Moving folded proteins across the bacterial cell membrane

Tracy Palmer¹,² and Ben C. Berks³

¹Centre for Metalloprotein Spectroscopy and Biology, School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ, UK
²Department of Molecular Microbiology, John Innes Centre, Colney Lane, Norwich NR4 7UH, UK
³Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK

The Tat protein export system is located in the bacterial cytoplasmic membrane and operates in parallel to the well-known Sec pathway. While the Sec system only transports unstructured substrates, the function of the Tat pathway is to translocate folded proteins. The Tat translocase thus faces the formidable challenge of moving structured macromolecular substrates across the bacterial cytoplasmic membrane without rendering the membrane freely permeable to protons and other ions. The substrates of the Tat pathway are often proteins that bind cofactor molecules in the cytoplasm, and are thus folded, prior to export. Such periplasmic cofactor-containing proteins are essential for most types of bacterial respiratory and photosynthetic energy metabolism. In addition, the Tat pathway is involved in outer membrane biosynthesis and in bacterial pathogenesis. Substrates are targeted to the Tat pathway by amino-terminal signal sequences harbouring consecutive, essentially invariant, arginine residues, and movement of proteins through the Tat system is energized by the transmembrane proton electrochemical gradient. The TatA protein probably forms the transport channel while the TatBC proteins act as a receptor complex that recognizes the signal peptide of the substrate protein.

Overview

Transport of proteins across biological membranes is one of the fundamental features of cellular life. Proteins that are destined for transmembrane transport are normally synthesized with amino-terminal signal sequences, termed signal peptides, which direct the protein to a specific transporter complex in the membrane. The signal sequence is subsequently cleaved from the passenger protein during the export event. In bacteria most proteins are exported across the cytoplasmic (inner) membrane by the well-characterized Sec apparatus (Pugsley, 1993; Manting & Driessen, 2000). In eukaryotes, protein export uses an homologous Sec system in the endoplasmic reticulum (Johnson & Haigh, 2000). A defining feature of both of these Sec systems is that the substrate proteins are threaded through the transporter, amino terminus first, and only fold upon reaching the far side of the membrane. Indeed, if transport is post-translational the substrate protein must be kept in an extended conformation to remain transport competent.

In bacteria, the generation of energy by respiratory or photosynthetic electron transfer chains takes place across the cytoplasmic membrane. In most cases a proportion of the redox-active components of these pathways are located at the periplasmic side of this membrane. Since the biologically useful redox chemistry of amino acids is limited, redox proteins normally contain a cofactor to carry out their catalytic or electron transfer functions. These cofactors include such things as nucleotides (NADP or FAD), simple metal ions (e.g. copper) and complex metal cofactors (e.g. haem or the molybdopterin cofactor). A consequence is that bacterial energy metabolism in most environments depends upon the bacterium being able to produce cofactor-containing proteins in the extracellular compartment. Biosynthesis of these proteins causes the cell a particular problem since, in most cases, cofactor binding requires protein folding and if the protein is folded it is no longer an acceptable substrate for the Sec system. For some types of cofactor this biosynthetic problem has been overcome by moving protein and cofactor separately to the periplasm before allowing cofactor insertion (Fig. 1). This is the system adopted, for example, in the biosynthesis of c-type cytochromes (Berks, 1996; Thony-Meyer, 2000). However, we and others have been able to show that for most types of cofactor-containing protein an alternative strategy is pursued in which the cofactor is inserted into the protein in the cytoplasm and the protein is then exported by a transporter that is completely distinct from the Sec apparatus (Berks, 1996; Settles et al., 1997;
Fig. 1. The two routes for the biosynthesis of cofactor-containing periplasmic proteins. For some classes of redox protein (left of figure) apoprotein and cofactor are transported to the periplasm separately. In this case the protein is transported by the Sec apparatus and cofactor insertion and protein folding occur in the periplasm. For other classes of redox protein (right of figure) cofactor insertion and protein folding take place in the cytoplasm. The assembled precursor protein is then transported to the periplasm by the Tat apparatus. The unfolded protein is represented by the grey zigzag line, with the signal peptide in black. The smaller circle represents a cofactor molecule.

Santini et al., 1998; Weiner et al., 1998; Sargent et al., 1998) (Fig. 1). Of necessity this transporter has to possess the remarkable property of moving folded proteins across a membrane. This pathway is normally termed the Tat system (though an Mtt designation has also been employed; Weiner et al., 1998). An analogous protein transport system is found in the thylakoid membrane of higher plant chloroplasts. This was formerly termed the APh-dependent protein transport system since it is energized exclusively by the transmembrane proton electrochemical gradient (Settles et al., 1997).

Here we describe the salient features of the bacterial Tat system. Comparison with the homologous thylakoid protein transport pathway is made where appropriate. Prior reviews include Berks (1996), Berks et al. (2000a, b), Vooroud (2000), Mori & Cline (2001), Yen et al. (2002) and Sargent et al. (2002).

The Tat pathway transports folded proteins

The defining feature of the Tat pathway is that it functions to transport folded proteins across the cytoplasmic membrane. The evidence supporting this assertion is now rather strong and originated with the observation that cofactor-containing Tat substrates acquire their cofactors, and therefore attain a folded conformation, in the cytoplasm (see Berks, 1996; Berks et al., 2000a for a detailed discussion). Indeed, it has been found that precursor proteins that accumulate in the cytoplasm of tat mutant strains contain cofactors (e.g. Sargent et al., 1998). The idea that the Tat system is capable of transporting folded proteins has been more directly tested by examining the transport behaviour of heterologous substrates of defined folding state. For example, it has been shown that folded green fluorescent protein (GFP) can be successfully exported via the Tat pathway if fused to a Tat signal peptide (Santini et al., 2001; Thomas et al., 2001). Another study employed a fusion between a Tat signal peptide and a c-type cytochrome (Sanders et al., 2001). Using a specially engineered strain it was possible to get haem inserted into the cytochrome in the cytoplasm rather than the periplasm. It was then shown that the folded, cofactor-containing cytochrome c could be exported by the Tat pathway. Transport of deliberately folded passenger proteins has also been demonstrated for the thylakoid Tat system (Clark & Theg, 1997; Hynds et al., 1998).

While it is now well established that the Tat system is capable of transporting folded proteins, a major open question is whether the Tat pathway actively discriminates against substrates that are not folded. Such a proofreading feature would be useful in preventing the interaction of immature or incorrectly assembled proteins with the Tat transport system. There is certainly reasonable evidence that cofactor-containing proteins are only exported once cofactor is bound. For example, strains of Escherichia coli that are blocked in the synthesis of the molybdopterin cofactor produce, but fail to export, periplasmic molybdoenzymes (Santini et al., 1998). Similarly, mutants of the Zymomonas mobilis glucose-fructose oxidoreductase engineered to weaken binding of its NADP$^+$ cofactor have dramatically slower rates of export than the wild-type protein (Halbig et al., 1999). What is less clear is whether unfolded proteins per se are incompatible with Tat transport. Certain data (Roffey & Theg, 1996; Halbig et al., 1999; Sanders et al., 2001; Stanley et al., 2002) can be interpreted as indicating selective transport of folded substrates by the Tat system but arguments to the contrary have also been advanced (Hynds et al., 1998).

A clue as to how the cell co-ordinates cofactor insertion and export events in the biosynthesis of a Tat substrate is provided by the recent work of Oresnik et al. (2001). These authors identified a protein that was specifically able to interact with the Tat signal peptide of the molybdopterin cofactor-binding DmsA subunit of the E. coli enzyme dimethylsulfoxide (DMSO) reductase. This Tat-signal-peptide-binding protein, designated DmsD, shows sequence similarity to the product of the torD gene involved in the biosynthesis of another Tat-dependent molybdoenzyme, trimethylamine N-oxide (TMAO) reductase, TorD is a cytoplasmically located chaperone protein that specifically interacts with the unfolded form of TMAO reductase prior to molybdenum cofactor insertion (Pommier et al., 1998). Taken together, these observations point to a mechanism whereby the signal peptide and unfolded mature region of the apo-molybdoenzyme are simultaneously bound by a specific chaperone protein which serves both to maintain the apoenzyme in a form competent for cofactor insertion and to shelter the signal peptide to prevent interaction with the Tat transporter (Sargent et al., 2002). Presumably once
the cofactor has bound to the enzyme, the chaperone is displaced, revealing the signal peptide and thus allowing export. It is not clear whether such a mechanism is also operational for other Tat substrate proteins. However, it is interesting to note that the operons encoding many redox enzymes exported by the Tat pathway contain accessory genes which might encode ‘proofreading’ chaperones.

A proofreading mechanism would be particularly pertinent to Tat substrate proteins that are exported as hetero-oligomers, where the signal sequence resides on just one of the subunits. The best-studied examples of such proteins are bacterial uptake hydrogenases. These minimally comprise an iron–sulfur cluster-containing small subunit and a large subunit which harbours the nickel–iron active site. A Tat signal peptide is found only on the small subunit but this directs export of both the large and small subunit to the periplasmic side of the membrane. Export does not occur until the small subunit has bound to the large subunit and the maturation of the large subunit [Ni–Fe] cofactor is complete. Moreover, in the absence of the large subunit, the small subunit accumulates in the cytoplasm in its precursor form (Rodri`que et al., 1999).

**Substrates of the Tat pathway**

The vast majority of currently identified Tat pathway substrates are proteins containing certain types of redox-active cofactor. Cofactor classes associated with Tat transport include those based on metal–sulfur clusters (including iron–sulfur clusters) or on nucleotides (including molybdopterin cofactors) (for a comprehensive analysis see Berks, 1996). It is becoming increasingly apparent, however, that some proteins without cofactors are also transported by the Tat pathway. One example of a Tat-dependent process that does not involve redox cofactors is phosphate acquisition by *Bacillus subtilis*. This utilizes the Tat-targeted PhoD phosphatase (Jongbloed et al., 2000). Another example is biosynthesis of the cell wall in *E. coli*, where two of the three amidase isoenzymes that cleave the division septum are substrates of the Tat pathway (Stanley et al., 2001; B. Ize, N. R. Stanley, G. Buchanan & T. Palmer, unpublished). *Pseudomonas aerugi-nos*a provides further examples of cofactor-independent Tat-linked processes. In this organism, the Tat system is required for the correct localization of proteins involved in iron siderophore biosynthesis and uptake (Ochsner et al., 2002). It also catalyses the first step in the export of phospholipases (Vou lhoux et al., 2001). Since these phospholipases are involved in the breakdown of eukaryotic cell membranes, and since iron acquisition is usually a pathogenicity factor in *P. aeruginosa* micro-organisms, it not surprising that the *tat* genes are upregulated in biofilms (Whiteley et al., 2001).

Heterologous proteins can be successfully targeted to the Tat pathway when fused to a Tat signal peptide. Examples include the normally Sec-dependent proteins β-lactamase and periplasmic P2 domain of signal peptidase as well as colicin V which would usually be transported to the periplasm by an ABC family transporter (Nivi``ere et al., 1992; Stanley et al., 2002; Cristóbal et al., 1999; Ize et al., 2002). Other examples are the cytoplasmic protein chloramphenicol acetyl transferase (Stanley et al., 2002) and, as discussed above, GFP (Thomas et al., 2001; Santini et al., 2001). These observations suggest that the Tat system may be a promising route for the export of heterologous proteins, particularly those which have proven refractory for secretion by the Sec system. However, one should be aware that not all proteins are competent for export by the Tat pathway. Both β-galactosidase and alkaline phosphatase cannot be exported by the *E. coli* Tat machinery indicating that the nature of the substrate protein has an important influence on its ability to be transported by the Tat system (Halbig et al., 1999; Stanley et al., 2002).

**The Tat signal peptide**

Proteins are specifically sorted to the Tat pathway by means of a signal peptide at the amino terminus of the precursor protein. These signal peptides in some ways resemble Sec signal peptides in that both have a tripartite structure in which a basic amino-terminal n-region is followed successively by a hydrophobic h-region and then a hydrophilic c-region containing the recognition site for the enzyme signal peptidase. However, in several critical respects Tat signal peptides differ from their Sec counterparts. Firstly, Tat signal peptides contain a conserved amino acid sequence motif at the n-region/h-region boundary. This motif can be defined as S-R-R-x-F-L-K, where the consecutive arginine residues are almost invariant, the frequency of the other motif residues is greater than 50% and x is normally a polar amino acid (Berks, 1996; Stanley et al., 2000). The conservation of the arginine pair is reflected in the designations ‘twin arginine’ for the signal peptides and Tat (twin arginine translocation) for the protein translocation pathway. A second distinguishing feature is that the h-regions of Tat signal peptides are significantly more hydrophobic than those of Sec targeting signals (Cristóbal et al., 1999). Thirdly, Tat signal peptides frequently contain basic residues in the c-region whereas such amino acids are almost never found at the equivalent position of Sec signal peptides (Wexler et al., 1998). A fourth feature is that Tat signal peptides are markedly longer than their Sec counterparts, in some cases reaching up to 58 aa in length (Berks, 1996). Much of this extra length can be attributed to a highly extended n-region (Cristóbal et al., 1999). Some examples of Tat targeting signals are shown in Fig. 2.

A number of site-directed mutagenesis experiments using natural substrates of the Tat pathway have confirmed that the consecutive arginine residues of the consensus motif play a key role in the export process (Dreusch et al., 1997; Gross et al., 1999; Halbig et al., 1999). However, there is growing evidence that Tat translocation does not always require the presence of an arginine pair. Substitution of either the first or
the second arginine residues of the signal peptide of *E. coli* pre-SufI protein with lysine (but not both) still allowed slow but significant export (Stanley et al., 2000) though an analogous alanine substitution did not. Further, while conservative replacement of the second arginine residue in the Tat signal peptide of *E. coli* TMAO reductase (TorA) abolished export, substitution of the first arginine with a lysine could be comfortably tolerated (Buchanan et al., 2001). Intriguingly, a naturally occurring Tat signal peptide has been identified that also has a lysine for arginine substitution at the first consensus position (Hinsley et al., 2001). Reversal of the arginine-lysine pairing in this signal peptide significantly slowed the rate of Tat-dependent export (Hinsley et al., 2001). Taken together these results indicate that in at least some signal peptide contexts it is possible to retain Tat transport function upon substitution of one (and especially the first) of the consensus arginine residues by lysine. More remarkably, a penicillin amidase present in some strains of *E. coli* was shown to be exported by the Tat pathway despite having a signal peptide without an obvious twin-arginine-like motif (Ignatova et al., 2002; Fig. 2). This suggests that our current understanding of the parameters for Tat targeting are insufficient to correctly identify all Tat substrates merely on the basis of sequence analysis.

The role and importance of the non-arginine residues of the consensus motif are less clear but have been experimentally addressed by Stanley et al. (2000) using the precursor proteins pre-SufI and pre-YacK as model substrates. Twin-arginine residues apart, the consensus phenylalanine is the most highly conserved amino acid in the Tat motif. Conservative substitution of this residue with leucine was without effect, but replacement with alanine or tyrosine led to a dramatic decrease in the rate of export, suggesting that the hydrophobicity of the amino acid at this position is important for the export process. The consensus leucine residue also contributes to efficient export since substitution with a non-hydrophobic side chain resulted in a small decrease in the export rate. Perhaps the most surprising observation was that substitution of the conserved lysine actually increased the rate of export. This has led to speculation that the role of the lysine may be to mediate interactions with accessory proteins necessary to ensure that cofactor insertion has been achieved prior to export (Stanley et al., 2000).

The signal peptide sequence requirements for Tat-dependent transport have recently been reassessed using a genetic selection to isolate signal peptide mutants with enhanced transport efficiency (DeLisa et al., 2002). These experiments made use of a chimera comprising GFP tagged both with the Tat signal peptide of the TorA protein and a marker directing the protein to a cytoplasmically located degradation pathway. This construct allows the efficiency of Tat transport to be measured as whole cell fluorescence since the fluorescence intensity is proportional to the amount of GFP that is transported out of the cytoplasm before it is digested. By means of a FACS scanner, large numbers of individual cells can be screened for changes in cellular fluorescence. Using this approach it was shown that a basic residue was necessary for Tat-dependent transport at the first arginine position of the Tat consensus motif in the TorA signal peptide but that any of a basic residue, asparagine or glutamine could be tolerated at the second arginine position. These observations may account for the Tat-dependent targeting seen with the penicillin amidase signal peptide, which plausibly has an arginine-glutamine motif at the two consensus arginine positions (Fig. 2; Ignatova et al., 2002). In addition to mutations at the consensus arginine residues, the fluorescence screen for enhanced Tat transport selected a number of substitutions in other amino acids in the consensus motif but did not isolate mutations elsewhere in the signal peptide. These observations reinforce the significance of the entire consensus motif region for interactions with the Tat export machinery.

The functional importance of the difference in hydrophobicity between Sec and Tat signal sequences has been addressed using a fusion between the P2 domain of signal peptidase, which is normally a substrate of the Sec pathway, and the Tat signal peptide of TorA (Cristóbal et al., 1999). Employing this fusion it was shown that the P2 domain can be successfully targeted through the Tat pathway. However, increasing the hydrophobicity of the signal peptide h-region in the chimeric precursor resulted in a rerouting of export from the Tat to the Sec pathway, even though the twin-arginine motif remained intact. Although the Sec-dependent export of P2 was at least an order of magnitude slower than that of authentic Sec substrates, subsequent substitution of the TorA signal peptide c-region positive charges resulted in very rapid Sec-dependent export (Cristóbal et al., 1999). Taken together with the observation that mutagenesis of the positive charge of the c-region of pre-SufI does not affect the export rate, this suggests that the positive charge acts as a ‘Sec-avoidance’ signal (Bogsch et al., 1997; Cristóbal et al., 1999; Stanley et al., 2000). While the TorA signal peptide is
able to direct export of the soluble P2 domain of signal peptidase, it is not able to target the full-length signal peptidase, which is an integral membrane protein, onto the Tat pathway (Cristóbal et al., 1999). This indicates that targeting information present in transmembrane helices overrides the Tat targeting properties of the TorA signal.

It is important to note that the choice of passenger protein can have a profound bearing on the transport phenotype of amino acid changes in Tat signal peptides. For example, conservative substitution of the twin-arginine residues of the TorA signal peptide with a lysine pair blocks transport when the passenger protein is GFP but not when it is colicin V (Ize et al., 2002). There is also some evidence that the Tat machinery from different bacteria may show specificity towards cognate signal peptides. For example, although the glucose–fructose oxidoreductase (GFOR) of Z. mobilis is a Tat substrate in its native organism, it is not exported when heterologously expressed in E. coli (Blaudeck et al., 2001). However, when the native GFOR signal sequence is precisely replaced by that of E. coli TorA, the hybrid protein is exported by E. coli in a Tat-dependent manner (Blaudeck et al., 2001). This result suggests that Tat signal peptides are not universally recognized by different Tat translocases. A similar observation was made by Pop et al. (2002). Here the authors demonstrated that the Tat signal peptide of the B. subtilis PhoD protein was not recognized by the E. coli Tat machinery, but could interact with heterologously expressed B. subtilis Tat components.

Strikingly, the signal peptides for proteins containing the same cofactor from different bacteria exhibit notable sequence conservation in addition to the twin-arginine motif. This marked sequence conservation often includes a highly extended n-region (for an example see the hydrogenase signal sequences aligned in Berks et al., 2000a). These regions of sequence conservation may mediate interactions with chaperones required for cofactor insertion and/or oligomerization (above). Such a scenario could provide an explanation for the inability of E. coli to recognize the extended Z. mobilis GFOR Tat signal sequence (above). That the signal sequence influences assembly of its passenger protein may also explain the observation that the replacement of the Tat signal sequence of the DmsA catalytic subunit of E. coli DMSO reductase with that of the E. coli TorA protein results in a markedly decreased level of membrane-bound DMSO reductase activity (Sambasivarao et al., 2000).

**Components of the Tat pathway**

Two genetic loci encoding components of the Tat machinery have been identified in E. coli (Bogsch et al., 1998; Sargent et al., 1998; Weiner et al., 1998). The tatA operon, which maps to 86 minutes on the E. coli chromosome, encodes the four genes tatABCtD while tatE, found at 14 minutes, is monocistronic (Fig. 3). Both transcription units are expressed constitutively in E. coli, indicating a requirement for the Tat pathway under all growth conditions (Jack et al., 2001).

**Fig. 3.** Chromosomal organization of genes encoding Tat components in E. coli. ORFs shown in grey have been demonstrated to participate in Tat-dependent protein transport. The termination symbol between tatC and tatD indicates the presence of a putative hairpin structure. The arrows underneath the tatA operon show the major transcripts identified by Northern blotting (Wexler et al., 2000).

TatA, TatB and TatE are sequence-related proteins with a common structure. As outlined in Fig. 4, each is predicted to comprise a membrane-spanning α-helix at the amino terminus, immediately followed by an amphipathic helix located at the cytoplasmic side of the membrane and a carboxy-terminal region of variable length. The homology amongst the TatA/B/E family of proteins is low. The only absolutely conserved residue is a glycine, which apparently forms an inter-domain hinge at the transmembrane helix/amphipathic helix boundary. Deletion analysis shows that the carboxy-terminal portions of TatA and TatB following the predicted amphipathic helix are not essential for transport by the Tat pathway (Lee et al., 2002). Several features of the TatA structural model have recently been experimentally tested. TatA has been confirmed to be a fully

**Fig. 4.** Predicted topological organization of the E. coli Tat components. The amino terminus of each polypeptide is marked with an ‘N’. Note that a recent study indicates that the fourth and fifth predicted transmembrane helices of TatC shown in the figure instead form a periplasmic domain (Gouffi et al., 2002).
integral membrane protein (de Leeuw et al., 2001) and to be accessible to protease digestion at the cytoplasmic, but not periplasmic, side of the membrane (Porcelli et al., 2002). The secondary structure content of purified TatA measured by circular dichroism spectroscopy is consistent with the presence of the predicted transmembrane and amphipathic helices with the remainder of the molecule being essentially unstructured (Porcelli et al., 2002). The role of the proposed amino-terminal transmembrane helix has been investigated by characterizing a truncated form of TatA in which this helix was deleted. These experiments show that the helix is required for membrane integration in vivo and for formation of secondary structure in the remainder of the molecule (de Leeuw et al., 2001; Porcelli et al., 2002). However, it was also shown that the truncated TatA molecule interacts with phospholipids in vitro and that this leads to α-helix formation (Porcelli et al., 2002). It is, therefore, reasonable to speculate that the predicted amphipathic helix of TatA forms when this region of the molecule is forced to the membrane by the adjacent transmembrane helix.

TatA and TatE display greater than 50% sequence identity and share overlapping functions in Tat-dependent protein export. Thus single deletions in tatA or tatE lead to a decrease in the export of a range of Tat substrate proteins, but deletion of both is required to observe a complete export block (Sargent et al., 1998). Generally the tatA mutant shows a more severe export defect than the tatE strain. This is consistent with genetic studies which indicate that tatA is transcribed and translated at a 50- to 100-fold higher level than tatE (Jack et al., 2001).

Despite sharing 20% sequence identity with TatA/TatE, TatB has a distinct function in the export pathway. An in-frame deletion in tatB alone is sufficient to abolish the export of endogenous Tat substrates although some Tat-dependent export of a TpA signal peptide–colin V chimera was still observed in a tatB deletion strain (Ize et al., 2002). tatB mutant strains cannot be complemented in trans by supplying extra copies of tatA or tatE, and likewise the tatA/tatE mutant strain cannot be rescued by plasmid-borne tatB (Sargent et al., 1999). However, the precise interrelationship between TatA/E and TatB proteins is not clear since it appears that some bacteria require, in addition to TatC, just one copy of a TatA/B/E-like protein for a functional Tat system (Berks et al., 2000a; Wu et al., 2000; Pop et al., 2002). Moreover, a hybrid protein comprising the amino-terminal transmembrane domain of TatA fused to the amphipathic and carboxy-terminal domains of TatB retains some properties of TatA protein function and some of TatB (Lee et al., 2002).

The TatC protein is highly hydrophobic and is predicted to have six transmembrane helices, with the amino and carboxy termini located at the cytoplasmic face of the membrane (Sargent et al., 1998; Drew et al., 2002). A recent study of the topology of TatC using compartment-sensitive marker fusions confirms the predicted model with the exception that only four transmembrane helices are detected (Gouffi et al., 2002). TatC is strictly required for protein export by the Tat pathway (Bogsch et al., 1998). Of all of the components of the Tat pathway it is TatC that shows the highest level of amino acid conservation. Twenty-one amino acids, several of which are polar residues, are strictly conserved amongst the eu-bacterial TatC proteins and seven of these are also conserved amongst the eukaryotic homologues (Buchanan et al., 2002; Allen et al., 2002). The majority of these conserved residues fall within predicted cytoplasmic loops of the protein. Recent site-directed mutagenesis experiments have confirmed an essential role for some of these residues in the operation of the Tat pathway (Allen et al., 2002; Buchanan et al., 2002).

The tatABC genes in E. coli and other closely related enteric bacteria are co-transcribed with a fourth gene, tatD, encoding a soluble cytoplasmic protein with nuclease activity (Wexler et al., 2000). Analysis of an E. coli strain carrying in-frame deletions in tatD and two further homologues, ycfH and yjJV, indicates that TatD family proteins play no obligate role in Tat-dependent protein export. Furthermore, Northern blot analysis demonstrates that although tatD is transcribed with tatABC, the presence of a putative stem–loop structure in the tatC–tatD intergenic region serves to greatly depress the level of TatD synthesis (Wexler et al., 2000; Jack et al., 2001; Fig. 3).

In summary, the genetic data indicate that the E. coli Tat pathway is comprised minimally of the three proteins TatA (or TatE), TatB and TatC. Phylogenetic analysis indicates that the structural genes for these proteins very frequently show genetic linkage but that they do not cluster with any further genes, suggesting that they probably form the core components of the Tat export system. Consistent with this proposal, overexpression of only the tatABC genes results in a marked increase in the in vivo and in vitro flux of SulF protein through the Tat pathway (Sargent et al., 2001; Yahr & Wickner, 2001; Alami et al., 2002).

Studies using chromosomal translational fusions indicate that E. coli produces TatA/TatB/TatC at a stoichiometry of 25:1:0.5 (Jack et al., 2001). This ratio is in reasonable agreement with the TatA/TatB ratio of around 20:1 determined in wild-type E. coli cells by quantitative immunoblotting (Sargent et al., 2001). Additional studies have shown that a TatB–TatC fusion protein is functional, suggesting that the two constituent Tat subunits are present at an equimolar ratio in wild-type cells (Bolhuis et al., 2001). Taken together, these data indicate that TatA is present at a high molar excess over the other components of the Tat pathway.

Two major complexes of high molecular mass containing varying ratios of the Tat proteins have been isolated from membranes of E. coli overproducing the known Tat components. One complex contains TatB and TatC together with a small proportion of the TatA present in the membrane (Bolhuis et al., 2001; de Leeuw et al., 2002). The complex has a molecular mass of approximately 600 kDa in detergent solution and appears to contain an equimolar
ratio of the three Tat proteins (Bolhuis et al., 2001). This complex specifically interacts with Tat signal peptides (de Leeuw et al., 2002). The major part of the TatA protein present in the membranes forms a separate high (approx. 600 kDa) molecular mass complex containing a very small proportion of TatB (Sargent et al., 2001; de Leeuw et al., 2002). The two types of purified bacterial Tat complex closely correspond to two Tat complexes identified in the plant thylakoid membrane by blue native PAGE analysis (Cline & Mori, 2001). One complex contains cpTatC together with Hcf106 protein, the chloroplast orthologue of bacterial TatB. The chloroplast TatA orthologue, Tha4, forms a separate entity. All of these bacterial and chloroplast Tat complexes are very considerably larger than the size of the individual Tat proteins, suggesting that Tat proteins are always present as high-order oligomers. Indeed, data from chemical cross-linking studies show that TatA and TatB have minimally tetrameric and dimeric (respectively) homo-oligomeric interactions (de Leeuw et al., 2001). For TatA it has been demonstrated that the predicted amino-terminal transmembrane helix is required for homooligomer formation (Porcelli et al., 2002). Recent genetic evidence indicates that TatC is also at least a functional dimer since single point mutations that inactivate TatC were able to restore Tat transport function when co-expressed on separate replicons (Buchanan et al., 2002).

**Functions of Tat components**

In contrast to other protein transport systems, protein translocation on the Tat pathway is energized exclusively by the transmembrane proton electrochemical gradient with no involvement of nucleotide hydrolysis (Mould & Robinson, 1991; Yahr & Wickner, 2001; Alami et al., 2002). Initial precursor binding to the Tat system is, however, independent of the protonmotive force (Ma & Cline, 2000; Musser & Theg, 2000; Alami et al., 2002; Cline & Mori, 2001). Studies with the thylakoid Tat system indicate that precursor proteins are initially recognized by a complex containing TatBC homologues (Cline & Mori, 2001). It has been argued that only TatC contains sufficient conserved amino acids to form a specific binding site for the twin-arginine motif of Tat signal peptides (Berks et al., 2000a). Further support for this contention comes from the study of Tat-dependent protein export in the Gram-positive bacterium *B. subtilis*. This organism contains two copies of a *tatC*-like gene, each of which is linked to a gene encoding a TatA/B/E homologue. However, export of the Tat pathway substrate PhoD only requires one of these *tatC* homologues and is independent of the second copy. These observations suggest that TatC acts as a specificity determinant for PhoD export, possibly at the level of signal peptide recognition (Jongbloed et al., 2000).

Tat substrates are presumably transported across the cytoplasmic membrane through a gated aqueous channel. A number of lines of circumstantial evidence point towards TatA playing the major role in forming this channel. Inspection of the structures of Tat substrates suggests that the channel must open wide enough to accommodate substrates having diameters of up to 70 Å (Berks et al., 2000a; Sargent et al., 2002). The number of α-helices required to enclose an aqueous channel of this diameter is greater than 20 (Berks et al., 2000a). Given the relatively small sizes of the Tat proteins, this would be consistent with a model where multiple copies of (at least) one of the Tat proteins would be required to form the channel. Since TatA is present in large molar excess over the other Tat components and is apparently not involved in signal peptide binding, this protein is an obvious candidate for the channel-forming component. Interestingly, negative-stain electron microscopy of an isolated TatAB complex, which contained TatA in an approximate 20-fold molar excess over TatB (Sargent et al., 2001), resulted in the visualization of annular structures with a central cavity of 70 Å in diameter, as shown in Fig. 5. Thus a complex containing predominantly TatA is forming a structure that has some of the characteristics expected of the transport pore. In thylakoids, antibodies to the TatA orthologue Tha4 prevent protein transport but do not block specific association of the precursor protein with the membrane (Cline & Mori, 2001). This suggests that TatA is required at a later stage of the transport cycle than signal peptide binding, and would be consistent with a role for TatA in the protein translocation step itself.

Mori & Cline (2002) have recently argued, on the basis of chemical cross-linking experiments, that thylakoid Tha4 protein only associates with the Hcf106–TatC receptor complex after the latter has bound precursor protein and that this association of Tat components requires the protonmotive force. The association is transient and disassembly correlates with completion of protein translocation. These observations point to a mechanism where assembly of the translocase might arise from the merging of two pre-existing subcomplexes and would again be consistent with the idea that TatA forms the transport channel. de Leeuw et al. (2002) recently reported the purification, albeit at very low yield, of an *E. coli* TatABC complex containing a high molar excess of the TatA component. It is possible that this corresponds to the proposed fully assembled Tat translocation site.

![Fig. 5. Averaged projection structures of negatively stained TatAB complex from *E. coli* solubilized and purified in the detergent CHAPS. Approximate dimensions are indicated. Reproduced with permission from Sargent et al. (2001).](http://mic.sgmjournals.org)
Future directions

We hope that this account has indicated to the reader the importance and intriguing mechanism of the Tat system. It should also have revealed that we are only at a beginning in understanding the operation of the Tat pathway at the molecular level. To conclude this review we will now briefly consider those aspects of the Tat system that we see as being of particular interest to address in the coming years.

Firstly, it is our view that it is highly unlikely that the mechanism of the Tat system will be elucidated without high resolution structures of the Tat components and we, at least, are expending a large amount of effort towards this goal. A more immediate structural question is whether there is a dynamic assembly and disassembly of the Tat translocation site. Data from the thylakoid system suggests a reversible association of TatA with the receptor TatBC complex (above). Can this be substantiated? If this is verified, is the free TatA present as a single type of pre-existing pore complex or do small TatA units assemble around the substrate to form the translocation pore? It will be necessary to define more precisely the signal peptide binding site on the TatBC complex. It is quite likely that there is more than one mode of binding since the conformation of a signal peptide that is compatible with a folded substrate protein at the cytoplasmic side of the membrane may no longer be suitable when the substrate protein (and the signal peptidase cleavage site of the signal peptide) has reached the opposite side of the membrane.

The establishment of an in vitro transport assay for the bacterial Tat system (Yahr & Wickner, 2001; Alami et al., 2002) together with the development of powerful genetic screens for Tat function (de Lisa et al., 2002; Ize et al., 2002) should open up mechanistic analysis of the bacterial Tat pathway. We need to establish how and when the proton-motive force is transduced to allow protein movement. In both bacteria and thylakoids the Tat system is located in an ionically tight coupling membrane. This means that protein transport must take place without allowing significant cotransport of ions including protons. This would appear to be a formidable challenge since the substrates to be transported are very large (up to 70 Å; above), and of highly variable size, shape and surface features (Berks et al., 2000a). Understanding the mechanism of sealing and gating will thus be of great interest. While the Tat system is capable of transporting folded proteins it also appears (though more work is needed here) that it actively discriminates against unfolded substrates. We do not understand how this proofreading occurs. One possible starting point is that, as described above, many cofactor-containing substrates appear to utilize dedicated assembly chaperones; some of these may be involved in mediating the interaction of the substrate with the pathway.

In addition to its role in protein transport, the Tat system is also involved in the biogenesis of membrane proteins. In E. coli, for example, a quarter of Tat substrates contain transmembrane domains (Sargent et al., 2002). How does the Tat system move these domains, not to the periplasm, but into the membrane bilayer?

Finally, applied aspects of the Tat pathway will become increasingly important. The involvement of the Tat system in bacterial pathogenesis should receive a lot of attention. Indeed, given that the Tat apparatus is well conserved among important bacterial pathogens but is absent from mammalian cells, the Tat system may represent an important new target for novel antimicrobial compounds. In addition it is likely that the unique operating features of the Tat system will be utilized for biotechnological purposes, most obviously for the attempted export of proteins that are incompatible with the Sec pathway.

Acknowledgements

We thank members of our laboratories, past and present for their contribution to our research on protein transport, and to our many colleagues with whom we have discussed the Tat system. Fig. 4 was provided by Dr Frank Sargent. Research in the authors’ laboratories has been or is currently supported by the Biotechnology and Biological Sciences Research Council, the Commission of the European Community, the Leverhulme Trust, the Wellcome Trust, the John Innes Centre, the University of East Anglia, the University of Oxford and The Royal Society. Tracy Palmer is a Royal Society University Research Fellow. Ben Berks is RJP Williams Senior Research Fellow at Wadham College, Oxford.

References


