Published in final edited form as:
*Microb Biotechnol*. 2008 September ; 1(5): 403–415. doi:10.1111/j.1751-7915.2008.00041.x.

# Exploration of twin-arginine translocation for expression and purification of correctly folded proteins in *Escherichia coli*

**Adam C. Fisher**[1,†], **Jae-Young Kim**[1,†], **Ritsdeliz Perez-Rodriguez**[1], **Danielle Tullman-Ercek**[2], **Wallace R. Fish**[3], **Lee A. Henderson**[3], and **Matthew P. DeLisa**[1,4]

[1]School of Chemical and Biomolecular Engineering, Cornell University, Ithaca NY 14853 USA

[2]Department of Pharmaceutical Chemistry and California Institute for Quantitative Biomedical Research (QB3), University of California, San Francisco, San Francisco, California 94143-2540, USA

[3]Vybion, Inc., Ithaca, NY 14850 USA

[4]Department of Biomedical Engineering, Cornell University, Ithaca, NY 14853 USA

## Summary

Historically, the general secretory (Sec) pathway of Gram-negative bacteria has served as the primary route by which heterologous proteins are delivered to the periplasm in numerous expression and engineering applications. Here we have systematically examined the twin-arginine translocation (Tat) pathway as an alternative, and possibly advantageous, secretion pathway for heterologous proteins. Overall, we found that: (1) export efficiency and periplasmic yield of a model substrate was affected by the composition of the Tat signal peptide; (2) Tat substrates were correctly processed at their N-termini upon reaching the periplasm; and (3) proteins fused to maltose-binding protein (MBP) were reliably exported by the Tat system, but only when correctly folded; aberrantly folded MBP fusions were excluded by the Tat pathway's folding quality control feature. We also observed that Tat export yield was comparable to Sec for relatively small, well-folded proteins, higher relative to Sec for proteins that required cytoplasmic folding, and lower relative to Sec for larger, soluble fusion proteins. Interestingly, the specific activity of material purified from the periplasm was higher for certain Tat substrates relative to their Sec counterparts, suggesting that Tat expression can give rise to relatively pure and highly active proteins in one step.

## Introduction

The export of proteins to the bacterial periplasm is a commonly used strategy in preparative protein expression (Georgiou and Segatori, 2005) and in many protein engineering applications (Chen *et al.*, 2001; Francisco *et al.*, 1993; Harvey *et al.*, 2004; Smith, 1985). Since its discovery more than 25 years ago, the general secretory (Sec) pathway has been the primary vehicle for transporting proteins to the periplasm in biotechnological applications (Baneyx, 1999; Baneyx and Mujacic, 2004; Swartz, 2001). However, since only extended polypeptides can fit through the 5–8 Å diameter pore of the Sec apparatus (Van den Berg *et al.*, 2004), it is imperative that protein substrates are maintained in an *unfolded* state via interactions with chaperones or by co-translational protein synthesis (Driessen *et al.*, 2001; Schatz and Dobberstein, 1996). For this reason, proteins exported via the Sec pathway attain

*Correspondence should be addressed to Matthew P. DeLisa, 254 Olin Hall, Cornell University, Ithaca, NY 14853 USA; phone: 607-254-8560; fax: 607-255-9166; md255@cornell.edu.
†Both authors contributed equally to this work.

their native, biologically active conformation only *after* they have been 'threaded' through the Sec pore. Consequently, numerous Sec-targeted proteins are not readily exported due to premature folding in the cytoplasm (Benson *et al.*, 1984; Schierle *et al.*, 2003). This often results in blocking or jamming of the secretion apparatus in a manner that leads to the toxic accumulation of precursors of all secreted substrates in the cytoplasm of cells (Benson *et al.*, 1984; Hayhurst and Georgiou, 2001; Kiino and Silhavy, 1984). Even when membrane translocation is successful, many proteins are susceptible to extensive degradation or cannot attain their native structure in the periplasmic space (Feilmeier *et al.*, 2000; Gentz *et al.*, 1988). Certain other complex proteins acquiring protein subunits or redox cofactors hinge on the formation of secondary or tertiary structure in the cytoplasm and thus are inherently incompatible with Sec export (DeLisa *et al.*, 2003; Rodrigue *et al.*, 1999).

Based on the above observations, it is clear that the Sec pathway is useful for the production of only a subset of proteins. As an alternative to the Sec system, we and others have explored the use of the twin-arginine translocation (Tat) pathway for exporting heterologous proteins across the cyptoplasmic membrane (Bruser, 2007; DeLisa *et al.*, 2003; Fisher *et al.*, 2006; Kim *et al.*, 2005). Proteins routed via the Tat pathway are synthesized with N-terminal signal peptides that bear a conserved, albeit non-essential (DeLisa *et al.*, 2002), S/T-R-R-x-F-L-K motif (Berks, 1996; Berks *et al.*, 2000a). Tat export in *E. coli* minimally requires three integral membrane proteins, namely TatA (or TatE), TatB and TatC, that are predicted to form the translocase (Berks *et al.*, 2000b; Settles *et al.*, 1997; Weiner *et al.*, 1998). While the precise mechanism of Tat-dependent transport remains unresolved, recent evidence suggests a "channel-less" transport model whereby TatC pulls substrates through a patch of TatA on the lipid bilayer (Cline and McCaffery, 2007). Most importantly, whereas Sec-dependent export is restricted to preproteins that have not yet reached a native conformation, the hallmark of the Tat pathway is its predilection for substrates that are correctly folded (Bruser *et al.*, 2003; DeLisa *et al.*, 2003; Fisher *et al.*, 2006) with few exceptions (Richter *et al.*, 2007). Based on this unique characteristic, it has been widely speculated that the Tat system will be a valuable tool for expressing heterologous proteins that attain a folded or partially folded conformation prior to reaching the translocation machinery. Moreover, since the Tat machinery does not typically accommodate unfolded proteins, poorly-folded polypeptides will be effectively eliminated during protein production or in combinatorial library screening experiments. However, owing to its relative novelty and lower reported throughput compared to Sec (Berks *et al.*, 2003; Settles *et al.*, 1997), the potential of the bacterial Tat pathway for preparative expression of secreted proteins has not been fully explored. Here, we address a number of unresolved issues related to Tat-dependent secretion of heterologous proteins.

## Results

### Identification of *E. coli* signal peptides that promote efficient Tat-dependent transport

The *E. coli* genome is predicted to encode up to 34 Tat substrates based on the presence of an N-terminal signal peptide containing a twin arginine motif (Bendtsen *et al.*, 2005; Dilks *et al.*, 2003; Robinson and Bolhuis, 2001). Of these, 27 have been validated experimentally based on their ability to export one or more reporter proteins through the Tat machinery (Tullman-Ercek *et al.*, 2007). Using maltose-binding protein (MBP, encoded by *malE*) as a genetic reporter (see Fig. 1) along with a simple MacConkey maltose agar plate strategy (Blaudeck *et al.*, 2003), we tested 29 Tat signal peptides for their ability to direct MBP to the periplasm. When *E. coli* HS3018 (a *malE*-negative derivative of the strain MC4100) expressed a fusion between MBP and the Tat-dependent signal peptide of trimethylamine *N*-oxide reductase plus 10 amino acids of the mature TorA domain (ssTorA(+10)-MBP) cells appeared as dark red colonies on MacConkey maltose agar plates (Supplementary Fig. 1a) indicating efficient periplasmic localization of functional MBP. On the contrary, plating

HS3018 cells that lacked the *tatC* gene resulted in white colonies indicating that ssTorA(+10)-MBP transport was dependent on a functional Tat pathway (Supplementary Fig. 1a). We also observed complete loss of export when the consensus Arg-Arg residues in the ssTorA signal were mutated to Lys-Lys (Table 1). Tat dependence for ssTorA(+10)-MBP was further confirmed by plating cells on minimal M9 agar supplemented with maltose as the sole carbon source (Supplementary Fig. 1a) and by Western blot analysis of subcellular fractions (Supplementary Fig. 1b).

Following the same strategy, we tested an additional 27 signal peptides for Tat export competence. As above, HS3018 cells were induced for expression of a putative Tat signal peptide plus 6–10 amino acids of the mature protein fused in-frame to MBP. Of these, a total of 14 (ssCueO, ssDmsA, ssFdnG, ssFdoG, ssHyaA, ssNapA, ssSufI, ssTorA, ssWcaM, ssYagT, ssYcbK, ssYedB, ssYdhX and ssYnfE) showed Tat-dependent export as indicated by plating on MacConkey maltose agar plates (Table 1). The rest showed Tat-independent transport in that they conferred a red appearance to both HS3018 and HS3018 Δ*tatC* cells, which was largely consistent with earlier analysis (Tullman-Ercek *et al.*, 2007). For the 14 Tat-dependent signals, we observed varying levels of red coloration that we suspected was indicative of the relative quantity of MBP localized in the periplasm. Indeed, Western blot analysis of subcellular fractions isolated from cells expressing ssTorA-MBP (bright red) and ssFdnG-MBP (pale red) revealed much greater accumulation of ssTorA-MBP in the periplasmic fraction (data not shown). Based on this analysis, 8/14 Tat-dependent signal peptides (CueO, HyaA, SufI, TorA, YcbK, YcdB, YdhX and YnfE) appeared to efficiently export MBP whereas the remaining 6 were less efficient targeting signals. Because ssTorA was found to efficiently export MBP and has been employed extensively in studies of the Tat system (Blaudeck *et al.*, 2001;Blaudeck *et al.*, 2003;Perez-Rodriguez *et al.*, 2007;Strauch and Georgiou, 2007;Tullman-Ercek *et al.*, 2007), we chose this signal peptide for all further studies.

## Translocation efficiency and processing of different MBP chimeras

Here and in many earlier studies employing ssTorA (DeLisa *et al.*, 2002; Thomas *et al.*, 2001; Tullman-Ercek *et al.*, 2007), a common practice for ensuring authentic and efficient substrate processing was to fuse the signal peptide plus several additional residues of the mature TorA protein to the recombinant protein-of-interest (POI). From a biotechnology standpoint, this will result in the undesirable outcome of a POI with an inauthentic N-terminus because these extra residues remain following proteolysis of the signal peptide by signal peptidase I (SPase I). To determine whether exclusion of these extra residues affected Tat transport efficiency or processing, we evaluated ssTorA fused to MBP with ten, six, and zero residues of mature TorA (ssTorA(+10)-MBP, ssTorA(+6)-MBP and ssTorA(+0)-MBP). Western blotting (Fig. 2a) and MacConkey maltose plating (Fig. 2b) revealed that each ssTorA signal was capable of efficiently routing MBP to the periplasm, with the +6 and +0 signals exhibiting slightly better export efficiency and significantly less insoluble accumulation compared to the +10 signal. For comparison, we also expressed MBP with its native Sec signal peptide and observed periplasmic accumulation that was comparable to that seen for the ssTorA constructs (Fig. 2a). One notable difference between the secretion pathways was that very little Sec-targeted MBP accumulated in the cytoplasm or insoluble fraction indicating that Sec export was more efficient than export via the Tat system, consistent with earlier findings (Berks *et al.*, 2003; Settles *et al.*, 1997). Despite this difference in transport efficiency, amylose affinity chromatography purification of the Tat- and Sec-targeted constructs from the periplasm of 50-mL flask cultures resulted in ~5–10 mg/L of soluble MBP for each (Fig. 2c and Table 2).

To assess whether the additional mature residues affected processing by SPase I, N-terminal sequencing of the purified proteins was performed. As expected, the sequence KIEEGKLV

corresponding to the authentic N-terminal residues of mature MBP was obtained for Sec-targeted MBP. Surprisingly, two sequences were obtained for each of the ssTorA constructs. The more abundant sequence corresponded to material that was processed after the predicted ATA cleavage site (Fig. 2d) yielding KIEEGKLV, AQAATDKI, and AQAATDAV for ssTorA(+0), ssTorA(+6) and ssTorA(+10), respectively. A second minor sequence of RATAAQAA for ssTorA(+6) and ssTorA(+10) or RATAKIEE for ssTorA(+0) was obtained. These minor sequences correspond to material that was apparently degraded between Arg35 and Arg36, a location immediately prior to the ATA cleavage site and in the middle of the charged c-region of the signal peptide.

## Exceptional total expression for Tat-targeted substrates

It was surprising to observe that the periplasmic accumulation of Sec- and Tat-mediated MBP was comparable, especially considering that MBP is a native Sec substrate. To determine if this was specific to MBP, we next compared Sec and Tat export of another native Sec substrate, *E. coli* alkaline phosphatase (PhoA). PhoA contains two disulfide bonds that are essential for folding into its native conformation (Sone *et al.*, 1997). These disulfide bonds cannot form within the reducing cytoplasm of *E. coli* (37) and, as a result, PhoA is incompatible with Tat transport under normal cellular conditions (DeLisa *et al.*, 2003). However, using engineered strains that have a more oxidizing cytoplasm, such as *E. coli* FÅ113 (Bessette *et al.*, 1999), protein disulfide bonds readily form in the cytoplasm of these cells resulting in correctly folded PhoA that can be exported via the Tat system (DeLisa *et al.*, 2003) (Fig. 3a and b). The efficiency of Tat-mediated PhoA localization was comparable to that observed for native PhoA exported via the Sec pathway in either FÅ113 cells (Fig. 3a and b) or wildtype DHB4 cells (data not shown).

Interestingly, inspection of whole cell lysates revealed a much greater total accumulation of ssTorA-PhoA relative to native PhoA or a version of PhoA expressed without an N-terminal export signal (Δss-PhoA, Fig. 3a). In fact, a similar phenomenon was observed above for MBP and previously for human tissue plasminogen activator (Kim *et al.*, 2005). Using the mfold program (Zuker, 2003), we found that the mRNA covering the signal peptide-coding region of *torA* folded into a stable stem-loop structure (Fig. 4a) ($\Delta G = -37.10$ kcal/mol at 37°C and physiological pH). Similar analysis of the 5' regions encoding native *phoA*, *Δss-phoA* or *malE* did not yield a stem-loop structural prediction (data not shown). Based on this analysis, we hypothesized that the difference in the amount of active protein could be explained by differences in the abundance of mRNA, akin to what was observed for the signal peptide coding region of native *E. coli* Tat substrate formate dehydrogenase N that folds into a stable secondary structure and mediates translational control (Punginelli *et al.*, 2004). To examine this, we performed a quantitative real-time PCR and discovered that the amount of mRNA was nearly identical in cells expressing ssTorA-PhoA, native PhoA or Δss-PhoA (Fig. 4b). Thus, we believe that the difference in PhoA protein levels is best explained by properties and interactions that occur at the protein level and not at the level of mRNA expression or stability as we had first hypothesized.

## MBP as a fusion partner for Tat-transported proteins

Many proteins that accumulate as insoluble aggregates when they are overproduced in *E. coli* can be rendered soluble simply by fusion to the C-terminus of MBP (Kapust and Waugh, 1999). Moreover, fusion of a POI to either terminus of MBP enables facile purification of the fusion based on the affinity of MBP to cross-linked amylose (di Guan *et al.*, 1988). To determine if the broad utility of MBP as both a fusion partner and a reporter protein could be coupled with Tat transport, we examined a collection of POI-MBP fusions for their ability to be routed to the periplasm by the Tat pathway. Several proteins known to be capable of soluble expression in the *E. coli* cytoplasm, namely *E. coli* glutathione S-

transferase (GST), *E. coli* thioredoxin (TrxA), green fluorescent protein (GFP) and the *de novo*-designed protein Top7, were fused in between ssTorA and MBP and expressed in *E. coli*. As expected, each of these constructs accumulated in the periplasm (Fig. 5a). It is noteworthy that most of the fusions accumulated to a lesser extent than ssTorA-MBP with the exception of the ssTorA-TrxA-MBP fusion that accumulated to a very high level in the periplasm. On the other hand, proteins that are known to misfold in the cytoplasm, such as *E. coli* PhoA or *Agrobacterium tumefaciens* TraR, were not localized in the periplasm (Fig. 5a). Transport to the periplasm for all of these constructs was blocked in a Δ*tatC* strain (shown in Fig. 5a for ssTorA-MBP), indicating that each was exported in a Tat-dependent manner. Thus, similar to our earlier observatons (Fisher *et al.*, 2006), export of these fusion proteins to the periplasm was dependent on correct folding and solubility of the POI in the cytoplasm and on an intact Tat system. To determine if transport of these MBP fusions conferred a *mal*+ phenotype to cells, we plated HS3018 cells expressing each of these constructs on M9 minimal maltose media. In general, the soluble POIs conferred growth in a Tat-dependent manner while the insoluble POIs were incapable of restoring growth to cells (Fig. 5b) indicating that a simple plating strategy can be used to screen for correctly folded POI-MBP fusions. The one exception to this was the ssTorA-GFP-MBP fusion that was localized to the periplasm (Fig. 5a) but incapable of conferring growth to HS3018 cells (Fig. 5b).

## Export and purification of recombinant proteins using the Tat system

Next, we sought to take advantage of the MBP fusion partner for purifying Tat transported POIs from the periplasm of *E. coli* cells. For these studies, we fused GFP to the C-terminus of either ssTorA-MBP or native MBP as this orientation was likely to capitalize on the solubility-enhancing properties of MBP (Kapust and Waugh, 1999). Earlier studies reported that Sec-targeted MBP-GFP fusions were non-fluorescent when exported via the Sec pathway (Feilmeier *et al.*, 2000); thus, GFP export represents an example where the choice of the Tat system was expected to be advantageous over the Sec pathway. Indeed, purification of ssTorA-GFP-MBP from the periplasm using IMAC yielded 2.6 mg/L of pure material compared to only 0.4 mg/L of Sec-targeted MBP-GFP (Table 2). More importantly, whereas the MBP-GFP was entirely non-fluorescent in all fractions, the purified ssTorA-GFP-MBP was highly fluorescent in the cytoplasmic, periplasmic and purified fractions (data not shown), indicating that active GFP could be purified from the periplasm using a Tat, but not a Sec, targeting signal and an MBP fusion tag.

To determine whether the MBP fusion approach could be applied generally to Tat transported proteins, we expressed and purified single-chain antibody fragments (scFvs) from the periplasm by fusion to the C-terminus of ssTorA-MBP. This presented a challenging scenario because scFvs typically misfold in the reducing environment of the *E. coli* cytoplasm where required disulfide bonds do not form. As a result, misfolded scFvs are not transported via the Tat system (DeLisa *et al.*, 2003). However, scFvs fused to the C-terminus of MBP can be expressed at high levels in the cytoplasm of *E. coli* as soluble and active proteins regardless of the redox state of the bacterial cytoplasm (Bach *et al.*, 2001). The scFvs chosen for these studies were each specific for β-galactosidase; the first was a wildtype scFv that folds poorly in the cytoplasm (scFv13) and the second was a variant derived from scFv13 (scFv13.R4) that exhibits increased solubility in the cytoplasm (Martineau *et al.*, 1998). Following expression of ssTorA-MBP-scFv13 and ssTorA-MBP-scFv13.R4 in wildtype *E. coli*, we observed strong accumulation in the cytoplasmic and periplasmic fractions for both scFvs (Fig. 6a). Thus, consistent with previous studies (Bach *et al.*, 2001), MBP-scFv fusions were soluble in the cytoplasm despite the fact that expression and folding occurred in a reducing environment. Moreover, as a result of this solubility, the scFv13 protein was competent for transport via the Tat pathway. Not

surprisingly, the level of periplasmic accumulation for the better folding scFv13.R4 protein was markedly greater than scFv13, consistent with our earlier findings (Fisher *et al.*, 2006). Since ssTorA-MBP-scFv13.R4 exhibited greater accumulation in the periplasm, we chose this construct for subsequent purification studies. From the periplasmic fraction, we recovered 0.5 mg/L of soluble MBP-scFv13.R4, which was slightly greater than the yield reported for a different scFv that had been exported via the Tat pathway (Ribnicky *et al.*, 2007). For comparison, we expressed and purified the scFv13.R4 as a C-terminal fusion to full-length MBP with its native Sec signal peptide (Fig. 6a) and recovered 8.4 mg/L. Thus, for certain proteins, Tat export is less effective than Sec-mediated export, which agrees with previous data comparing expression yields for scFvs and certain other enzymes exported via the Tat and Sec pathways (Table 2). Interestingly, despite the overall lower yield resulting from Tat export compared to Sec export, the specific activity of purified fusion protein recovered using the Tat pathway was ~30% greater on average relative to that recovered using the Sec pathway (Fig. 6b).

Finally, it is noteworthy that when the same scFv13.R4 protein was expressed without the MBP fusion partner, the yield was only 0.01 mg/L for Sec export compared to 0.06 mg/L for Tat export (Table 2). The low yield of Sec-exported scFv13.R4 was accompanied by a marked decrease in culture growth immediately following induction that was not seen when the same protein was exported via the Tat pathway (data not shown). Thus, similar to the results for GFP above, the inability to localize scFv13.R4 in the periplasm using the Sec pathway can be overcome by employing the Tat pathway. Also, the specific activity of the Tat-exported scFv13.R4 was 2–3 times greater than that observed for the MBP fusion (Fig. 6b) suggesting that for some proteins exported via the Tat pathway, lower yields may be offset by greater activity.

## Discussion

It has been previously reported that secretion via the Tat pathway results in lower levels of secreted protein as compared to Sec-dependent secretion. This behavior has been observed for a relatively small handful of secreted substrates expressed in *E. coli* or *Streptomyces lividans* (see Table 2), leading to the generally accepted statement that “the Tat apparatus is…orders-of-magnitude less efficient than the Sec system” (Bruser 2007). Surprisingly, we show comparable or improved periplasmic yield (Tat/Sec ratio > 47%; see Table 2) for a variety of substrates including MBP, PhoA, GFP and an engineered scFv specific for β-galactosidase. In fact, the only case where we observed significantly decreased export efficiency (~10-fold) for the Tat system was following expression of larger proteins comprised of a fusion between MBP and POIs such as scFv13. Collectively, these results suggest that smaller, well-folded substrates are more efficiently processed by the Tat pathway relative to larger, bulkier substrates. This observation is in line with the “channel-less” model of transport that suggests protein fusions exceeding a characteristic or threshold length might be limited in Tat export (Cline et al 2007). Another explanation for the less efficient export of larger proteins is that faster folding, as would be expected for smaller proteins, results in greater export efficiency (Ribnicky *et al.*, 2007). It should be noted that our protein expression experiments were performed without any significant optimization such as the coexpression of factors that are known to increase Tat throughput. Thus, we are optimistic that greater levels of export may be attained using high-expression vectors and/or by co-expressing factors such as TatABC (Alami *et al.*, 2002), PspA (DeLisa *et al.*, 2004), TorD (Li *et al.*, 2006) or DnaK (Perez-Rodriguez *et al.*, 2007), all of which have been shown to increase Tat translocation efficiency. This optimization will be especially important for commercial applications where titers in the 10–100 mg/L range or above are the benchmark. In that sense, it remains to be seen whether the Tat pathway will be

competitive with the Sec pathway. Nonetheless, it is clear that the Tat pathway allows for the transport of folded proteins thereby opening up a new window of secreted products.

Our studies revealed that Tat-targeted MBP is an extremely versatile tool for expression and purification studies. For instance, Tat export of MBP was useful as a: (1) genetic reporter for screening or optimizing Tat export (e.g., via mutations to signal peptides or the TatABC machinery); (2) non-enzymatic reporter of fusion protein solubility; (3) a solubility-enhancing fusion partner, and (4) a facile affinity partner for protein purification. As a reporter for screening putative Tat signal peptides, we tested the *mal* phenotype on MacConkey medium containing maltose as a more rapid (overnight vs. 2–3 days) and semi-quantitative indicator of MBP transport via the Tat system that proved to be relatively consistent with earlier analysis (Tullman-Ercek *et al.*, 2007).

We further examined the N-terminal processing of secreted MBP. Since export is posttranslational, properties of the mature polypeptide are important in routing the protein to the appropriate secretory apparatus and in signal peptide cleavage. For example, a simple ≥ +1 charge to the N terminus of a mature protein abolished or drastically reduced routing through the Sec pathway without affecting the ability to transit the Tat pathway (Tullman-Ercek *et al.*, 2007). Further, these studies revealed that mutations in and around the signal peptide significantly altered substrate expression levels (Pratap and Dikshit, 1998; Robbens *et al.*, 2006). Thus, to ensure fidelity with respect to routing and processing, it is common practice to include amino acids of the mature protein following the A-X-A SPaseI cleavage site. However, our results indicated that similar N-terminal processing by SPaseI occurred irrespective of whether 0, 6, or 10 residues of mature TorA were included in the fusion to MBP. Interestingly, two sequences were obtained for each of the ssTorA-MBP constructs. The more abundant sequence corresponded to protein that was processed after the predicted ATA cleavage site while a less abundant product corresponded to cleavage at a position 5 amino acids upstream of the ATA cleavage site (−5). Such proteolytic processing of ssTorA was observed previously (Genest *et al.*, 2006) and while the cause of this is currently unknown, it is possible that redundant information in the signal peptide could lead to alternative cleavage sites processed by SPase I (Fikes *et al.*, 1990). An alternate possibility is that this arginine pair is proteolytically cleaved by a protease such as OmpT which exhibits a strong preference for cleavage between two basic residues (Lys and, especially, Arg) in the P1 and P'1 positions of the substrate (Dekker *et al.*, 2001).

We also demonstrated MBP to be an effective non-enzymatic protein fusion reporter of Tat transport and as such, protein solubility. This was validated by our ability to recover HS3018 cells on maltose minimal medium *only* when these cells expressed correctly folded POIs (e.g., GST, TrxA, Top7) fused to MBP; the same cells expressing MBP fusions with misfolded or unstable POIs (e.g., PhoA, TraR) exhibited a *mal*− phenotype. A notable exception was ssTorA-GFP-MBP, which was localized in the periplasm but did not confer growth to HS3018 cells. The ssTorA-GFP-MBP localized in the periplasm was fluorescent (data not shown), thus it appears that the GFP moiety folded correctly but somehow interfered with the function of MBP. Since MBP must effectively bind maltose, change conformation and then engage the $MalFGK_2$ complex in order to confer a *mal*+ phenotype, we suspect that GFP interferes with one or more of these steps and renders the screen ineffective.

Finally, we employed fusions in which MBP was positioned N-terminally to ensure the solubility of the POI and determined that the yield of material recovered from the periplasm varied for different POIs. For instance, Tat-secreted MBP-GFP yield was 6.5-fold *greater* than that recovered using Sec export, whereas the yield of Tat-secreted MBP-scFv13.R4 was 16.8-fold *less* compared to Sec. Remarkably, for MBP-scFv13.R4, the specific activity of

the purified material generated using the Tat pathway was 30% greater than for material obtained using the Sec pathway. We reported a similar phenomenon for export of a truncated version of human tPA when affixed with a either a Tat or Sec signal peptide (Kim *et al.*, 2005). While the reasons for this difference in specific activity are not currently known, we speculate that periplasmic folding of certain Sec-targeted substrates may be inefficient, resulting in a large quantity of partially folded, inactive protein in the periplasm. On the contrary, because the Tat system usually exports correctly folded proteins (Bruser *et al.*, 2003; DeLisa *et al.*, 2003; Fisher *et al.*, 2006; Richter *et al.*, 2007), the activity of material isolated from the periplasm represents proteins that have folded completely. Thus, despite the lower expression yields for certain substrates, a potential advantage of the Tat system is that purification from the periplasm can give rise to relatively pure and highly active proteins in one step.

## Experimental Procedures

### Bacterial strains and plasmids

All strains and plasmids used in this study are listed in Table 3. Expression of ssTorA-MBP and all related variants was from plasmid pTrc99A unless otherwise noted. The signal peptide of TorA with an additional 0, 6 or 10 residues from the beginning of the mature portion of the TorA enzyme was PCR-amplified was *E. coli* MC4100 genomic DNA and cloned into vector pTrc99A yielding pssTorA(+0)-MBP, pssTorA(+6)-MBP and pssTorA(+10)-MBP, respectively. Plasmids for the expression of ssTat-MBP fusions are described elsewhere (Tullman-Ercek *et al.*, 2007) as are constructs for expressing ssTorA-PhoA (DeLisa *et al.*, 2003). Plasmids pPhoA and Δss-PhoA were constructed by PCR-amplification of native *E. coli* PhoA or PhoA lacking residues 1–22 and insertion of the resulting PCR products into pTrc99A. For expression of different ssTorA-POI-MBP fusion proteins, DNA encoding ssTorA and MBP was cloned separately in pTrc99A yielding plasmid pTMM. Next, genes encoding various POIs were PCR-amplified and cloned in between ssTorA and MBP in pTMM yielding the final ssTorA-POI-MBP expression plasmids. PCR products were generated using the following as template: MC4100 genomic DNA for TrxA, GST and PhoA; plasmid pTGS encoding GFP (DeLisa *et al.*, 2002); plasmid DNA encoding Top7 (Kuhlman *et al.*, 2003) and *A. tumefaciens* TraR (Zhu and Winans, 2001). Plasmids pMBP-GFP and pMBP-scFv13 were created by first cloning native MBP into pTrc99A followed by in-frame insertion of the gene encoding GFP or scFv13.R4 (Martineau *et al.*, 1998). Plasmids pssTorA-MBP-GFP, pssTorA-MBP-scFv13 and pssTorA-MBP-scFv13.R4 were constructed by inserting the gene encoding GFP, scFv13 or scFv13.R4 after ssTorA-MBP (lacking a stop codon at the C-terminus of MBP) in pTrc99A.

### Cell growth and protein expression

To test growth on maltose, *E. coli* HS3018 or its derivatives containing plasmids encoding MBP fusions were grown overnight, diluted, plated on either MacConkey agar plates containing 0.4% maltose or M9 minimal medium containing 0.4% maltose, and incubated at 37 °C overnight or 48 hours, respectively. For liquid cultures, cells containing pTrc99A derivatives were grown overnight and then subcultured into fresh LB supplemented with 100 μg/ml ampicillin to a starting $OD_{600}$ = 0.1 followed by incubation for 2–3 hr at 30 or 37°C with shaking. When cells reached $OD_{600}$ = 0.5, IPTG was added to a final concentration of 1 mM to induce protein synthesis and cells were grown an additional 4–6 hr. Antibiotic selection was maintained for all markers on plasmids at the following concentrations: ampicillin, 100 μg/ml; chloramphenicol, 25 μg/ml.

### Cell fractionation

An equivalent number of cells were harvested following recombinant protein induction, pelleted by centrifugation and fractionated into cytoplasmic and periplasmic fractions by the ice-cold osmotic shock procedure (Bogsch *et al.*, 1998; DeLisa *et al.*, 2003). The protocol was modified for the expression of disulfide bond-containing proteins (e.g., PhoA) by treating intact cells with 100 mM iodoacetamide for 30 min prior to centrifugation to prevent spontaneous activation of free thiols (Derman and Beckwith, 1995). To analyze total cellular proteins, collected cells were centrifuged, resuspended in cold PBS containing 100 mM iodoacetamide and lysed by sonication. The insoluble fractions were removed by centrifugation (12,000 × *g*, 10 min, 4°C), and soluble protein was quantified by the Bio-Rad protein assay with BSA as standard. The quality of all fractionation experiments was confirmed by analyzing the localization of a cytoplasmic marker protein, namely GroEL, using Western blot analysis.

### Quantitative real-time PCR

For RNA expression analysis, total RNA was extracted from cells using an RNeasy Mini Kit (Qiagen) according to the manufacturer's instructions. The RNA extracted was digested with RNase-free DNaseI for 15 min to remove contaminating DNA. The total mRNA of native PhoA, ssTorA-PhoA and Δss-PhoA was detected by RT-PCR using TaqMan® One-Step RT-PCR Master Mix Reagents Kit (Applied Biosystems). Two microliters of total RNA was used as a template for the amplification reaction. Probes and primer sets for PhoA were designed using Primer Express software (Applied Biosystems). The primers sequences specific for PhoA were: (a) Forward: 5'-AAAGGCGGAAAAGGATCGAT-3'; (b) Reverse: 5'- GCCGCCAAGCGTAACATC-3'; and the probe sequence specific for PhoA was: 6FAM-CCGAACAGCTGCTTAACGCTCGTG–TAMRA. For normalizing purposes, primers and probe sequences specific for 16s rRNA were also designed: (a) Forward: 5'-CCAGCAGCCGCGGTAAT-3', (b) Reverse: 5'- TGCGCTTTACGCCCAGTAAT-3'; and the probe specific for 16srRNA sequence was: 6FAM-CCGATTAACGCTTGCACCCTCCG-TAMRA. The efficiency of the primers for amplification of both PhoA and 16srRNA sequences was tested by electrophoresis before performing the RT-PCR. The program of RT-PCR was 50°C (2 min), 95°C (10 min), followed by 40 cycles of 95°C (15 s), 60°C (1min).

### Protein analysis

Western blotting was performed as previously described (DeLisa *et al.*, 2003). All lanes of SDS-12% polyacrylamide gels were loaded with samples generated from an equivalent number of cells harvested from each experiment. The following primary antibodies were used: monoclonal mouse anti-MBP (Sigma), monoclonal mouse anti-PhoA (Sigma) and polyclonal rabbit anti-GroEL (Sigma). The secondary antibody was goat anti-mouse or goat anti-rabbit horseradish peroxidase. All primary and secondary antibodies were diluted according to manufacturer's specifications. Membranes were first probed with primary and secondary antibodies and, following development, were stripped in Tris-buffered saline/2% SDS/0.7 M β-mercaptoethanol. Stripped membranes were re-blocked and probed with anti-GroEL antibody. PhoA activity was measured as described previously (Derman *et al.*, 1993) while GFP activity was measured according to (DeLisa *et al.*, 2002). The scFv13 activity was quantified by measuring in vitro activation of a normally inactive β-galactosidase variant known as AMEF β-gal (Martineau *et al.*, 1998). Briefly, lysate was prepared from 50 ml cultures of *E. coli* strain AMEF 959 (Martineau *et al.*, 1998) that express AMEF β-gal. Next, 10 μl of lysate containing AMEF β-gal was incubated with 10 μl of a purified scFv13.R4 chimera (e.g., ssTorA-MBP-scFv13.R4) for 3 h at 37°C in a 96-well plate. To this, 100 μl of substrate (o-nitrophenyl-β-D-galactoside (ONPG) in Z buffer) was added to

each well and incubated for 10 min at 37°C with shaking. The reaction was stopped by the addition of 50 µl of 1 M Na2CO3 and absorbance readings at 420 nm were recorded.

### Purification and N-terminal sequencing

For purification of his- tagged proteins, standard IMAC was used according to manufacturer's specifications (Qiagen). For purification of all MBP constructs, periplasmic fractions were isolated from 50-mL bacterial cultures as described above and added to amylose affinity chromatography columns. Columns were prepared by adding 2 mL of amylase resin (New England Biolabs) to empty columns and washing 3 times with 5 mL of MBP buffer (20 mM Tris (pH 7.4), 200 mM NaCl, 1 mM EDTA, 10 mM 2-mercaptoethanol). Following addition of periplasmic fractions, columns were washed 3 times with 5 mL of MBP buffer. Bound proteins were eluted from the resin with elution buffer (MBP buffer + 10 mM Maltose). For N-terminal sequencing, purified samples were separated using standard SDS-PAGE protocols and electroblotted according to protocols found elsewhere (http://cpmcnet.columbia.edu/dept/protein/ss/n_terminal.htmlprepared). Briefly, proteins were transferred to polyvinyldifluoride (PVDF) membranes from SDS-PAGE gels. Following Coomassie Blue staining of PVDF membranes, bands of interest corresponding in size to ssTorA-MBP or native MBP were excised with clean razor blades, transferred to sterile microcentrifuge tubes and stored at −70°C until sent to the Columbia University Medical Center Protein Core Facility for N-terminal sequencing.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

The authors would like to thank David Baker (Top7), Pierre Martineau (scFv13 and scFv13.R4) and Steve Winans (TraR) for plasmid DNA used in this study and Tracy Palmer and Howard Shuman for strains B1LK0 and HS3018, respectively. Funding was provided by an NSF CAREER Award (CBET #0449080) and a NYSTAR James D. Watson Award (both to MPD).

## References

Alami M, Trescher D, Wu LF, Muller M. Separate analysis of twin-arginine translocation (Tat)-specific membrane binding and translocation in *Escherichia coli*. J Biol Chem 2002;277:20499–20503. [PubMed: 11923313]

Bach H, Mazor Y, Shaky S, Shoham-Lev A, Berdichevsky Y, Gutnick DL, Benhar I. *Escherichia coli* maltose-binding protein as a molecular chaperone for recombinant intracellular cytoplasmic single-chain antibodies. J Mol Biol 2001;312:79–93. [PubMed: 11545587]

Baneyx F. Recombinant protein expression in *Escherichia coli*. Curr Opin Biotechnol 1999;10:411–421. [PubMed: 10508629]

Baneyx F, Mujacic M. Recombinant protein folding and misfolding in *Escherichia coli*. Nat Biotechnol 2004;22:1399–1408. [PubMed: 15529165]

Barrett CM, Ray N, Thomas JD, Robinson C, Bolhuis A. Quantitative export of a reporter protein, GFP, by the twin-arginine translocation pathway in *Escherichia coli*. Biochem Biophys Res Commun 2003;304:279–284. [PubMed: 12711311]

Bendtsen JD, Nielsen H, Widdick D, Palmer T, Brunak S. Prediction of twin-arginine signal peptides. BMC Bioinformatics 2005;6:167. [PubMed: 15992409]

Benson SA, Bremer E, Silhavy TJ. Intragenic regions required for LamB export. Proc Natl Acad Sci U S A 1984;81:3830–3834. [PubMed: 6374667]

Berks BC. A common export pathway for proteins binding complex redox cofactors? Mol Microbiol 1996;22:393–404. [PubMed: 8939424]

Berks BC, Sargent F, De Leeuw E, Hinsley AP, Stanley NR, Jack RL, Buchanan G, Palmer T. A novel protein transport system involved in the biogenesis of bacterial electron transfer chains. Biochim Biophys Acta 2000a;1459:325–330. [PubMed: 11004447]

Berks BC, Sargent F, Palmer T. The Tat protein export pathway. Mol Microbiol 2000b;35:260–274. [PubMed: 10652088]

Berks BC, Palmer T, Sargent F. The Tat protein translocation pathway and its role in microbial physiology. Adv Microb Physiol 2003;47:187–254. [PubMed: 14560665]

Bessette PH, Aslund F, Beckwith J, Georgiou G. Efficient folding of proteins with multiple disulfide bonds in the *Escherichia coli* cytoplasm. Proc Natl Acad Sci U S A 1999;96:13703–13708. [PubMed: 10570136]

Blaudeck N, Sprenger GA, Freudl R, Wiegert T. Specificity of signal peptide recognition in Tat-dependent bacterial protein translocation. J Bacteriol 2001;183:604–610. [PubMed: 11133954]

Blaudeck N, Kreutzenbeck P, Freudl R, Sprenger GA. Genetic analysis of pathway specificity during posttranslational protein translocation across the *Escherichia coli* plasma membrane. J Bacteriol 2003;185:2811–2819. [PubMed: 12700260]

Bogsch EG, Sargent F, Stanley NR, Berks BC, Robinson C, Palmer T. An essential component of a novel bacterial protein export system with homologues in plastids and mitochondria. J Biol Chem 1998;273:18003–18006. [PubMed: 9660752]

Bruser T, Yano T, Brune DC, Daldal F. Membrane targeting of a folded and cofactor-containing protein. Eur J Biochem 2003;270:1211–1221. [PubMed: 12631279]

Bruser T. The twin-arginine translocation system and its capability for protein secretion in biotechnological protein production. Appl Microbiol Biotechnol 2007;76:35–45. [PubMed: 17476499]

Chen G, Hayhurst A, Thomas JG, Harvey BR, Iverson BL, Georgiou G. Isolation of high-affinity ligand-binding proteins by periplasmic expression with cytometric screening (PECS). Nat Biotechnol 2001;19:537–542. [PubMed: 11385457]

Cline K, McCaffery M. Evidence for a dynamic and transient pathway through TAT protein transport machinery. EMBO J 2007;26:3039–3049. [PubMed: 17568769]

Dekker N, Cox RC, Kramer RA, Egmond MR. Substrate specificity of the integral membrane protease OmpT determined by spatially addressed peptide libraries. Biochemistry 2001;40:1694–1701. [PubMed: 11327829]

DeLisa MP, Samuelson P, Palmer T, Georgiou G. Genetic analysis of the twin arginine translocator secretion pathway in bacteria. J Biol Chem 2002;277:29825–29831. [PubMed: 12021272]

DeLisa MP, Tullman D, Georgiou G. Folding quality control in the export of proteins by the bacterial twin-arginine translocation pathway. Proc Natl Acad Sci U S A 2003;100:6115–6120. [PubMed: 12721369]

DeLisa MP, Lee P, Palmer T, Georgiou G. Phage shock protein PspA of *Escherichia coli* relieves saturation of protein export via the Tat pathway. J Bacteriol 2004;186:366–373. [PubMed: 14702305]

Derman AI, Puziss JW, Bassford PJ Jr, Beckwith J. A signal sequence is not required for protein export in *prlA* mutants of *Escherichia coli*. EMBO J 1993;12:879–888. [PubMed: 8458344]

Derman AI, Beckwith J. *Escherichia coli* alkaline phosphatase localized to the cytoplasm slowly acquires enzymatic activity in cells whose growth has been suspended: a caution for gene fusion studies. J Bacteriol 1995;177:3764–3770. [PubMed: 7601842]

di Guan C, Li P, Riggs PD, Inouye H. Vectors that facilitate the expression and purification of foreign peptides in *Escherichia coli* by fusion to maltose-binding protein. Gene 1988;67:21–30. [PubMed: 2843437]

Dilks K, Rose RW, Hartmann E, Pohlschroder M. Prokaryotic utilization of twin-arginine translocation pathway: a genomic survey. J Bacteriol 2003;185:1478–1483. [PubMed: 12562823]

Driessen AJ, Manting EH, van der Does C. The structural basis of protein targeting and translocation in bacteria. Nat Struct Biol 2001;8:492–498. [PubMed: 11373615]

Feilmeier BJ, Iseminger G, Schroeder D, Webber H, Phillips GJ. Green fluorescent protein functions as a reporter for protein localization in *Escherichia coli*. J Bacteriol 2000;182:4068–4076. [PubMed: 10869087]

Fikes JD, Barkocy-Gallagher GA, Klapper DG, Bassford PJ Jr. Maturation *Escherichia coli* maltose-binding protein by signal peptidase I *in vivo*. Sequence requirements for efficient processing and demonstration of an alternate cleavage site. J Biol Chem 1990;265:3417–3423. [PubMed: 2406254]

Fisher AC, Kim W, DeLisa MP. Genetic selection for protein solubility enabled by the folding quality control feature of the twin-arginine translocation pathway. Protein Sci 2006;15:449–458. [PubMed: 16452624]

Francisco JA, Campbell R, Iverson BL, Georgiou G. Production and fluorescence-activated cell sorting of *Escherichia coli* expressing a functional antibody fragment on the external surface. Proc Natl Acad Sci U S A 1993;90:10444–10448. [PubMed: 8248129]

Gauthier C, Li H, Morosoli R. Increase in xylanase production by *Streptomyces lividans* through simultaneous use of the Sec- and Tat-dependent protein export systems. Appl Environ Microbiol 2005;71:3085–3092. [PubMed: 15933005]

Genest O, Seduk F, Ilbert M, Mejean V, Iobbi-Nivol C. Signal peptide protection by specific chaperone. Biochem Biophys Res Commun 2006;339:991–995. [PubMed: 16337610]

Gentz R, Kuys Y, Zwieb C, Taatjes D, Taatjes H, Bannwarth W, Stueber D, Ibrahimi I. Association of degradation and secretion of three chimeric polypeptides in *Escherichia coli*. J Bacteriol 1988;170:2212–2220. [PubMed: 3129403]

Georgiou G, Segatori L. Preparative expression of secreted proteins in bacteria: status report and future prospects. Curr Opin Biotechnol 2005;16:538–545. [PubMed: 16095898]

Harvey BR, Georgiou G, Hayhurst A, Jeong KJ, Iverson BL, Rogers GK. Anchored periplasmic expression, a versatile technology for the isolation of high-affinity antibodies from *Escherichia coli*-expressed libraries. Proc Natl Acad Sci U S A 2004;101:9193–9198. [PubMed: 15197275]

Hayhurst A, Georgiou G. High-throughput antibody isolation. Curr Opin Chem Biol 2001;5:683–689. [PubMed: 11738179]

Kapust RB, Waugh DS. *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. Protein Sci 1999;8:1668–1674. [PubMed: 10452611]

Kiino DR, Silhavy TJ. Mutation *prlF1* relieves the lethality associated with export of beta-galactosidase hybrid proteins in *Escherichia coli*. J Bacteriol 1984;158:878–883. [PubMed: 6233268]

Kim J-Y, Fogarty EA, Lu FJ, Zhu H, Henderson LA, DeLisa MP. Twin-arginine translocation of active human tissue plasminogen activator in *Escherichia coli*. Appl Environ Microbiol 2005;71:8451–8459. [PubMed: 16332834]

Kuhlman B, Dantas G, Ireton GC, Varani G, Stoddard BL, Baker D. Design of a novel globular protein fold with atomic-level accuracy. Science 2003;302:1364–1368. [PubMed: 14631033]

Li SY, Chang BY, Lin SC. Coexpression of TorD enhances the transport of GFP via the TAT pathway. J Biotechnol 2006;122:412–421. [PubMed: 16253369]

Martineau P, Jones P, Winter G. Expression of an antibody fragment at high levels in the bacterial cytoplasm. J Mol Biol 1998;280:117–127. [PubMed: 9653035]

Messer, W.; Melchers, F. The activation of mutant beta-galactosidase by specific antibodies. In: Miller, JH.; Reznikoff, WS., editors. The Operon. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 1978. p. 305-315.

Perez-Rodriguez R, Fisher AC, Perlmutter JD, Hicks MG, Chanal A, Santini CL, Wu LF, Palmer T, DeLisa MP. An essential role for the DnaK molecular chaperone in stabilizing over-expressed substrate proteins of the bacterial twin-arginine translocation pathway. J Mol Biol 2007;367:715–730. [PubMed: 17280684]

Pimienta E, Ayala JC, Rodriguez C, Ramos A, Van Mellaert L, Vallin C, Anne J. Recombinant production of *Streptococcus equisimilis* streptokinase by *Streptomyces lividans*. Microb Cell Fact 2007;6:20. [PubMed: 17610745]

Pratap J, Dikshit KL. Effect of signal peptide changes on the extracellular processing of streptokinase from *Escherichia coli*: requirement for secondary structure at the cleavage junction. Mol Gen Genet 1998;258:326–333. [PubMed: 9648736]

Punginelli C, Ize B, Stanley NR, Stewart V, Sawers G, Berks BC, Palmer T. mRNA secondary structure modulates translation of Tat-dependent formate dehydrogenase N. J Bacteriol 2004;186:6311–6315. [PubMed: 15342602]

Ribnicky B, Van Blarcom T, Georgiou G. A scFv antibody mutant isolated in a genetic screen for improved export via the twin arginine transporter pathway exhibits faster folding. J Mol Biol 2007;369:631–639. [PubMed: 17462668]

Richter S, Lindenstrauss U, Lucke C, Bayliss R, Bruser T. Functional Tat transport of unstructured, small, hydrophilic proteins. J Biol Chem 2007;282:33257–33264. [PubMed: 17848553]

Robbens J, De Coen W, Fiers W, Remaut E. Improved periplasmic production of biologically active murine interleukin-2 in *Escherichia coli* through a single aminoacid change at the cleavage site. Proc Biochem 2006;41:1343–1346.

Robinson C, Bolhuis A. Protein targeting by the twin-arginine translocation pathway. Nat Rev Mol Cell Biol 2001;2:350–356. [PubMed: 11331909]

Rodrigue A, Chanal A, Beck K, Muller M, Wu LF. Co-translocation of a periplasmic enzyme complex by a hitchhiker mechanism through the bacterial Tat pathway. J Biol Chem 1999;274:13223–13228. [PubMed: 10224080]

Schaerlaekens K, Lammertyn E, Geukens N, De Keersmaeker S, Anne J, Van Mellaert L. Comparison of the Sec and Tat secretion pathways for heterologous protein production by *Streptomyces lividans*. J Biotechnol 2004;112:279–288. [PubMed: 15313005]

Schatz G, Dobberstein B. Common principles of protein translocation across membranes. Science 1996;271:1519–1526. [PubMed: 8599107]

Schierle CF, Berkmen M, Huber D, Kumamoto C, Boyd D, Beckwith J. The DsbA signal sequence directs efficient, cotranslational export of passenger proteins to the *Escherichia coli* periplasm via the signal recognition particle pathway. J Bacteriol 2003;185:5706–5713. [PubMed: 13129941]

Settles AM, Yonetani A, Baron A, Bush DR, Cline K, Martienssen R. Sec-independent protein translocation by the maize Hcf106 protein. Science 1997;278:1467–1470. [PubMed: 9367960]

Shuman HA. Active transport of maltose in *Escherichia coli* K12. Role of the periplasmic maltose-binding protein and evidence for a substrate recognition site in the cytoplasmic membrane. J Biol Chem 1982;257:5455–5461. [PubMed: 7040366]

Smith GP. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. Science 1985;228:1315–1317. [PubMed: 4001944]

Sone M, Kishigami S, Yoshihisa T, Ito K. Roles of disulfide bonds in bacterial alkaline phosphatase. J Biol Chem 1997;272:6174–6178. [PubMed: 9045630]

Strauch EM, Georgiou G. A bacterial two-hybrid system based on the twin-arginine transporter pathway of *E. coli*. Protein Sci 2007;16:1001–1008. [PubMed: 17456749]

Swartz JR. Advances in *Escherichia coli* production of therapeutic proteins. Curr Opin Biotechnol 2001;12:195–201. [PubMed: 11287237]

Thomas JD, Daniel RA, Errington J, Robinson C. Export of active green fluorescent protein to the periplasm by the twin-arginine translocase (Tat) pathway in *Escherichia coli*. Mol Microbiol 2001;39:47–53. [PubMed: 11123687]

Tullman-Ercek D, DeLisa MP, Kawarasaki Y, Iranpour P, Ribnicky B, Palmer T, Georgiou G. Export pathway selectivity of *Escherichia coli* twin arginine translocation signal peptides. J Biol Chem 2007;282:8309–8316. [PubMed: 17218314]

Van den Berg B, Clemons WM Jr, Collinson I, Modis Y, Hartmann E, Harrison SC, Rapoport TA. X-ray structure of a protein-conducting channel. Nature 2004;427:36–44. [PubMed: 14661030]

Weiner JH, Bilous PT, Shaw GM, Lubitz SP, Frost L, Thomas GH, Cole JA, Turner RJ. A novel and ubiquitous system for membrane targeting and secretion of cofactor-containing proteins. Cell 1998;93:93–101. [PubMed: 9546395]

Zhu J, Winans SC. The quorum-sensing transcriptional regulator TraR requires its cognate signaling ligand for protein folding, protease resistance, and dimerization. Proc Natl Acad Sci U S A 2001;98:1507–1512. [PubMed: 11171981]

Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 2003;31:3406–3415. [PubMed: 12824337]

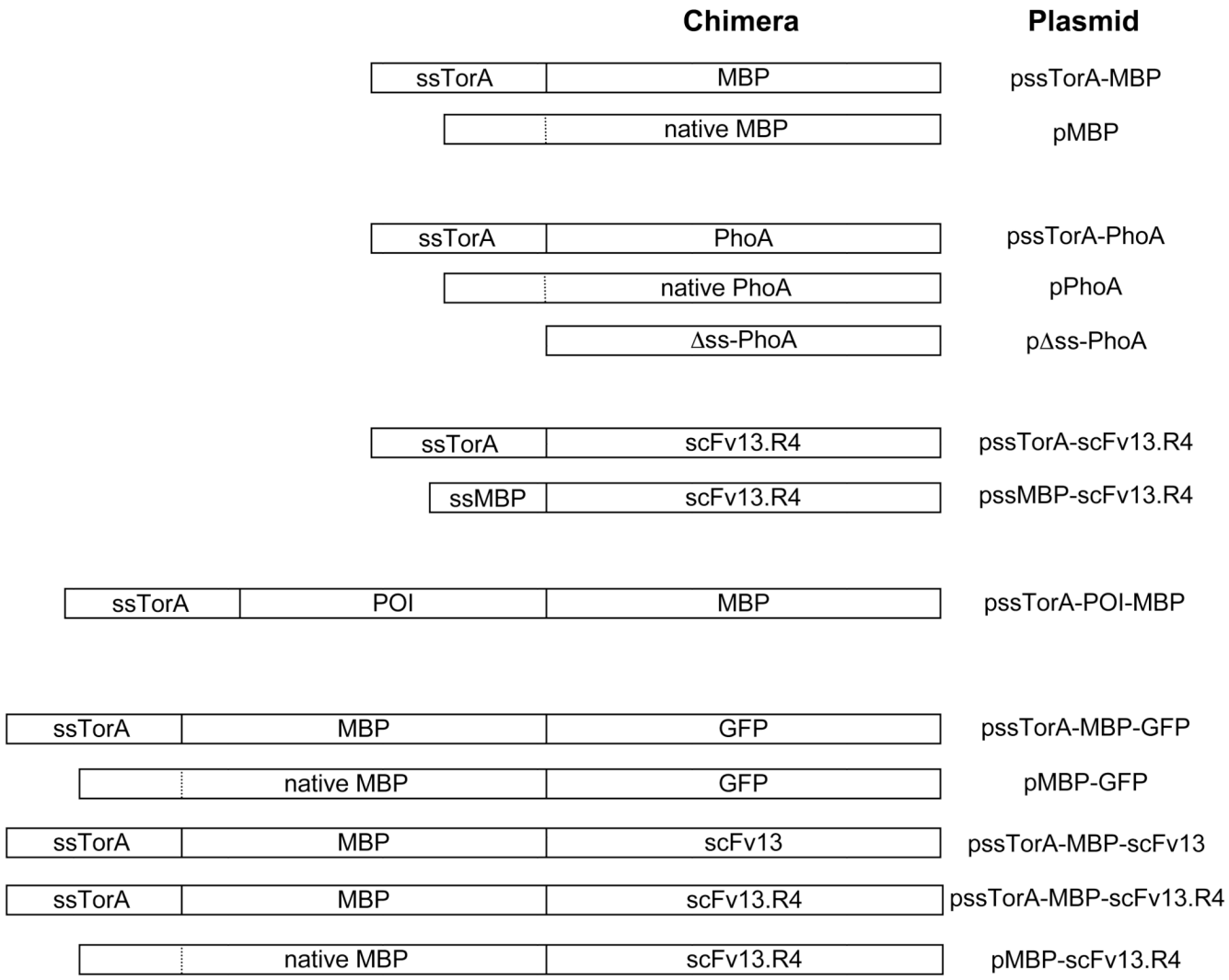


**Figure 1. Schematic of expression constructs used in this study**

Shown are the various chimeras generated for these studies along with the corresponding plasmid name (see Materials and Methods and Table 3 for more details). Dashed lines indicate cleavage site between native signal and mature domain for MBP and PhoA. POI refers to protein of interest and includes one of the following: TrxA, GST, PhoA, GFP, Top7, and TraR. Drawing is not to scale.

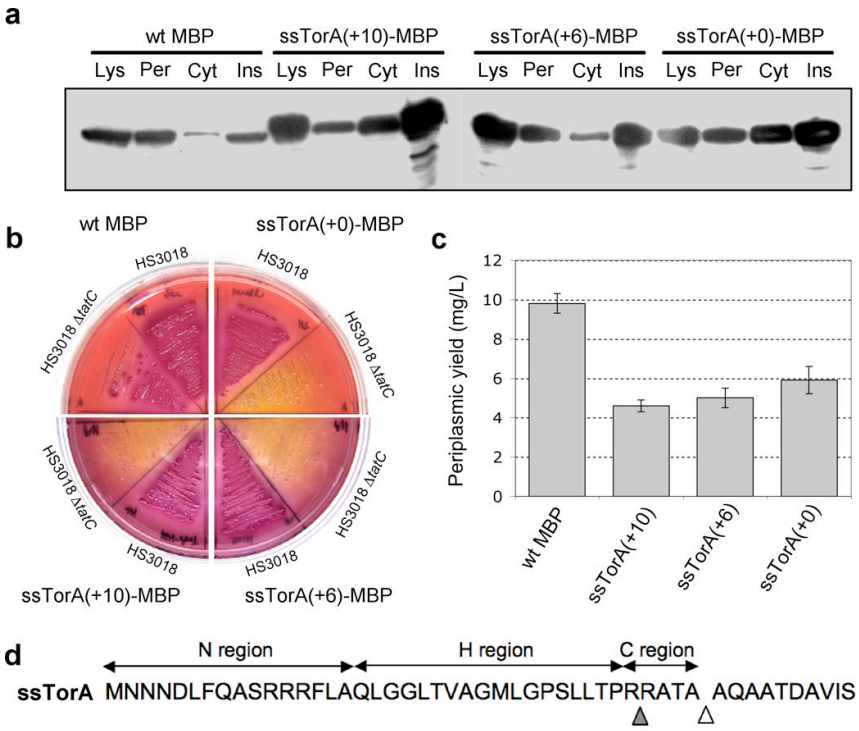


**Figure 2.**

Intracellular localization, phenotype, and yield of Tat-targeted MBP. (a) Western blot analysis of total soluble lysate (Lys), cytoplasmic (Cyt), periplasmic (Per) and insoluble (Ins) fractions from HS3018 (*malE*-deficient derivative of MC4100) cells expressing wildtype (wt) MBP and ssTorA-MBP with 0, 6 or 10 additional residues of mature TorA (+0, +6, +10). Western blots were probed with anti-MBP primary antibody. (b) HS3018 and HS3018 Δ*tatC* cells expressing the above constructs streaked on MacConkey agar media supplemented with 0.4% maltose. (c) Yield of the above constructs following IMAC purification from the periplasm of HS3018 cells. Data represents the average of 3 replicate experiments. (d) Amino acid sequence of the TorA (+10) signal peptide highlighting the primary signal peptidase cleavage site (white triangle) and the second cleavage site (grey triangle) as reported by N-terminal sequencing.

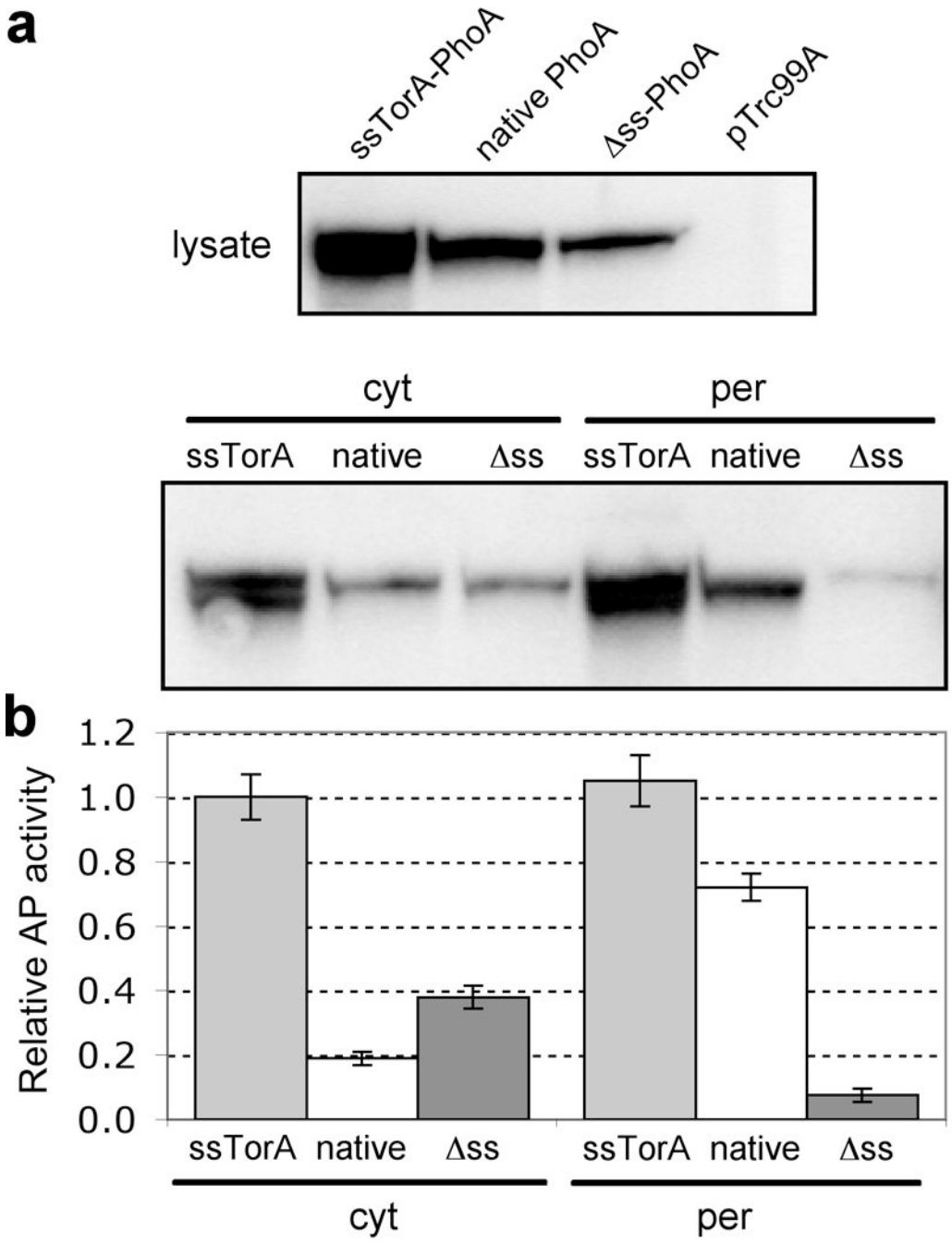


**Figure 3.**

Intracellular localization and activity of ssTorA-PhoA. (a) Western blot analysis of total soluble lysate (top panel) and cytoplasmic (cyt) and periplasmic (per) fractions (bottom panel) from FÅ113 cells expressing Tat-targeted (ssTorA), Sec-targeted (native), and cytoplasmic (Δss) PhoA and empty vector pTrc99A control cells. Western blots were probed with anti-PhoA primary antibody. (b) PhoA activity in the above cells and constructs assayed in both the cytoplasmic and periplasmic fractions. All data normalized to the activity measured for ssTorA-PhoA in the periplasmic fraction. Data represents the average of 3 replicate experiments.

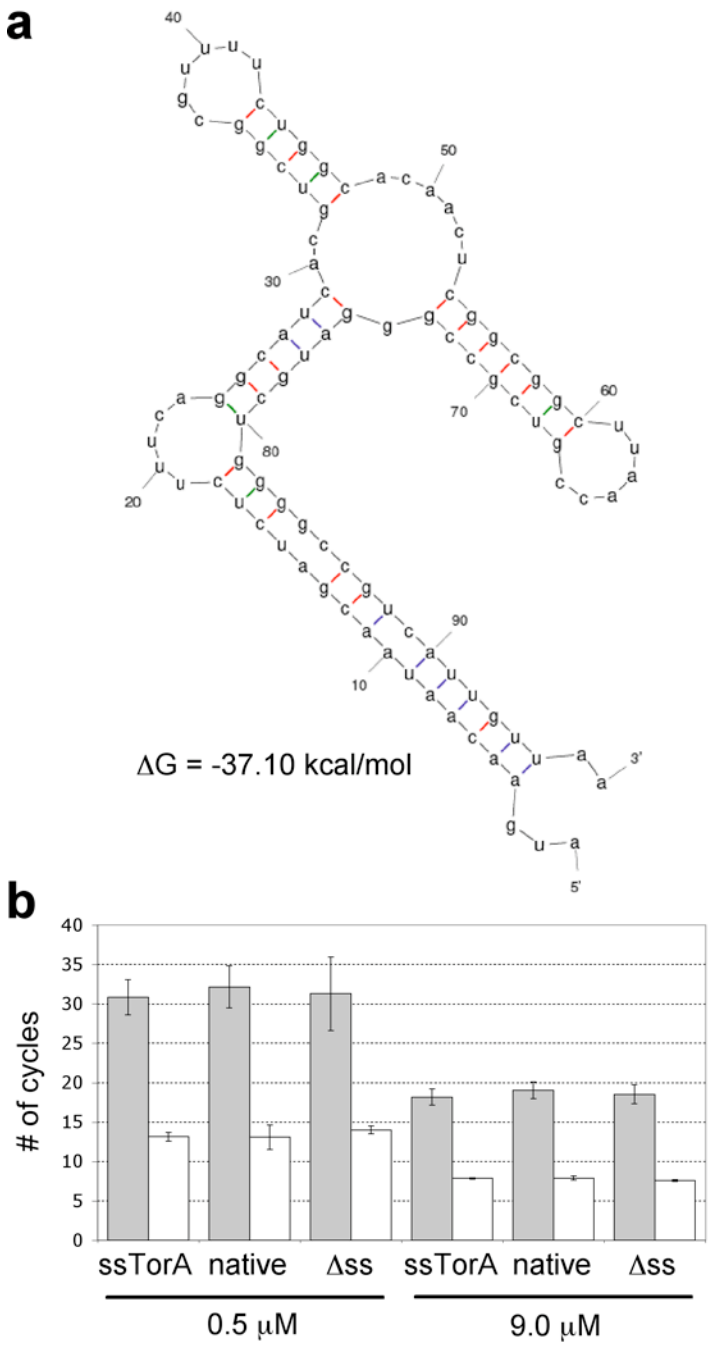


**Figure 4.**

mRNA stability analysis of PhoA constructs. (a) Predicted mRNA secondary structure and thermodynamic stability of *sstorA* transcript using the mfold program (Zuker, 2003). (b) Quantitative real-time PCR analysis of mRNA isolated from FÅ113 cells expressing Tat-targeted (ssTorA), Sec-targeted (native), and cytoplasmic (Δss) PhoA. Gray bars indicate *phoA* levels while white bars correspond to 16s rRNA controls. Data for two different primer concentrations (0.5 and 9.0 μM) is shown and is the average of 3 replicate experiments.

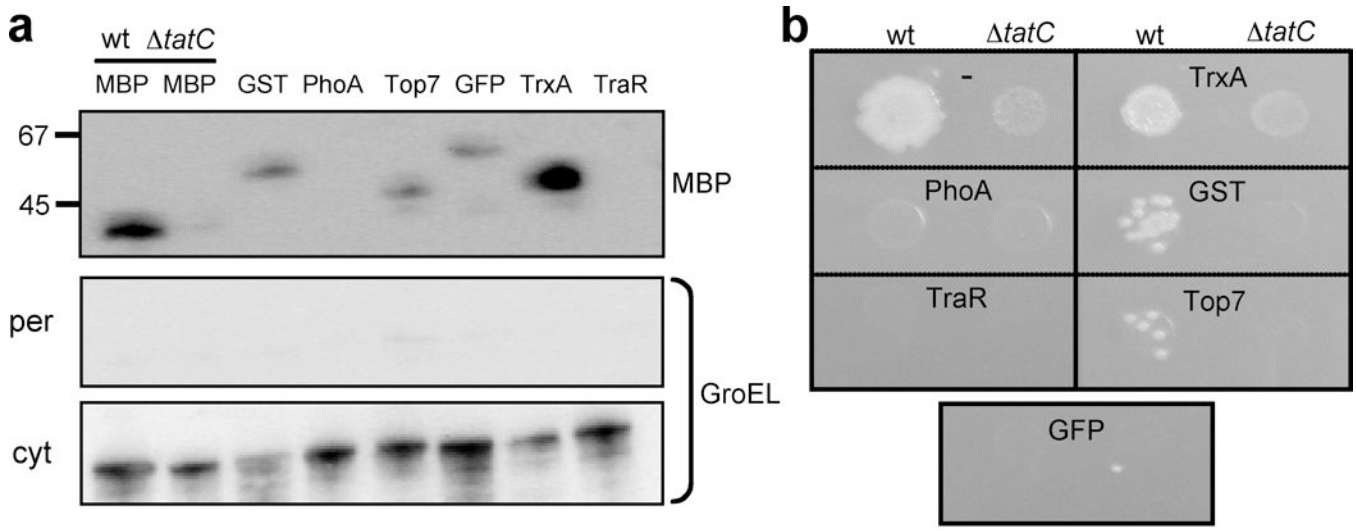


**Figure 5.**

Intracellular localization and phenotype of Tat-targeted MBP sandwich fusions. (a) Western blot analysis of periplasmic (per; top panel) and cytoplasmic (cyt; bottom panel) fractions from HS3018 and HS3018 Δ*tatC* (lane 2 only) cells expressing ssTorA-MBP (lanes 1 and 2) and ssTorA-POI-MBP sandwich fusions of the following POIs: GST, PhoA, Top7, GFP, TrxA, and TraR (lanes 3–8). Western blots were probed with anti-MBP and anti-GroEL primary antibody as indicated. GroEL was used to indicate the quality of each fractionation. (b) Spot plating of cells and constructs described above on M9 minimal agar media supplemented with 0.4% maltose.

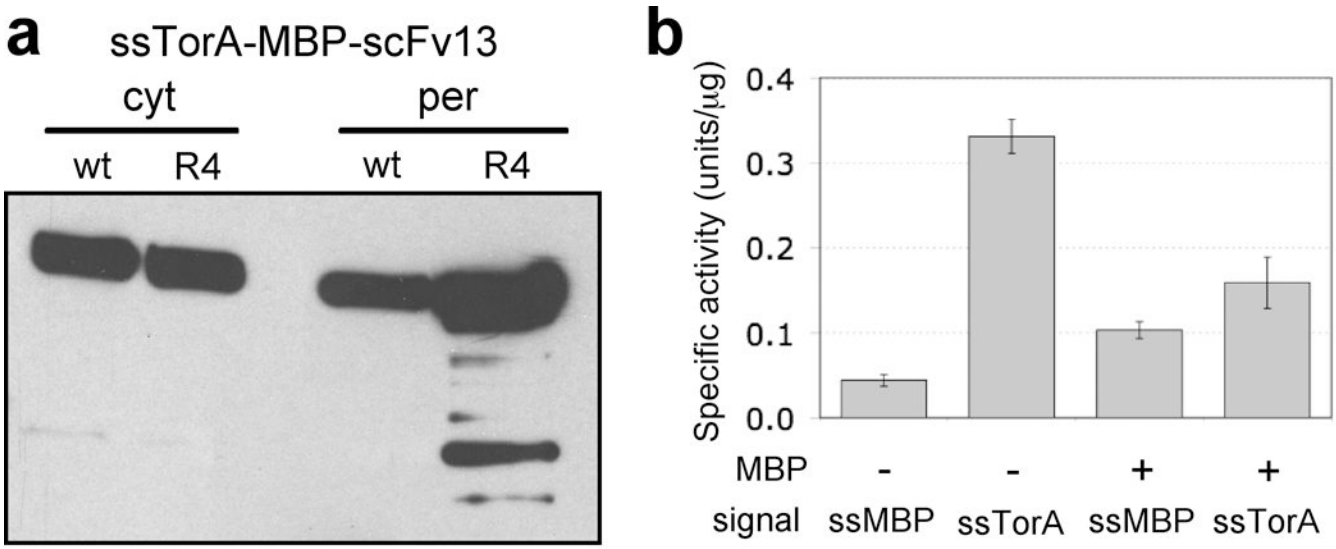


**Figure 6.**

Intracellular localization and activity of Tat-targeted MBP-scFv fusions. (a) Western blot analysis of periplasmic (per) and cytoplasmic (cyt) fractions from HS3018 cells expressing ssTorA-MBP-scFv13 and ssTorA-MBP-scFv13.R4. Western blots were probed with anti-c-myc primary antibody. (b) *In vitro* β-galactosidase activity of whole-cell lysates from *E. coli* 959 cells incubated with purified periplasmic scFv13 proteins targeted through either Sec (ssMBP) or Tat (ssTorA) as an unfused protein (−) or as a fusion to MBP (+). Each sample was assayed in triplicate and data represents the average of 3 replicate purification and activity experiments.

**Table 1**

Tat signal peptides that promote transport of MBP

| **Signal peptide**[†] | ***E. coli* HS3018** | ***E. coli* HS3018 Δ*tatC*** |
|---|---|---|
| AmiA | ++ | ++ |
| AmiC | ++ | ++ |
| CueO | ++ | − |
| DmsA | + | − |
| FdnG | + | − |
| FdoG | + | − |
| FhuD | ++ | ++ |
| HyaA | ++ | − |
| HybA | + | + |
| HybO | + | + |
| NapA | + | − |
| NapG | + | + |
| NrfC | + | + |
| SufI | ++ | − |
| TorA (RR) | ++ | − |
| TorA (KK) | − | − |
| TorZ | + | + |
| WcaM | + | − |
| YaeI | ++ | ++ |
| YahJ | + | + |
| YagT | + | − |
| YcbK | ++ | − |
| YcdB | ++ | − |
| YdcG | ++ | ++ |
| YdhX | ++ | − |
| YedY | + | + |
| YfhG | + | + |
| YnfE | ++ | − |
| YnfF | + | + |

[†] expression of each ssTat-MBP construct was from pBAD18-Cm (Tullman-Ercek *et al.*, 2007)

[‡] as determined by Tullman-Ercek et al., 2007

++ = bright red colonies; equivalent to *malE*+ cells streaked on MacConkey maltose agar plates

+ = pale red colonies

− = white colonies

Bold font = Tat specific

## Table 2

Comparison of expression yields using the Tat or Sec export pathways

| Protein of interest | Signal peptide[1] | Yield (mg/L) | Tat/Sec ratio (%)[2] | Reference |
|---|---|---|---|---|
| *E. coli* expression[3] | | | | |
| MBP | *MBP* | *9.8* | - | This study |
| | TorA(+10) | 4.6 | 47 | This study |
| | TorA(+6) | 5.0 | 51 | This study |
| | TorA(+0) | 5.9 | 60 | This study |
| GFP | *PelB* | nm | nd | This study |
| | TorA | 10–15 | nd | (Barrett *et al.*, 2003) |
| MBP-GFP | *MBP* | *0.4* | - | This study |
| | TorA | 2.6 | 650 | This study |
| scFv 26.10 | *PelB* | *1.0* | - | (Ribnicky *et al.*, 2007) |
| | TorA[4] | 0.1 | 12 | (Ribnicky *et al.*, 2007) |
| scFv13.R4 | *MBP* | *0.01* | - | This study |
| | TorA | 0.06 | 600 | This study |
| MBP-scFv13.R4 | *MBP* | *8.4* | - | This study |
| | TorA | 0.5 | 6 | This study |
| human tPA (truncated) | *StII* | *140.0*[5] | - | (Kim *et al.*, 2005) |
| | TorA | 44.8[5] | 29 | (Kim *et al.*, 2005) |
| *S. lividans* expression[6] | | | | |
| *S. lividans* XlnB1 | *XlnB1*[7] | nr | - | (Gauthier *et al.*, 2005) |
| | XlnC[8] | nr | 33 | (Gauthier *et al.*, 2005) |
| *S. lividans* XlnB2 (truncated) | *XlnB1* | nr | - | (Gauthier *et al.*, 2005) |
| | XlnC | nr | 70 | (Gauthier *et al.*, 2005) |
| *S. equisimilis* streptokinase | *Vsi*[9] | *4* | - | (Pimienta *et al.*, 2007) |
| | XlnC | 0.1–0.2 | 4 | (Pimienta *et al.*, 2007) |
| human TNFa | *Vsi* | *23.0* | - | (Schaerlaekens *et al.*, 2004) |
| | XlnC | 1.6 | 7 | (Schaerlaekens *et al.*, 2004) |
| | XyyZ[10] | 1.4 | 6 | (Schaerlaekens *et al.*, 2004) |
| human IL-2 | *Vsi* | *10.4* | - | (Schaerlaekens *et al.*, 2004) |
| | XlnC | 4.8 | 46 | (Schaerlaekens *et al.*, 2004) |

[1] All signal peptides derived from *E. coli* unless otherwise noted. Sec-dependent signal peptides are listed in italics.

[2] Tat/Sec ratio calculated by dividing yield measured for Tat expression by the value measured for Sec expression. All values represent yield obtained from periplasmic or supernatant fraction.

[3] All proteins expressed in *E. coli* were purified from the periplasm.

[4] Yields obtained using high-cell *E. coli* density fermentations.

[5] Expression via the Tat system was achieved using strain FÅ113 that has a more oxidizing cytoplasm than wildtype strains of *E. coli*

[6] All proteins expressed in *Streptomyces lividans* were secreted and purified from the medium.

[7] Sec signal peptide derived from *S. venezuelae* xylanase B1 (XlnB1).

[8] Tat-dependent signal peptide derived from *S. lividans* xylanase C (XlnC).

[9] Signal peptide derived from *S. venezuelae* CBS762.70 subtilisin inhibitor (Vsi).

[10] Signal peptide derived from *S. lividans* tyrosinase (XyyZ).

nd = not determined; nm = not measurable (protein is completely undetectable); nr = not reported.

**Table 3**

Bacterial strains and plasmids used in this study

| Strain or plasmid | Relevant genotype or features | Reference or source |
|---|---|---|
| MC4100 | F- *araD139* Δ(*argF-lac*)*U169 flbB5301 deoC1 ptsF25 relA1 rbsR22 rpsL150 thiA* | Laboratory stock |
| B1LK0 | MC4100 Δ*tatC* | (Bogsch *et al.*, 1998) |
| HS3018 | MC4100 *malT*(Con)-*1* Δ*malE444* | (Shuman, 1982) |
| HS3018 Δ*tatC* | HS3018 *tatC*::Spec | (Tullman-Ercek *et al*., 2007) |
| DHB4 | F' *lacI*$^{q}$ *pro*/λ- Δ*lacX74 galE galK thi rpsL phoR* Δ*phoA*(PvuII) Δ*malF3* | Laboratory stock |
| FÅ113 | DHB4 *gor522*…mini-Tn*10*Tet *trxB*::Kan *ahpC** | (Bessette *et al.*, 1999) |
| 959 | AMEF *lacZ* gene | (Messer and Melchers, 1978) |
| pTrc99A | *trc* promoter, ColE1 ori, Amp$^{r}$ | Amersham Biosciences |
| pMBP | Native MBP in pTrc99A | This study |
| pssTorA(+10)-MBP | ssTorA plus residues 1–10 of mature TorA fused to Δ(1–26)MBP in pTrc99A | This study |
| pssTorA(+6)-MBP | ssTorA plus residues 1–6 of mature TorA fused to Δ(1–26)MBP in pTrc99A | This study |
| pssTorA(+0)-MBP | ssTorA fused to Δ(1–26)MBP in pTrc99A | This study |
| pssTat-MBP | ssTat-Δ(1–26)MBP in pBAD18-Cm where ssTat is any putative *E. coli* Tat signal peptide | (Tullman-Ercek *et al.*, 2007) |
| pssTorA-PhoA | ssTorA-Δ(1–22)PhoA in pTrc99A | (DeLisa *et al.*, 2003) |
| pPhoA | native PhoA in pTrc99A | This study |
| pΔss-PhoA | Δ(1–22)PhoA in pTrc99A | This study |
| pssTorA-scFv13.R4 | ssTorA fused to scFv13.R4 in pTrc99A | This study |
| pssMBP-scFv13.R4 | ssMBP fused to scFv13.R4 in pTrc99A | This study |
| pTMM | ssTorA fused to Δ(1–26)MBP with a mini MCS in between the signal and MBP, cloned in pTrc99A | This study |
| pssTorA-POI-MBP | ssTorA signal peptide-POI-Δ(1–26)MBP fusion in pTrc99A | This study |
| pssTorA-MBP-GFP | ssTorA-Δ(1–26)MBP-GFP fusion in pTrc99A | This study |
| pMBP-GFP | native MBP fused to GFP in pTrc99A | This study |
| pssTorA-MBP-scFv13 | ssTorA-Δ(1–26)MBP-scFv13 in pTrc99A | This study |
| pssTorA-MBP-scFv13.R4 | ssTorA-Δ(1–26)MBP-scFv13.14 in pTrc99A | This study |
| pMBP-scFv13.R4 | native MBP fused to scFv13.R4 in pTrc99A | This study |