

# Competition between Sec- and TAT-dependent protein translocation in *Escherichia coli*

Susana Cristóbal, Jan-Willem de Gier, Henrik Nielsen<sup>1</sup> and Gunnar von Heijne<sup>2</sup>

Department of Biochemistry, Stockholm University, S-106 91 Stockholm, Sweden and <sup>1</sup>Center for Biological Sequence Analysis, Department of Biotechnology, Technical University of Denmark, DK-2800 Lyngby, Denmark

<sup>2</sup>Corresponding author  
e-mail: gunnar@biokemi.su.se

Recently, a new protein translocation pathway, the twin-arginine translocation (TAT) pathway, has been identified in both bacteria and chloroplasts. To study the possible competition between the TAT- and the well-characterized Sec translocon-dependent pathways in *Escherichia coli*, we have fused the TorA TAT-targeting signal peptide to the Sec-dependent inner membrane protein leader peptidase (Lep). We find that the soluble, periplasmic P2 domain from Lep is re-routed by the TorA signal peptide into the TAT pathway. In contrast, the full-length TorA–Lep fusion protein is not re-routed into the TAT pathway, suggesting that Sec-targeting signals in Lep can override TAT-targeting information in the TorA signal peptide. We also show that the TorA signal peptide can be converted into a Sec-targeting signal peptide by increasing the hydrophobicity of its h-region. Thus, beyond the twin-arginine motif, the overall hydrophobicity of the signal peptide plays an important role in TAT versus Sec targeting. This is consistent with statistical data showing that TAT-targeting signal peptides in general have less hydrophobic h-regions than Sec-targeting signal peptides.

**Keywords:** membrane protein/protein secretion/Sec pathway/TAT pathway

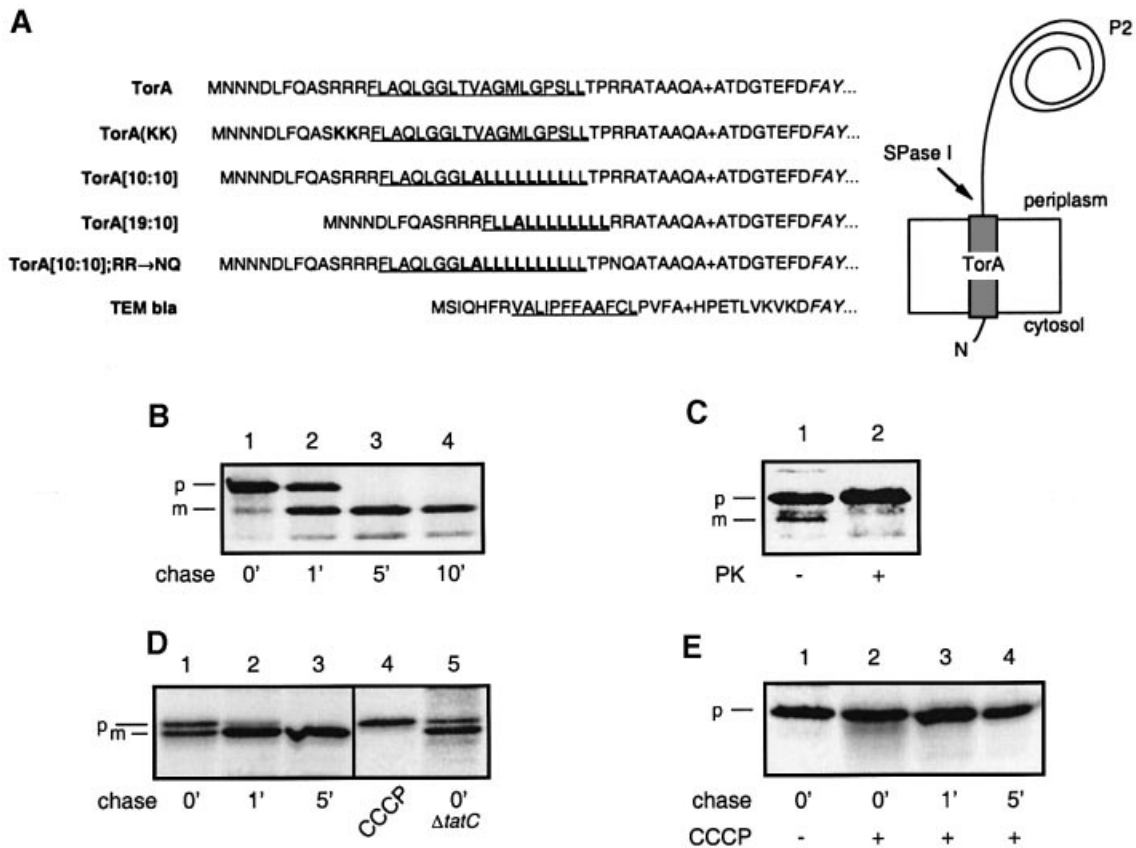
## Introduction

Although most bacterial periplasmic and membrane proteins use the well-characterized Sec machinery for translocation across or insertion into the inner membrane (Bernstein, 1998; Bibi, 1998), there is now compelling evidence for the existence of a novel translocation system in bacteria, the TAT (twin-arginine translocation) pathway, which is structurally and mechanistically similar to the thylakoid  $\Delta$ pH/TAT pathway (Robinson *et al.*, 1998; Settles and Martienssen, 1998; Dalbey and Robinson, 1999). The observation that TAT signal peptides from bacteria and thylakoids appear to be interchangeable stresses the conserved nature of the TAT system (Mori and Cline, 1998; Settles and Martienssen, 1998; Wexler *et al.*, 1998). Pre-proteins transported by the TAT pathway in most cases bind redox cofactors and seem to fold or even oligomerize before translocation across the membrane

(Santini *et al.*, 1998; Settles and Martienssen, 1998), whereas the Sec translocation machinery can only act on unfolded polypeptide chains. Interestingly, no clear example of an integral inner membrane protein targeted through the TAT pathway has been found so far.

Mutational analysis and database searches have led to the identification of genes that seem to be required only for the TAT pathway. Hcf106 from maize was the first TAT component to be identified. Hcf106 has homology to open reading frames in all fully sequenced bacterial genomes and in other organisms ranging from archaea to higher plants (Settles *et al.*, 1997; Sargent *et al.*, 1998; Settles and Martienssen, 1998; Weiner *et al.*, 1998). As far as is known, most bacterial species have two Hcf106 homologues, except for *Escherichia coli* and *Bacillus subtilis* which seem to have three (Chanal *et al.*, 1998; Sargent *et al.*, 1998). Components of the *E.coli* TAT system are encoded by the *tatABCD* operon and by *tatE*, which appears not to be part of an operon (Sargent *et al.*, 1998). TatA, TatB and TatE are the *E.coli* homologues of Hcf106 (Chanal *et al.*, 1998; Sargent *et al.*, 1998; Settles and Martienssen, 1998). Substrates of the TAT pathway have a characteristic, unusually long so-called twin-arginine signal peptide with the consensus sequence S/T-R-R-X- $\phi$ - $\phi$  ( $\phi$  is a hydrophobic residue and the twin arginines are completely invariant) in the N-terminal region (Berks, 1996). The importance of the hydrophobic residues in positions +2 and +3 relative to the two arginines has been demonstrated in thylakoid twin-arginine signal peptides (Brink *et al.*, 1998). In addition, most twin-arginine signal peptides have one or more positively charged residues, a ‘Sec-avoidance’ signal, in the C-terminal region just upstream of the signal peptidase cleavage site (Bogsch *et al.*, 1997).

Considering that signal peptides appear to be the sole arbiters of TAT versus Sec targeting, we have asked if a TAT signal peptide can re-route a typical inner membrane protein from the SecAYEG translocon-dependent pathway into the TAT pathway in *E.coli*. Using the Sec-dependent inner membrane protein leader peptidase (Lep) (Wolfe *et al.*, 1985) as a model, we show that the twin-arginine signal peptide from the TAT-dependent TorA protein targets the soluble, periplasmic P2 domain of Lep to the TAT pathway. In contrast, a fusion between the TorA signal peptide and full-length Lep is not routed through the TAT pathway, and the presence of the TorA signal peptide does not affect the membrane topology of either Lep or a Lep derivative with an ‘inverted’ topology. Thus, a TAT-targeting signal peptide can be overridden by downstream, Sec-targeting hydrophobic segments. We also show that TAT-targeting signal peptides tend to be less hydrophobic than Sec-targeting signal peptides, and that the TorA twin-arginine signal peptide can be converted into a Sec-targeting signal peptide by increasing the



**Fig. 1.** Expression of TorA/P2 and  $\beta$ -lactamase/P2 fusion proteins. (A) Signal peptides used in this study. The hydrophobic h-regions are underlined, and the signal peptidase I cleavage sites are marked by '+'. Residues in italics are from the Lep P2 domain (starting with Lep residue Phe106). The cartoon shows the orientation of the TorA/P2 fusion protein in the inner membrane. (B) MC4100 cells expressing TorA/P2 were pulse labelled with [<sup>35</sup>S]methionine for 15 s and chased with non-radioactive methionine for the indicated times. p = precursor, m = mature protein resulting from cleavage of the TorA signal peptide. (C) MC4100 cells expressing TorA/P2 were pulse labelled with [<sup>35</sup>S]methionine for 15 s, converted to spheroplasts and either incubated with proteinase K (lane 2) or left untreated (lane 1). (D) Lanes 1–3: MC4100 cells expressing  $\beta$ -lactamase/P2 were pulse labelled with [<sup>35</sup>S]methionine for 15 s and chased with non-radioactive methionine for the indicated times. Lane 4: MC4100 cells were pulse labelled as in lane 1 in the presence of CCCP. Lane 5:  $\Delta$ tatC cells were pulse labelled as in lane 1. p = precursor, m = mature protein. (E) MC4100 cells expressing TorA/P2 were pulse labelled with [<sup>35</sup>S]methionine for 15 s and chased for the indicated times in the presence (lanes 2–4) or absence (lane 1) of CCCP.

hydrophobicity of its central h-region. It thus appears that, in addition to the twin-arginine motif, overall hydrophobicity is an important determinant of TAT versus Sec targeting.

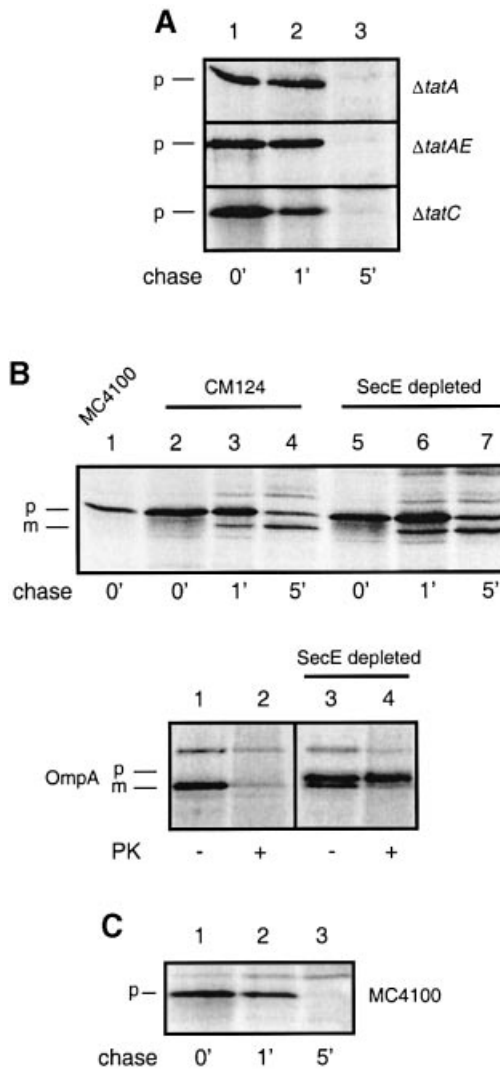
## Results

### The TorA twin-arginine signal peptide targets the Lep P2 domain to the TAT pathway

Initially, we set out to study the translocation pathway followed by a model periplasmic protein constructed by joining the periplasmic P2 domain of Lep (Wolfe *et al.*, 1983) to the twin-arginine signal peptide from the TAT-dependent TorA protein (M $\acute{e}$ jean *et al.*, 1994) and, as a control, to the signal peptide from the Sec-dependent protein  $\beta$ -lactamase (Figure 1A). In the wild-type *E. coli* strain MC4100, the TorA/P2 precursor was processed slowly (Figure 1B), presumably by signal peptidase I. Slow processing appears to be a general characteristic of the TAT pathway and has been observed previously for, for example, wild-type TorA (Santini *et al.*, 1998). The cleaved but not the uncleaved form of the protein was susceptible to proteinase K digestion of spheroplasts (Figure 1C), demonstrating the periplasmic location of the

P2 domain. Thus, the P2 domain can be targeted to the periplasm by the TorA twin-arginine signal peptide. Quantitation of the data in Figure 1B (lane 3 versus lane 1) suggests that ~35% of the molecules were exported to the periplasm and cleaved during a 5 min chase, while the remaining uncleaved precursor was degraded rapidly. The  $\beta$ -lactamase/P2 fusion protein was processed rapidly when expressed in MC4100 (Figure 1D), and the mature form but not the precursor was susceptible to proteinase K digestion of spheroplasts (data not shown), demonstrating efficient export to the periplasm. The export of both the TorA/P2 and the  $\beta$ -lactamase/P2 fusions was blocked completely by the protonophore carbonyl cyanide *m*-chlorophenylhydrazine (CCCP), which is known to inhibit both the Sec and TAT export machineries (Driessen, 1992; Santini *et al.*, 1998) (Figure 1D–E).

To determine the export pathway used by the TorA/P2 and  $\beta$ -lactamase/P2 fusion proteins, we expressed TorA/P2 in the  $\Delta$ tatA,  $\Delta$ tatAE and  $\Delta$ tatC mutant strains and in the SecE depletion strain CM124, and  $\beta$ -lactamase/P2 in the  $\Delta$ tatAE and  $\Delta$ tatC mutant strains. Export of TorA/P2 was completely blocked in the *tat* strains (Figure 2A), whereas export of the  $\beta$ -lactamase/P2 fusion was not affected in the  $\Delta$ tatC strain (Figure 1D, lane 5) or in the



**Fig. 2.** The TorA/P2 fusion protein is routed via the TAT pathway. (A)  $\Delta tatA$  (top),  $\Delta tatAE$  (middle) and  $\Delta tatC$  (bottom) cells expressing TorA/P2 were pulse labelled with [ $^{35}$ S]methionine for 15 s and chased for the indicated times. p = uncleaved precursor protein. (B) CM124 cells expressing TorA/P2 were pulse labelled with [ $^{35}$ S]methionine for 15 s and chased for the indicated times, either in the presence of arabinose to induce expression of SecE (upper panel, lanes 2–4), or after growth to mid-logarithmic phase in medium lacking arabinose to deplete SecE (upper panel, lanes 5–7). Translocation of the Sec-dependent protein OmpA after a 15 s pulse labelling was monitored in parallel by proteinase K (PK) treatment of spheroplasts (lower panel). p = precursor, m = mature protein. (C) MC4100 cells expressing TorA[KK]/P2 were pulse labelled with [ $^{35}$ S]methionine for 15 s and chased for the indicated times.

$\Delta tatAE$  strain (data not shown). We also tested a  $\Delta tatE$  strain, though, as found previously for other TAT-dependent proteins (Chanal *et al.*, 1998), the effect on export of TorA/P2 was very weak in this case (data not shown). Translocation of TorA/P2 was somewhat slowed down in CM124 (Figure 2B, upper panel), though SecE depletion per se had no effect on the translocation kinetics (lanes 2–4 versus 5–7). In contrast, translocation and signal peptide cleavage of the Sec-dependent outer membrane protein OmpA was almost completely blocked by SecE depletion (Figure 2B, lower panel). These observations strongly suggest that the TorA/P2 fusion is exported via the TAT pathway, in contrast to the  $\beta$ -lactamase/P2 fusion.

To confirm the TAT dependence of TorA/P2, we replaced the conserved arginines in the twin-arginine motif by lysines (Figure 1A), a mutation known to block TAT-dependent export (Chaddock *et al.*, 1995). As seen in Figure 2C, the TorA[KK]/P2 fusion was not exported even in the wild-type strain MC4100. Interestingly, the uncleaved precursor proteins, both in the TorA/P2 and TorA[KK]/P2 fusions, were degraded rapidly under all conditions, except when export was blocked by CCCP (Figures 1 and 2 and results not shown); we have no explanation for this phenomenon at present.

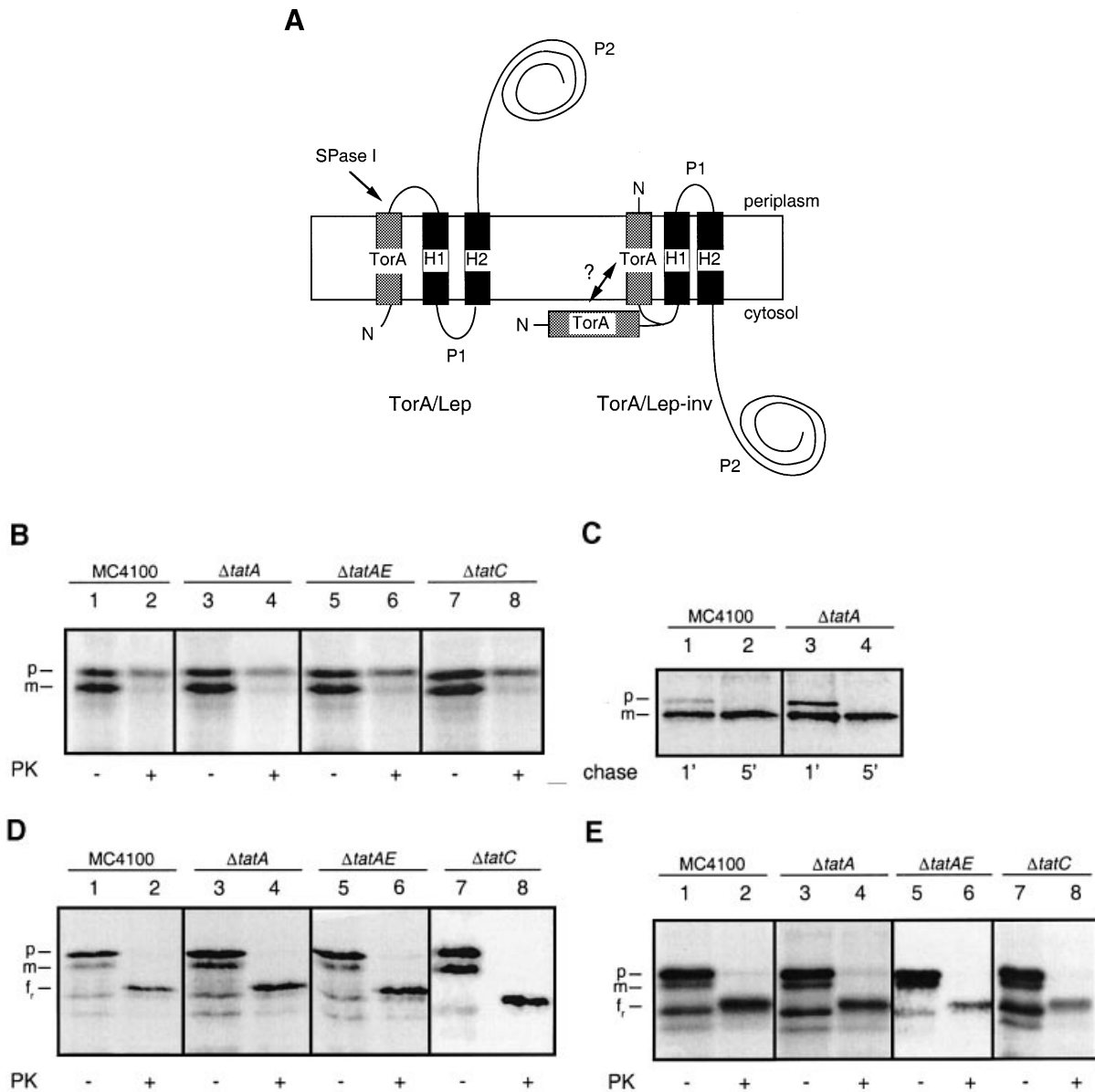
We conclude that the TorA twin-arginine signal peptide targets the P2 domain to the TAT pathway and that any non-exported precursor protein is degraded rapidly, except when the protonmotive force across the inner membrane is dissipated.

### The TorA twin-arginine signal peptide does not target full-length Lep to the TAT pathway

To assess the possible role of the TAT pathway in the assembly of inner membrane proteins, we constructed fusions between the TorA twin-arginine signal peptide and both full-length Lep and a Lep derivative with an 'inverted' membrane topology (Lep-inv) (von Heijne, 1989) (Figure 3A). As seen in Figure 3B (odd-numbered lanes), TorA/Lep was processed rapidly in both MC4100 and the different *tat* strains, demonstrating that the TAT pathway is not used in this case. This was confirmed by using the TorA[KK] mutant signal peptide; again, rapid processing of TorA[KK]/Lep was observed (Figure 3C), in contrast to what was found for TorA[KK]/P2 (Figure 2C).

Using a protease accessibility protocol, we determined the topology of TorA/Lep and TorA/Lep-inv. In spheroplasts, proteinase K degrades the periplasmic P2 domain of Lep and the periplasmic P1 loop of Lep-inv, respectively (von Heijne, 1989); in the latter case, a diagnostic protease-resistant fragment corresponding to the H2–P2 domain is produced. The presence of the TAT twin-arginine signal peptide was found not to affect the topology of either Lep or Lep-inv: the P2 domain in TorA/Lep was fully degraded (Figure 3B, even-numbered lanes), whereas a protease-resistant fragment was evident after proteolysis of spheroplasts containing TorA/Lep-inv (Figure 3D). The topologies were the same also in the various *tat* strains and for the TorA[KK]/Lep and TorA[KK]/Lep-inv constructs (Figure 3E, and data not shown). The TorA signal peptide was cleaved only in a small fraction of the TorA/Lep-inv molecules (Figure 4D–E), consistent with most of the molecules having the 'inverted'  $N_{in}$ – $C_{in}$  topology shown in Figure 3A. The small fraction of TorA/Lep-inv molecules with a cleaved signal peptide most probably have the  $N_{out}$ – $C_{out}$  topology; this fraction would have escaped detection in previous studies of Lep-inv (von Heijne, 1989; Nilsson and von Heijne, 1990).

We conclude that the TorA twin-arginine signal peptide does not target Lep and Lep-inv to the TAT pathway. Interestingly, however, the signal peptide is nevertheless cleaved, presumably by signal peptidase I, in TorA/Lep, strongly suggesting that it inserts across the inner membrane even when the TAT machinery is not involved (cf. Figure 3A).

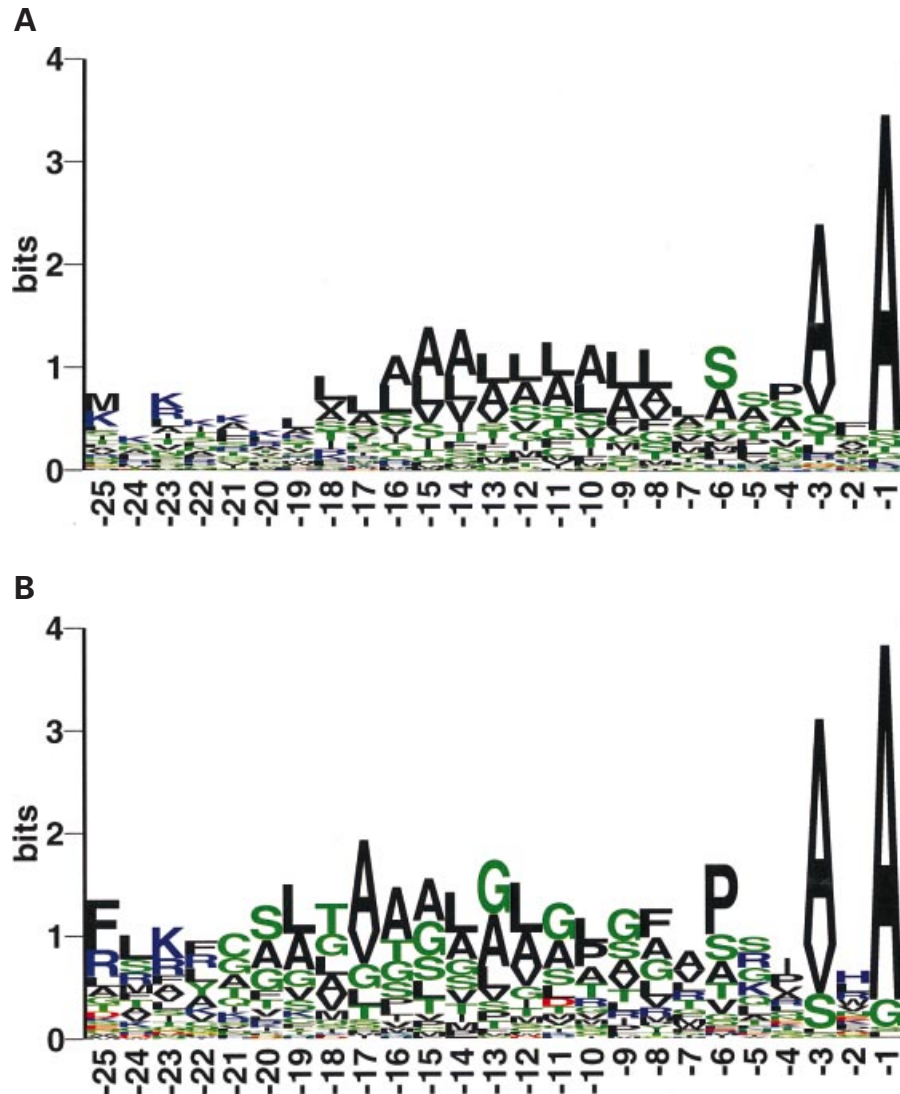


**Fig. 3.** The TorA/Lep and TorA/Lep-inv fusion proteins are not routed via the TAT pathway. (A) Membrane topology of TorA/Lep (right) and TorA/Lep-inv (left). Note that the orientation of the TorA signal peptide in TorA/Lep-inv is not known. (B) MC4100,  $\Delta tatA$ ,  $\Delta tatAE$  and  $\Delta tatC$  cells expressing TorA/Lep were pulse labelled with [ $^{35}$ S]methionine for 15 s, converted to spheroplasts and either incubated with proteinase K (even-numbered lanes) or left untreated (odd-numbered lanes). p = precursor, m = mature protein. (C) MC4100 and  $\Delta tatA$  cells expressing TorA[KK]/Lep were pulse labelled with [ $^{35}$ S]methionine for 15 s and chased for the indicated times. (D) MC4100,  $\Delta tatA$  and  $\Delta tatAE$  cells expressing TorA/Lep-inv were pulse labelled with [ $^{35}$ S]methionine for 15 s, converted to spheroplasts and either incubated with proteinase K (even-numbered lanes) or left untreated (odd-numbered lanes).  $f_r$  = protease-resistant H2-P2 fragment. (E) MC4100,  $\Delta tatA$ ,  $\Delta tatAE$  and  $\Delta tatC$  cells expressing TorA[KK]/Lep-inv were pulse labelled with [ $^{35}$ S]methionine for 15 s, converted to spheroplasts and either incubated with proteinase K (even-numbered lanes) or left untreated (odd-numbered lanes).

### **Twin-arginine signal peptides contain more glycine and less leucine than Sec-targeting signals**

The failure to target Lep to the TAT pathway by use of the TorA twin-arginine signal peptide suggested that Sec-targeting information present in the hydrophobic H1 and H2 regions in Lep might override the information in the twin-arginine signal. This prompted us to check whether Sec- and TAT-targeting signal peptides might differ in respects other than the twin-arginine motif and the Sec-avoidance signal. We thus calculated amino acid usage statistics on a data set of the two groups of signal peptides, collected as described in Materials and methods.

The difference between twin-arginine and Sec-targeting signal peptides is visualized in the sequence logos (Figure 4) where the signal peptides have been aligned from their signal peptidase I cleavage site. In the region between -18 and -8, there is a clearly visible difference: while Sec-targeting signal peptides have leucine and alanine as the most abundant residues, the twin-arginine signal peptides have a less hydrophobic region with much higher contents of glycine, but less leucine. The -3 and -1 positions next to the cleavage site are very similar, but a high occurrence of proline in position -6 seems to be a distinctive feature of the twin-arginine signal peptides.



**Fig. 4.** Sequence logos (see Materials and methods) of a set of Sec-targeting signal peptides (**A**) and twin-arginine signal peptides (**B**) from Gram-negative bacteria. Position  $-1$  is the last residue before the signal peptidase I cleavage site. In (**A**), only signal peptides longer than 24 residues were included from the data set described in Materials and methods. The height of each stack represents the degree of sequence conservation (information content) and the size of each letter is proportional to the frequency of the corresponding amino acid in that position. Amino acids are colour coded as follows: black, hydrophobic; green, polar; red, acidic; blue, basic.

Average amino acid frequencies were calculated for the three regions of the signal peptides: the N-terminal n-region, hydrophobic h-region and the c-region between the h-region and the cleavage site. The borders between the regions were assigned according to a rule allowing for variable lengths of the regions (see Materials and methods for details). In the calculation of these statistics, the twin-arginine signal peptides were weighted in order to avoid over-representation of families containing many homologous sequences. The results are shown in Table I.

Again, the most prominent feature of the twin-arginine signal peptides is the 3-fold higher proportion of glycine in the h-region, and a correspondingly lower proportion of leucine. The threonine content is also slightly higher. In the c-region, twin-arginine signal peptides have more lysine and arginine, corresponding to the ‘Sec-avoidance’ signal. They also contain 2-fold more proline and, although the difference is not significant when measured over the entire

c-region, there is significantly more proline in position  $-6$  in the twin-arginine signal peptides (data not shown).

The increased length of the twin-arginine signal peptides (38 versus 24 residues) is largely contributed by the n-region; the h-regions have the same length and the c-region is only slightly longer. The long n-regions cannot be explained solely as an effect of the twin-arginine motif being added—even the portion of the n-region upstream of the motif is approximately two residues longer than n-regions of Sec-targeting signal peptides (results not shown).

In the n-region, the twin-arginine signal peptides show a higher content of arginine, glycine and negatively charged residues, but a lower content of lysine. The arginine enrichment can be accounted for by the two arginine residues in the motif: disregarding the motif positions, both lysine and arginine are found in lower percentages than in the Sec-targeting n-regions. This is

**Table I.** Amino acid composition (in percent) of TAT- and Sec-targeting signal peptides

Residue	n-region (TAT/Sec)	h-region (TAT/Sec)	c-region (TAT/Sec)
A	6.0/4.3	29.1/25.5	24.0/30.9
C	1.1/1.0	2.0/1.7	0.6/1.5
D	4.4/0.8 <sup>a</sup>	0.0/0.1	1.6/0.3
E	4.0/0.9 <sup>a</sup>	0.0/0.1	1.2/0.5
F	5.6/3.8	4.1/7.0	2.0/3.7
G	9.1/2.8 <sup>a</sup>	13.3/4.2 <sup>a</sup>	11.3/7.5
H	0.4/2.2	0.0/0.0	3.1/2.5
I	4.2/5.0	3.6/7.9	4.1/1.5
K	8.5/23.2 <sup>a</sup>	0.1/0.1	2.2/0.5
L	7.2/6.7	14.7/26.3 <sup>a</sup>	5.9/5.8
M	1.5/2.1	3.6/4.0	0.8/1.9
N	3.5/6.4	0.1/0.5	1.5/3.2
P	1.6/3.0	1.9/1.6	9.7/4.5
Q	2.4/3.3	1.1/0.3	2.9/4.0
R	16.8/10.8	0.1/0.1	6.5/0.8 <sup>a</sup>
S	11.4/8.8	5.2/6.3	9.1/16.6
T	7.4/8.5	8.6/4.0 <sup>a</sup>	3.8/7.1
V	4.0/3.5	12.1/9.1	8.0/5.1
W	0.4/0.8	0.6/0.8	0.9/0.4
Y	0.8/2.3	0.0/0.3	0.7/1.8
Mean length	17.1/5.9	12.0/11.4	7.7/5.8

<sup>a</sup>Differences that are significant on the 0.1% level as assessed by  $\chi^2$  analysis.

related to the length difference: the approximate numbers of lysine and arginine per sequence (upstream of the motif) show no significant differences. The higher occurrence of aspartate and glutamate might simply reflect a higher tolerance for negative charge when balanced by the two arginines in the motif.

### **An increase in the hydrophobicity of the TorA twin-arginine signal peptide converts it to a Sec-targeting signal**

As shown above, twin-arginine signal peptides tend to have significantly less hydrophobic h-regions than Sec-targeting signal peptides. To test if TAT versus Sec targeting might depend on the hydrophobicity of the signal peptide, we modified the TorA signal peptide by replacing either the entire h-region with a LAL<sub>8</sub> stretch (construct TorA[19:10]/P2; see Figure 1A) or only the 10 central residues in the h-region with the same LAL<sub>8</sub> stretch (construct TorA[10:10]/P2) (Doud *et al.*, 1993). In both constructs, the critical twin-arginine motif (RRxΦΦ) and the C-terminal Sec-avoidance motif (RR) were retained. The signal peptides in TorA[19:10]/P2 and TorA[10:10]/P2 were both cleaved (Figure 5A) and the P2 domain was located in the periplasm as seen by its protease sensitivity in spheroplasts. For both constructs, ~60% of the precursor protein was converted to the cleaved form during a 5 min chase. In contrast to what was found for the TorA/P2 fusion, however, the uncleaved precursor was not degraded during the chase. Strikingly, none of the *tat* deletion mutants affected the export kinetics to any significant degree (Figure 5B). Upon depletion of SecE, on the other hand, translocation was completely eliminated (Figure 5C; compare lanes 3 and 6). As a last control, we changed the two arginines in the twin-arginine motif in TorA[10:10]/P2 to lysines in order to exclude further the possibility of routing via the TAT pathway (Chaddock *et al.*, 1995). The

TorA(KK-[10:10]) signal peptide was as effective as the original TorA[10:10] signal peptide in exporting the P2 domain of Lep in MC4100 (Figure 5D).

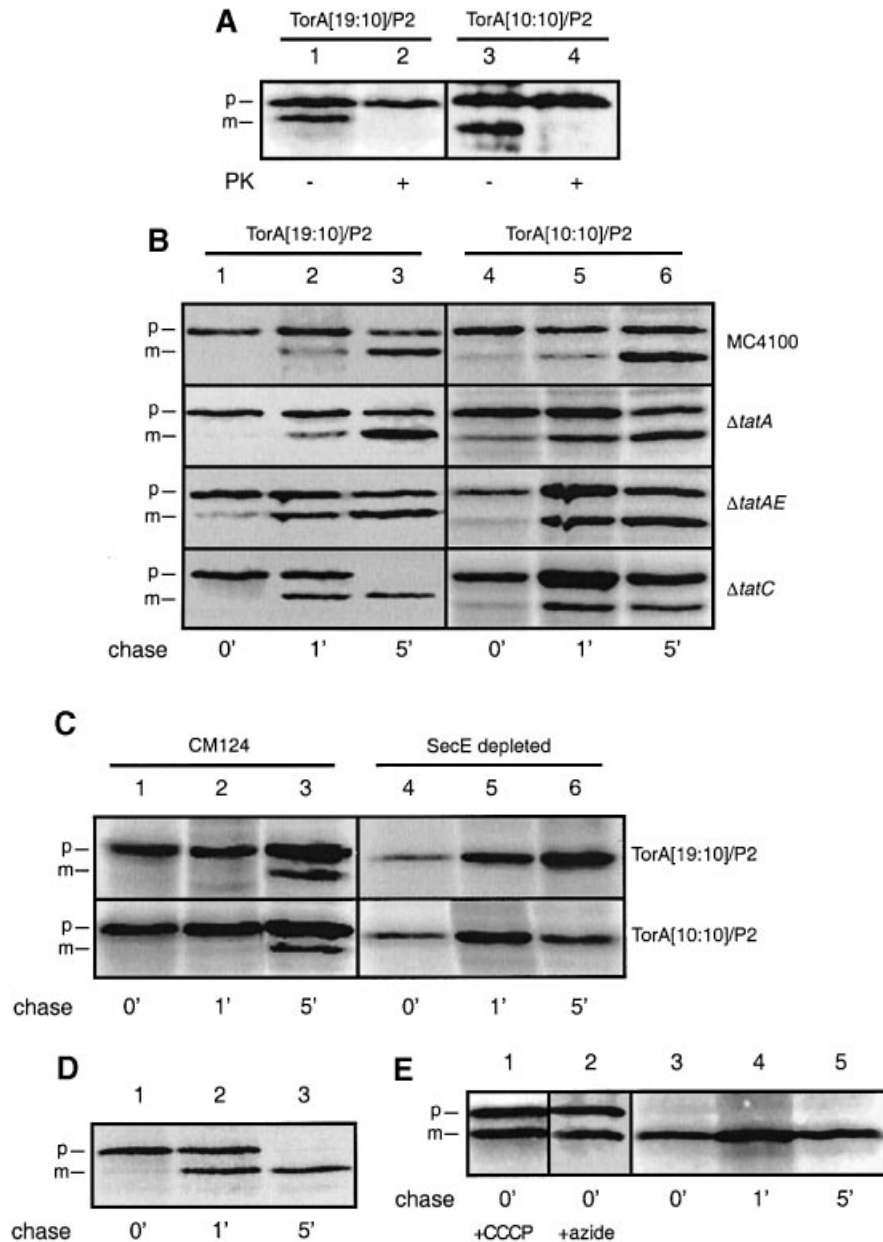
Finally, we considered the possibility that the persistence of some uncleaved, non-exported precursor protein even in MC4100 might be caused by the 'Sec-avoidance' signal still present at the C-terminal end of the TorA[10:10] signal peptide. The two arginines in positions -7 and -8 upstream of the signal peptidase cleavage site were thus changed to asparagine-glutamine in TorA[10:10]/P2 (see Figure 1A). In this case, rapid and complete export was seen (Figure 5E). Both dissipation of the protonmotive force by CCCP and inhibition of the SecA ATPase by sodium azide (Oliver *et al.*, 1990) partially blocked export, consistent with Sec translocon-dependent translocation.

Taken together, these observations clearly demonstrate that an increase in the hydrophobicity of the TorA signal peptide is sufficient to convert it from a TAT-targeting to a Sec-targeting signal peptide, and that removal of the 'Sec-avoidance' signal near its C-terminal end further improves the translocation efficiency.

## **Discussion**

In the present study, we have analysed the competition between the TAT and Sec translocon-dependent protein translocation pathways for a substrate protein containing both a twin-arginine signal peptide and a Sec-targeting signal. The TAT pathway in *E. coli* so far is defined by the TatA, TatB, TatE and TatC proteins (Chanal *et al.*, 1998), while Sec-dependent proteins all use the SecAYEG translocon and are targeted by either the cytoplasmic chaperone SecB or the Ffh-4.5S RNA signal recognition particle (SRP) (de Gier *et al.*, 1997). Using  $\Delta$ *tatA*,  $\Delta$ *tatAE* and  $\Delta$ *tatC* deletion strains, we have found that the TorA twin-arginine signal peptide can direct the soluble, periplasmic P2 domain of the normally Sec-dependent inner membrane protein Lep to the TAT translocation pathway (Figures 1 and 2). Further, mutation of the two critical arginines in the TorA signal peptide to lysines completely blocked export of TorA/P2 (Figure 2C). The TorA twin-arginine signal peptide is thus unable to target the P2 domain to the Sec pathway, even under conditions where the TAT pathway is non-functional. In contrast, the signal peptide from the Sec-dependent protein  $\beta$ -lactamase does not route the P2 domain into the TAT pathway (Figure 1), showing that the P2 domain itself does not influence the choice of translocation pathway.

While the TorA twin-arginine signal peptide thus targets the P2 domain to the TAT pathway, similar fusions to two full-length Lep constructs where both transmembrane domains are present resulted in proteins that do not depend on the TAT machinery for insertion (Figure 3). Nevertheless, it is clear from the efficient cleavage of the TorA signal peptide in the TorA/Lep fusion that the signal peptide is inserted across the membrane such that it can be processed by signal peptidase I also in this case. Further, the membrane topology of the two mature fusion proteins (N<sub>out</sub>-C<sub>out</sub> for TorA/Lep and N<sub>in</sub>-C<sub>in</sub> for TorA/Lep-inv) is the same as when these proteins lack the TorA signal peptide. Thus, the Sec-targeting information present in the two hydrophobic transmembrane segments in Lep (Wolfe *et al.*, 1985) apparently overrides the TAT-targeting



**Fig. 5.** The TorA[19:10]/P2 and TorA[10:10]/P2 fusion proteins are not routed via the TAT pathway. **(A)** MC4100 cells expressing TorA[19:10]/P2 (lanes 1 and 2) and TorA[10:10]/P2 (lanes 3 and 4) were pulse labelled with [<sup>35</sup>S]methionine for 15 s, converted to spheroplasts and either incubated with proteinase K (even-numbered lanes) or left untreated (odd-numbered lanes). p = precursor, m = mature protein. **(B)** MC4100,  $\Delta$ tatA,  $\Delta$ tatAE and  $\Delta$ tatC cells expressing TorA[19:10]/P2 (lanes 1–3) and TorA[10:10]/P2 (lanes 4–6) were pulse labelled with [<sup>35</sup>S]methionine for 15 s and chased for the indicated times. **(C)** CM124 cells expressing TorA[19:10]/P2 (top) and TorA[10:10]/P2 (bottom) were pulse labelled with [<sup>35</sup>S]methionine for 15 s and chased for the indicated times, either in the presence of arabinose to induce expression of SecE (lanes 1–3), or after growth to mid-logarithmic phase in medium lacking arabinose to deplete SecE (lanes 4–6). **(D)** MC4100 cells expressing TorA(KK-[10:10])/P2 were pulse labelled with [<sup>35</sup>S]methionine for 15 s and chased for the indicated times. **(E)** MC4100 cells expressing TorA[10:10]; RR→NQ/P2 were pulse labelled with [<sup>35</sup>S]methionine for 15 s and chased for the indicated times (lanes 3–5). In lanes 1 and 2, cells were treated with CCCP and sodium azide, respectively, before pulse labelling.

information in the TorA twin-arginine signal peptide, showing that it is not necessarily the most N-terminal signal that determines into which targeting pathway a protein is funnelled. Our findings also raise the question of whether it is at all possible to target integral inner membrane proteins to the TAT pathway; indeed, we know of no clear case of a TAT-dependent inner membrane protein.

Given the above observations and the fact that twin-arginine signal peptides in general have less hydrophobic h-regions than Sec-targeting signal peptides (Figure 4),

we also tested whether an increase in the hydrophobicity of the TorA signal peptide had any effect on the choice of translocation pathway. Surprisingly, using two differently modified versions of the TorA signal peptide, we found that translocation of the P2 domain of Lep was not affected in *tat* deletion strains, whereas its export was blocked by SecE depletion (which also results in strongly reduced levels of SecY; Yang *et al.*, 1997) (Figure 5). Substitution of the two arginines in the modified TorA signal peptide by two lysines did not change the translocation characteristics, indicating that the twin-arginine motif fails to target to

the TAT pathway if the h-region of the signal peptide is too hydrophobic.

We conclude that the twin-arginine motif and a weakly hydrophobic h-region are essential characteristics of a TAT signal peptide, and that an increase in the hydrophobicity of a twin-arginine signal peptide can convert it into a substrate for the Sec pathway while at the same time preventing export through the TAT pathway. Possibly, the increase in hydrophobicity may weaken the effect of the 'Sec-avoidance' signal (Bogsch *et al.*, 1997) contributed by the two C-terminal arginines in the TorA signal peptide, thus allowing export through the Sec pathway; indeed, the additional substitution of these two arginines by uncharged polar residues leads to rapid and complete Sec-dependent export (Figure 5).

It is interesting to note the importance of the hydrophobicity of the signal peptide for TAT versus Sec targeting, as overall signal peptide hydrophobicity has also been found to affect whether targeting to the Sec translocon is mediated by SecB or SRP (de Gier *et al.*, 1997). Although we have not addressed the issue of whether the Sec-dependent proteins studied here use SecB or SRP for targeting to the SecAYEG translocon, it would appear that a previously unsuspected level of fine tuning of signal peptide hydrophobicity can have profound consequences for protein targeting in *E. coli*.

## Materials and methods

### Enzymes and chemicals

Unless otherwise stated, all enzymes were from Promega. For PCR, the DNA polymerase long expand template system from Boehringer Mannheim was used. Deoxyribonucleotides, CCCP and [<sup>35</sup>S]methionine were from Amersham-Pharmacia. Proteinase K was from Gibco-BRL. Oligonucleotides were from Kebo Lab. Hen egg white lysozyme and phenylmethylsulfonyl fluoride (PMSF) were from Sigma.

### Strains, plasmids and growth conditions

Strain MC4100 (Casadaban, 1976) and its derivative  $\Delta$ tatA,  $\Delta$ tatE,  $\Delta$ tatAE and  $\Delta$ tatC mutant strains (Bogsch *et al.*, 1998; Sargent *et al.*, 1998) were cultured in M9 minimal medium supplemented with 0.2% glucose. The SecE depletion strain CM124 (Traxler and Murphy, 1996) was cultured in M9 minimal medium with 0.4% glucose and 0.2% L-arabinose (de Gier *et al.*, 1998). Overnight cultures of CM124 were washed once with M9 medium and backdiluted 1:20. To deplete the CM124 cells of SecE, cells were grown to mid-logarithmic phase in the absence of L-arabinose (Traxler and Murphy, 1996; Yang *et al.*, 1997). Depletion of SecE was checked by monitoring the accumulation of the pro-form of outer membrane protein A (pro-OmpA) during a short pulse labelling with [<sup>35</sup>S]methionine (results not shown). Where appropriate, ampicillin (final concentration 100 µg/ml) and kanamycin (final concentration, 50 µg/ml) were added to the medium.

TorA/Lep, TorA/Lep-inv and TorA/P2 fusions were constructed using an overlap PCR approach. The region coding for the TorA signal peptide plus the first eight amino acids of mature TorA were fused to the complete coding regions of Lep, Lep-inv and the P2 domain of Lep (Lep<sub>106-323</sub>). The β-lactamase/P2 fusion was generated by fusing the region coding for the β-lactamase signal peptide plus the first 10 amino acids of the mature protein to the coding region of the P2 domain of Lep (Lep<sub>106-323</sub>). The fusions were cloned into the expression vectors pBAD18, pBAD24 (Guzman *et al.*, 1995) and pDHB5700 (de Gier *et al.*, 1998).

The two invariant arginines in the n-region of the TorA signal peptide were changed to lysines by means of site-directed mutagenesis using the Quickchange approach (Stratagene) yielding the TorA[KK] signal peptide. The same technique was used to make the TorA[10:10]; RR→NQ mutant.

Using a PCR overlap approach, TorA signal peptide mutants with an increased hydrophobicity were constructed by replacing either the entire 19 residue h-region or only the 10 central residues from the h-region

with a strongly hydrophobic h-region consisting of one alanine and nine leucines (LAL<sub>8</sub>) (Doud *et al.*, 1993), yielding TorA[19:10] and TorA[10:10], respectively (Figure 1A).

### Assay for membrane targeting

For all experiments, cells were grown to mid-logarithmic phase. In all strains except CM124, expression was induced from the pBAD vectors with L-arabinose (final concentration 0.2%). In strain CM124, expression was induced from the pDHB5700 vector with isopropyl-β-D-thiogalactopyranoside (IPTG, final concentration 1 mM). Cells were labelled with [<sup>35</sup>S]methionine (150 µCi/ml, Ci = 37 GBq) for 15 s whereupon non-radioactive methionine was added (final concentration 500 µg/ml). In experiments where the role of the protonmotive force was studied, CCCP (0.1 mM final concentration) was added 1 min before labelling with [<sup>35</sup>S]methionine. For inhibition of the SecA ATPase, sodium azide (2 mM final concentration) was added 1 min before labelling. After labelling, cells were either precipitated directly with trichloroacetic acid (TCA; final concentration 10%) or converted to spheroplasts. For spheroplasting, cells were collected at 14 000 r.p.m. for 2 min in a microfuge, resuspended in ice-cold buffer (40% w/v sucrose, 33 mM Tris pH 8.0) and incubated with lysozyme (final concentration 5 µg/ml) and 1 mM EDTA for 15 min on ice. Aliquots of the spheroplast suspension were incubated on ice for 1 h either in the presence or absence of proteinase K (final concentration 0.3 mg/ml). PMSF was added to the spheroplast suspensions (final concentration 0.33 mg/ml) and subsequently the spheroplasts were TCA precipitated (final concentration 10%). After TCA precipitation, the pellet was resuspended in 10 mM Tris/2% SDS, immunoprecipitated with antisera to Lep, OmpA (a periplasmic control), AraB/bandX [a cytoplasmic control (de Gier *et al.*, 1996), results not shown], washed and analysed by standard SDS-PAGE (Laemmli, 1970). Gels were scanned on a Fuji BAS1000 phosphorimager and quantitated using the MacBAS software (version 2.31).

### Sequence analysis

The data set of TAT-targeting signal peptides was an extended version of the list given by Berks (1996), provided as supplementary material at <http://www.blackwell-science.com/products/journals/contents/berks.htm>. From those 90 sequences, we removed the proposed membrane-bound Rieske proteins (group f), the sequences not belonging to Gram-negative bacteria, and one unusually short transfer peptide (*Synechococcus* NrTA), leaving 72 sequences. Within this set, there are several pairs of close homologues from different species, so to avoid heavily biased amino acid usage statistics (Table I), the amino acid counts from each sequence were divided by the number of sequences in the corresponding sequence family as given by Berks (1996).

The data set of Sec-targeting signal peptides from Gram-negative bacteria was extracted from SWISS-PROT release 35 (Bairoch and Apweiler, 1997) as described in Nielsen *et al.* (1996). To avoid putative TAT transfer peptides, any sequences with two consecutive arginines in the signal peptide were removed, leaving 320 sequences. Since a rigorous homology reduction was used while extracting this data set, no weighting scheme was necessary. Compositional differences between TAT- and Sec-targeting signal peptides were identified by  $\chi^2$  analysis. First, the compositional differences for the n-, h- and c-regions were tested in a 20×2 contingency table and found to be strongly significant for all three regions ( $P < 0.001$ , 19 d.f.). Then, the residues contributing to the difference were identified by a series of 2×2 contingency tables where, for each amino acid type  $i$  ( $i = 1-20$ ), the number of residues of type  $i$  and non- $i$  in the two sets were compared with the number expected from the null hypothesis. Only strongly significant differences ( $P < 0.001$ , 1 d.f.) were reported. For the TAT-targeting signal peptides, the weighted counts were used, reducing the effective number of sequences in this category to 21 (if calculated without weighting, the redundancy would lead to strongly overestimated  $\chi^2$  values).

Sequence logos were constructed according to Schneider and Stephens (1990). Briefly, sequence logos combine the information contained in consensus sequences with a quantitative measure of information, by representing each position in an alignment by a stack of letters. The height of the stack is a measure of the non-randomness (i.e. information content) at the position, while the height of a letter corresponds to its frequency. The logos were made with the original data set, i.e. no weighting was employed.

In each sequence, the n-, h- and c-regions were defined according to the following set of rules, designed to correspond roughly to an intuitive concept of 'hydrophobic region': (i) place a pointer at the -1 position (immediately before the cleavage site), set the assignment to c-region and scan the sequence upstream towards the N-terminus; (ii) move the



pointer three positions upstream (assigning -1 through -3 as a minimal c-region); (iii) set the assignment to h-region at the first occurrence of at least two consecutive hydrophobic residues (alanine, isoleucine, leucine, methionine, phenylalanine, tryptophan or valine); (iv) move the pointer six positions upstream; (v) set the assignment to n-region at the first occurrence of either a charged residue or at least three consecutive non-hydrophobic residues; (vi) if the N-terminal end of the h-region is not a hydrophobic residue, move the pointer back downstream, changing the assignment to n-region until a hydrophobic residue is found.

## Acknowledgements

The *tat* deletion strains were kindly provided by Drs C.Robinson and T.Palmer. This work was supported by grants from the Swedish Natural and Technical Sciences Research Councils, the Swedish Cancer Foundation and the Göran Gustafsson Foundation to G.v.H. S.C. is the recipient of a postdoctoral fellowship from Basque Country Government. J.W.d.G. is the recipient of a TMR fellowship from the EC. H.N. is supported by the Danish National Research Foundation.

## References

- Bairoch,A. and Apweiler,R. (1997) The SWISS-PROT protein sequence data bank and its supplement TrEMBL. *Nucleic Acids Res.*, **25**, 31–36.
- Berks,B.C. (1996) A common export pathway for proteins binding complex redox cofactors? *Mol. Microbiol.*, **22**, 393–404.
- Bernstein,H. (1998) Membrane protein biogenesis: the exception explains the rules. *Proc. Natl Acad. Sci. USA*, **95**, 14587–14589.
- Bibi,E. (1998) The role of the ribosome–translocon complex in translation and assembly of polytopic membrane proteins. *Trends Biochem. Sci.*, **23**, 51–55.
- Bogsch,E., Brink,S. and Robinson,C. (1997) Pathway specificity for a ΔpH-dependent precursor thylakoid lumen protein is governed by a ‘Sec-avoidance’ motif in the transfer peptide and a ‘Sec-incompatible’ mature protein. *EMBO J.*, **16**, 3851–3859.
- Bogsch,E.G., Sargent,F., Stanley,N.R., Berks,B.C., Robinson,C. and Palmer,T. (1998) An essential component of a novel bacterial protein export system with homologues in plastids and mitochondria. *J. Biol. Chem.*, **273**, 18003–18006.
- Brink,S., Bogsch,E.G., Edwards,W.R., Hynds,P.J. and Robinson,C. (1998) Targeting of thylakoid proteins by the ΔpH-driven twin-arginine translocation pathway requires a specific signal in the hydrophobic domain in conjunction with the twin-arginine motif. *FEBS Lett.*, **434**, 425–430.
- Casadaban,M. (1976) Transposition and fusion of the *lac* genes to selected promoters in *Escherichia coli* using bacteriophage λ and Mu. *J. Mol. Biol.*, **104**, 541–555.
- Chaddock,A.M., Mant,A., Karnauchoy,I., Brink,S., Herrmann,R.G., Klösgen,R.B. and Robinson,C. (1995) A new type of signal peptide: central role of a twin-arginine motif in transfer signals for the ΔpH-dependent thylakoidal protein translocase. *EMBO J.*, **14**, 2715–2722.
- Chanal,A., Santini,C. and Wu,L. (1998) Potential receptor function of three homologous components, TatA, TatB and TatE, of the twin-arginine signal sequence-dependent metalloenzyme translocation pathway in *Escherichia coli*. *Mol. Microbiol.*, **30**, 674–676.
- Dalbey,R. and Robinson,C. (1999) Protein translocation into and across the bacterial plasma membrane and the plant thylakoid membrane. *Trends Biochem. Sci.*, **24**, 17–22.
- de Gier,J.-W.L., Mansournia,P., Valent,Q., Phillips,G.J., Luirink,J. and von Heijne,G. (1996) Assembly of a cytoplasmic membrane protein in *Escherichia coli* is dependent on the signal recognition particle. *FEBS Lett.*, **399**, 307–309.
- de Gier,J.-W.L., Valent,Q., von Heijne,G. and Luirink,J. (1997) The *E.coli* SRP: preferences of a targeting factor. *FEBS Lett.*, **408**, 1–4.
- de Gier,J.-W.L., Scotti,P.A., Sääf,A., Valent,Q.A., Kuhn,A., Luirink,J. and von Heijne,G. (1998) Differential use of the signal recognition particle translocase targeting pathway for inner membrane protein assembly in *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **95**, 14646–14651.
- Doud,S.K., Chou,M.M. and Kendall,D.A. (1993) Titration of protein transport activity by incremental changes in signal peptide hydrophobicity. *Biochemistry*, **32**, 1251–1256.
- Driessen,A.J.M. (1992) Bacterial protein translocation—kinetic and thermodynamic role of ATP and the protonmotive force. *Trends Biochem. Sci.*, **17**, 219–223.
- Guzman,L.M., Belin,D., Carson,M.J. and Beckwith,J. (1995) Tight regulation, modulation and high-level expression by vectors containing the arabinose pBAD promoter. *J. Bacteriol.*, **177**, 4121–4130.
- Laemmli,U.K. (1970) Cleavage of structural proteins during assembly of bacteriophage T4. *Nature*, **227**, 680–685.
- Méjean,V., Iobbi-Nivol,C., Lepelletier,M., Giordano,G., Chippaux,M. and Pascal,M.C. (1994) TMAO anaerobic respiration in *Escherichia coli*: involvement of the *tor* operon. *Mol. Microbiol.*, **11**, 1169–1179.
- Mori,H. and Cline,K. (1998) A signal peptide that directs non-Sec transport in bacteria also directs efficient and exclusive transport on the thylakoid ΔpH pathway. *J. Biol. Chem.*, **273**, 11405–11408.
- Nielsen,H., Engelbrecht,J., von Heijne,G. and Brunak,S. (1996) Defining a similarity threshold for a functional protein sequence pattern: the signal peptide cleavage site. *Protein Struct. Funct. Genet.*, **24**, 165–177.
- Nilsson,I.M. and von Heijne,G. (1990) Fine-tuning the topology of a polytopic membrane protein. Role of positively and negatively charged residues. *Cell*, **62**, 1135–1141.
- Oliver,D.B., Cabelli,R.J., Dolan,K.M. and Jarosik,G.P. (1990) Azide-resistant mutants of *Escherichia coli* alter the SecA protein, an azide-sensitive component of the protein export machinery. *Proc. Natl Acad. Sci. USA*, **87**, 8227–8231.
- Robinson,C., Hynds,P.J., Robinson,D. and Mant,A. (1998) Multiple pathways for the targeting of thylakoid proteins in chloroplasts. *Plant Mol. Biol.*, **38**, 209–221.
- Santini,C.L., Ize,B., Chanal,A., Müller,M., Giordano,G. and Wu,L.F. (1998) A novel Sec-independent periplasmic protein translocation pathway in *Escherichia coli*. *EMBO J.*, **17**, 101–112.
- Sargent,F., Bogsch,E.G., Stanley,N.R., Wexler,M., Robinson,C., Berks,B.C. and Palmer,T. (1998) Overlapping functions of components of a bacterial Sec-independent protein export pathway. *EMBO J.*, **17**, 3640–3650.
- Schneider,T. and Stephens,R. (1990) Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res.*, **18**, 6097–6100.
- Settles,A. and Martienssen,R. (1998) Old and new pathways of protein export in chloroplasts and bacteria. *Trends Cell Biol.*, **8**, 494–501.
- Settles,A., Yonetani,A., Baron,A., Bush,D., Cline,K. and Martienssen,R. (1997) Sec-independent protein translocation by the maize Hcf106 protein. *Science*, **278**, 1467–1470.
- Traxler,B. and Murphy,C. (1996) Insertion of the polytopic membrane protein MalF is dependent on the bacterial secretion machinery. *J. Biol. Chem.*, **271**, 12394–12400.
- von Heijne,G. (1989) Control of topology and mode of assembly of a polytopic membrane protein by positively charged residues. *Nature*, **341**, 456–458.
- Weiner,J., Bilous,P., Shaw,G., Lubitz,S., Frost,L., Thomas,G., Cole,J. and Turner,R. (1998) A novel and ubiquitous system for membrane targeting and secretion of cofactor-containing proteins. *Cell*, **93**, 93–101.
- Wexler,M., Bogsch,E.G., Klösgen,R.B., Palmer,T., Robinson,C. and Berks,B.C. (1998) Targeting signals for a bacterial Sec-independent export system direct plant thylakoid import by the ΔpH pathway. *FEBS Lett.*, **431**, 339–342.
- Wolfe,P.B., Wickner,W. and Goodman,J.M. (1983) Sequence of the leader peptidase gene of *Escherichia coli* and the orientation of leader peptidase in the bacterial envelope. *J. Biol. Chem.*, **258**, 12073–12080.
- Wolfe,P.B., Rice,M. and Wickner,W. (1985) Effects of two *sec* genes on protein assembly into the plasma membrane of *Escherichia coli*. *J. Biol. Chem.*, **260**, 1836–1841.
- Yang,Y.B., Yu,N.J. and Tai,P.C. (1997) SecE-depleted membranes of *Escherichia coli* are active—SecE is not obligatorily required for the *in vitro* translocation of certain protein precursors. *J. Biol. Chem.*, **272**, 13660–13665.

Received February 1, 1999; revised March 18, 1999;  
accepted April 8, 1999