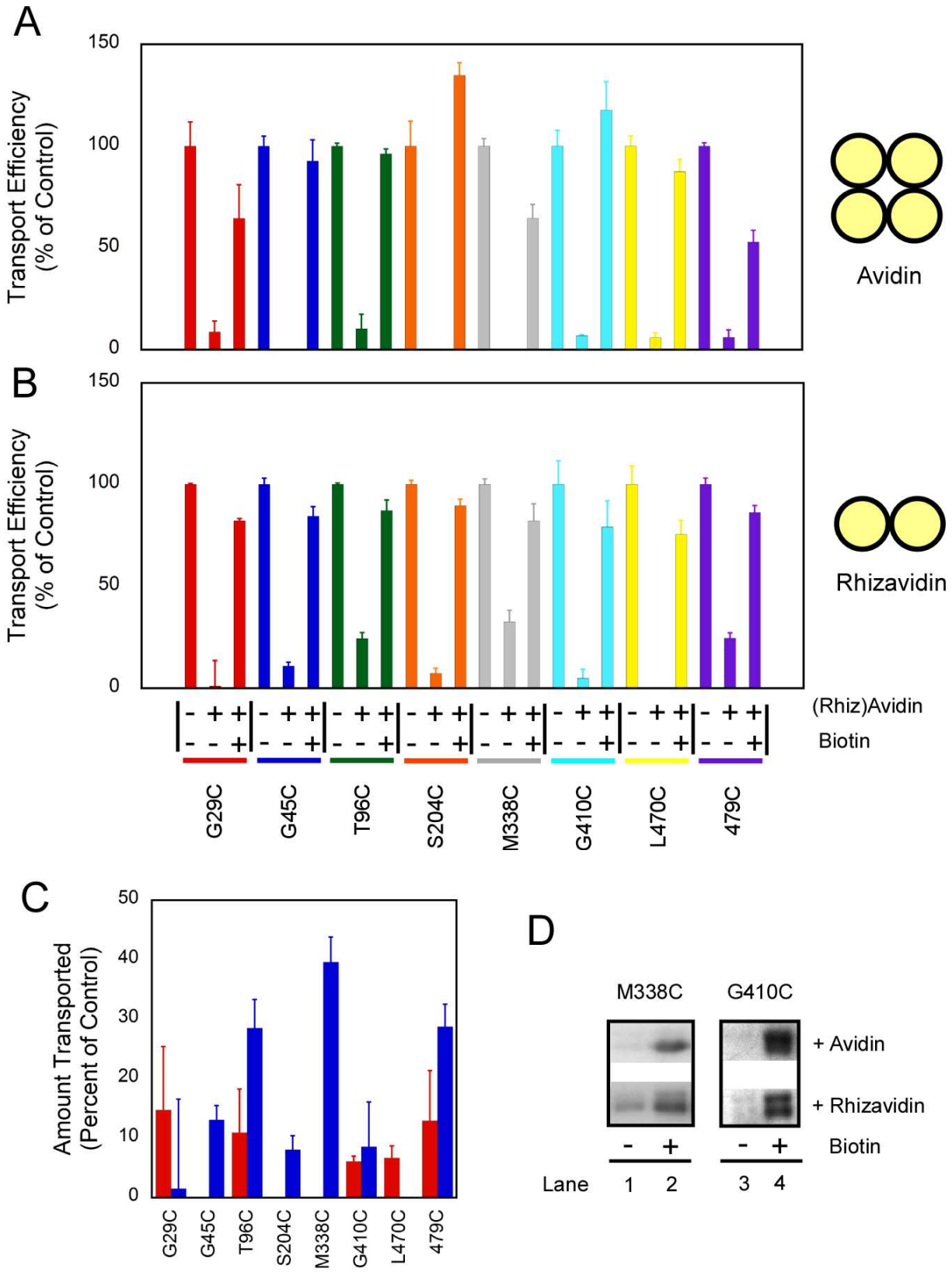


Figure 3.2. Pre-SufI-GFP transports with limited efficiency. (A) *In vitro* transport assays were performed with pre-SufI (lanes 1-5) or pre-SufI-GFP (lanes 6-10). Limited transport (17% of pre-SufI) was observed (compare lane 10 to lane 5). Lanes 1-3 and 6-8 are concentration standards. Reactions were performed with 50 nM of cargo. Lanes 4 and 9 are negative controls. Bands were visualized by anti-SufI immunoblotting. (B) Experiment similar to (A) but with pre-SufI-2xGFP labeled with Alexa647 at the C-terminus as the substrate. No transport was detected. Visualization of protein bands was performed by direct in-gel fluorescence imaging (model FX PhosphorImager, Bio-Rad Laboratories).

Figure 3.3. Pre-SufI^{biotin} transports in the presence of rhizavidin with limited and shape dependent efficiency. (A) Pre-SufI^{biotin} transport was strongly inhibited when bound to avidin tetramers. Proteins were detected by Western blotting using avidin-HRP. When present, biotin (36 μM) was in large excess over neutravidin (3.6 μM), and neutravidin was in large excess over the precursor protein (90 nM, 3.1 pmol). Thus, the binding interactions were saturated. The graph represents an average of three reactions, normalized to transport in the absence of avidin (wildtype). (B) In the presence of rhizavidin pre-SufI(T96C)^{biotin}, pre-SufI(M338C)^{biotin}, and pre-SufI(479C)^{biotin} were found to transport with the greatest efficiency efficiency. All other pre-SufI^{biotin} mutants failed to transport more than 12% of wildtype when bound to rhizavidin. Experiments performed identically to those in (A) with rhizavidin substituted for avidin. (C) The percent of transported pre-Suf^{Ibiotin} in the presence of avidin (red) and rhizavidin (blue) was divided by the amount transported in the presence of both (rhiz)avidin and excess biotin. (D) Example reactions of a pre-SufI^{biotin} mutant that was found to transport in the presence of rhizavidin, but not avidin (M338C) and a mutant that did not transport under either condition (G410C). Lanes 2 and 4 are in the presence of 36 μM biotin. All reactions were energized with NADH.


A
Transport Efficiency (% of Control)
150
100
50
0
Avidin
B
Transport Efficiency (% of Control)
150
100
50
0
Rhizavidin
(Rhiz)Avidin
Biotin
G29C
G45C
T96C
S204C
M338C
G410C
L470C
479C
C
Amount Transported (Percent of Control)
50
40
30
20
10
0
G29C
G45C
T96C
S204C
M338C
G410C
L470C
479C
D
M338C
G410C
+ Avidin
+ Rhizavidin
Biotin
Lane
1
2
3
4

Unlike previous studies with GFP fused to the signal peptide of SufI (spSufI-GFP) (Bageshwar and Musser, 2007), full length SufI fused to GFP was transportable, albeit with a lower (17% of wild-type SufI) efficiency (Figure 3.2A), indicating that the mature domain of pre-SufI plays an important role in the ability of the translocon to transport the cargo. A similar construct with two copies of the GFP moiety fused to the C-terminus of SufI, however, failed to transport to any detectable degree (Figure 3.2B).

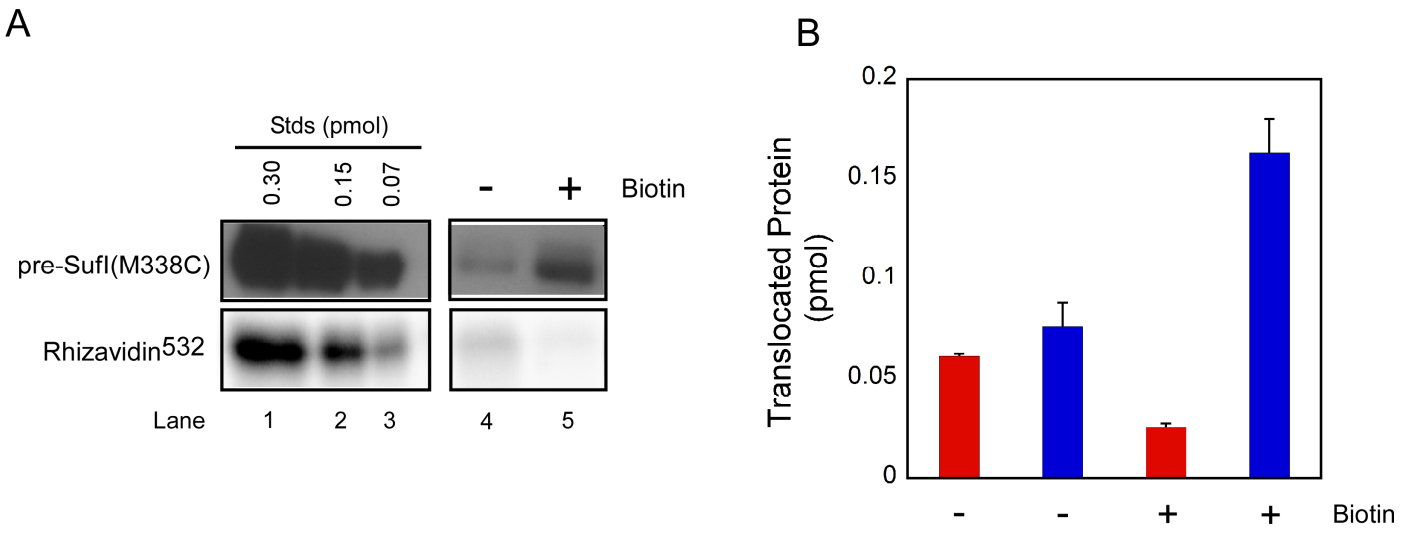


Figure 3.4. Rhizavidin is translocated into the lumen of IMVs. (A) Reactions were performed essentially as in Figure 3.2 but without proteinase K digestion for the rhizavidin^{532} gel. Instead, the reactions were pelleted and washed with translocation buffer containing 1 M KCl and 2 M urea to remove untransported rhizavidin. Lanes 1-3 are standards. Lanes 4 and 5 contain rhizavidin, with saturating levels of biotin present in lane 5. Pre-SufI(M338C)^{biotin} was detected by Western blotting using avidin-HRP. Rhizavidin^{532} was detected by in-gel fluorescence imaging (model FX PhosphorImager, Bio-Rad Laboratories). (B) Bar graph of data from (A). Total rhizavidin transported is shown in red, total pre-SufI transport is shown in blue (N = 3).

Comparative transport efficiencies were also determined for the avidin and rhizavidin pre-SufI complexes (Figure 3.3). Similar to the pre-SufI-IAC^{biotin} experiments shown in Figure 3.1, pre-SufI^{biotin} failed to transport in significant amounts when bound to the avidin tetramer (Figure 3.3A). When bound to the dimeric rhizavidin, however, limited transport was observed for pre-SufI(T96C)^{biotin} , pre-SufI(M338C)^{biotin}, and pre-SufI(T479C), which yielded transport efficiencies of 28%, 39% and 28%, respectively (compared to when biotin was present) for the pre-SufI(M338C)^{biotin} mutant (Figures 3.3B and 3.3C). The presence of saturating amounts of biotin was found to largely restore transport, indicating that the inhibition of transport was due primarily to the specific avidin-biotin interaction, and not due to any unanticipated effect of the avidin on Tat protein transport.

For the pre-SufI mutants that transport in the presence of rhizavidin, it is necessary to determine if the “hitchhiking” rhizavidin dimer actually transports into the vesicle lumen. There is the possibility, however unlikely, that either the pre-SufI protein crosses the membrane while leaving the attached rhizavidin outside (like in the case of the membrane spanning intermediate observed by (Cline and McCaffery, 2007) or that the rhizavidin becomes disassociated during the transport process, despite its high affinity. To this end, rhizavidin was labeled with the fluorescent dye AlexaFlour532 and translocated into vesicles under the same conditions as in Figure 3.3. Untransported rhizavidin could not be removed by proteinase K digestion, due to the proteins resistance to

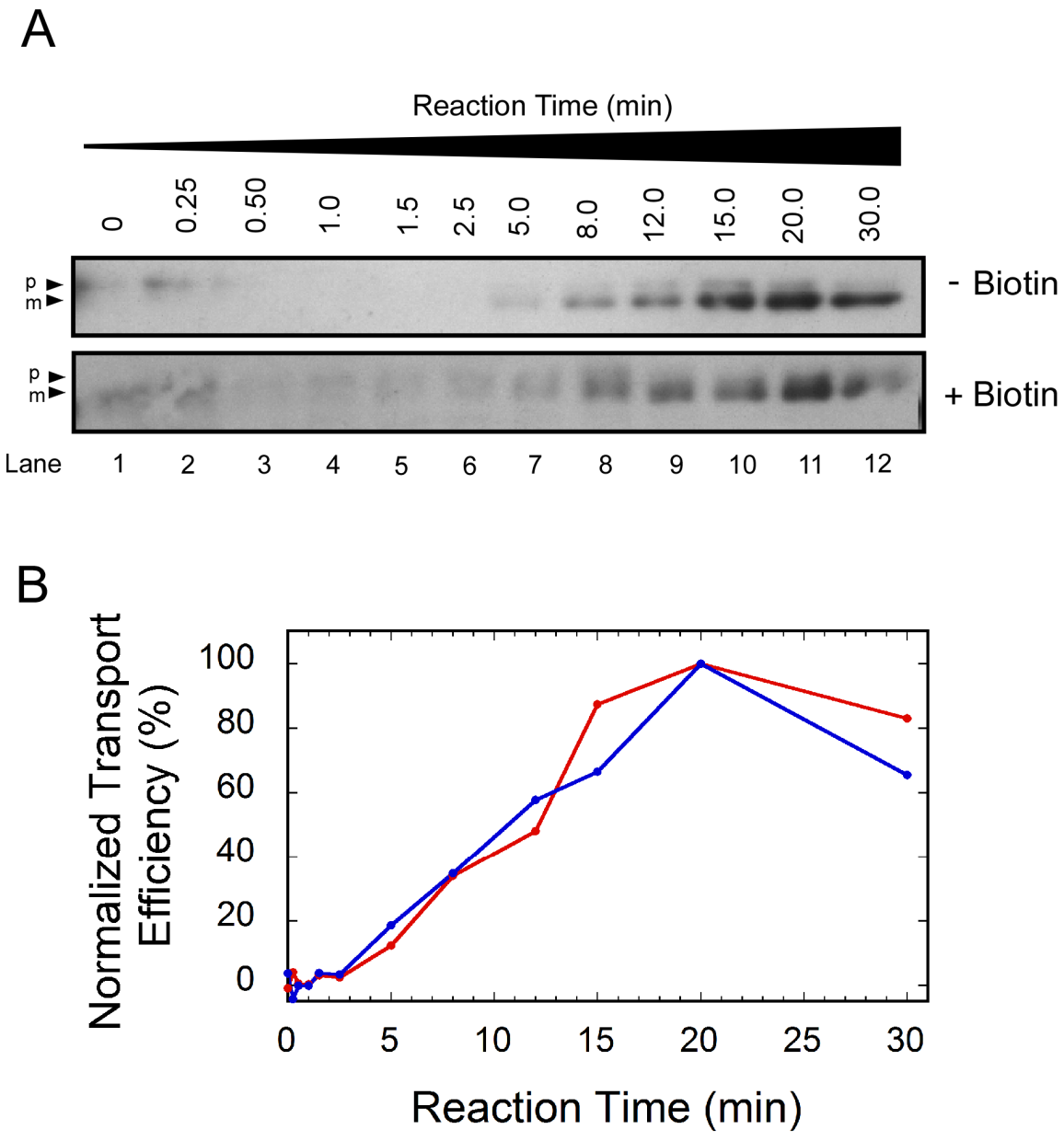


Figure 3.5. Pre-SufI(M338C)^{biotin} transports with similar kinetics in both rhizavidin bound and unbound states. (A) Time course of pre-SufI(M338C)^{biotin} translocation in the presence of rhizavidin and either in the presence (bottom gel) or absence (top gel) of a saturating amount of biotin. The reaction was performed in a heated (37ºC) cuvette with stirring, as in the kinetics experiments in CHAPTER II. Aliquots were removed at the time points indicated and added to chilled tubes containing nigericin and valinomycin to quench to reaction. Gels were was detected by Western blotting using avidin-HRP. (B) Quantification of the data from (A). Transport efficiencies were normalized based on the 20 min time point. Note that the blots were exposed for different amounts of time, and intensities between the two sets of bands are not directly comparable.

proteolytic digestion (Helppolainen et al., 2007), so it was removed by pelleting and washing with translocation buffer containing 2 M urea and 1 M KCl to remove precursor bound rhizavidin from the vesicles. The total amount of translocated rhizavidin was found to be approximately the same as the total amount of transported pre-SufI (Figure 3.4).

The lower apparent transport efficiencies for the avidin and rhizavidin bound pre-SufI proteins could be due to the larger, heterodimeric proteins transporting at a slower rate than the unbound pre-SufI. To test this, two time course reactions were performed: one with pre-SufI(M338C)^{biotin} bound to rhizavidin and one without (Figure 3.5A). Although the overall kinetics appeared slower than those of similar experiments reported in the previous chapter, they were similar to each other, indicating that the lower transport efficiencies observed for the avidin and rhizavidin bound forms of pre-SufI are likely not a consequence of slower transport kinetics (Figure 3.5B).

DISCUSSION

All studies thus far on the Tat system that have sought to determine limitations on cargo translocation have focused on either the molecular weight of the substrate, or the total length (i.e., the number of amino acids) of a linker region between two folded moieties. These are related but not identical parameters since a folded domain yields a shorter length than a flexible linker. Defining Tat substrates via these parameters are likely inadequate given that the

Tat machinery transports fully folded and assembled cargos, and folded proteins of similar molecular weights may have considerably different dimensions.

Additionally, several other properties of proteins that are directed to the Tat machinery may affect their ability to be translocated. These could include the overall charge of the protein, the surface charge, the properties of its specific signal peptide, or the ability of the translocon to bind to the signal peptide. This study has for the first time utilized a system that both probes the translocon's ability to handle unnatural substrates in terms of both size and shape, while attempting to change the properties listed above as little as possible.

In all cases, the presence of a bound avidin tetramer substantially inhibited pre-SufI transport (Figure 3.3A). In the cases were a small amount was translocated (such as pre-SufI-T96C), more or at least equal amounts (within error) of protein was transported when the pre-SufI mutant was bound to rhizavidin instead. Pre-SufI-2xGFP, which is similar in molecular weight to pre-SufI bound to an avidin tetramer, also failed to transport to a detectable degree (Figure 3.2B). While this does not prove a molecular weight cutoff for the Tat translocon, it certainly suggests that larger proteins, despite significant similarity, are more difficult to transport than smaller ones. It is important to consider that the duration, and possibly the magnitude, of the membrane potential are not as high or as long as they would be under *in vivo* conditions. Presumably, in live *E. coli* cells, the membrane potential is essentially constant, and this sustained potential may allow larger proteins to be translocated. Furthermore, using

purified IMVs may exclude some essential soluble factors that would allow these larger proteins to translocate. However, this is unlikely in this case since it has been determined that no soluble factor is required for translocation of pre-SufI (Holzapfel et al., 2009). Importantly, evidence exists that suggest that proteins of similar, and even larger, molecular weights are transported under natural conditions (Palmer et al., 2005). The catalytic subunit of the *E. coli* Tat substrate TMAO reductase, for instance, has a molecular weight of approximately 94 kDa. This discrepancy could be due to a specific function of the signal peptide attached to the cargo (i.e., a TorA signal peptide in place of the SufI signal peptide may allow for transport of some of the untransportable cargos used in this study), or it could simply be due to some unknown limitation of using this *in vitro* system.

Interestingly, some pre-SufI mutants were transported when bound to rhizavidin and some were not (Figure 3.3B, C and D). The hetero-dimeric cargos have different shapes and the only other difference in these substrates is the single point mutations used for labeling. However, these mutations were shown to not significantly affect the transportability of the proteins in the majority of cases (Figure 2.1B), and the adverse effects on transport efficiency were not nearly as significant as those observed in Figure 3.3.

Based on the transport efficiency of the various biotinylated precursors in the presence of (rhiz)avidin, it is expected that something could be learned about the shape requirements for Tat cargos. However, as can be seen in

Figure 2.1A, the two proteins that transported with the highest efficiency when bound to rhizavidin (pre-SufI(M338C) and pre-SufI(T96C)) have mutations on opposite sides of the protein's mature domain. Furthermore, pre-SufI(479C)^{biotin} and the similar pre-SufI-GFP precursor were found to transport significantly better in the presence of rhizavidin than pre-SufI(L470C)^{biotin} despite the fact that the two cysteine mutations are very close to each other. The reason for this difference in transport efficiencies is unknown. One possibility is that the flexible linker region in pre-SufI(479C)^{biotin} may allow the cargo to assume a more transportable shape, but this has not been shown explicitly.

It is possible that the hetero-dimeric proteins simply transport slower than unbound pre-SufI. Rhizavidin bound and unbound pre-SufI(M338C)^{biotin} was found to transport with similar kinetics (Figure 3.5), so it is unlikely that this is the case. Either only a certain fraction of Rhizavidin bound pre-SufI(M338C)^{biotin} is transport capable (implying that the cargo population is somehow not uniform), or that every transport event has a certain chance of aborting, and the cargo's dimensions (the size and shape) affect this chance.

The Tat machinery's ability to translocate folded and oligomeric proteins will likely continue to attract attention for the production and purification of proteins that cannot be secreted by the more commonly used Sec system. This study implies that not every tightly folded protein can be secreted by the Tat machinery when coupled to readily recognizable transportable carrier cargo.

Thus, we still have much to learn about the recognition, size, and shape requirements of Tat cargos.

METHODS

Bacterial Strains, Plasmids, and Growth Conditions

E. coli strain MC4100ΔTatABCDE was used for growth and isolation of IMVs as described in Chapter II. *E. coli* TatA, TatB and TatC was induced by addition of 0.6% arabinose to cultures of MC4100ΔTatABCDE containing the plasmid pTatABC, as previously described (Bageshwar and Musser, 2007) (Chapter II). Expression conditions for production of the pre-SufI cysteine mutants were also as described in Chapter II.

The plasmid pPre-SufI-GFP was generated by amplifying the GFP gene from spTorA-GFP (Bageshwar and Musser, 2007) with PCR using primers 5'-AAAACTCGAGGATCGGCGCAAGCGGCGCGTAAAGGA-3' (forward) and 5'-AAAAGCTAGCTTTGTATAGTTCATCCATGACCATGCCATGTC-3' (reverse), which introduced two flanking restriction sites (NheI and XhoI) to the GFP gene. These sites were used to insert GFP into pSufI-MCC (Bageshwar et al., 2009) between the pre-SufI gene and the sequence for the 6xHis tag. The resultant ligated plasmid contained a frame shift error that was corrected by XhoI digestion and a subsequent Klenow fragment fill-in reaction, which also destroyed the XhoI site after the plasmid was re-ligated (Figure A16). The plasmid pPre-SufI-GFPx2 was similarly produced by amplifying the GFP gene

from spTorA-GFP with PCR using 5'-GGGGACTAGTTTTGTATAGTTCATCCATGCCA-3' (forward) and 5'-GGGGACTAGTAAGCGTAAAGGAGAAGAACTTTA-3' (reverse) which introduced flanking SpeI and NheI restriction enzyme sites. The amplified second GFP gene was then inserted into a NheI site, destroying the site upon religation (Figure A17). Because the pre-SufI-GFP fusion protein is too large to be readily distuinguished from the mature, SufI-GFP protein, pre-SufI-GFP was inserted into the vector pTYB11 (New England Biochem) to ensure the full length precursor would be purified, generating pIntein-pre-SufI-GFP (Figure A18). The cargo coding region of plasmid pPre-SufI-2xGFP was also inserted into the pTYB11 vector, but expression of the resultant fusion protein produced very little soluble precursor protein. It was therefore not used in this study. Instead, pPre-SufI-2xGFP was expressed from plasmid pPre-SufI-2xGFP using the pre-SufI expression protocols described previously (Bageshwar and Musser, 2007). Growth and expression conditions for pre-SufI-GFP from pIntein-pre-SufI-GFP are identical to conditions used for expression from pIntein-preSufI-MA as described in Chapter II.

Plasmids used for generating the proteins necessary for the pre-SufI^{biotin} experiments were produced as described in Chapter II. The plasmid used for rhizavidin expression was a gift from Barbara Niederhauser. (Helppolainen et al., 2007).

Protein Purification and Labeling

Labeling and purification of the pre-SufI cysteine mutants was performed as described in Chapter II, except that the labeling reactions were performed with N-(3-maleimidyl propionyl)biocytin (Invitrogen). Purification of pre-SufI-GFP was performed essentially by the same method, except without the labeling step. Pre-SufI-2xGFP and rhizavidin were purified as described (Bageshwar and Musser, 2007; Helppolainen et al., 2007). Rhizavidin was labeled with AlexaFluor532 carboxylic acid, succinimidyl ester. The reaction was performed (1 hour at room temperature) sodium bicarbonate buffer (pH 8.25). The reaction was quenched by addition of 100 mM Tris-HCl (pH 8.0). Excess dye was removed by spin filtration with addition of fresh bicarbonate buffer after each spin (8 consecutive 5 minute spins at 10,000 rcf) using Pall Nanosep Omega spin-filters (MW cutoff = 3,000 KDa). The term “avidin” refers to the glycosylation free variant NeutrAvidin, purchased from Invitrogen.

Transport Reactions

Protein translocation assays were performed essentially as described (Bageshwar and Musser, 2007) (Chapter II). Unless otherwise noted, initial precursor concentrations were 90 nM. All errors bars represent the standard error of the mean. The amount of IMVs in the reactions was estimated by the absorbance at 280 nm in 2% SDS. All IMV optical densities (OD) in the translocation reactions in this study were A_{280} = 5.

CHAPTER IV

CONCLUSIONS

In the first chapter, over two decades of research on the twin arginine translocase were summarized. Despite the numerous studies which have sought to elucidate aspects of the Tat machinery, much of the translocation process is still unknown. For instance, we know very little about the substeps of the translocation process, and the limitations of the Tat system in terms of what types of cargos it can successfully transport are poorly understood. This dissertation has reported on two studies which have attempted to address these discrepancies, and the following are the major conclusions.

Chapter II reported on the development of a real-time FRET based assay that allowed for the observation of the dynamic interactions of cargo with the Tat translocon. To date, no other study has been able to observe any step of Tat dependent translocation with higher time resolution. The data from this chapter support the following conclusions. First, the cargo's binding affinity to the translocon under the currently described conditions is reversible and has a relatively high affinity. The estimated dissociation constant ($K_D \approx$ 7-23 nM) is the first such measurement made for a functional Tat system. Second, cargo was found to first bind to the TatBC complex, confirming one of the early steps in the popular model for Tat transport. Then, in the presence of a membrane potential,

the mature domain migrated away from its initial binding site. The substrate's migration and subsequent translocation occur on the minute timescale in these experiments. Third, cargo (or at least pre-SufI) appears to bind closer to the C-terminus of TatC than the C-terminus of TatB. Fourth, there is a delay in cargo migration from its binding site after energization, which lasts for approximately 20-45 seconds. The fact that it occurs after the translocon is fully energized indicates that the membrane potential does not directly promote cargo migration, but rather facilitates some process that prepares the translocon to allow for the migration to take place. Fifth, in the absence of TatA, energized TatBC membranes lose affinity for bound cargo when energized.

In Chapter III, a series of different sized and shaped cargos was used to probe the Tat translocon's ability to accommodate a range of unnatural cargos. The study used the natural Tat substrate pre-SufI as a scaffold for constructing eighteen different cargos. The data support the following conclusions. First, of the two size classes of substrates used (~85 kDa and ~110 kDa) the four that were found to transport were ~85 kDa. Second, when the same protein moiety (rhizavidin) was bound to different locations on pre-SufI, different transport efficiencies were observed. Third, the transport efficiencies of the newly tested cargo were lower than that of wild-type pre-SufI, but were found to transport with approximately the same kinetics.

In conclusion, these studies have provided two distinct methods for probing the mechanism of Tat translocation. The method utilized in CHAPTER

II was able to observe a sub-step in translocation. This substep appears to take up a significant amount (~40%) of the total transport time. The fact that this apparent migration step is not directly caused by the membrane potential, but still requires it to occur, implies that there is an unknown $\Delta\psi$–dependent process that “primes” the translocon for the next stage in the transport cycle. Whether this is an oligomerization step (i.e., the formation of a TatA pore), a conformational change, or something else is unclear. Additionally, the finding that TatA plays a role in the early steps of transport is an important distinction from previous models. The data reported in CHAPTER III address the long standing goal of adapting the Tat machinery for transport of tightly folded protein products that cannot be transported with the more commonly used Sec system. The results indicate that the limitations of the translocon are not only limited to the overall size (molecular weight) of the potential substrates, but also their dimensions.

The dynamic nature of the Tat machinery has made observing substeps with commonly used biochemical methods that provide only static snapshots of transport (such as chemical crosslinking) or data that pertain only to non-functional translocons (such as EM structures of detergent solubilized Tat proteins) have yet to shed sufficient light on transport as a dynamic process. Hopefully, by optimization and the labeling of other proteins (different cargos, TatA, or TatE for example) a clearer picture of the process may be uncovered. Additionally, the reasons for the apparent shape dependence on cargo efficiency

remain unclear. For instance, in the presence of a longer lasting $\Delta\psi$, previously untransportable cargo may be successfully translocated. An alternative explanation for this would be that there is a certain chance for every translocation event to abort, and this is dependent on the properties of the substrate. If this is the case, at what step in the translocation process does this occur? A combination of the two techniques described in this dissertation, such as utilizing fluorescent rhizavidin in FRET based kinetics experiments, may serve to answer this question.

REFERENCES

Alami, M., Luke, I., Deitermann, S., Eisner, G., Koch, H.G., 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.

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

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

Bageshwar, U.K., Whitaker, N., Liang, F.C., and Musser, S.M. (2009). Interconvertibility of lipid- and translocon-bound forms of the bacterial Tat precursor pre-SufI. Mol Microbiol *74*, 209-226.

Baglieri, J., Beck, D., Vasisht, N., Smith, C.J., and Robinson, C. (2012). Structure of TatA paralog, TatE, suggests a structurally homogeneous form of Tat protein translocase that transports folded proteins of differing diameter. J Biol Chem *287*, 7335-7344.

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

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

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

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.

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

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

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

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

Blaudeck, N., Kreutzenbeck, P., Freudl, R., and Sprenger, G.A. (2003). Genetic analysis of pathway specificity during posttranslational protein translocation across the *Escherichia coli* plasma membrane. J Bacteriol *185*, 2811-2819.

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.G., Sargent, F., Stanley, N.R., Berks, B.C., 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. (2002). Protein transport in the halophilic archaeon *Halobacterium sp*. NRC-1: a major role for the twin-arginine translocation pathway? Microbiology *148*, 3335-3346.

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

Braun, N.A., Davis, A.W., and Theg, S.M. (2007). The chloroplast Tat pathway utilizes the transmembrane electric potential as an energy source. Biophys J *93*, 1993-1998.

Braun, N.A., and Theg, S.M. (2008). The chloroplast Tat pathway transports substrates in the dark. J Biol Chem *283*, 8822-8828.

Brehmer, T., Kerth, A., Graubner, W., Malesevic, M., Hou, B., Bruser, T., and Blume, A. (2012). Negatively charged phospholipids trigger the interaction of a bacterial Tat substrate precursor protein with lipid monolayers. Langmuir *28*, 3534-3541.

Bronstein, P.A., Marrichi, M., Cartinhour, S., Schneider, D.J., and DeLisa, M.P. (2005). Identification of a twin-arginine translocation system in Pseudomonas syringae pv. tomato DC3000 and its contribution to pathogenicity and fitness. J Bacteriol *187*, 8450-8461.

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

Caldelari, I., Mann, S., Crooks, C., and Palmer, T. (2006). The Tat pathway of the plant pathogen *Pseudomonas syringae* is required for optimal virulence. Mol Plant Microbe Interact *19*, 200-212.

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

Celedon, J.M., and Cline, K. (2012). Stoichiometry for binding and transport by the twin arginine translocation system. J Cell Biol *197*, 523-534.

Chaddock, A.M., Mant, A., Karnauchov, I., Brink, S., Herrmann, R.G., Klosgen, R.B., 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.S., Haney, E.F., Vogel, H.J., and Turner, R.J. (2011). Towards understanding the Tat translocation mechanism through structural and biophysical studies of the amphipathic region of TatA from *Escherichia coli*. Biochim Biophys Acta *1808*, 2289-2296.

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

Chong, S., Montello, G.E., Zhang, A., Cantor, E.J., Liao, W., Xu, M.Q., and Benner, J. (1998). Utilizing the C-terminal cleavage activity of a protein splicing element to purify recombinant proteins in a single chromatographic step. Nucleic Acids Res *26*, 5109-5115.

Clark, S.A., and Theg, S.M. (1997). A folded protein can be transported across the chloroplast envelope and thylakoid membranes. Mol Biol Cell *8*, 923-934.

Cline, K., Ettinger, W.F., and Theg, S.M. (1992). Protein-specific energy requirements for protein transport across or into thylakoid membranes. Two

lumenal proteins are transported in the absence of ATP. J Biol Chem *267*, 2688-2696.

Cline, K., and McCaffery, M. (2007). Evidence for a dynamic and transient pathway through the TAT protein transport machinery. EMBO J *26*, 3039-3049.

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

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.

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

De Buck, E., Hoper, D., Lammertyn, E., Hecker, M., and Anne, J. (2008a). Differential 2-D protein gel electrophoresis analysis of *Legionella pneumophila* wild type and Tat secretion mutants. Int J Med Microbiol *298*, 449-461.

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

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

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

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

de Leeuw, E., Granjon, T., Porcelli, I., Alami, M., Carr, S.B., Muller, M., Sargent, F., Palmer, T., and Berks, B.C. (2002). Oligomeric properties and signal peptide binding by *Escherichia col*i Tat protein transport complexes. J Mol Biol *322*, 1135-1146.

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

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

Di Cola, A., Bailey, S., and Robinson, C. (2005). The thylakoid delta pH/delta psi are not required for the initial stages of Tat-dependent protein transport in tobacco protoplasts. J Biol Chem *280*, 41165-41170.

Di Cola, A., and Robinson, C. (2005). Large-scale translocation reversal within the thylakoid Tat system in vivo. J Cell Biol *171*, 281-289.

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

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

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

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

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

Fisher, A.C., and DeLisa, M.P. (2009). Efficient isolation of soluble intracellular single-chain antibodies using the twin-arginine translocation machinery. J Mol Biol *385*, 299-311.

Fisher, A.C., Haitjema, C.H., Guarino, C., Celik, E., Endicott, C.E., Reading, C.A., Merritt, J.H., Ptak, A.C., Zhang, S., and DeLisa, M.P. (2011a). Production of secretory and extracellular N-linked glycoproteins in *Escherichia coli*. Appl Environ Microbiol *77*, 871-881.

Fisher, A.C., Rocco, M.A., and DeLisa, M.P. (2011b). Genetic selection of solubility-enhanced proteins using the twin-arginine translocation system. Methods Mol Biol *705*, 53-67.

Frielingsdorf, S., Jakob, M., and Klosgen, R.B. (2008). A stromal pool of TatA promotes Tat-dependent protein transport across the thylakoid membrane. J Biol Chem *283*, 33838-33845.

Frielingsdorf, S., and Klosgen, R.B. (2007). Prerequisites for terminal processing of thylakoidal Tat substrates. J Biol Chem *282*, 24455-24462.

Frobel, J., Rose, P., and Muller, M. (2011). Early contacts between substrate proteins and TatA translocase component in twin-arginine translocation. J Biol Chem *286*, 43679-43689.

Frobel, J., Rose, P., and Muller, M. (2012). Twin-arginine-dependent translocation of folded proteins. Philos Trans R Soc Lond B Biol Sci *367*, 1029-1046.

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

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

Gerard, F., and Cline, K. (2007). The thylakoid proton gradient promotes an advanced stage of signal peptide binding deep within the Tat pathway receptor complex. J Biol Chem *282*, 5263-5272.

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

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

Gralnick, J.A., Vali, H., Lies, D.P., and Newman, D.K. (2006). Extracellular respiration of dimethyl sulfoxide by *Shewanella oneidensis* strain MR-1. Proc Natl Acad Sci U S A *103*, 4669-4674.

Greene, N.P., Porcelli, I., Buchanan, G., Hicks, M.G., Schermann, S.M., Palmer, T., and Berks, B.C. (2007). Cysteine scanning mutagenesis and disulfide

mapping studies of the TatA component of the bacterial twin arginine translocase. J Biol Chem *282*, 23937-23945.

Halbig, D., Wiegert, T., Blaudeck, N., Freudl, R., and Sprenger, G.A. (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. Eur J Biochem *263*, 543-551.

Helppolainen, S.H., Nurminen, K.P., Maatta, J.A., Halling, K.K., Slotte, J.P., Huhtala, T., Liimatainen, T., Yla-Herttuala, S., Airenne, K.J., Narvanen, A.*, et al.* (2007). Rhizavidin from *Rhizobium etli*: the first natural dimer in the avidin protein family. Biochem J *405*, 397-405.

Hicks, M.G., de Leeuw, E., Porcelli, I., Buchanan, G., Berks, B.C., 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.

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

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 targeting motif. FEBS Lett *497*, 45-49.

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

Holzapfel, E., Moser, M., Schiltz, E., Ueda, T., Betton, J.M., and Muller, M. (2009). Twin-arginine-dependent translocation of SufI in the absence of cytosolic helper proteins. Biochemistry *48*, 5096-5105.

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

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 Am Chem Soc *132*, 15942-15944.

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

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

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

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

Ize, B., Porcelli, I., Lucchini, S., Hinton, J.C., Berks, B.C., and Palmer, T. (2004). Novel phenotypes of *Escherichia coli* tat mutants revealed by global gene expression and phenotypic analysis. J Biol Chem *279*, 47543-47554.

Jack, R.L., 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.L., Sargent, F., Berks, B.C., 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.

Ki, J.J., Kawarasaki, Y., Gam, J., Harvey, B.R., Iverson, B.L., 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.

Klosgen, R.B., Brock, I.W., Herrmann, R.G., 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.

Koch, S., Fritsch, M.J., Buchanan, G., and Palmer, T. (2012). *Escherichia coli* TatA and TatB proteins have N-out, C-in topology in intact cells. J Biol Chem *287*, 14420-14431.

Kreutzenbeck, P., Kroger, C., Lausberg, F., Blaudeck, N., Sprenger, G.A., and Freudl, R. (2007). *Escherichia coli* twin arginine (Tat) mutant translocases possessing relaxed signal peptide recognition specificities. J Biol Chem *282*, 7903-7911.

Lavander, M., Ericsson, S.K., Broms, J.E., and Forsberg, A. (2006). The twin arginine translocation system is essential for virulence of *Yersinia pseudotuberculosis*. Infect Immun *74*, 1768-1776.

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

Lee, P.A., Buchanan, G., Stanley, N.R., Berks, B.C., 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.A., 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.M., and Turner, R.J. (2010). DmsD, a Tat system specific chaperone, interacts with other general chaperones and proteins involved in the molybdenum cofactor biosynthesis. Biochim Biophys Acta *1804*, 1301-1309.

Lindenstrauss, U., and Bruser, T. (2009). Tat transport of linker-containing proteins in *Escherichia coli*. FEMS Microbiol Lett *295*, 135-140.

Lindenstrauss, U., Matos, C.F., Graubner, W., Robinson, C., and Bruser, T. (2010). Malfolded recombinant Tat substrates are Tat-independently degraded in *Escherichia coli*. FEBS Lett *584*, 3644-3648.

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

Maldonado, B., Buchanan, G., Muller, M., Berks, B.C., and Palmer, T. (2011). Genetic evidence for a TatC dimer at the core of the *Escherichia coli* twin arginine (Tat) protein translocase. J Mol Microbiol Biotechnol *20*, 168-175.

Matos, C.F., 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 Rep *10*, 474-479.

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

Maurer, C., Panahandeh, S., Jungkamp, A.C., Moser, M., and Muller, M. (2010). TatB functions as an oligomeric binding site for folded Tat precursor proteins. Mol Biol Cell *21*, 4151-4161.

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

Meir, A., Helppolainen, S.H., Podoly, E., Nordlund, H.R., Hytonen, V.P., Maatta, J.A., Wilchek, M., Bayer, E.A., Kulomaa, M.S., and Livnah, O. (2009). Crystal structure of rhizavidin: insights into the enigmatic high-affinity interaction of an innate biotin-binding protein dimer. J Mol Biol *386*, 379-390.

Mergulhao, F.J., Summers, D.K., and Monteiro, G.A. (2005). Recombinant protein secretion in *Escherichia coli*. Biotechnol Adv *23*, 177-202.

Molik, S., Karnauchov, I., Weidlich, C., Herrmann, R.G., and Klosgen, R.B. (2001). The Rieske Fe/S protein of the cytochrome b6/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. (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.

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

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

Natale, P., Bruser, T., and Driessen, A.J. (2008). Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane--distinct translocases and mechanisms. Biochim Biophys Acta *1778*, 1735-1756.

Niviere, V., Wong, S.L., and Voordouw, G. (1992). Site-directed mutagenesis of the hydrogenase signal peptide consensus box prevents export of a beta-lactamase fusion protein. J Gen Microbiol *138*, 2173-2183.

Oates, J., Barrett, C.M., Barnett, J.P., Byrne, K.G., 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 *346*, 295-305.

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

Ochsner, U.A., Snyder, A., Vasil, A.I., and Vasil, M.L. (2002). Effects of the twin-arginine translocase on secretion of virulence factors, stress response, and pathogenesis. Proc Natl Acad Sci U S A *99*, 8312-8317.

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

Palmer, T., and Berks, B.C. (2012). The twin-arginine translocation (Tat) protein export pathway. Nat Rev Microbiol.

Palmer, T., Sargent, F., and Berks, B.C. (2005). Export of complex cofactor-containing proteins by the bacterial Tat pathway. Trends Microbiol *13*, 175-180.

Panahandeh, S., Maurer, C., Moser, M., DeLisa, M.P., 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.

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

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

Posey, J.E., Shinnick, T.M., and Quinn, F.D. (2006). Characterization of the twin-arginine translocase secretion system of *Mycobacterium smegmatis*. J Bacteriol *188*, 1332-1340.

Pradel, N., Ye, C., Livrelli, V., Xu, J., Joly, B., and Wu, L.F. (2003). Contribution of the twin arginine translocation system to the virulence of enterohemorrhagic *Escherichia coli* O157:H7. Infect Immun *71*, 4908-4916.

Prickril, B.C., Czechowski, M.H., Przybyla, A.E., Peck, H.D., Jr., and LeGall, J. (1986). Putative signal peptide on the small subunit of the periplasmic hydrogenase from *Desulfovibrio vulgaris*. J Bacteriol *167*, 722-725.

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

Ray, N., Oates, J., Turner, R.J., 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.

Reynolds, M.M., Bogomolnaya, L., Guo, J., Aldrich, L., Bokhari, D., Santiviago, C.A., McClelland, M., and Andrews-Polymenis, H. (2011). Abrogation of the twin arginine transport system in *Salmonella enterica* serovar Typhimurium leads to colonization defects during infection. PLoS One *6*, e15800.

Ribnicky, B., Van Blarcom, T., and Georgiou, G. (2007). A scFv antibody mutant isolated in a genetic screen for improved export via the twin arginine transporter pathway exhibits faster folding. J Mol Biol *369*, 631-639.

Robinson, C., and Bolhuis, A. (2004). Tat-dependent protein targeting in prokaryotes and chloroplasts. Biochim Biophys Acta *1694*, 135-147.

Robinson, C., and Klosgen, R.B. (1994). Targeting of proteins into and across the thylakoid membrane--a multitude of mechanisms. Plant Mol Biol *26*, 15-24.

Robinson, C., Klosgen, R.B., Herrmann, R.G., and Shackleton, J.B. (1993). Protein translocation across the thylakoid membrane--a tale of two mechanisms. FEBS Lett *325*, 67-69.

Rocco, M.A., Waraho-Zhmayev, D., and Delisa, M.P. (2012). Twin-arginine translocase mutations that suppress folding quality control and permit export of misfolded substrate proteins. Proc Natl Acad Sci U S A *109*, 13392-13397.

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

Rose, R.W., Bruser, T., Kissinger, J.C., 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.

Rossier, O., and Cianciotto, N.P. (2005). The *Legionella pneumophila* tatB gene facilitates secretion of phospholipase C, growth under iron-limiting conditions, and intracellular infection. Infect Immun *73*, 2020-2032.

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

Sambasivarao, D., Turner, R.J., Simala-Grant, J.L., Shaw, G., Hu, J., and Weiner, J.H. (2000). Multiple roles for the twin arginine leader sequence of dimethyl sulfoxide reductase of *Escherichia coli*. J Biol Chem *275*, 22526-22531.

Sanders, C., Wethkamp, N., and Lill, H. (2001). Transport of cytochrome c derivatives by the bacterial Tat protein translocation system. Mol Microbiol *41*, 241-246.

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

Sargent, F. (2007a). Constructing the wonders of the bacterial world: biosynthesis of complex enzymes. Microbiology *153*, 633-651.

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

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

Sargent, F., Stanley, N.R., Berks, B.C., 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.

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

Schreiber, S., Stengel, R., Westermann, M., Volkmer-Engert, R., Pop, O.I., and Muller, J.P. (2006). Affinity of TatCd for TatAd elucidates its receptor function in the *Bacillus subtilis* twin arginine translocation (Tat) translocase system. J Biol Chem *281*, 19977-19984.

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

Shaner, N.C., Campbell, R.E., Steinbach, P.A., Giepmans, B.N., Palmer, A.E., and Tsien, R.Y. (2004). Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma sp*. red fluorescent protein. Nat Biotechnol *22*, 1567-1572.

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

Shruthi, H., Anand, P., Murugan, V., and Sankaran, K. (2010a). Twin arginine translocase pathway and fast-folding lipoprotein biosynthesis in *E. coli*: interesting implications and applications. Mol Biosyst *6*, 999-1007.

Shruthi, H., Babu, M.M., and Sankaran, K. (2010b). TAT-pathway-dependent lipoproteins as a niche-based adaptation in prokaryotes. J Mol Evol *70*, 359-370.

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

Strambio-De-Castillia, C., Niepel, M., and Rout, M.P. (2010). The nuclear pore complex: bridging nuclear transport and gene regulation. Nat Rev Mol Cell Biol *11*, 490-501.

Strauch, E.M., and Georgiou, G. (2007a). A bacterial two-hybrid system based on the twin-arginine transporter pathway of *E. coli*. Protein Sci *16*, 1001-1008.

Strauch, E.M., and Georgiou, G. (2007b). *Escherichia coli* tatC mutations that suppress defective twin-arginine transporter signal peptides. J Mol Biol *374*, 283-291.

Studier, F.W., Rosenberg, A.H., Dunn, J.J., and Dubendorff, J.W. (1990). Use of T7 RNA polymerase to direct expression of cloned genes. Methods Enzymol *185*, 60-89.

Tarry, M., Arends, S.J., Roversi, P., Piette, E., Sargent, F., Berks, B.C., Weiss, D.S., and Lea, S.M. (2009a). The *Escherichia coli* cell division protein and model Tat substrate SufI (FtsP) localizes to the septal ring and has a multicopper oxidase-like structure. J Mol Biol *386*, 504-519.

Tarry, M.J., Schafer, E., Chen, S., Buchanan, G., Greene, N.P., Lea, S.M., Palmer, T., Saibil, H.R., and Berks, B.C. (2009b). 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.

Theg, S.M., Cline, K., Finazzi, G., and Wollman, F.A. (2005). The energetics of the chloroplast Tat protein transport pathway revisited. Trends Plant Sci *10*, 153-154.

Thomas, J.D., Daniel, R.A., 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.

Tombola, F., Pathak, M.M., and Isacoff, E.Y. (2005). Voltage-sensing arginines in a potassium channel permeate and occlude cation-selective pores. Neuron *45*, 379-388.

Tranier, S., Mortier-Barriere, I., Ilbert, M., Birck, C., Iobbi-Nivol, C., Mejean, V., and Samama, J.P. (2002). Characterization and multiple molecular forms of TorD from *Shewanella massilia*, the putative chaperone of the molybdoenzyme TorA. Protein Sci *11*, 2148-2157.

Turner, R.J., Papish, A.L., and Sargent, F. (2004). Sequence analysis of bacterial redox enzyme maturation proteins (REMPs). Can J Microbiol *50*, 225-238.

van Dongen, W., Hagen, W., van den Berg, W., and Veeger, C. (1988). Evidence for an unusual mechanism of membrane translocation of the periplasmic hydrogenase of Desulfovibrio vulgaris (Hildenborough), as derived from expression in *Escherichia coli*. FEMS Microbiology Letters *50*, 5-9.

von Heijne, G., and Gavel, Y. (1988). Topogenic signals in integral membrane proteins. Eur J Biochem *174*, 671-678.

Walther, T.H., Grage, S.L., Roth, N., and Ulrich, A.S. (2010). Membrane alignment of the pore-forming component TatA(d) of the twin-arginine translocase from *Bacillus subtilis* resolved by solid-state NMR spectroscopy. J Am Chem Soc *132*, 15945-15956.

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

Wang, X., and Lavrov, D.V. (2007). Mitochondrial genome of the homoscleromorph *Oscarella carmela* (Porifera, Demospongiae) reveals unexpected complexity in the common ancestor of sponges and other animals. Mol Biol Evol *24*, 363-373.

Weiner, J.H., Bilous, P.T., Shaw, G.M., Lubitz, S.P., Frost, L., Thomas, G.H., Cole, J.A., and Turner, R.J. (1998). A novel and ubiquitous system for

membrane targeting and secretion of cofactor-containing proteins. Cell *93*, 93-101.

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

Westermann, M., Pop, O.I., Gerlach, R., Appel, T.R., Schlormann, W., Schreiber, S., and Muller, J.P. (2006). The TatAd component of the *Bacillus subtilis* twin-arginine protein transport system forms homo-multimeric complexes in its cytosolic and membrane embedded localisation. Biochim Biophys Acta *1758*, 443-451.

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

Widdick, D.A., Eijlander, R.T., van Dijl, J.M., Kuipers, O.P., 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.

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

Yamada, H., Yamagata, H., and Mizushima, S. (1984). The major outer membrane lipoprotein and new lipoproteins share a common signal peptidase that exists in the cytoplasmic membrane of *Escherichia coli*. FEBS Lett *166*, 179-182.

Yanisch-Perron, C., Vieira, J., and Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene *33*, 103-119.

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

Yoon, S.H., Kim, S.K., and Kim, J.F. (2010). Secretory production of recombinant proteins in *Escherichia coli*. Recent Pat Biotechnol *4*, 23-29.

Zoufaly, S., Frobel, J., Rose, P., Flecken, T., Maurer, C., Moser, M., and Muller, M. (2012). Mapping precursor-binding site on TatC subunit of twin arginine-

specific protein translocase by site-specific photo cross-linking. J Biol Chem *287*, 13430-13441.

APPENDIX

GENERATION OF TATA WITH A C-TERMINAL 6XHIS-TAG FOR PURIFICATION AND LABELING

Purification of the channel forming unit of the Tat translocase has been reported in multiple studies (Gohlke et al., 2005; Mehner et al., 2012). To allow for rapid purification of TatA, as well as to create a C-terminal site for fluorescent labeling, a 6xHis-tag and a cysteine (together called a HisC-tag) were attached to its C-terminus. These were created by inverse PCR using the primers 5'-ATTAGCAATGATGATGATGATGCACCGCTCTTTATCGCGCTTCG-3' and 5'-CCATGGTTGATATCGGTTTTAGCGAACTGCTATTGGTGTTCATCATCGG-3' to amplify pTatABC (described in (Bageshwar and Musser, 2007)). The amplified product was then self-ligated generating the plasmid pTatA-H6-BC, which encodes TatA-HisC, TatB, an TatC. TatA-HisC was found to be as active as wild-type TatA when tested *in vitro* (Figure A1).

The plasmid pTatA-H6-BC uses the arabinose inducible pBAD promoter. In order to increase expression yield for potential TatA purification, the TatA-HisC coding region of pTatA-H6-BC was inserted into the plasmid pACYCDuet-1 at the NcoI site, generating plasmid pDUET-TatA-H6 (Figure A2).

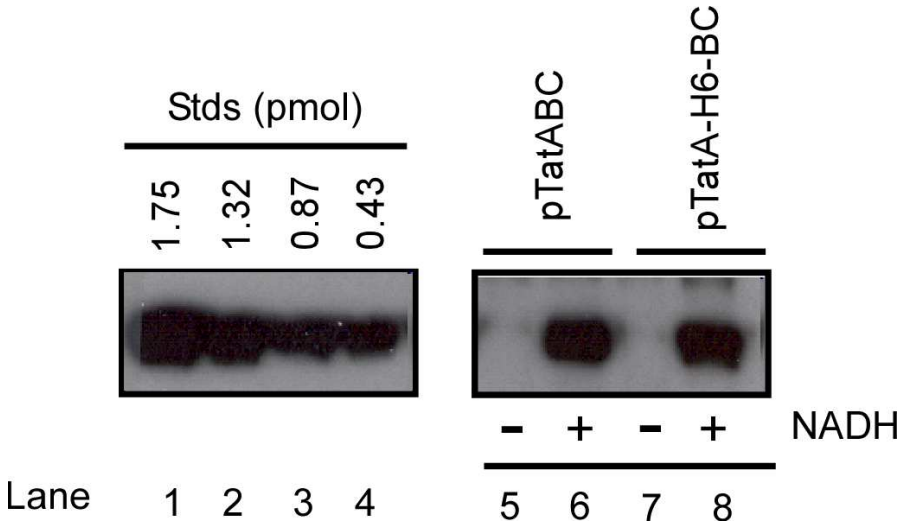


Figure A1. TatA is functional with a C-terminal 6xHis-tag and cysteine. Tat function was tested using the protease protection assay described previously, using 30 min transport reactions with IMVs (A_{280} = 5), 4 mM NADH and 50 nM pre-SufI. Lanes 5 and 6 use "wild-type" IMVs that were generated by expression of the Tat machinery from the plasmid pTatABC (Bageshwar and Musser, 2007). Lanes 7 and 8 use IMVs with Tat machinery expressed from plasmid pTatA-H6-BC in *E. coli* strain MC4100, induced by addition of 0.7% arabinose as described in Bageshwar and Musser (2007). Lanes 1-4 are standards.

Labeling of TatA-HisC was performed with IMVs produced from *E. coli* strain BL21 containing pDUET-TatA-H6 essentially as described in Chapter II, except TatA-HisC expression was induced with 1 mM ITPG. The labeling reaction was performed with AlexaFluor568 maleimide using the standard protocol as described in Chapter II with dye concentrations 10x higher than the protein concentration in the IMVs estimated by A_{280} in 2% SDS. When resolved via SDS-PAGE and visualized by Alexa568 florescence, a major band was found at the molecular weight at which TatA typically migrates (~18 kDa) (Figure A3).

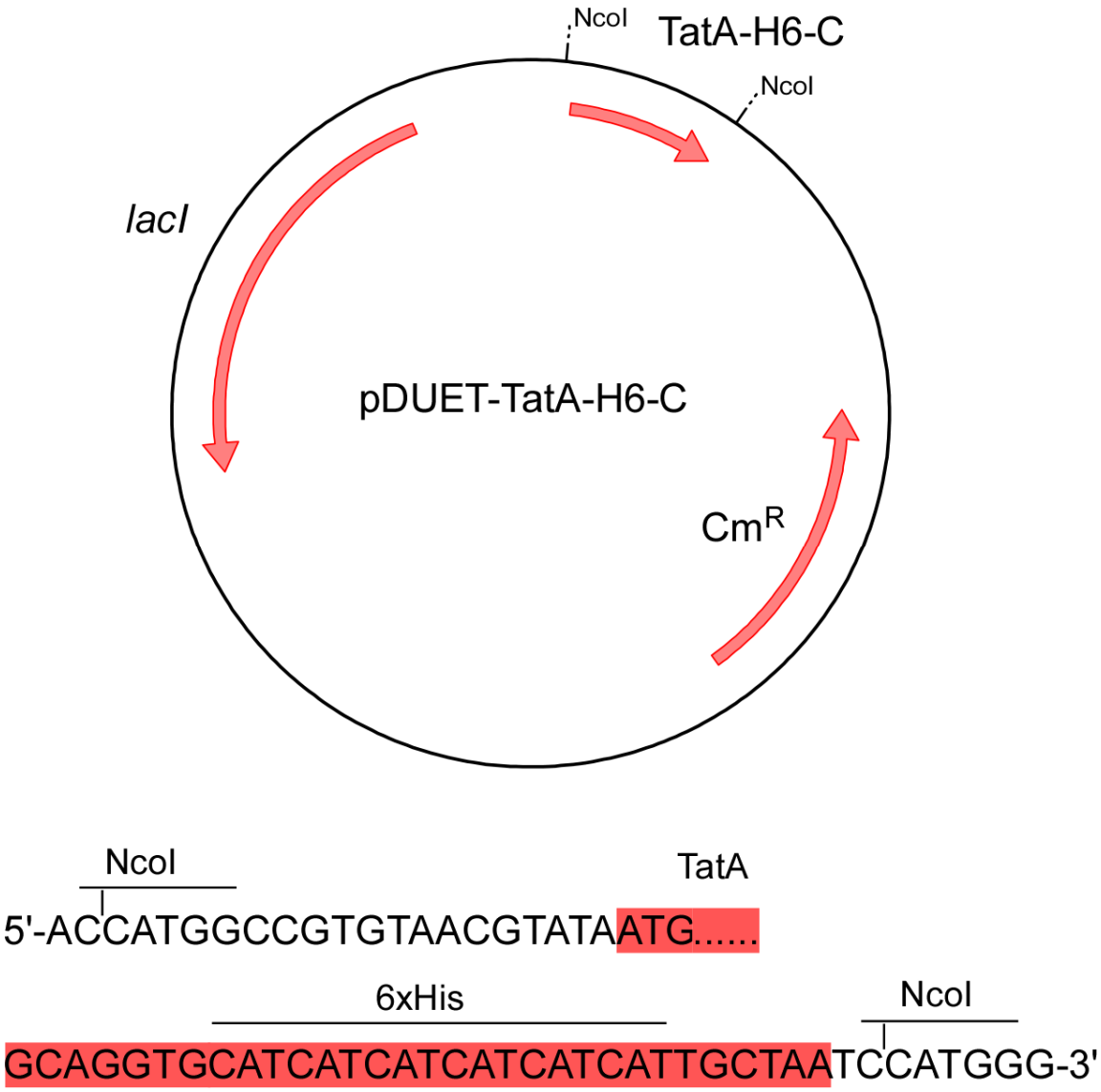


Figure A2. Plasmid map of pDUET-TatA-H6-C. The coding region for TatA-H6-C was excised from from pTatA-H6-C and inserted into the NcoI site of pACYCDuet-1.

The putative TatA band appears as a doublet when visualized by fluorescence for unknown reasons.

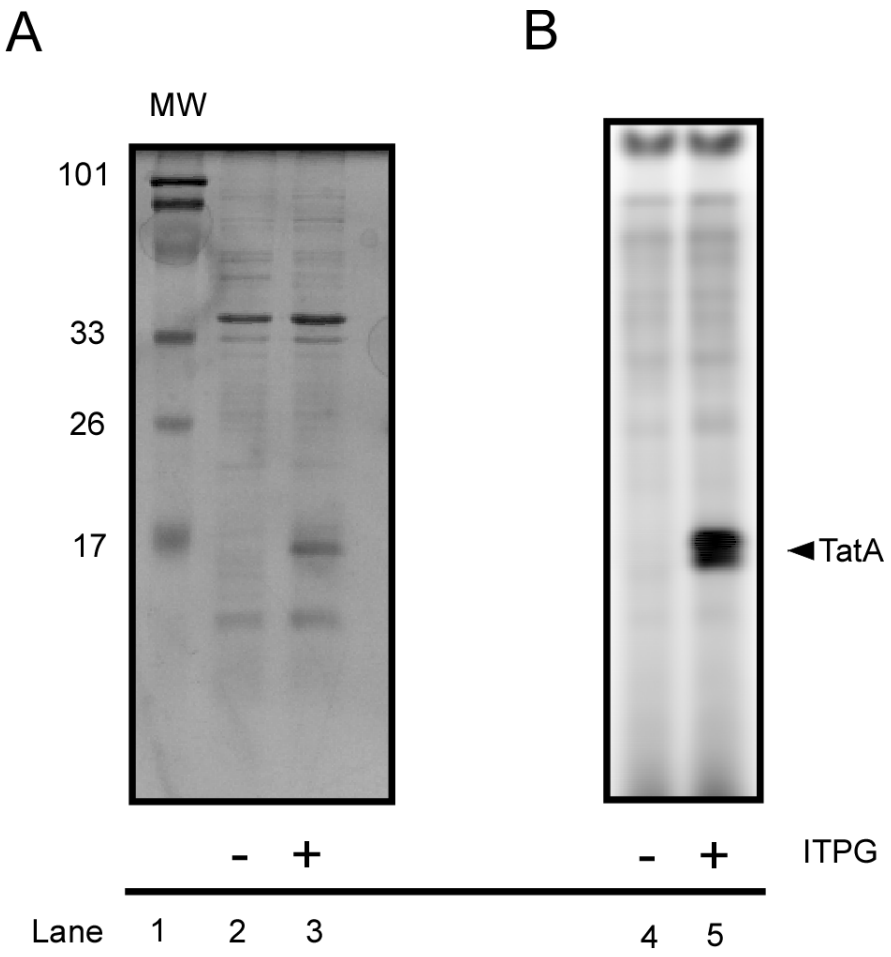


Figure A3. TatA-HisC was labeled with Alexa568. IMVs isolated from *E. coli* (pDuet-TatA-H6) in the presence of (lanes 3 and 5) or absence of (lanes 2 and 4) Tat-HisC induction were resolved by SDS-PAGE. The gel was visualized both by Coomassie blue staining (A) and by Alexa568 fluorescence (B).

In one study, it was reported that TatA was purified from a soluble fraction of lysed cells by affinity purification (Mehner et al., 2012), despite other studies stating that TatA requires detergent solubilization to be isolated (Gohlke et al., 2005; Porcelli et al., 2002). A similar approach was attempted by over-expressing the Tat machinery in *E. coli* strain MC4100 cells containing the plasmid pTatA-H6-BC. Cells were induced with 1.2% arabinose for 5 hours at 37ºC. After lysis by French press, the lysate was pelleted and the supernatant was run over a Ni-NTA purification column. No TatA-HisC was found to have bound to the column, or to be present in any part of the lysate other than the pellet. This result appears contrary to the data in (Mehner et al., 2012) despite similar protocols. However, use of a 6xHis tag rather than a Strep tag may explain the different results.

TAT INHIBITOR INCREASES THE LENGTH OF THE LAG PHASE IN FRET ASSAY

Recently, a Tat inhibitor was identified by a colleague in the lab, Umesh Bageshwar, by high through-put screening. The inhibitor, termed TatBlock-1, was added to experiments essentially identical as the FRET assays as described in Chapter II. In the presence of TatBlock-1, the lag phase was extended. Additionally, at higher concentrations of the inhibitor the lag was further extended (Figure A4). TatBlock-1 was added after addition of cargo and

did not result in a decrease in FRET signal, indicating that the inhibitor does not act by directly competing for the cargo's binding site on TatBC.

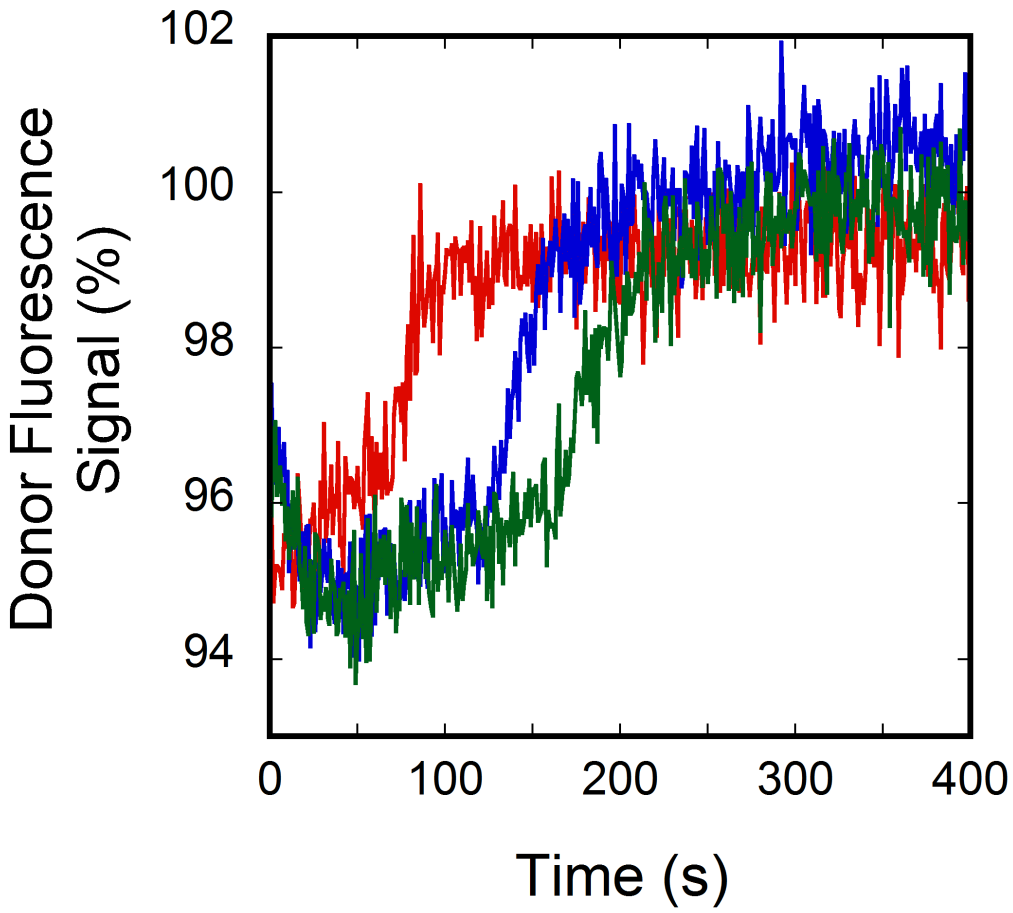


Figure A4. TatBlock-1 delays cargo migration from TatBC. Reactions were performed with no TatBlock-1 (red), 40 μM TatBlock-1 (blue), or 100 μM TatBlock-1 (green). The fluorescence signal is emission at 550 nm (donor) upon membrane energization with excitation at 500 nm. TatBlock-1 was allowed to incubate with membranes for 300 seconds before addition of NADH. Reactions contained pre-SufI(T96C)^{Alexa532} (20 nM) and IMVs (A_{280} = 2) with TatAB^{cherry}C. NADH (4 mM) was added at t = 0 s.

DERIVATION OF THE BINDING EQUATION INCLUDING THE RECEPTOR CONCENTRATION (EQUATION 2 FOR FIGURE 2.2D)

If it can be assumed that the TatBC complex concentration is much less than the precursor K_D, the binding data in Figure 2.2D can be fit to a typical single site binding equation:

$$y = \frac{B_{max} x}{(K_D + x)} \qquad \text{(equation 3)}$$

However, if the TatBC concentration and K_D are of similar magnitude, this equation is invalid. The appropriate fitting equation is derived as follows.

The relevant binding equilibrium is:

$$P + T \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} PT \qquad K_D = \frac{k_{-1}}{k_1} = \frac{[P][T]}{[PT]} \qquad \text{(equation 4)}$$

where P is the precursor and T is the Tat(A)BC receptor complex. Assuming that P_a is the total amount of precursor added and that T_0 is the concentration of receptor binding sites, mass balance yields:

$$P_a = [P] + [PT]$$
$$T_0 = [T] + [PT]$$

[PT] can then be obtained from equation 4 after substitution and use of the quadratic formula,

$$K_D[PT] = [P][T]$$

$$K_D[PT] = \left(P_a - [PT]\right)\left(T_0 - [PT]\right)$$

$$0 = [PT]^2 - (P_a + T_0 + K_D)[PT] + P_a T_0$$

$$[PT] = \frac{(P_a + T_0 + K_D) \pm \sqrt{(P_a + T_0 + K_D)^2 - 4P_a T_0}}{2}$$

The proper solution is given by the difference in the numerator. Since T_0 is unknown, dividing by T_0 yields the percent occupancy of the available binding sites,

$$\frac{[PT]}{T_0} = \frac{(P_a + T_0 + K_D) - \sqrt{(P_a + T_0 + K_D)^2 - 4P_a T_0}}{2T_0}$$

(equation 5)

A multiplicative scaling factor, C, was then used to fit the data in Figure 2.2D with C, T_0 and K_D as fitting parameters,

$$y = C * \frac{(x + T_0 + K_D) - \sqrt{(x + T_0 + K_D)^2 - 4(x)T_0}}{2T_0}$$

(equation 6)

Equation 6 is the formula given in the Experimental Procedures section (equation 2).

Note that if $T_0 << K_D$, then [P] can be approximated by P_a (i.e., the bound precursor is always a small fraction of the total precursor added). Under these conditions,

$$K_D \approx \frac{P_a(T_0 - [PT])}{[PT]} \qquad \text{or,} \qquad \frac{[PT]}{T_0} \approx \frac{P_a}{K_D + P_a}$$

(equation 7)

With the exception of a scaling factor, equation 7 is identical to equation 3.

Unfortunately, fitting the data in Figure 2.2D with equation 6 does not yield a robust, unique solution for T_0 and K_D. However a range of possible values can be determined by determining the extremes of possible values. In principle, T_0 could be infinitely small. However, the K_D determined from the fit is essentially identical (= 23 nM) for $T_0 \leq 0.1$ nM. For T_0 = 20 nM, the K_D = 7.5 nM. For $T_0 >$ 20 nM, the fit becomes significantly worse. Thus, a range of possible values consistent with the data in Figure 2.2D is $T_0 \approx$ 0-20 nM, and $K_D \approx$ 7-23 nM. Due to the high mCherry concentration present in our samples, which reports on the receptor concentration, a high T_0 (~20 nM) is more likely, which corresponds to the lower end of the K_D range (i.e., ~7 nM).

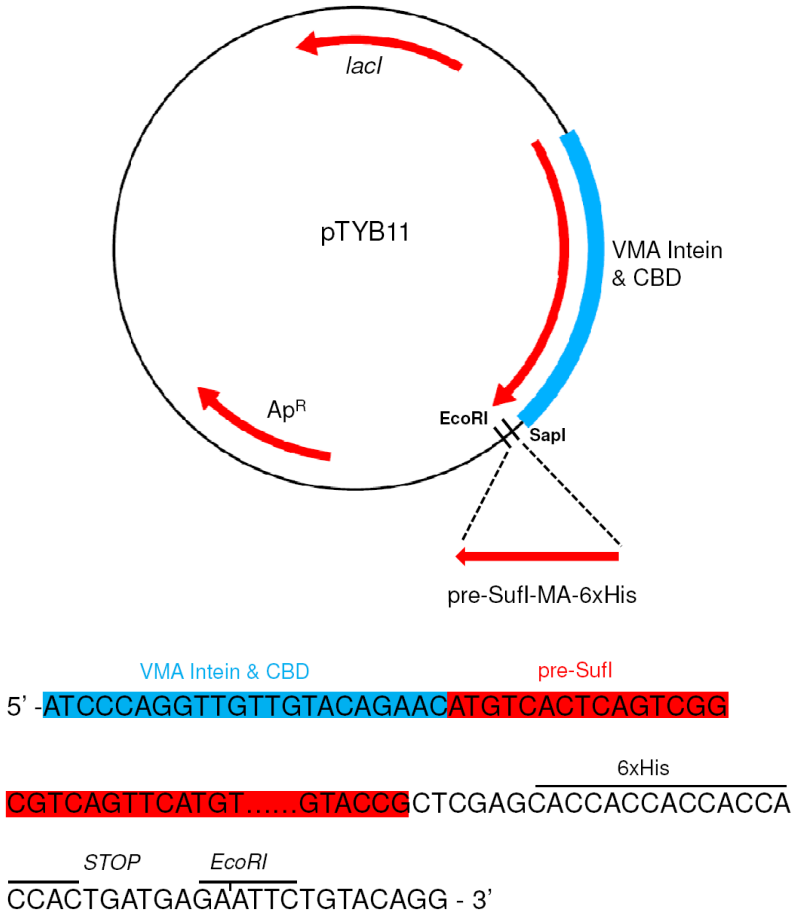
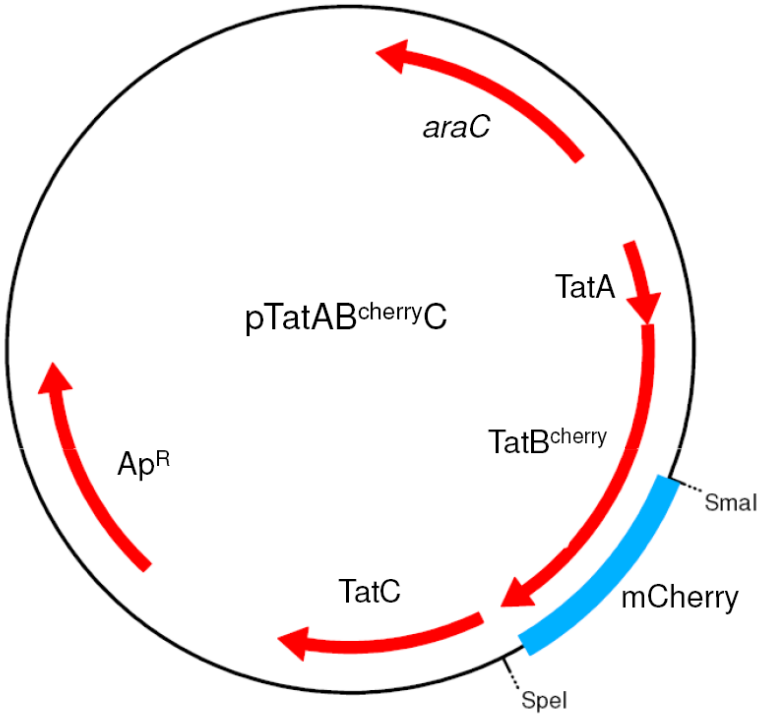


Figure A5. Generation of pIntein-preSufI-MA from pTYB11. The plasmid pIntein-preSufI-MA was generated by inserting a cysteine-free form of pre-SufI with a 6xHis tag into the vector pTYB11 (New England Biolabs) at restriction enzyme sites SapI and EcoRI. Since SapI does not cleave at the restriction enzyme recognition site, the recognition sequence was lost after insertion of the pre-SufI sequence. The pre-SufI gene was obtained from pET-SufI (Yahr and Wickner, 2001) and the cysteines were removed with mutations C17M and C295A. Site directed mutagenesis was used to generate plasmids pIntein-preSufI-MA(G29C), pIntein-preSufI-MA(G45C), pIntein-preSufI-MA(T96C), pIntein-preSufI-MA(S204C), pIntein-preSufI-MA(M338C), pIntein-preSufI-MA(G410C), the names of which reflect single cysteine mutations in the mature region of the pre-SufI protein. Plasmids pIntein-preSufI-MA(L470C) and pIntein-preSufI-MA(479C) were made by mutating a leucine in the short linker region between the mature domain and the 6xHis tag (CTC codon), and by adding a cysteine residue to the C-terminus, respectively.



TatB SmaI mCherry

5’ -CCGTATACCCGGGTACCGGTCGCCACCATGGTG……

STOP

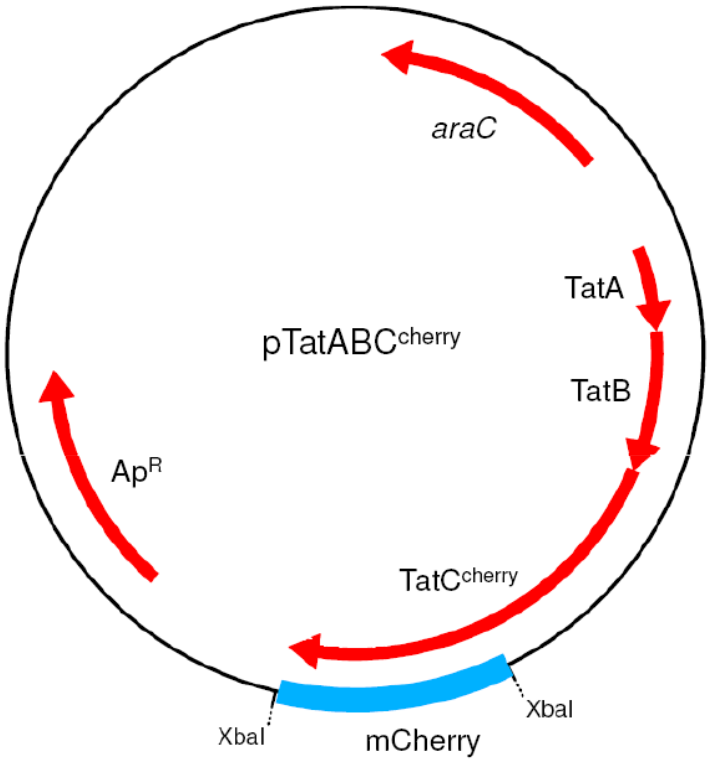
TACAAGTAGCGGCCGCGACTCTAGAATTCCAACTGAGCGC

CGGTCGCTACCATTACCAACTTGTCTGGTGTCAAAAA

SpeI TatC

TAATAGGCCTACTAGTCATGTCT- 3’

Figure A6. Plasmid map of pTatAB^{cherry}C. Plasmid pTatAB^{cherry}C was generated by insertion of the mCherry sequence (plasmid pmCherry, Clontech) between restriction sites SmaI and SpeI. These restriction sites were inserted into pTatABC (Yahr and Wickner, 2001) immediately after the coding region of TatB by site-directed mutagenesis, as indicated, deleting the stop codon.



TatC XbaI

5’ -GAATATATTCAACCGCCTCTAGAGGATCCCCGGGTACC

mCherry STOP

GGTCGCCACCATGGTG……TACAAGTAGCGGCCGCGAC

XbaI

TCTAGA- 3’

Figure A7. Plasmid map of pTatABC^{cherry}. Plasmid pTatABC^{cherry} was generated by insertion of the mCherry sequence into pTatABC immediately after the coding region of TatC by insertion at an XbaI site. The TatC stop codon was removed by site-directed mutagenesis.

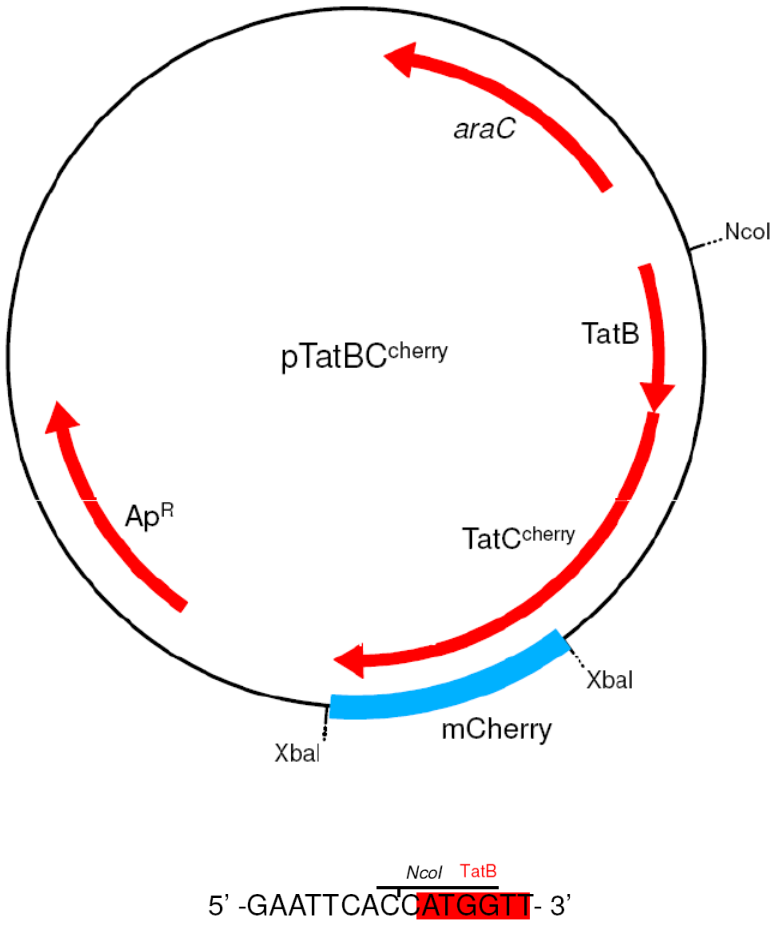


Figure A8. Plasmid map of pTatBC^{cherry}. Plasmids pTatBC and pTatBC^{cherry} were generated by engineering two NcoI sites flanking the TatA coding sequence and excising *tatA* from pTatABC and pTatABC^{cherry}, respectively.

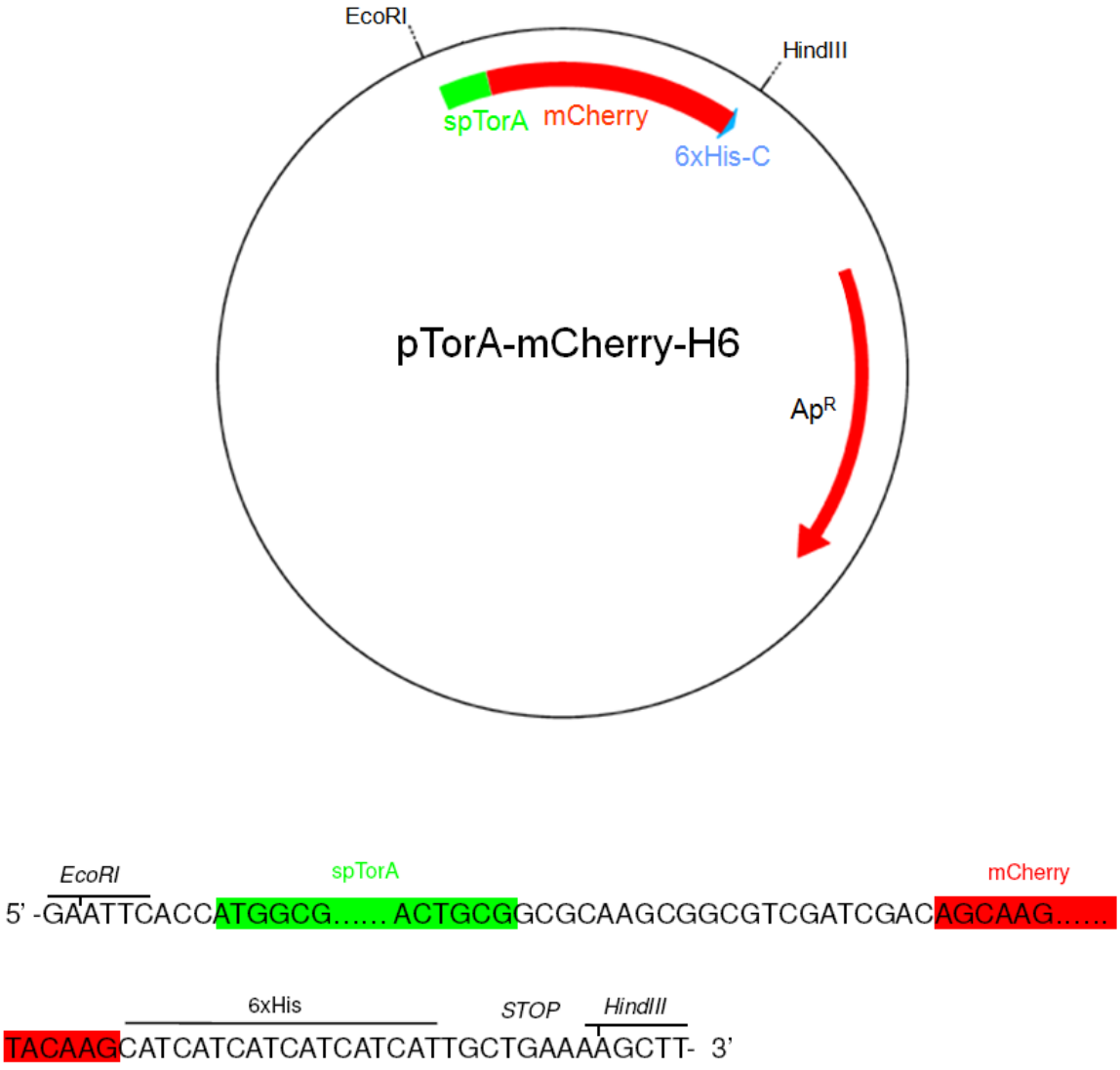


Figure A9. Plasmid map of pTorA-mCherry-H6. The coding sequence for the TorA signal peptide (from plasmid pTorA-GFP (Bageshwar and Musser, 2007)) was fused with the coding sequence for mCherry, as indicated, between restriction sites EcoRI and HindIII of pBAD24 (ATCC). The first four amino acids following spTorA (AQAA) are from the N-terminus of the TorA mature domain.

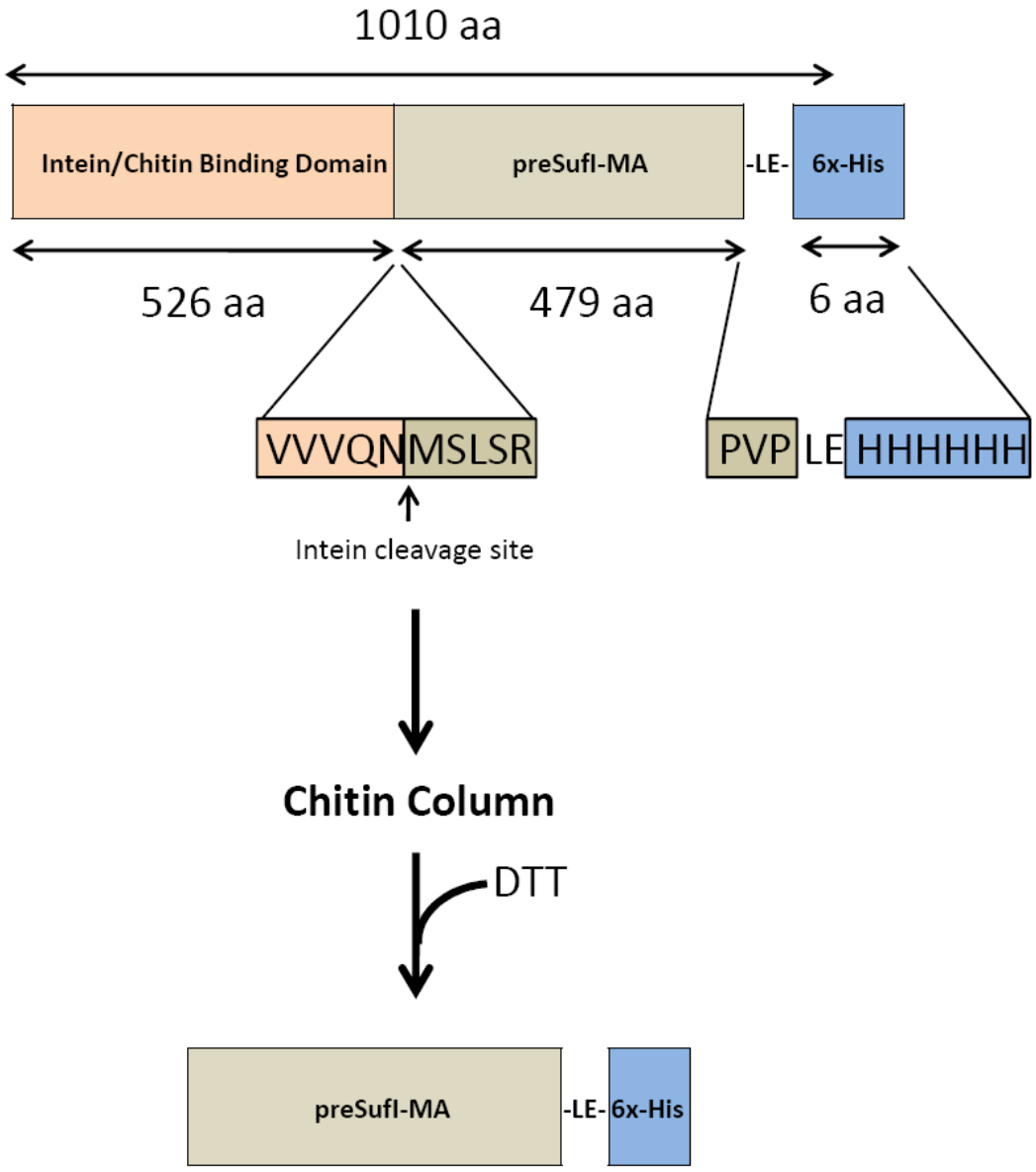


Figure A10. Purification of pre-SufI. In order to obtain pure precursor protein free from mature protein, pre-SufI was purified using the IMPACT-CN system according to the manufacturer (New England Biolabs). In brief, a bifunctional tag, consisting of intein and chitin binding domains, was fused to the N-terminus of pre-SufI. Crude cell extracts were added directly to a chitin column. Pre-SufI was cleaved off the column by activating intein cleavage by the addition of DTT. The protein was further purified by Ni-NTA chromatography, as described (Bageshwar and Musser, 2007).

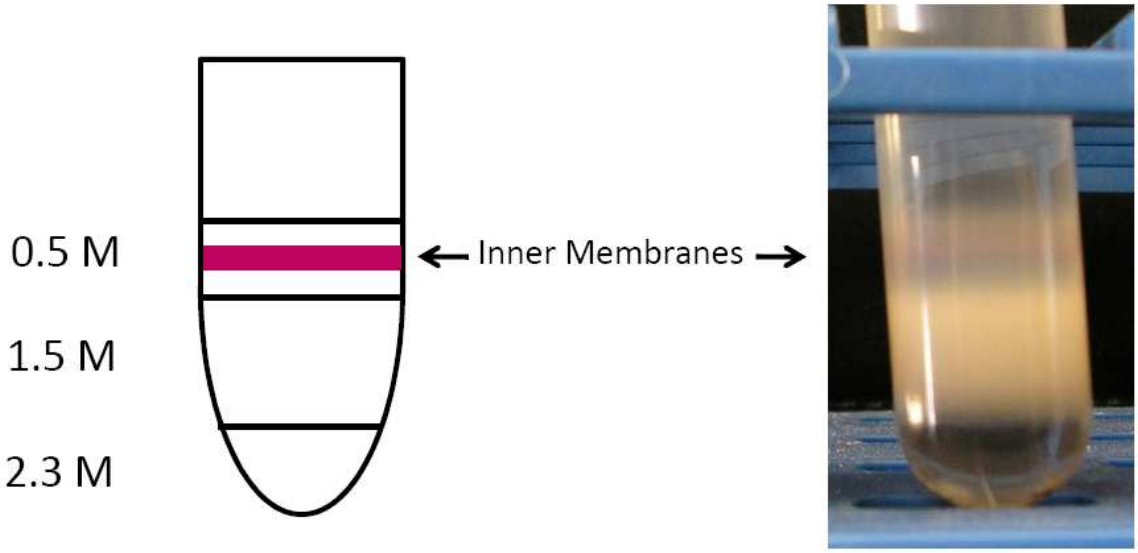


Figure A11. Sucrose gradient for IMV isolation. A three step (0.5 M, 1.5 M, and 2.3 M) sucrose gradient was used for IMV purification in Buffer B (1 mM KCl, 1 mM $MgSO_4$, 2 mM DTT, 10 mM Hepes, pH 7.0) (Bageshwar and Musser, 2007). IMVs typically sedimented within the 0.5 M sucrose fraction. A sucrose gradient with membranes containing $TatABC^{cherry}$ is shown on the right. We previously reported that IMV preparations with *E. coli* strain JM109 consistently yielded high transport efficiencies, but attempts with strain MC4100 were less successful, for unknown reasons (Bageshwar and Musser, 2007). The Tat proteins in this study were all expressed in strain MC4100ΔTatABCDE in order to obtain IMVs in which all TatB or TatC molecules were labeled with mCherry. The pink color of proteins labeled with mCherry served as a convenient marker during purification of inner membranes on a sucrose gradient as shown. Consequently, the recovered membranes were significantly more highly purified. Moreover, the IMVs displayed more consistent transport activity both in terms of transport efficiencies and transport rate. Transport rates were about an order of magnitude faster (Fig. 1*D*) than reported earlier (Bageshwar and Musser, 2007; Bageshwar et al., 2009). We expect that this is a consequence of the higher purity of the IMV preparations.

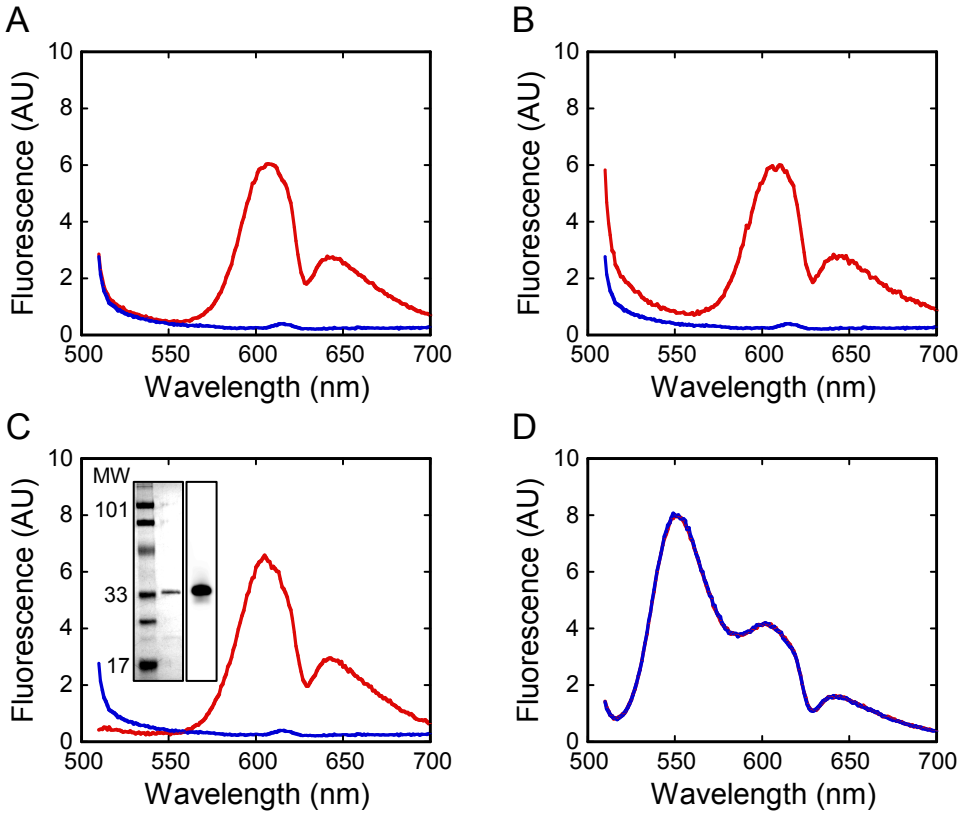


Figure A12. Fluorescence emission of mCherry-tagged proteins and lack of FRET for non-interacting proteins. (A and B) Emission spectra of IMVs containing TatAB^{cherry}C (A, *red*), TatABC^{cherry} (B, *red*) and TatABC (*blue*). (C) Emission spectra of mCherry (*red*) and TatABC IMVs (*blue*). Note that the mCherry protein exhibits two emission peaks. The mCherry protein was obtained by combining fractions of mature-length protein obtained from anion-exchange chromatography of spTorA-mCherry-H6. (inset, left) Coomassie blue stained SDS-PAGE gel of mCherry, with molecular weight ladder in the left lane. (inset, right) Fluorescence of the same gel before staining. The native structure of mCherry is retained on this gel because the sample was not boiled. As further tests to identify the reason behind the two well-resolved peaks for mCherry in our experiments, we tested phosphate buffer (eliminating effects of a component in the transport buffer), excitation wavelength (eliminating an off-peak excitation effect), and slit width (eliminating resolution due to a narrow emission window). None of these different conditions eliminated the two well-resolved peaks for mCherry emission. (D) Emission spectra of a mixture of pre-SufI(T96C)^{Alexa532} (20 nM) and spTorA-mCherry (300 nM) before (*red*) and after (*blue*) addition of 200 nM unlabeled pre-SufI(T96C). Note that there is no difference in the intensities of any of the peaks (compare with Fig. 2*A*). This confirms that the increase in the 550 nm peak observed in Fig. 2*A* results from loss of FRET. EX = 500 nm for all spectra.

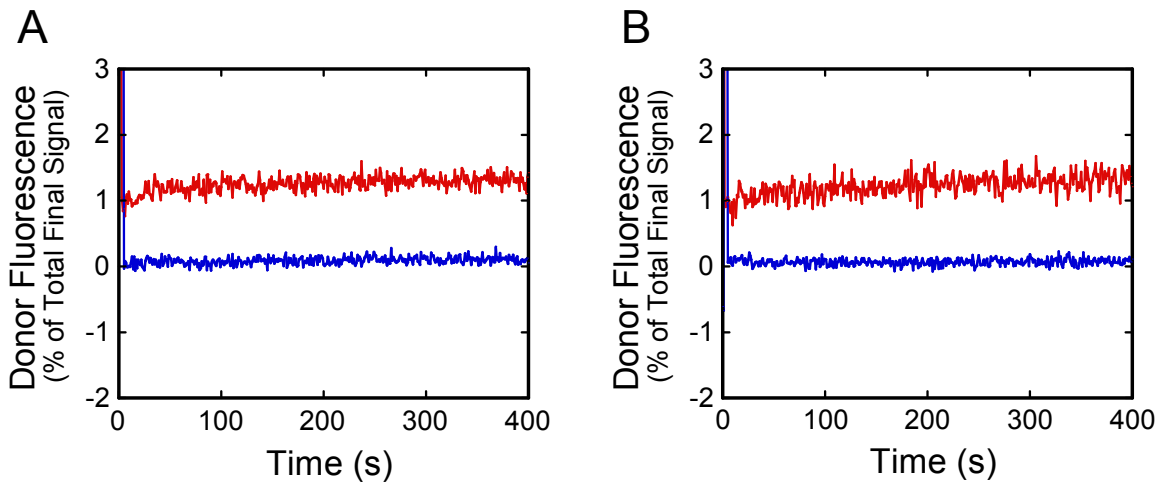


Figure A13. Control experiments for FRET kinetics. (A) Donor emission at 550 nm (EX = 500 nm) after addition of 200 nM unlabeled pre-SufI(T96C) at *t* = 0 s. Conditions are as in Fig. 2B except with TatABC IMVs (*red*) (i.e., with no mCherry acceptor) or with unlabeled pre-SufI(T96C) (*blue*) (i.e., with no Alexa532 donor). These data indicate that no observable signal change was detectable without both donor and acceptor fluorophores (compare with Fig. 2*B*). (B) Same as (A) except that reactions were initiated with 4 mM NADH at *t* = 0 s instead of competitor. Again, no observable signal change was detectable without both donor and acceptor fluorophores (compare with Fig. 4). For both panels, the same ordinate scale was used as in Figs. 2*B* and 4. Traces were zeroed and the upper traces (*red*) were displaced for clarity.

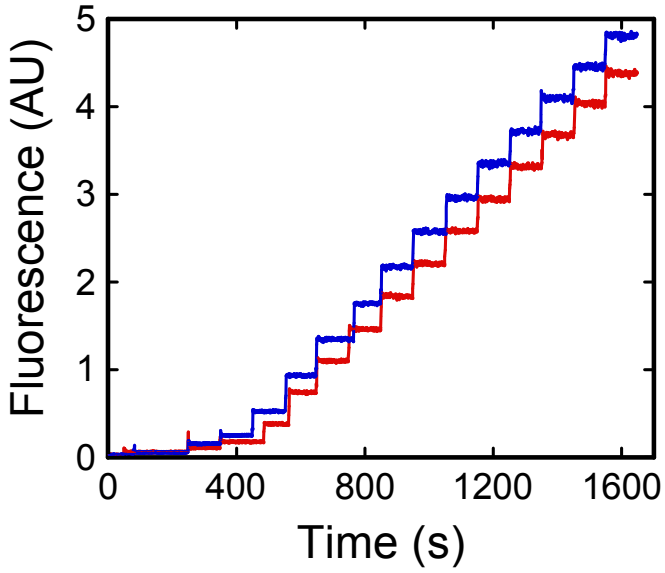


Figure A14. Titrations used to determine binding affinity. Sample titration of TatABC^{cherry} IMVs with pre-SufI(T96C)^{Alexa532} in the presence (*blue*) and absence (*red*) of 300 nM pre-SufI(T96C) as a competitive inhibitor. Each step results from the addition of pre-SufI(T96C)^{Alexa532}. One of the data sets used to generate Fig. 2*D*. EX = 500 nm; EM = 550 nm.

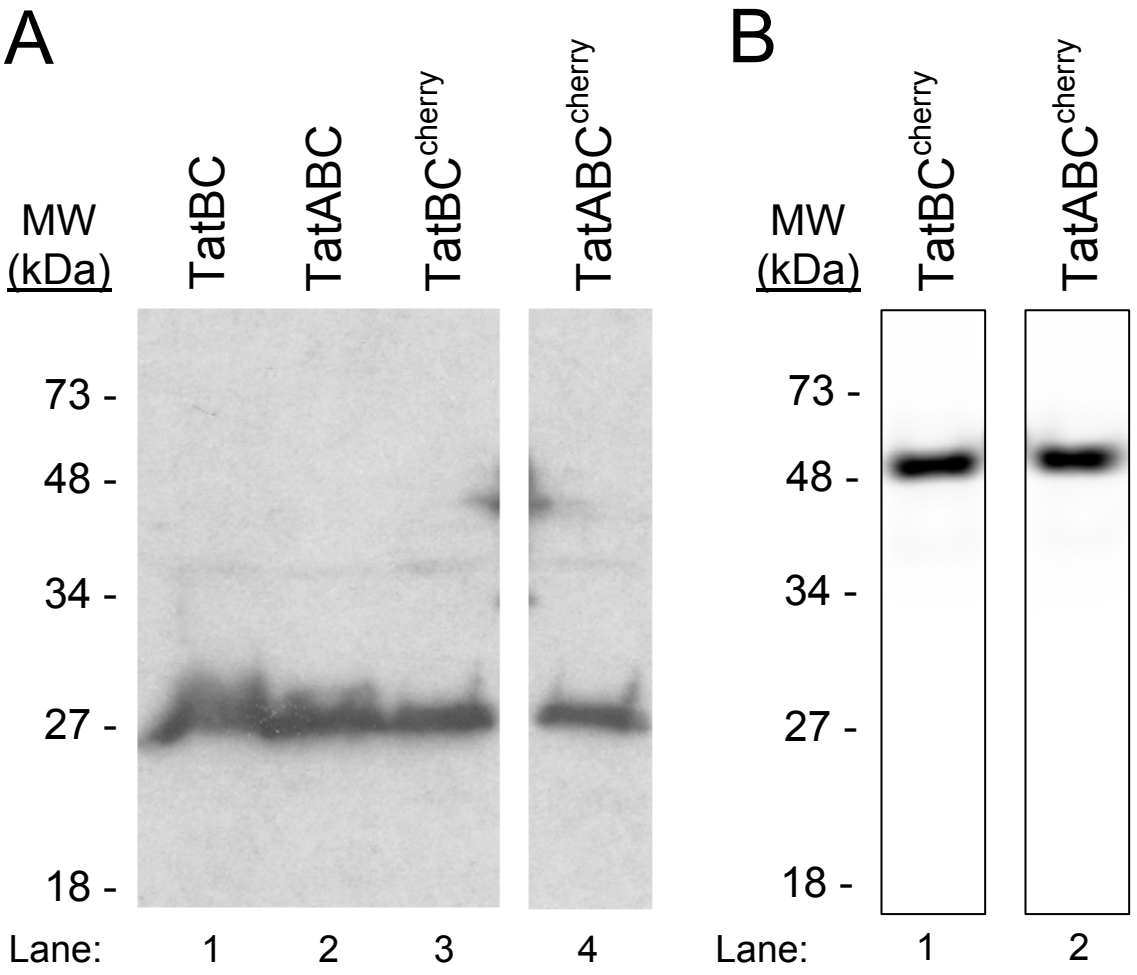


Figure A15. Expression of Tat proteins. (A) Western blot of the indicated IMV preparations using TatB antibodies. TatB protein levels are approximately the same in IMVs with or without TatA. The SDS-PAGE gel was electroblotted onto PVDF membranes and immunoblotted with rabbit polyclonal TatB antibodies (Yahr and Wickner, 2001) (1:5,000 dilution in 2% nonfat dry milk, 0.1% Triton X-100, and 0.1% Tween). Goat polyclonal anti–rabbit IgG-HRP conjugate (1:15,000; Santa Cruz Biotechnology, Inc.) was used as the secondary antibody, and bands were visualized by chemiluminescence. (B) $TatC^{cherry}$ fusions in the indicated IMV prreparations visualized by mCherry fluorescence. These data indicate that the mCherry protein was only present as part of the full-length fusion protein. The predicted molecular weights of TatB, TatC and $TatC^{cherry}$ are 18, 29, and 57 kDa, respectively. Note, however, that TatB and TatC migrate abnormally on gels (Sargent et al., 1998).

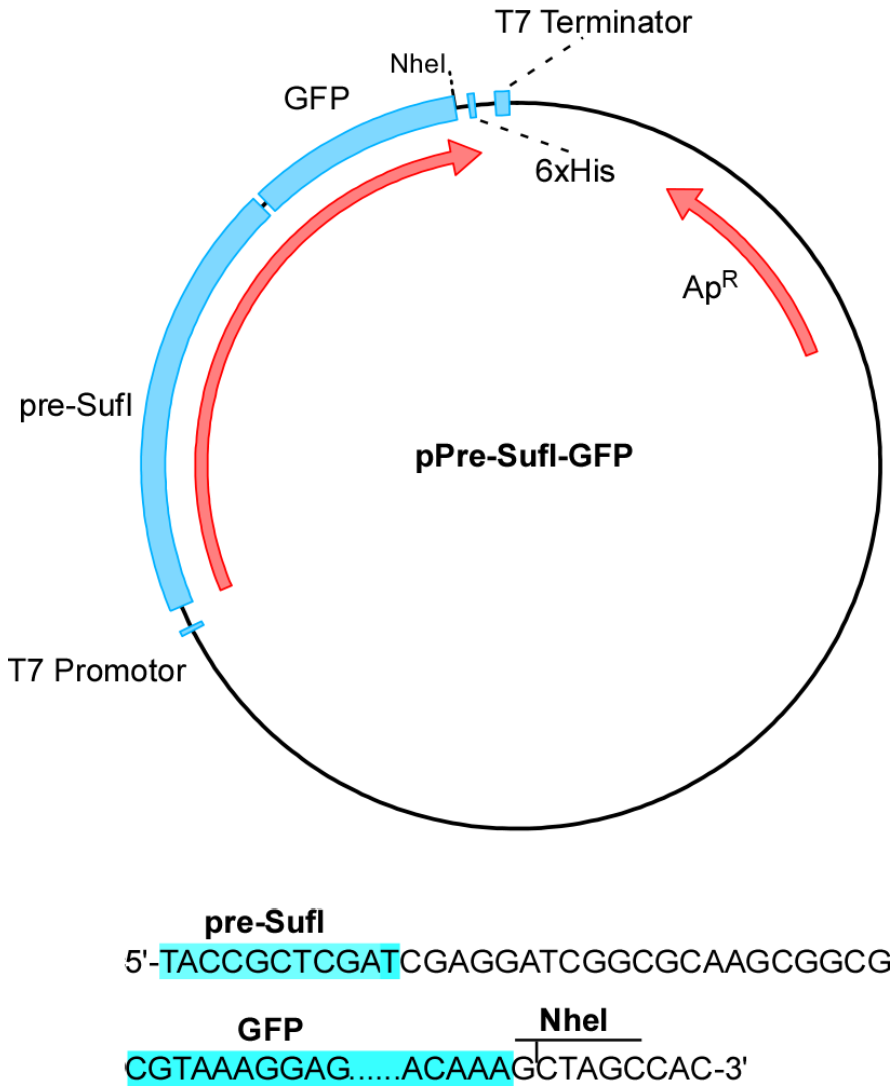


Figure A16. Plasmid map of pPre-SufI-GFP. The GFP gene was inserted behind the C-terminus of pre-SufI with a 21 bp linker, and before the 6xHis tag sequence.

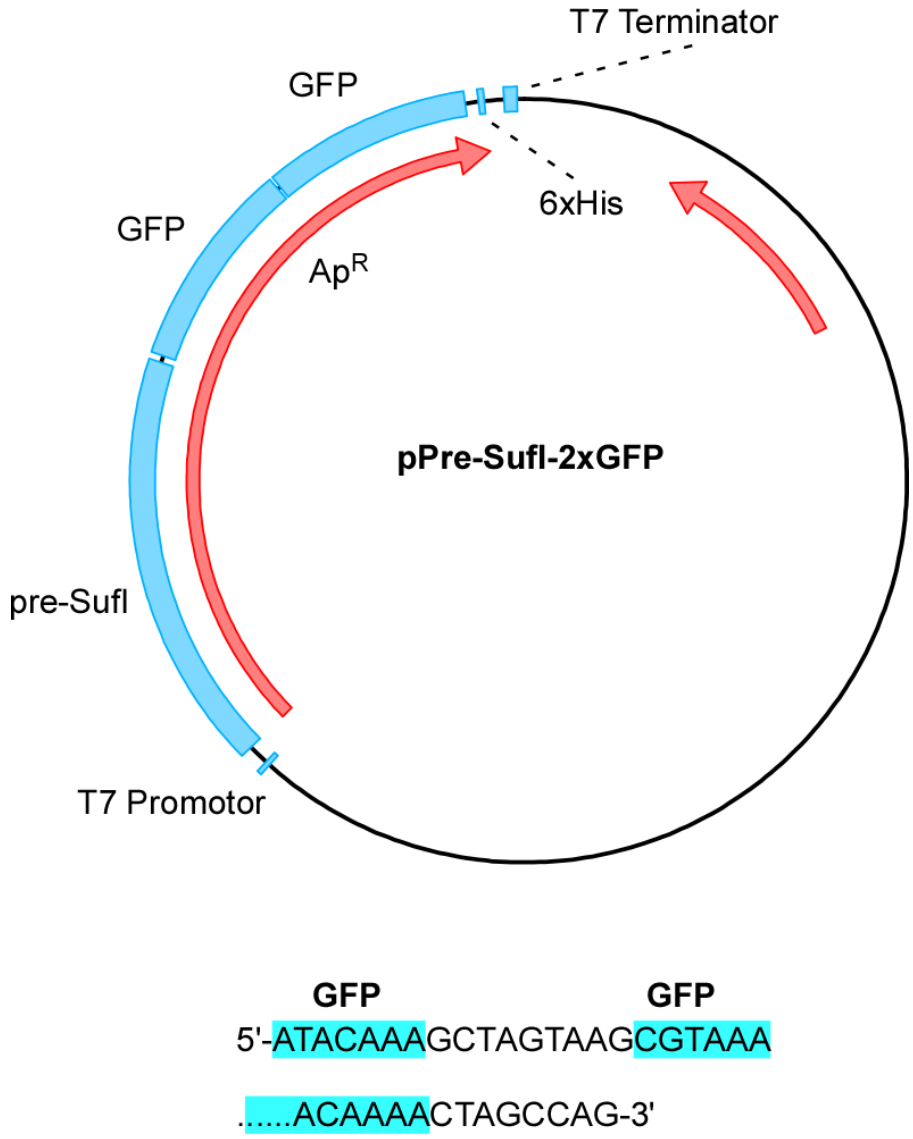


Figure A17. Plasmid map of pPre-SufI-2xGFP. Amplified GFP gene was inserted between the GFP and the 6xHis sequences at a NheI site by SpeI ligation, destroying the site and resulting a 9 bp pair linkager.

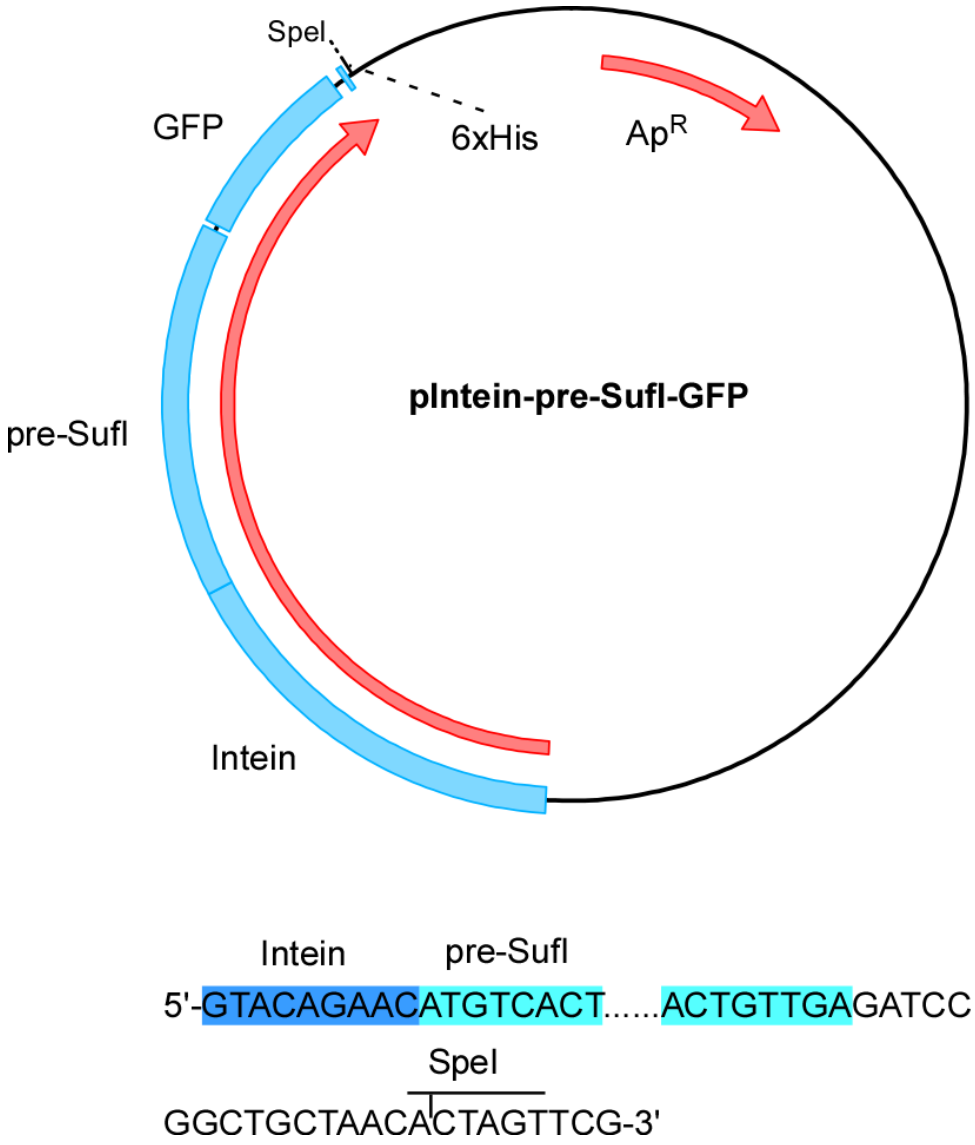


Figure A18. Plasmid map of pIntein-pre-SufI-GFP. The plasmid pIntein-pre-SufI-GFP was generated by inserting pre-SufI-GFP from plasmid pPre-SufI-GFP into the vector pTYB11 (New England Biolabs) at restriction enzyme sites SapI and SpeI. Since SapI does not cleave at the restriction enzyme recognition site, the recognition sequence was lost after insertion of the pre-SufI sequence.