Chapter 16
Mechanistic Challenges and Engineering Applications of Protein Export in *E. coli*

Eva-Maria Strauch and George Georgiou

Contents

Abstract Protein secretion and subcellular localization in *E. coli* has been under investigation for more than 60 years. While many details about the molecular mechanisms of these processes have been revealed, several facets of protein translocation still remain unclear. Bacteria secrete numerous proteins such as pathogenicity factors, toxins or degradative enzymes (Fernandez and Berenguer 2000). Six different secretion mechanisms for extruding proteins into the extracellular environment have been identified to-date. In Gram-negative bacteria such as *E. coli*, secretion into the extracellular medium requires crossing of two biological membranes, the inner and outer membranes of the cell. However, systems for protein translocation into the extracellular medium are generally highly protein-specific and with very few exceptions have not yet been engineered for the efficient export of recombinant proteins. More relevant from a technical and engineering standpoint, is the translocation of polypeptides from the cytoplasm into the periplasmic space, the main secretory compartment which is equivalent to the endoplasmic reticulum of eukaryotic cells.

G. Georgiou (✉)
Department of Chemical Engineering, Biomedical Engineering, Molecular Genetics and Microbiology and the Institute for Cellular and Molecular Biology, University of Texas at Austin, Austin, TX, USA
e-mail: gg@che.utexas.edu

S.Y. Lee (ed.), *Systems Biology and Biotechnology of Escherichia coli*,
DOI 10.1007/978-1-4020-9394-4_16,

In the first part of this chapter, we discuss export via the general Sec pathway and the Twin-Arginine Translocase (Tat) pathway. Compartmentalized molecular chaperones facilitate folding, impose a quality control step on the maturation of certain secreted proteins, especially those exported via Tat, and further facilitate the decision which protein export route should be chosen. The second part of this chapter focuses on the design of genetic screens or selections that capitalize on protein secretion to aid the screening of libraries of protein variants for molecular recognition or catalysis. We will briefly summarize the major *E. coli*-based display technologies and introduce new methodologies particularly those utilizing the Twin-Arginine Translocase pathway.

16.1 Protein Transport in *E. coli*

Proteins are charged, bulky heteropolymers of which transport across, or insertion into, the low dielectric barrier of a lipid bilayer membrane is thermodynamically highly unfavorable. Hence, there are several different transport pathways that expend metabolic energy to overcome this physical barrier. The Sec protein translocase utilizes energy mainly generated by the hydrolysis of nucleoside triphosphates. However, there are several different transport processes across biological membranes which solely rely on ion gradients.

In gamma-proteobacteria, the main route for protein transport across the cytoplasmic membrane is through the Sec translocon, a set of transmembrane proteins which form a hydrophilic channel. The SecYEG translocon is the bacterial homologue of the Sec61$\alpha\beta\gamma$ in eukaryotic cells. Sec protein translocation can be summarized in a basic set of rules (Schatz and Dobberstein 1996): A precursor protein containing a targeting sequence which is typically N-terminal, is maintained in an unfolded state prior to export. This unfolded, export-competent state of the precursor protein is either achieved by cytoplasmic chaperones (post-translational) or by an immediate association of the protein-synthesizing ribosome with a receptor in the membrane (co-translational export). For Sec substrates, folding and cofactor assembly occur in the periplasm. Folding is assisted by periplasmic chaperones such as DegP which functions as either protease or chaperone depending on the growth temperature, Skp that binds to non-native forms of periplasmic or outer membrane proteins preventing their aggregation, and four peptidyl-proline *cis-trans* isomerases, PpiA, PpiD, SurA and FkpA. The periplasm is an oxidizing environment mainly due to the action of DsbA which oxidizes free cysteines, whereas the isomerase DsbC rearranges disulfide bridges to their native conformation (Georgiou and Segatori 2005).

In contrast, proteins exported via the Tat pathway first fold within the cytoplasm (DeLisa et al. 2003). The cytoplasmic folding environment contains the general chaperones GroEL, DnaK/DnaJ/GrpE, the trigger factor, ClpB and the small heat-shock chaperones (IbpAB) among others. DnaK and possibly other chaperons play a role in the folding of some Tat substrate proteins prior to export (Graubner et al.

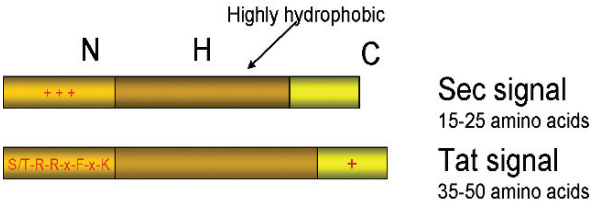

Fig. 16.1 Signal peptides. Comparison between a Sec and Tat signal peptides. N, H and C region are indicated. The positive charge within the C-region of a Tat signal peptide marks the Sec avoidance signal

2007, Perez-Rodriguez et al. 2007). Further, for the incorporation of metal cofactors into Tat substrates, complex protein maturation pathways involving several maturation enzymes are necessary. For example, the incorporation of iron-sulfur cluster requires at least 8 proteins (Tokumoto et al. 2002).

N-terminal signal peptides are evolutionarily well conserved. Signal peptides have three distinct regions (Fig. 16.1). The N-region harbors a positive charge, whereas the hydrophobic H-region comprises the center and the longest part of the signal peptide. The H-region of Tat signal peptides has a less hydrophobic character and is typically longer than the one of Sec signal peptides. The N-terminal positively charged region of Tat signal peptides contains the hallmark twin-arginine amino acids with the consensus sequence S/T-R-R-x-F-x-K. The C-region of both Sec and Tat signal peptides bears the signal peptidase cleavage site which is recognized by the type I signal peptidase. However, signal peptides of lipoproteins in *E. coli*, which so far have been only found to be exported by the Sec pathway, are cleaved by the type II signal peptidase (Paetzel et al. 2002). In general, Tat signal peptides typically contain a lysine or arginine residue within the C-terminal region which serves as Sec avoidance signal (Blaudeck et al. 2003). The more positively charged the C-region together with the beginning part of the mature protein is, the lower the likelihood that the precursor protein will be targeted to the Sec pathway (Tullman-Ercek et al. 2007).

16.1.1 The General Pathway for Secretion: The Sec Pathway

The Sec apparatus transports substrate polypeptides in an unfolded state through a narrow pore of about 5–8 Å minimally established by the SecY/E/G membrane proteins. Translocation can occur in two different ways which have been referred to as the co-translational and the post-translational export modes. Co-translational export involves the signal recognition particle (SRP) which is composed of the protein Ffh (fifty four homologue, based on its similarity to the eukaryotic SRP version in the endoplasmatic reticulum) and the 4.5S RNA unit (Luirink et al. 1992). The co-translational mode is often also referred to as the SRP pathway and ubiquitously found in all three kingdoms of life. The SRP complex binds to either a signal peptide or to a highly hydrophobic peptide stretch corresponding to a transmembrane domain of a integral membrane protein, as it exits the ribosome (Kim et al. 2001) (Fig. 16.2A). The loaded SRP complex then binds to the membrane-bound receptor

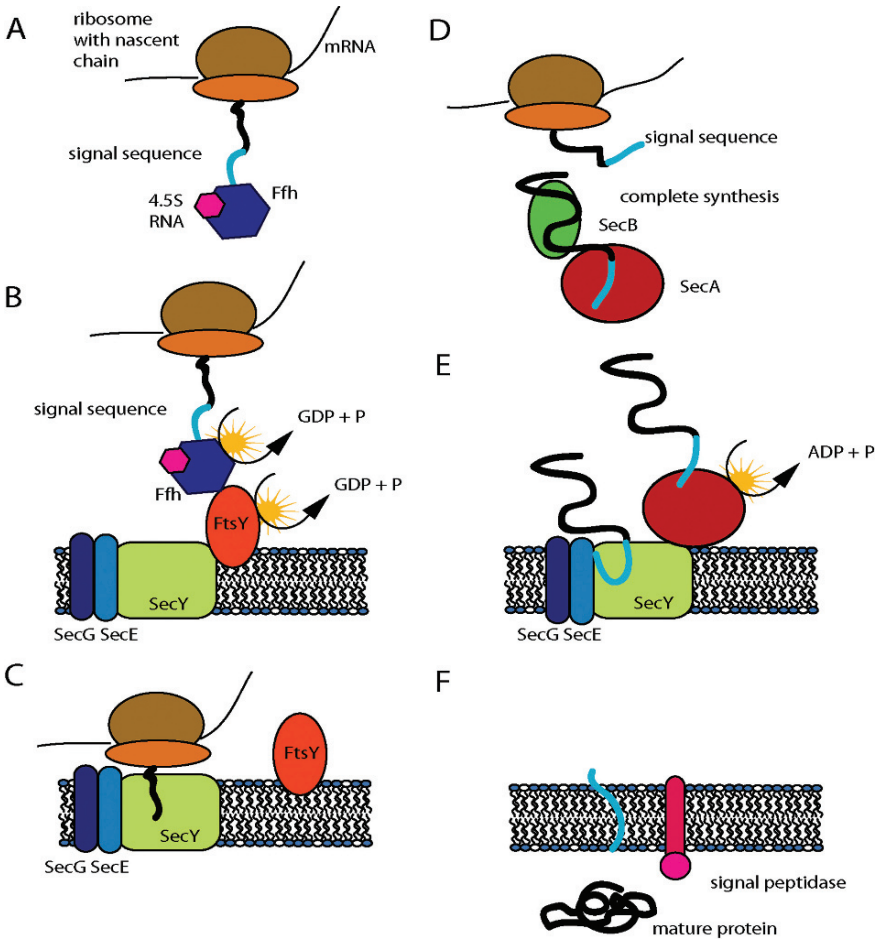

Fig. 16.2 Overview of protein targeting to the Sec translocon via the co-translational (**A–C**) and post-translational route (**D–E**). (**A**) A hydrophobic signal peptide or transmembrane domain of the nascent polypeptide chain is recognized by SRP (Ffh protein and 4.5S RNA unit). (**B**) SRP guides the ribosome with the nascent chain to the membrane-embedded receptor FtsZ which ensures the transfer of the nascent chain to the Sec apparatus. GTP hydrolysis is required for the release of SRP and the receptor. (**C**) The membrane-associated ribosome proceeds to synthesize the protein directly into the Sec system. (**D**) The signal peptide exiting the ribosome as a nascent chain is recognized by SecB which prevents its folding. SecA associates with SecB. (**E**) SecB transfers the preprotein to SecA and dissociates since it is not necessary for the translocation step. SecA associates with the Sec apparatus and proceeds to insert around 20 amino acids at a time into the translocation machinery; ATP hydrolysis is necessary for this motion. (**F**) Upon completion of translocation signal peptidase I cleaves off the signal peptide

FtsY which mediates the interaction of Ffh with the Sec translocon (Fig. 16.2B); GTP binding to both SRP and the receptor FtsY is a prerequisite for their interaction. GTP hydrolysis is precisely timed to transfer the ribosome nascent chain to the Sec translocon releasing the SRP from its receptor (Bange et al. 2007). In this

manner the ribosome can resumes protein synthesis (Fig. 16.2C). Proteins following the co-translational export route are typically inner membrane proteins, but a few soluble proteins utilize this route as well. Probably the best studied SRP substrate is the disulfide oxidase DsbA. DsbA appears to fold too rapidly to be maintained in an unfolded state which is required for post-translational Sec transport (Schierle et al. 2003). Other proteins that utilize the cotranslational route include TorT, TolB or FlgI (Huber et al. 2005a). It is conceivable that the co-translational route could be utilized for the expression of any protein which is otherwise prone to aggregation in the cytoplasm.

The post-translational secretion mode normally involves the tetrameric chaperone SecB which binds to a nascent chain exiting the ribosome to prevent its immediate folding (Fig. 16.2D). In a SecB mutant, other general chaperones such as GroEL and/or DnaK can compensate for the loss of SecB (Kumamoto 1991). The association with SecB maintains the protein in a transport-compatible state since only the unfolded protein can be threaded through the membrane. SecB typically transfers its substrate directly to SecA to which it binds asymmetrically in its dimeric form. Binding of SecB presumably results in the dissociation of one SecA monomer, which may be important for the transfer step of the precursor protein to SecA (Fekkes et al. 1997). The translocation through the Sec pore is a step-wise event in which ATP hydrolysis by SecA allows the threading of around 20–30 amino acids of the polypeptide at a time through the SecYEG pore (van der Wolk et al. 1997), (Fig. 16.2E). However, once the preprotein is inserted into the membrane, translocation can be completed in the presence of solely a electrochemical potential, even without SecA (Duong and Wickner 1997, Schiebel et al. 1991).

16.1.2 Transporting Folded Proteins: The Twin-Arginine Translocase

The Tat pathway was discovered in bacteria 12 years ago (Berks 1996), and in the thylakoids membrane of plants 16 years ago (Cline et al. 1992). Little is known about the detailed molecular mechanism of protein translocation via Tat. The most remarkable feature of the Tat pathway is that it exports completely folded and assembled protein substrates. An unknown step in the translocation process serves as proof-reading function allowing only native proteins to be exported (DeLisa et al. 2003). The energy for translocation is derived solely from the proton motive force (Alder and Theg 2003, Bageshwar and Musser 2007) and so far no ATP requirement has been demonstrated. The Tat pathway is highly conserved in archaea, in most bacteria and in the chloroplasts of plants. Tat signal sequences can be found in most organisms and are partially interchangeable. Only a few protozoa encode proteins with homology to Tat components in their mitochondrial genome (Gray et al. 2004). The majority of the protein substrates for this pathway function in alternative anaerobic respiration pathways and catalyze redox reactions. They typically require the assembly of a set of complex cofactors and are often composed

of multiple polypeptide subunits. The incorporation of these cofactors often necessitates specialized chaperones which are only available in the cytoplasm. For example, trimethylamine N-oxide reductase (TorA) contains a Fe-S cluster and a bis-molybdopterin guanine dinucleotide (MGM) cofactor (Mejean et al. 1994). Next to the advantage of having cytoplasmic chaperones assisting their folding maturation, iron-sulfur clusters are sensitive to oxidants which can be easier avoided in the reducing environment of the cytoplasm. Particularly, folding in the cytoplasm is of great advantage for halophilic organisms, which had they relied on the Sec pathway would have to fold their extracytoplasmic proteins under higher salt concentrations that favor protein aggregation. Hence these organisms often solely rely on Tat-mediated export (Dilks et al. 2005, Rose et al. 2002).

The minimal composition of the Tat translocon consists of the membrane proteins TatA, TatB and TatC. TatB is dispensable in some Gram-positive bacteria, and most archaea or can be replaced by mutated TatA variants (Blaudeck et al. 2005). Tat components can be found in two distinct subcomplexes in resting membranes (Orriss et al. 2007): a receptor complex composed of stoichiometric amounts of TatB and TatC, which is responsible for the recognition of Tat signal peptides (Alami et al. 2003, Kreutzenbeck et al. 2007, Strauch and Georgiou 2007b) and a second subcomplex containing high-molecular weight complexes of TatA. TatA forms pore-like structures of varying sizes in certain detergents, leading to the hypothesis that it mediates the actual translocation step (Gohlke et al. 2005). Whether TatA actually forms a channel or whether it is involved in lipid rearrangements that in turn mediate translocation has yet to be clarified. TatC interacts with the twin-arginine motif, whereas TatB associates with the hydrophobic stretch within the signal peptide. Currently, the most favored model for Tat export proposes a "handing-over" mechanism in which the signal peptide is first recognized by the TatB/C complex followed by the recruitment of several TatA oligomers; the substrate is then "handed over" to the TatA complex which mediates the actual translocation event (Fig. 16.3).

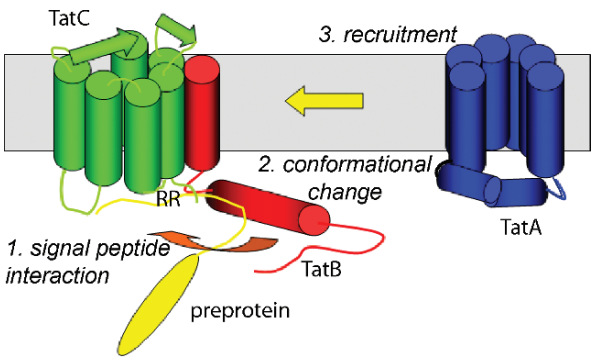

Fig. 16.3 Current translocation model for the export of proteins via the Tat pathway. TatB and TatC establish the signal peptide recognition complex. Upon interaction with the signal peptide (1), a possible conformational change occurs (2), followed by the recruitment of several TatA oligomers (3). TatA oligomers assemble in complexes of variable sizes which might depend on the dimensions of the substrate. TatA putatively mediates the actual translocation event

The division of the translocation step into two separate events is reflected in the biophysical properties of the translocation event (Bageshwar and Musser 2007). The question remains how does the receptor complex signal its interaction with the substrate to TatA? A conformational change is probably the most plausible explanation for this event. Gerard and coworkers suggested a pulling mechanism of the precursor by TatC based on the observation that proteins can be exported when covalently linked to the plant TatC homologue in the thylakoid Tat pathway (Gerard and Cline 2006). The fact that the translocon remains functional even when TatC is fused to its substrate could indicate that TatC at least remains in close proximity to the translocation step. Conformational changes could thus either serve as a trigger for the onset of the translocation event or could provide an actual pulling mechanism.

The fact that the Tat pathway exports folded proteins, begs the question whether there is a relationship between folding quality and export competence. Does the pathway discriminate between folded and unfolded proteins? DeLisa and coworkers (2003) showed that alkaline phosphatase, in which two intramolecular disulfide bridges must form to assume its active dimer conformation, cannot be exported when expressed in cells with a reducing cytoplasm that prevents disulfide bond formation (DeLisa et al. 2003). Deletions of *gor* and *trxB* which inactivate the thioredoxin and the glutathione reduction pathways that normally maintain the cytoplasm under reducing conditions allow disulfide bond formation in alkaline phosphatase and result in export via Tat. Similarly, Fisher et al. (2006) reported that the export rates of maltose binding protein variants (MBP-G32D, MBP-I33P, and MalE31-G32D/I33P) correlates with their solubility and the *in vitro* folding kinetics. These observations further support the notion that some step in the Tat pathway functions as a filter to prevent the export of misfolded proteins. Richter and coworkers proposed that it is the exposure of hydrophobic patches in unfolded proteins which allows the pathway to determine whether a protein is folded or misfolded (Richter et al. 2007). Notably they showed that an intrinsically disordered protein could be translocated though Tat but insertion of short hydrophobic stretches in this protein abolished export.

Several cofactor-containing Tat substrates have their own dedicated chaperone that are referred to as redox enzyme maturation proteins (REMPs). REMPs behave as specific proofreading chaperones escorting various oxido-reductases to the Tat apparatus. The enzyme trimethylamine N-oxide reducatase, TorA, for instance, has its own chaperone, TorD, which greatly facilitates the incorporation of its cofactors and retards the export process (Pommier et al. 1998). TorD binds specifically to the core region of the TorA signal peptide, but also to some parts of the mature enzyme (Hatzixanthis et al. 2005). Once TorD is bound to the signal peptide, its affinity for GTP increases. No GTP hydrolysis could be detected *in vitro*, indicating that the role of GTP might be more regulatory than catalytic. Similar to TorA, the DMSO reductase DmsA and the nitrate reductase NarG contain a molybdenum cofactor and an iron-sulfur cluster, respectively. The insertion of the cofactor and folding maturation is assisted by the small chaperones DmsD and NarJ (Chan et al. 2006, Oresnik et al. 2001). Based on phylogenetic analyses, TorD and NarJ have been classified as belonging to one group of maturation chaperones, whereas DmsD and NapD, which assists the folding of the nitrate reductase NapA (Maillard et al. 2007)

have been assigned to a second group. At least for NapD, it was recently demonstrated that its molecular role is not only to camouflage the signal peptide, but also to actively inhibit transport before folding maturation has been completed (Maillard et al. 2007). A third group for Tat chaperones, the small chaperones HyaE and HybE assist the folding maturation of the [NiFe]-containing hydrogenase 1 (HyaA) and hydrogenase 2 (HybO/HybC), respectively (Dubini and Sargent 2003).

16.2 Expression and Folding of Exported Recombinant Proteins in *E. coli*

E. coli is widely used as the host organism for preparative protein expression in the laboratory and in the biotechnology industry (Baneyx and Mujacic 2004, de Marco 2007). Expression of heterologous proteins in secreted form is desirable when the heterologous protein contains disulfide bonds or otherwise cannot fold in the cytoplasm and when periplasmic localization confers protection against proteolysis or provides an advantage for downstream processing. Only a small set of *E. coli* proteins are secreted into the extracellular space primarily by pathogenic strains (Lawley et al. 2003, Pallen et al. 2003, Sandkvist 2001). Most naturally transported proteins of non-pathogenic *E. coli* are localized either in the periplasmic space, or associate with the outer membrane. Heterologous proteins secreted via Sec can be expressed at very high levels in the periplasmic space of bacteria (Choi and Lee 2004, Mergulhao et al. 2005). Early reports suggested that Tat-mediated transport results in lower protein yields than Sec transport (Berks et al. 2003, Sargent et al. 1998). However, it now appears that the efficiency of expression via Tat is dependent on the protein of interest. Fisher et al. (2008) reported that the periplasmic accumulation of different proteins such as alkaline phosphatase, GFP and a scFv antibody fragment fused to MBP were comparable or at most two fold lower for Tat-mediated export. Interestingly, the purity and activity levels of Tat exported proteins in the osmotic shock fraction were higher than those exported via Sec (Fisher et al. 2008). On the other hand, Tat export resulted in higher periplasmic yields of thioredoxin variants compared to export via the Sec pathway (Masip et al. 2008). This is presumably because the rapid folding kinetics of thioredoxin render it incompetent for Sec export, but favor secretion through Tat. Upon overexpression of a protein, misfolding leading to polypeptide degradation, aggregation or cell toxicity can occur. Precursor proteins that are exported slowly or that jam the Sec translocon result in cell toxicity and accumulation of the precursor protein in the cytoplasm (Feilmeier et al. 2000, Kiino and Silhavy 1984). However, this does not seem to be the case for export via Tat. Unlike the Sec translocon, the Tat apparatus does not appear to be prone to jamming, since the components of the translocon dissociate from any stuck precursor polypeptides (Cline and McCaffery 2007). Ultimately, in order to achieve higher yields of active proteins, the export pathway has to be chosen carefully. Export rates, yields, purity and yield of active proteins depend strictly on the amino acid sequence of the protein of interest.

Protein folding is a central issue in the expression of secreted proteins. As was discussed above, proteins secreted via Tat must attain a native-like conformation in the cytoplasm whereas proteins exported by the Sec apparatus have to fold within the periplasmic space. The co-expression of the proper set of endogenous chaperones can facilitate the expression of secreted proteins for both transport pathways (de Marco 2007). Cytosolic chaperones are often classified as folding, holding and disaggregation chaperones. The first class includes the ribosome-associated trigger factor (TF), the DnaK system (DnaK with its DnaJ and GrpE co-chaperones; KJE), and the GroEL system (GroEL with its GroES co-chaperone; ELS). Collectively these chaperones assist *de novo* protein folding. Both DnaK and GroEL are capable of refolding host proteins that become unfolded under environmental stress. The second class of cytosolic chaperones comprises of the holdases (the small heat-shock chaperones IbpA and IbpB), the redox-regulated Hsp33 and the "emergency" chaperone Hsp31. Holdases are active during severe stress and bind to early folding intermediates to prevent overloading of the KJE and ELS system (Mujacic et al. 2004). If folding and holding of proteins fail to deter protein aggregation, the third class of chaperones kicks in. Chaperones of this class promote aggregate solubilization and include ClpB ClpA, ClpX and ClpY with the latter three being involved in directing proteins to degradation. Disaggregation chaperones do not participate in the refolding of solubilized proteins, but rather transfer them to DnaK.

Several cytoplasmic chaperones and other cytosolic factors have been shown to increase the efficiency of export via the Tat pathway. For example the general chaperone DnaK aids the folding of several Tat substrates (Graubner et al. 2007, Perez-Rodriguez et al. 2007), resulting in increased export. Improved export of fusions to the TorA signal peptide has been observed upon overexpression of the chaperone TorD (Hatzixanthis et al. 2005, Jack et al. 2004, Li et al. 2006). In addition, overexpression of proteins that do not have a chaperone function, including the Tat pathway components TatABCE (Alami et al. 2003) and PspA (DeLisa et al. 2004) that possibly affects the electron gradient, have been shown to enhance Tat export. On the other hand, proteins translocated via the Sec pathway are released into the periplasm in an unfolded conformation and must attain their native state in that compartment. One of the major folding chaperones in the periplasmic space is DegP, which exhibits two functions: At lower growth temperatures, this protein typically acts as a molecular chaperone whereas at elevated temperatures its function as a degrading enzyme becomes more pronounced (Spiess et al. 1999). It recognizes unfolded proteins presumably via its PDZ domain (Iwanczyk et al. 2007, Wilken et al. 2004). Further, protein degradation in the periplasm can also involve the protease III and Tsp and thus may be alleviated in strains carrying deletions of the respective genes (Meerman and Georgiou 1994).

Many secreted proteins contain disulfide-bridges which need to be correctly formed for the polypeptide to attain its native conformation. In *E. coli* periplasmic protein thiol oxidation is catalyzed by the enzyme DsbA whereas isomerization of misfolded disulfide bonds is mediated by DsbC and to a lesser extend by DsbG (Bessette et al. 1999, Rietsch et al. 1996). Overexpression of DsbA and DsbC can result in a marked increase in the yield of complex recombinant proteins such as

the human plasminogen activator (Bessette et al. 1999, Qiu et al. 1998), human nerve growth factor (Kurokawa et al. 2001), insulin-like growth factor-I (Joly et al. 1998) or horseradish peroxidase (Kurokawa et al. 2000). Additionally, overexpression of periplasmic chaperones such as Skp or the peptidyl-proline *cis-trans* isomerases, PpiA, PpiD, SurA or FkpA (Arie et al. 2001, Bothmann and Pluckthun 2000, Missiakas et al. 1996) have been shown to enhance the soluble yield of numerous proteins including antibody fragments (Choi and Lee 2004, Hayhurst et al. 2003). Combinations of overexpressed periplasmic chaperones have been shown to aid the folding of human plasma retinol-binding protein and of the extracellular carbohydrate recognition domain of the dendritic cell membrane receptor DC-SIGN (Schlapschy et al. 2006); for review (Choi and Lee 2004) (Fig. 16.4).

16.2.1 Protein Secretion and Display in Combinatorial Library Screening

16.2.1.1 Phage Display

The display of proteins on the surface of viral particles or cells constitutes the foundation of high throughput screening technologies for protein engineering purposes. Display technologies describe a variety of methodologies for the presentation of biomolecules onto a virus or cell. Protein display allows the screening of large combinatorial protein libraries for the isolation of ligand binding proteins, the engineering of protein stability (Kotz et al. 2004) and catalytic activity (Fernandez-Gacio et al. 2003), the detection of interacting proteins, determining the substrate specificity of proteases (Matthews and Wells 1993) and for several other applications (Hwang et al. 2007, Li et al. 2008, Matthews and Wells 1993). Viral, cell-based and *in vitro* display systems, such as ribosome display (Lipovsek and Pluckthun 2004) have been developed, but for the purposes of the present review we will focus only on viral (bacteriophage) and bacterial cell display methodologies.

Phage display is the first genetic strategy developed for the isolation of ligand-binding proteins from combinatorial libraries (Smith 1985). For filamentous phage display, the phage particles harboring the protein of interest and the gene that encodes it are continuously secreted into the growth medium. The protein of interest is typically displayed as a fusion to one of the coat proteins. Normally, the displayed protein fusion is secreted via the Sec pathway and is incorporated onto the virion during phage assembly in the periplasm. While lytic phages (such as T7 or lambda) have been used for display, non-lytic filamentous phages such as f1, M13 or fd are much more commonly employed. The displayed protein is either encoded in a phagemid, a plasmid containing both an *E. coli* and a phage origin of replications or it is directly integrated into the phage genome. With filamentous phage, the protein of interest is typically fused to the N-terminus of protein pIII (Fig. 16.5A) allowing the presentation of up to 5 copies, or to the major coat protein pVIII (Fig. 16.5A) which allows more than 2700 copies to be displayed on the phage particle (Gucksman et al. 1992, Malik et al. 1996, Zwick et al. 2000).

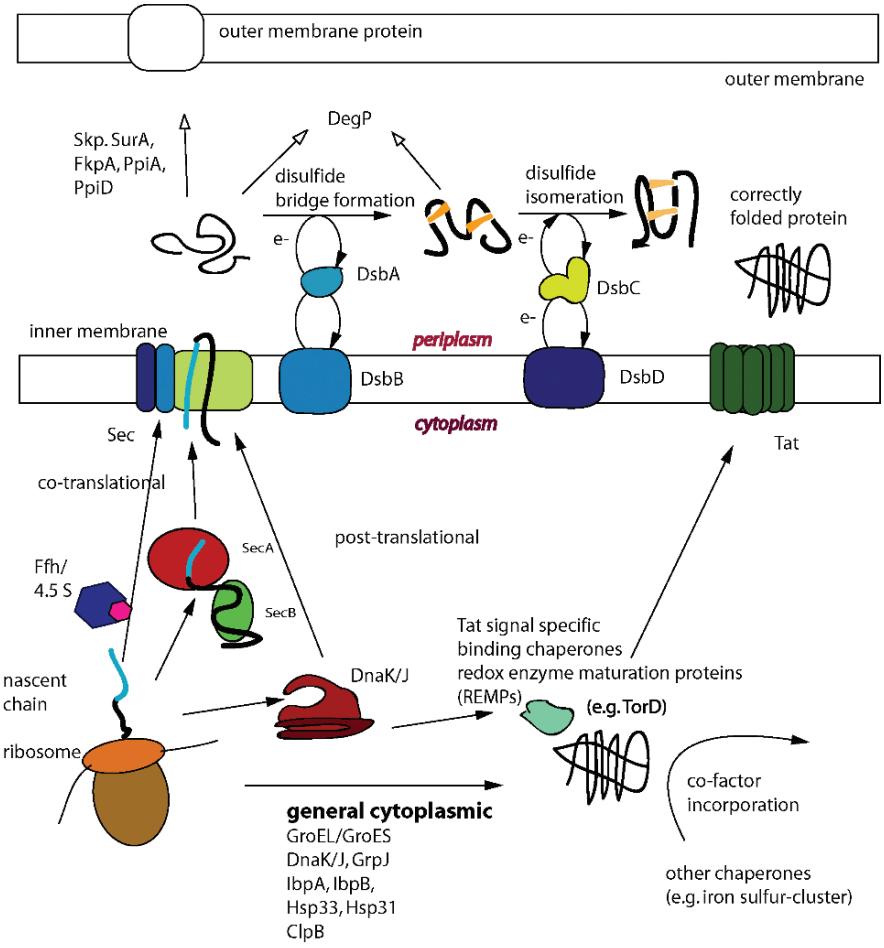

Fig. 16.4 Compartmental chaperones and their contribution to different export pathways. Cytoplasmic chaperones facilitate the export process of most precursor proteins. SecB and SecA bind to the fully synthesized polypeptide to maintain it in its export-competent, unfolded state and the guide it to the Sec translocon (post-translational). On the other hand, the signal recognition particle, composed of Ffh and 4.5S RNA, associates with the nascent chain extruding from the ribosome. The latter allows the recruitment of the translating ribosome to the membrane where it resumes synthesis of the polypeptide directly into the Sec pore (co-translational Sec transport). The presence of the general chaperone DnaK/J can be beneficial for the export of both Sec and Tat substrates. Tat substrates fold prior to export and therefore their folding maturation can be improved by the overexpression of various cytoplasmic chaperones. Additionally, the folding of many Tat substrates requires the participation of specialized cytoplasmic redox enzyme maturation proteins (REMPs). Tat substrates are translocated while in their correctly folded state, whereas folding of Sec substrates takes place in the periplasmic environment. General periplasmic chaperones, such as Skp, SurA etc. improve the solubility or folding of the Sec secreted polypeptides. DegP switches between its protease and foldase function depending on the temperature. DsbA and DsbC catalyze oxidative protein folding in the periplasm. DsbA introduces disulfide bridges, whereas DsbC re-shuffles their confirmation. DsbB and DsbD maintain these two crucial proteins in their appropriate oxidation state.

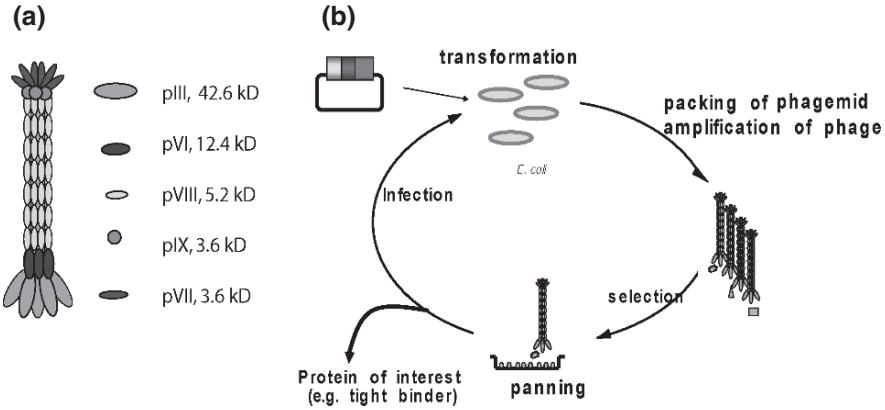

Fig. 16.5 (**A**) Filamentous phage particle, (**B**) panning cycle for enrichment of binders. A phagemid encoding the genetic information for the protein of interest is transformed into *E. coli* cells, which amplify the virion. After purification of the phage particles from the culture supernatant, protein displaying phage particles are applied to the immobilized ligand. Non-binding phage particles are eliminated by washing. Bound phage is eluted, pooled and used to infect new *E. coli* cells allowing a repetition of the panning cycle until interesting ligand binders are identified

Phage displaying polypeptides that bind to a desired ligand are enriched by several rounds of panning onto immobilized ligand. The ligand can be immobilized either directly by adsorption onto a plastic surface or indirectly, e.g. by using a biotin conjugate together with streptavidin-coated beads (Blazek et al. 2004).

Normally, proteins displayed on filamentous phage are secreted by the post-translational Sec pathway (Rapoza and Webster 1993). Often however, limitations associated with the post-translational Sec apparatus restrict the ability to display certain kinds of polypeptides, especially proteins that fold quickly in the cytoplasm or the export of which can block the early steps in the secretion process, e.g. by tight binding to SecA. Employing co-translational Sec export using an appropriate signal peptide such as the one for DsbA can alleviate these problems. For example, the fast-folding designed ankyrin-repeat proteins (DARPins) can be transported with high efficiency when switching the export signal to the co-translational DsbA signal peptide resulting in a 700-fold increase in their display on filamentous phage (Steiner et al. 2006). Export following this route prevents premature cytoplasmic folding prior translocation.

For reasons that probably relate to the fact that coat proteins are embedded in the membrane before they assemble onto the phage particle, it is not possible to export pIII or pV fusions via the Tat apparatus. Export via Tat is desirable for the display of proteins that require the incorporation of cytoplasmic cofactors, for proteins that are unable to fold into the periplasm for other reasons (Feilmeier et al. 2000) or for those that might fold too fast and cannot be maintained in a Sec competent state by the cytoplasmic chaperone machinery. A system that capitalizes on the Tat pathway for protein display is shown in Fig. 16.6 (Paschke and Hohne 2005; Strauch and

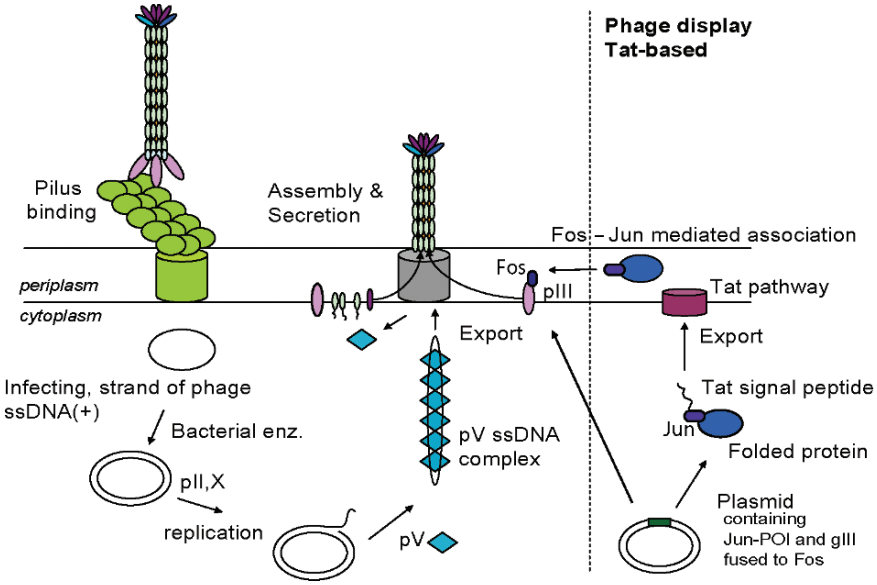

Fig. 16.6 Display of proteins exported via Tat on phage. The phage particle binds to the F-pilus and inserts its single stranded DNA into the bacterium where it uses bacterial enzymes and its own proteins for second strand synthesis and replication. The phage pV protein sequesters the + single strand away to enable its packaging into the phage particle that is assembled in its own secretion apparatus and extruded through the outer membrane via a phage encoded channel formed by the pIV protein. In the Tat-based phage display, pIII is expressed as a fusion of a leucine zipper domain (here Fos) and exported via Sec. The protein of interest (POI) is expressed a fusion to the complementary leucine zipper (here Jun) and a Tat signal peptide. After folding in the cytoplasm, the POI is exported via Tat and associates with the phage by non-covalently binding to the heterodimerizing leucine zipper sequence fused to pIII

Tullman-Ercek, unpublished results). pIII with a N-terminal Sec signal peptide is fused to half of a heterodimerizing leucine zipper sequence whereas the protein of interest is fused to an N-terminal Tat signal peptide and to the complementary leucine zipper sequence. The two gene constructs are expressed from a bicistronic operon. The pIII fusion and the protein of interest are exported via separate routes, namely Sec and Tat respectively, but once in the periplasm, their association is ensured by the leucine zipper dimerization and thus the target protein becomes non-covalently attached to pIII on the surface of phage. Cysteine residues may be placed at the ends of the leucine zipper halves to allow covalent disulfide linkages within the oxidizing environment of the periplasm. This strategy has been successfully employed to display fluorescent GFP (Paschke and Hohne 2005, Strauch 2007) that necessitates folding in the cytoplasm to form the active chromophore.

16.2.1.2 *E. coli*-Based Protein Display

Bacterial display offers several distinct advantages relative to phage: (i) it is possible to display many more protein copies on a bacterium compared to a phage particle:

(ii) complex proteins consisting of multiple polypeptides or proteins containing cofactors are more easy to display on cells; (iii) components of cells surfaces can be exploited for the retention of fluorescent products of enzymatic reactions and (iv) finally but most importantly, because of their larger size, bacteria are compatible with methodologies that utilize fluorescent activated cell sorting (FACS). Using multi-color fluorescence labeling strategies, it is possible to interrogate every clone in a library for the level of expression of a target protein, ligand binding or catalytic activity in a quantitative fashion. The advantages of FACS as a library screening tool have been instrumental in the isolation of very high affinity (picomolar) binding polypeptides and enzymes with high catalytic activity and selectivity from libraries displayed on bacteria.

In *E. coli* proteins can be displayed either on the surface or on a subcellular location that can be made accessible to extracellularly added fluorescently conjugated molecules following chemical treatment. A variety of protein fusions have been used for protein display on the surface of *E. coli* and the topic has been reviewed recently (Daugherty 2007, Lee et al. 2003, Samuelson et al. 2002). Several native outer membrane proteins (OMPs) such as FhuA, OmpA, OmpS, OmpX, and its circular permutated variant CPX, have been utilized for the display of short peptides with varying sizes typically between 12 and 28 aa long. In addition flagellar proteins, such as the commercially available recombinant constructs FliTrx (Lu et al. 1995, Westerlund-Wikstrom 2000) have been used for peptide display whereas Lpp-OmpA fusions and autotransporter proteins from pathogenic *E. coli* have been exploited for the display of several small proteins for ligand binding and enzymatic activity selections (Becker et al. 2007, 2004, Jose et al. 2005, Wentzel et al. 1999). However, surface display of intact proteins is often accompanied by changes in outer membrane permeability and loss of viability. In addition, the display of multi subunit proteins or proteins that contain multiple disulfide bonds is problematic since there is no folding machinery on the surface of the cell (Stathopoulos et al. 1996). Finally, it is not known whether large heterologous polypeptides fused to outer membrane protein targeting sequences can engage the periplasmic folding chaperones (Adams et al. 2005, Bos et al. 2007, Veiga et al. 2002) and the YaeT outer membrane protein localization machinery that might be required for surface display (Kim et al. 2007).

Proteins anchored on the inner membrane or expressed in the periplasm are of course not exposed to the extracellular fluid because the outer membrane of *E. coli* presents a formidable diffusion barrier that excludes molecules larger than 600 kDa. However, various chemical treatments can be used to increase the permeability of, or to completely remove, the outer membrane thus allowing access to periplasmic proteins with externally added ligands. Typically such ligands are fluorescently labeled so that upon binding to an *E. coli* displayed protein they render the cell fluorescent allowing its isolation by flow cytometry (Chen et al. 2001, Harvey et al. 2004). Libraries of scFv antibodies expressed in soluble form in the periplasmic space have been screened for binding to fluorescently labeled low molecular weight ligands that gain access into that compartment by incubating the cells in a high salt environment. Incubation in a hypertonic solution allowed molecules up to 10–15 KDa to diffuse into the cell without the release of the scFv proteins from the periplasm

(Chen et al. 2001). Using the same approach (Ribnicky et al. 2007) isolated a mutant scFv antibody that exhibited improved export via the Tat pathway, leading to greater accumulation of functional protein and therefore increased binding of fluorescently labeled antigen. Interestingly, the selected scFv variant exhibited faster folding kinetics *in vitro*, indicating that the rate of folding within the cytoplasm correlates with competence for translocation via the Tat pathway (Ribnicky et al. 2007).

Access of larger ligands into the periplasm requires rupture of the outer membrane and can be accomplished easily by treating the cells with a combination of chelating agents and lysozyme. However, under these conditions soluble secreted proteins, including the proteins to be displayed, are released from the periplasm either partially or completely. To avoid this problem Harvey et al developed the Anchored Periplasmic Expression (APEx) display system, in which the protein of interest is tethered to the inner membrane by fusing it genetically to an appropriate anchoring sequence. In principle any transmembrane α-helix can be used as an anchoring sequence (Ki et al. 2004). However, fusion to a targeting sequence comprised of a signal sequence followed by the first few amino acids of an inner membrane lipoprotein can be employed to convert the protein of interest into a lipoprotein. This is advantageous because the fusion tag required for display is very short and the expression of lipoproteins is better tolerated by the cell compared to integral membrane proteins. APEx has been used for the isolation of proteins that bind to extracellular ligands, for the engineering of variants that express better in the periplasm and for the detection of protein-protein interactions (Jeong et al. 2007). For the latter application, a bait protein is expressed in membrane-tethered form whereas the prey is expressed solubly in the periplasm. Following permeabilization of the outer membrane, the prey is released from the cell unless it captured by the inner membrane-tethered bait. The resulting complex can be detected by fluorescent anti-prey antibodies allowing isolation of the respective cell by FACS. Recently, a variation of APEx was employed to screen libraries of full length IgG antibodies in bacteria (Mazor et al. 2007). The ability to isolate and express full length IgG in bacteria may allow the rapid generation of antibodies for many therapeutic and diagnostic purposes. Finally our lab recently demonstrated that APEx can be carried out in *dsbA* strains where the formation of disulfide bonds is compromised. In this manner we were able to isolate mutant scFv antibody fragments that are stable and can fold in the absence of disulfide bonds (Seo et al. unpublished). Such antibody fragments are desired for gene therapy applications in which they would be expressed in the cytoplasm (where disulfide bonds cannot normally form) and can be used to disrupt the function of proteins associated with disease (Fig. 16.7).

16.2.2 Exploiting the Secretion Machinery as a Solubility and Folding Filter

As already mentioned, the folding state of a polypeptide is a major determinant of export competence. Huber and coworkers demonstrated that thioredoxin, which

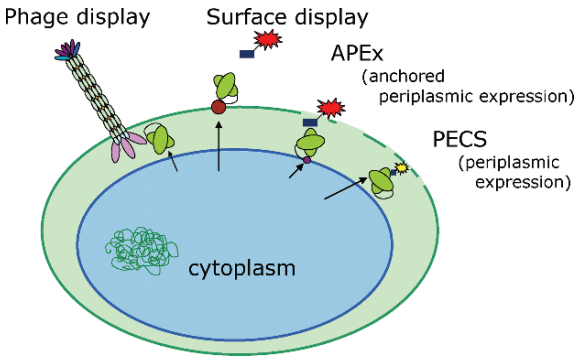

Fig. 16.7 *E. coli*-based display techniques. For simplicity lipopolysaccharides (LPS) and the peptidoglycan layer are not shown. For filamentous phage display, the protein of interest (POI) is exported into the periplasm of *E. coli* before it can be assembled on the phage particle. For surface display, outer membrane proteins or autotransporters can be utilized as carrier proteins for the polypeptide of interest (POI). For anchored periplasmicexpression (APEx), the POI can be either tethered to the inner membrane by a N-terminal NlpA signal peptide fusion or by a fusion to a transmembrane helix. In soluble periplasmic expression (PECS), any transport pathway can be used to secrete the POI into the periplasm. The latter two display technologies require the fracture of the outer membrane via lysozyme-EDTA treatment (APEx) or high salt concentrations (PECS)

folds very rapidly *in vitro*, cannot be secreted via the post-translational Sec pathway but is efficiently translocated into the periplasm when fused to a signal peptide that mediates co-translational export (Huber et al. 2005a). Huber et al. then selected for thioredoxin variants that can be exported post-translationally and showed that these proteins exhibit up to 30 fold slower rates in one of the critical steps of folding (Huber et al. 2005b). In effect, this is the opposite selection to that of Ribnicky who isolated faster folding proteins based on Tat export competence (Ribnicky et al. 2007). In other studies, the quality control feature of the Tat pathway was exploited to select for variant proteins displaying greater solubility (Fisher et al. 2006). In that system the protein of interest is expressed as a tripartite fusion with an N-terminal Tat signal peptide and β-lactamase fused to the C-terminus. Growth on ampicillin is employed to select for mutations in the protein of interest that allow export of the fusion and localization of the β-lactamase moiety in the periplasm. Since the export rates correlate with the solubility of a protein, fusion constructs that conferred ampicillin resistance carried variants with better folding abilities. Utilizing this approach, Fisher and coworkers were able to isolate higher solubility variants of the aggregation-prone amyloid precursor protein Aβ42, a primary constituent of the toxic plaques in Alzheimer disease (Fisher et al. 2006).

The Tat pathway is capable of co-transporting at least two folded protein subunits at a time, only one of which has a signal sequence, via a “hitchhiker export” mechanism (Rodrigue et al. 1999). This observation was recently exploited to develop a Tat based 2-hybrid system in which one protein (bait) is expressed as a fusion to a Tat signal peptide whereas the second protein (prey) is fused to a protein reporter that can confer a phenotype only after export into the bacterial periplasmic space. Since

the prey-reporter fusion lacks a signal peptide, it can only be exported as a complex with the bait-signal peptide fusion which is capable of targeting the Tat translocon. Using maltose-binding protein as the reporter, clones expressing interacting proteins could be identified on maltose minimal media or on MacConkey plates. Alternatively, using cysteine disulfide oxidase DsbA as reporter, export of a signal peptide-prey:bait-DsbA complex into the periplasm allowed complementation of $dsbA^-$ mutants. The prey:bait-DsbA complex was able to restore the formation of active alkaline phosphatase, an enzyme that can be easily detected by a chromogenic assay (Strauch and Georgiou 2007a). The Tat two-hybrid system can be used as a new tool to identify protein-protein interaction on a genomic scale by including two libraries as bait and prey, or it can be utilized to identify protein interaction partners for a protein of interest. Additionally, a 2-hybrid assay may be utilized as a tool for the *in vivo* co-evolution of interacting protein pairs as has been demonstrated by DeLisa and coworkers (private communication).

16.3 Conclusions

Protein secretion in *E. coli* is of enormous significance in biotechnology for applications ranging from preparative protein production to combinatorial library screening. After more than 30 years of study many of the mechanistic details of Sec protein translocation have been elucidated. In contrast, the sequence of events that lead to the export of proteins via the Tat pathway is not completely understood. A significant difference between Sec and Tat pathways is that the former exports proteins that are unfolded whereas the latter accepts only proteins that have attained a native-like conformation. The Sec pathway has been used to express recombinant proteins at high levels that can exceed 5 g/L. Periplasmic expression allows the engagement of the post-translational modification apparatus enabling the introduction and refinement of disulfide bridges which can be additionally optimized by the overexpression of endogenous chaperones. Recent evidence suggests that high yields may also be attained with the Tat pathway provided that the protein is compatible for export via this route. However, g/L expression of Tat proteins needs to be demonstrated.

Protein secretion is an essential step for display and the screening of combinatorial libraries. The export pathway inflicts an additional filter step onto the general screening or selection scheme. For example the use of signal peptide that target different export pathways can lead to the isolation of distinct pools of protein variants that exhibit different folding characteristics (Table 16.1).

Despite the fact that little is known about the actual molecular process of targeting to and transport through the Twin-Arginine Translocase, it is a promising candidate for periplasmic expression of heterologous proteins, including those that otherwise would be incompatible with the Sec transport pathway. The use of the Tat pathway does not only allow increased periplasmic yields of actual active proteins, specifically those that fold fast, it further enables the refinement of existing screens or selections, and even allows the launching of novel protein engineering platforms

Table 16.1 Comparisons of current applications of different routes across the cytoplasmic membrane

	Post-translational Sec export	Co-translational	Tat export
Typical signal sequence	ssPelB, ssPhoA, ssOmpA,	ssDsbA, ssTorT	ssTorA,
Folding preference	slow folding	fast folding, aggregation-prone proteins	fast folding, containing cytoplasmically inserted cofactors,
Maturation	disulfide bridges, fatty acylation (lipoproteins), heme insertion (requires periplasmic reduction)	disulfide bridges	cofactors, other protein subunits, disulfide bridges[1]
Current applications	phage display, bacterial display technologies, protein expression, screen for slower folding variants, protein expression	phage display, cell display on the inner membrane	screen for solubility increase or faster folders, Tat two-hybrid, Tat-based phage display, protein expression

[1] Non-native protein substrates containing disulfide bonds must be expressed in strains having an oxidizing cytoplasm.

that capitalize on its proofreading mechanism and the possibility to fold the protein of interest in the cytoplasmic environment of the cell.

References

Adams TM, Wentzel A, Kolmar H (2005) Intimin-mediated export of passenger proteins requires maintenance of a translocation-competent conformation. J Bacteriol 187(2):522–33

Alami M, Luke I, Deitermann S et al. (2003) Differential interactions between a twin-arginine signal peptide and its translocase in *Escherichia coli*. Mol Cell 12(4):937–46

Alder NN, Theg SM (2003) Energetics of protein transport across biological membranes. a study of the thylakoid DeltapH-dependent/cpTat pathway. Cell 112(2):231–42

Arie JP, Sassoon N, Betton JM (2001) Chaperone function of FkpA, a heat shock prolyl isomerase, in the periplasm of *Escherichia coli*. Mol Microbiol 39(1):199–210

Bageshwar UK, Musser SM (2007) Two electrical potential dependent steps are required for transport by the *Escherichia coli* Tat machinery. J Cell Biol 179(1):87–99

Baneyx F, Mujacic M (2004) Recombinant protein folding and misfolding in *Escherichia coli*. Nat Biotechnol 22(11):1399–408

Bange G, Wild K, Sinning I (2007) Protein translocation: checkpoint role for SRP GTPase activation. Curr Biol 17(22):R980–2

Becker S, Michalczyk A, Wilhelm S et al. (2007) Ultrahigh-throughput screening to identify *E. coli* cells expressing functionally active enzymes on their surface. Chembiochem 8(8):943–9

Becker S, Schmoldt HU, Adams TM et al. (2004) Ultra-high-throughput screening based on cell-surface display and fluorescence-activated cell sorting for the identification of novel biocatalysts. Curr Opin Biotechnol 15(4):323–9

Berks BC (1996) A common export pathway for proteins binding complex redox cofactors? Mol Microbiol 22(3):393–404

Berks BC, Palmer T, Sargent F (2003) The Tat protein translocation pathway and its role in microbial physiology. Adv Microb Physiol 47:187–254

Bessette PH, Aslund F, Beckwith J et al. (1999) Efficient folding of proteins with multiple disulfide bonds in the *Escherichia coli* cytoplasm. Proc Natl Acad Sci USA 96(24):13703–8

Blaudeck N, Kreutzenbeck P, Freudl R et al. (2003) Genetic analysis of pathway specificity during posttranslational protein translocation across the *Escherichia coli* plasma membrane. J Bacteriol 185(9):2811–9

Blaudeck N, Kreutzenbeck P, Muller M et al. (2005) Isolation and characterization of bifunctional *Escherichia coli* TatA mutant proteins that allow efficient tat-dependent protein translocation in the absence of TatB. J Biol Chem 280(5):3426–32

Blazek D, Celer V, Navratilova I et al. (2004) Generation and characterization of single-chain antibody fragments specific against transmembrane envelope glycoprotein gp46 of maedi-visna virus. J Virol Methods 115(1):83–92

Bos MP, Robert V, Tommassen J (2007) Biogenesis of the gram-negative bacterial outer membrane. Annu Rev Microbiol 61:191–214

Bothmann H, Pluckthun A (2000) The periplasmic *Escherichia coli* peptidylprolyl cis,trans-isomerase FkpA. I. Increased functional expression of antibody fragments with and without cis-prolines. J Biol Chem 275(22):17100–5

Chan CS, Howell JM, Workentine ML et al. (2006) Twin-arginine translocase may have a role in the chaperone function of NarJ from *Escherichia coli*. Biochem Biophys Res Commun 343(1):244–51

Chen G, Hayhurst A, Thomas JG et al. (2001) Isolation of high-affinity ligand-binding proteins by periplasmic expression with cytometric screening (PECS). Nat Biotechnol 19(6):537–42

Choi JH, Lee SY (2004) Secretory and extracellular production of recombinant proteins using *Escherichia coli*. Appl Microbiol Biotechnol 64(5):625–35

Cline K, Ettinger WF, Theg SM (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(4):2688–96

Cline K, McCaffery M (2007) Evidence for a dynamic and transient pathway through the TAT protein transport machinery. Embo J 26(13):3039–49

Daugherty PS (2007) Protein engineering with bacterial display. Curr Opin Struct Biol 17(4): 474–80

de Marco A (2007) Protocol for preparing proteins with improved solubility by co-expressing with molecular chaperones in *Escherichia coli*. Nat Protoc 2(10):2632–9

DeLisa MP, Lee P, Palmer T et al. (2004) Phage shock protein PspA of *Escherichia coli* relieves saturation of protein export via the Tat pathway. J Bacteriol 186(2):366–73

DeLisa MP, Tullman D, Georgiou G (2003) Folding quality control in the export of proteins by the bacterial twin-arginine translocation pathway. Proc Natl Acad Sci USA 100(10):6115–20

Dilks K, Gimenez MI, Pohlschroder M (2005) Genetic and biochemical analysis of the twin-arginine translocation pathway in halophilic archaea. J Bacteriol 187(23):8104–13

Dubini A, Sargent F (2003) Assembly of Tat-dependent [NiFe] hydrogenases: identification of precursor-binding accessory proteins. FEBS Lett 549(1–3):141–6

Duong F, Wickner W (1997) The SecDFyajC domain of preprotein translocase controls preprotein movement by regulating SecA membrane cycling. Embo J 16(16):4871–9

Feilmeier BJ, Iseminger G, Schroeder D et al. (2000) Green fluorescent protein functions as a reporter for protein localization in *Escherichia coli*. J Bacteriol 182(14):4068–76

Fekkes P, van der Does C, Driessen AJ (1997) The molecular chaperone SecB is released from the carboxy-terminus of SecA during initiation of precursor protein translocation. Embo J 16(20):6105–13

Fernandez-Gacio A, Uguen M, Fastrez J (2003) Phage display as a tool for the directed evolution of enzymes. Trends Biotechnol 21(9):408–14

Fernandez LA, Berenguer J (2000) Secretion and assembly of regular surface structures in Gram-negative bacteria. FEMS Microbiol Rev 24(1):21–44

Fisher AC, Kim J-Y, Perez-Rodriguez R et al. (2008) Exploration of twin-arginine translocation for the expression and purification of correctly folded proteins in *Escherichia coli*. Microbial Biotechnol 1(5):403–415.

Fisher AC, Kim W, DeLisa MP (2006) Genetic selection for protein solubility enabled by the folding quality control feature of the twin-arginine translocation pathway. Protein Sci 15(3):449–58

Georgiou G, Segatori L (2005) Preparative expression of secreted proteins in bacteria: status report and future prospects. Curr Opin Biotechnol 16(5):538–45

Gerard F, Cline K (2006) Efficient twin arginine translocation (Tat) pathway transport of a precursor protein covalently anchored to its initial cpTatC binding site. J Biol Chem 281(10):6130–5

Glucksman MJ, Bhattacharjee S, Makowski L (1992) Three-dimensional structure of a cloning vector. X-ray diffraction studies of filamentous bacteriophage M13 at 7 A resolution. J Mol Biol 226(2):455–70

Gohlke U, Pullan L, McDevitt CA et al. (2005) The TatA component of the twin-arginine protein transport system forms channel complexes of variable diameter. Proc Natl Acad Sci USA 102(30):10482–6

Graubner W, Schierhorn A, Bruser T (2007) DnaK plays a pivotal role in Tat targeting of CueO and functions beside SlyD as a general Tat signal binding chaperone. J Biol Chem 282(10): 7116–24

Gray MW, Lang BF, Burger G (2004) Mitochondria of protists. Annu Rev Genet 38:477–524

Harvey BR, Georgiou G, Hayhurst A et al. (2004) Anchored periplasmic expression, a versatile technology for the isolation of high-affinity antibodies from *Escherichia coli*-expressed libraries. Proc Natl Acad Sci USA 101(25):9193–8

Hatzixanthis K, Clarke TA, Oubrie A et al. (2005) Signal peptide-chaperone interactions on the twin-arginine protein transport pathway. Proc Natl Acad Sci USA 102(24):8460–5

Hayhurst A, Happe S, Mabry R et al. (2003) Isolation and expression of recombinant antibody fragments to the biological warfare pathogen Brucella melitensis. J Immunol Methods 276(1–2):185–96

Huber D, Boyd D, Xia Y et al. (2005a) Use of thioredoxin as a reporter to identify a subset of *Escherichia coli* signal sequences that promote signal recognition particle-dependent translocation. J Bacteriol 187(9):2983–91

Huber D, Cha MI, Debarbieux L et al. (2005b) A selection for mutants that interfere with folding of *Escherichia coli* thioredoxin-1 in vivo. Proc Natl Acad Sci USA 102(52):18872–7

Hwang BY, Varadarajan N, Li H et al. (2007) Substrate specificity of the *Escherichia coli* outer membrane protease OmpP. J Bacteriol 189(2):522–30

Iwanczyk J, Damjanovic D, Kooistra J et al. (2007) Role of the PDZ domains in *Escherichia coli* DegP protein. J Bacteriol 189(8):3176–86

Jack RL, Buchanan G, Dubini A et al. (2004) Coordinating assembly and export of complex bacterial proteins. Embo J 23(20):3962–72

Jeong KJ, Seo MJ, Iverson BL et al. (2007) APEx 2-hybrid, a quantitative protein-protein interaction assay for antibody discovery and engineering. Proc Natl Acad Sci USA 104(20): 8247–52

Joly JC, Leung WS, Swartz JR (1998) Overexpression of *Escherichia coli* oxidoreductases increases recombinant insulin-like growth factor-I accumulation. Proc Natl Acad Sci USA 95(6):2773–7

Jose J, Betscheider D, Zangen D (2005) Bacterial surface display library screening by target enzyme labeling: Identification of new human cathepsin G inhibitors. Anal Biochem 346(2):258–67

Ki JJ, Kawarasaki Y, Gam J et al. (2004) A periplasmic fluorescent reporter protein and its application in high-throughput membrane protein topology analysis. J Mol Biol 341(4):901–9

Kiino DR, Silhavy TJ (1984) Mutation prlF1 relieves the lethality associated with export of beta-galactosidase hybrid proteins in *Escherichia coli*. J Bacteriol 158(3):878–83

Kim J, Rusch S, Luirink J et al. (2001) Is Ffh required for export of secretory proteins? FEBS Lett 505(2):245–8

Kim S, Malinverni JC, Sliz P et al. (2007) Structure and function of an essential component of the outer membrane protein assembly machine. Science 317(5840):961–4

Kotz JD, Bond CJ, Cochran AG (2004) Phage-display as a tool for quantifying protein stability determinants. Eur J Biochem 271(9):1623–9

Kreutzenbeck P, Kroger C, Lausberg F et al. (2007) *Escherichia coli* twin arginine (Tat) mutant translocases possessing relaxed signal peptide recognition specificities. J Biol Chem 282(11):7903–11

Kumamoto CA (1991) Molecular chaperones and protein translocation across the *Escherichia coli* inner membrane. Mol Microbiol 5(1):19–22

Kurokawa Y, Yanagi H, Yura T (2000) Overexpression of protein disulfide isomerase DsbC stabilizes multiple-disulfide-bonded recombinant protein produced and transported to the periplasm in *Escherichia coli*. Appl Environ Microbiol 66(9):3960–5

Kurokawa Y, Yanagi H, Yura T (2001) Overproduction of bacterial protein disulfide isomerase (DsbC) and its modulator (DsbD) markedly enhances periplasmic production of human nerve growth factor in *Escherichia coli*. J Biol Chem 276(17):14393–9

Lawley TD, Klimke WA, Gubbins MJ et al. (2003) F factor conjugation is a true type IV secretion system. FEMS Microbiol Lett 224(1):1–15

Lee SY, Choi JH, Xu Z (2003) Microbial cell-surface display. Trends Biotechnol 21(1):45–52

Li HX, Hwang BY, Laxmikanthan G et al. (2008) Substrate specificity of human kallikreins 1 and 6 determined by phage display. Protein Sci 17(4):664–72

Li SY, Chang BY, Lin SC (2006) Coexpression of TorD enhances the transport of GFP via the TAT pathway. J Biotechnol 122(4):412–21

Lipovsek D, Pluckthun A (2004) In-vitro protein evolution by ribosome display and mRNA display. J Immunol Methods 290(1–2):51–67

Lu Z, Murray KS, Van Cleave V et al. (1995) Expression of thioredoxin random peptide libraries on the *Escherichia coli* cell surface as functional fusions to flagellin: a system designed for exploring protein-protein interactions. Biotechnology (NY) 13(4):366–72

Luirink J, High S, Wood H et al. (1992) Signal-sequence recognition by an *Escherichia coli* ribonucleoprotein complex. Nature 359(6397):741–3

Maillard J, Spronk CA, Buchanan G et al. (2007) Structural diversity in twin-arginine signal peptide-binding proteins. Proc Natl Acad Sci USA 104(40):15641–6

Malik P, Terry TD, Gowda LR et al. (1996) Role of capsid structure and membrane protein processing in determining the size and copy number of peptides displayed on the major coat protein of filamentous bacteriophage. J Mol Biol 260(1):9–21

Masip L, Klein-Marcuschamer D, Quan S et al. (2008) Laboratory evolution of *Escherichia coli* thioredoxin for enhanced catalysis of protein oxidation in the periplasm reveals a phylogenetically conserved substrate specificity determinant. J Biol Chem 283(2):840–8

Matthews DJ, Wells JA (1993) Substrate phage: selection of protease substrates by monovalent phage display. Science 260(5111):1113–7

Mazor Y, Van Blarcom T, Mabry R et al. (2007) Isolation of engineered, full-length antibodies from libraries expressed in *Escherichia coli*. Nat Biotechnol 25(5):563–5

Meerman HJ, Georgiou G (1994) Construction and characterization of a set of *E. coli* strains deficient in all known loci affecting the proteolytic stability of secreted recombinant proteins. Biotechnology (NY) 12(11):1107–10

Mejean V, Iobbi-Nivol C, Lepelletier M et al. (1994) TMAO anaerobic respiration in *Escherichia coli*: involvement of the tor operon. Mol Microbiol 11(6):1169–79

Mergulhao FJ, Summers DK, Monteiro GA (2005) Recombinant protein secretion in *Escherichia coli*. Biotechnol Adv 23(3):177–202

Missiakas D, Betton JM, Raina S (1996) New components of protein folding in extracytoplasmic compartments of *Escherichia coli* SurA, FkpA and Skp/OmpH. Mol Microbiol 21(4):871–84

Mujacic M, Bader MW, Baneyx F (2004) *Escherichia coli* Hsp31 functions as a holding chaperone that cooperates with the DnaK-DnaJ-GrpE system in the management of protein misfolding under severe stress conditions. Mol Microbiol 51(3):849–59

Oresnik IJ, Ladner CL, Turner RJ (2001) Identification of a twin-arginine leader-binding protein. Mol Microbiol 40(2):323–31

Orriss GL, Tarry MJ, Ize B et al. (2007) TatBC, TatB, and TatC form structurally autonomous units within the twin arginine protein transport system of *Escherichia coli*. FEBS Lett 581(21): 4091–7

Paetzel M, Karla A, Strynadka NC et al. (2002) Signal peptidases. Chem Rev 102(12):4549–80

Pallen MJ, Chaudhuri RR, Henderson IR (2003) Genomic analysis of secretion systems. Curr Opin Microbiol 6(5):519–27

Paschke M, Hohne W (2005) A twin-arginine translocation (Tat)-mediated phage display system. Gene 350(1):79–88

Perez-Rodriguez R, Fisher AC, Perlmutter JD et al. (2007) An essential role for the DnaK molecular chaperone in stabilizing over-expressed substrate proteins of the bacterial twin-arginine translocation pathway. J Mol Biol 367(3):715–30

Pommier J, Mejean V, Giordano G et al. (1998) TorD, a cytoplasmic chaperone that interacts with the unfolded trimethylamine N-oxide reductase enzyme (TorA) in *Escherichia coli*. J Biol Chem 273(26):16615–20

Qiu J, Swartz JR, Georgiou G (1998) Expression of active human tissue-type plasminogen activator in *Escherichia coli*. Appl Environ Microbiol 64(12):4891–6

Rapoza MP, Webster RE (1993) The filamentous bacteriophage assembly proteins require the bacterial SecA protein for correct localization to the membrane. J Bacteriol 175(6):1856–9

Ribnicky B, Van Blarcom T, Georgiou G (2007) A scFv antibody mutant isolated in a genetic screen for improved export via the twin arginine transporter pathway exibits faster folding. J Mol Biol 369(3):631–9

Richter S, Lindenstrauss U, Lucke C et al. (2007) Functional Tat transport of unstructured, small, hydrophilic proteins. J Biol Chem 282(46):33257–64

Rietsch A, Belin D, Martin N et al. (1996) An in vivo pathway for disulfide bond isomerization in *Escherichia coli*. Proc Natl Acad Sci USA 93(23):13048–53

Rodrigue A, Chanal A, Beck K et al. (1999) Co-translocation of a periplasmic enzyme complex by a hitchhiker mechanism through the bacterial tat pathway. J Biol Chem 274(19):13223–8

Rose RW, Bruser T, Kissinger JC et al. (2002) Adaptation of protein secretion to extremely high-salt conditions by extensive use of the twin-arginine translocation pathway. Mol Microbiol 45(4):943–50

Samuelson P, Gunneriusson E, Nygren PA et al. (2002) Display of proteins on bacteria. J Biotechnol 96(2):129–54

Sandkvist M (2001) Type II secretion and pathogenesis. Infect Immun 69(6):3523–35

Sargent F, Bogsch EG, Stanley NR et al. (1998) Overlapping functions of components of a bacterial Sec-independent protein export pathway. Embo J 17(13):3640–50

Schatz G, Dobberstein B (1996) Common principles of protein translocation across membranes. Science 271(5255):1519–26

Schiebel E, Driessen AJ, Hartl FU et al. (1991) Delta mu H+ and ATP function at different steps of the catalytic cycle of preprotein translocase. Cell 64(5):927–39

Schierle CF, Berkmen M, Huber D et al. (2003) The DsbA signal sequence directs efficient, cotranslational export of passenger proteins to the *Escherichia coli* periplasm via the signal recognition particle pathway. J Bacteriol 185(19):5706–13

Schlapschy M, Grimm S, Skerra A (2006) A system for concomitant overexpression of four periplasmic folding catalysts to improve secretory protein production in *Escherichia coli*. Protein Eng Des Sel 19(8):385–90

Smith GP (1985) Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. Science 228(4705):1315–7

Spiess C, Beil A, Ehrmann M (1999) A temperature-dependent switch from chaperone to protease in a widely conserved heat shock protein. Cell 97(3):339–47

Stathopoulos C, Georgiou G, Earhart CF (1996) Characterization of *Escherichia coli* expressing an Lpp'OmpA(46–159)-PhoA fusion protein localized in the outer membrane. Appl Microbiol Biotechnol 45(1–2):112–9

Steiner D, Forrer P, Stumpp MT et al. (2006) Signal sequences directing cotranslational translocation expand the range of proteins amenable to phage display. Nat Biotechnol 24(7):823–31

Strauch EM, (2007) Characterization and Applications of the Twin-Arginine Transporter pathway. Department of Biochemistry, pp. 188. University of Texas at Austin

Strauch EM,Georgiou G (2007a) A bacterial two-hybrid system based on the twin-arginine transporter pathway of *E. coli*. Protein Sci 16(5):1001–8

Strauch EM, Georgiou G (2007b) *Escherichia coli* tatC mutations that suppress defective twin-arginine transporter signal peptides. J Mol Biol 374(2):283–91

Tokumoto U, Nomura S, Minami Y et al. (2002) Network of protein-protein interactions among iron-sulfur cluster assembly proteins in *Escherichia coli*. J Biochem (Tokyo) 131(5): 713–9

Tullman-Ercek D, DeLisa MP, Kawarasaki Y et al. (2007) Export pathway selectivity of *Escherichia coli* twin arginine translocation signal peptides. J Biol Chem 282(11):8309–16

van der Wolk JP, de Wit JG, Driessen AJ (1997) The catalytic cycle of the *Escherichia coli* SecA ATPase comprises two distinct preprotein translocation events. Embo J 16(24):7297–304

Veiga E, Sugawara E, Nikaido H et al. (2002) Export of autotransported proteins proceeds through an oligomeric ring shaped by C-terminal domains. Embo J 21(9):2122–31

Wentzel A, Christmann A, Kratzner R et al. (1999) Sequence requirements of the GPNG beta-turn of the *Ecballium elaterium* trypsin inhibitor II explored by combinatorial library screening. J Biol Chem 274(30):21037–43

Westerlund-Wikstrom B (2000) Peptide display on bacterial flagella: principles and applications. Int J Med Microbiol 290(3):223–30

Wilken C, Kitzing K, Kurzbauer R et al. (2004) Crystal structure of the DegS stress sensor: How a PDZ domain recognizes misfolded protein and activates a protease. Cell 117(4):483–94

Zwick MB, Shen J,Scott JK (2000) Homodimeric peptides displayed by the major coat protein of filamentous phage. J Mol Biol 300(2):307–20