**The EMBO Journal** Vol.17 No.13 pp.3640–3650, 1998

# Overlapping functions of components of a bacterial Sec-independent protein export pathway

**Frank Sargent[1,2], Erik G.Bogsch[3], Nicola R.Stanley[1,2], Margaret Wexler[1,2], Colin Robinson[3], Ben C.Berks[2] and Tracy Palmer[1,2,4]**

[1]Nitrogen Fixation Laboratory, John Innes Centre, Colney, Norwich NR4 7UH, [2]Centre for Metalloprotein Spectroscopy and Biology, School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ and [3]Department of Biological Sciences, University of Warwick, Coventry CV4 7AL, UK

[4]Corresponding author
e-mail: PALMER@BBSRC.AC.UK

**We describe the identification of two *Escherichia coli* genes required for the export of cofactor-containing periplasmic proteins, synthesized with signal peptides containing a twin arginine motif. Both gene products are homologous to the maize HCF106 protein required for the translocation of a subset of lumenal proteins across the thylakoid membrane. Disruption of either gene affects the export of a range of such proteins, and a complete block is observed when both genes are inactivated. The Sec protein export pathway was unaffected, indicating the involvement of the gene products in a novel export system. The accumulation of active cofactor-containing proteins in the cytoplasm of the mutant strains suggests a role for the gene products in the translocation of folded proteins. One of the two HCF106 homologues is encoded by the first gene of a four cistron operon, *tatABCD*, and the second by an unlinked gene, *tatE*. A mutation previously assigned to the hcf106 homologue encoded at the *tatABCD* locus, *mttA*, lies instead in the *tatB* gene.**

*Keywords*: hydrogenases/molybdoenzymes/periplasm/Sec-independent protein export/twin arginine transfer peptide

## Introduction

A large proportion of newly synthesized proteins require translocation across cellular membranes before they reach their sites of function. In prokaryotes, proteins that are targeted to the periplasm and outer cell membrane are normally transferred across the cytoplasmic membrane by the Sec system (Pugsley, 1993). The Sec apparatus comprises a membrane-embedded translocation complex, consisting minimally of SecY, SecE and SecG, together with a peripheral SecA subunit. Protein export through the membrane complex is by a threading mechanism and is driven by SecA-catalysed ATP hydrolysis. The threading mechanism precludes the transfer of substrate polypeptides with significant tertiary structure (Pugsley, 1993; Schatz and Dobberstein, 1996).

Proteins destined for export by the Sec pathway are synthesized with N-terminal signal sequences which are removed at a late stage in the export process by a signal peptidase whose active site is on the periplasmic face of the membrane. These signal peptides lack sequence similarity but share overall physical–chemical properties (von Heijne, 1985; Izard and Kendall, 1994). Recently it was noted that a subset of unusually long signal sequences contain a consensus S/T-R-R-x-F-L-K motif in which the arginine residues are invariant (Berks, 1996). Such 'twin arginine' signal peptides are particularly associated with periplasmic proteins that bind certain redox cofactors. There is experimental evidence that these proteins acquire their cofactors in the cytosol, and are thus folded, prior to export (Berks, 1996; Santini *et al*., 1998). Export of folded proteins would require a mechanism fundamentally different from that of the Sec apparatus, and it was therefore suggested that twin arginine signal peptides mediate the export of folded proteins by a novel, Sec-independent translocation system. Consistent with this proposal, it has been demonstrated recently that the export of trimethylamine *N*-oxide (TMAO) reductase, a periplasmic molybdopterin cofactor-containing enzyme with a twin arginine signal sequence, is Sec independent in *Escherichia coli* (Santini *et al*., 1998).

The use of twin arginine signal peptides for the targeting of these bacterial proteins was viewed as potentially important for a second reason: plant thylakoids have already been shown to possess a Sec-independent import pathway in which protein transport is driven by the transmembrane ΔpH (reviewed by Robinson and Mant, 1997). The substrates for this pathway are synthesized with cleavable targeting signals (termed 'transfer peptides') with an essential twin arginine motif (Chaddock *et al*., 1995), and there is evidence that the pathway may be capable of transporting folded proteins (Creighton *et al*., 1995; Roffey and Theg, 1996; Clark and Theg, 1997). For these reasons, it was suggested that bacteria may possess a Sec-independent exporter that is structurally and mechanistically related to the thylakoid ΔpH-dependent transporter (Berks, 1996).

Recently, the gene *Hcf106* has been identified that codes for a component of the ΔpH-dependent thylakoid import pathway of maize (Voelker and Barkan, 1995; Settles *et al*., 1997). If the suggested correspondence between the thylakoid and bacterial Sec-independent transport processes holds, then one would expect bacteria to encode HCF106 homologues. Such proteins are indeed encoded in bacterial genomes (Settles *et al*., 1997), and a correlation with the bacterial Sec-independent export pathway is strengthened by the observation that genes encoding HCF106 homologues are found in the complete genome sequences of all prokaryotes (including the Archaeaon *Archaeoglobus fulgidus*) in which proteins are targeted to the periplasm or cell exterior by twin arginine signal

peptides. On the other hand, such genes are absent from the genomes of organisms (mycoplasmas, *Borrelia burgdorferi*, methanogens) which lack such proteins. Further, a *Pseudomonas stutzeri* gene cluster involved in the maturation of nitrous oxide reductase, a copper-containing periplasmic enzyme with a twin arginine signal peptide (Dreusch *et al*., 1997), codes for an HCF106 homologue (*orf57*; Glockner and Zumft, 1996). Thus an HCF106-like protein is implicated in the post-translational processing of a twin arginine signal peptide-dependent enzyme.

Weiner *et al.* (1998) have recently examined this issue using an *E.coli* mutant that was found to be defective in the export of a predicted substrate for this pathway, dimethyl sulfoxide (DMSO) reductase. The mutation was found to lie in a previously unassigned gene, designated *mttA* (for membrane targeting and translocation). The *mttA* gene product exhibited apparent homology to HCF106, and *mttA* was predicted to form an operon with two further unassigned genes; this operon was designated *mttABC*. In this report, we have tested directly the hypothesis that HCF106 homologues are involved in a novel export system by disrupting the two genes encoding HCF106 homologues in *E.coli.* One of the genes does indeed lie in the operon studied by Weiner *et al*. (1998), but we show that this operon contains not three but four distinct genes, due to the presence of a stop codon in the gene identified by Weiner *et al*. (1998) as *mttA*. The first gene product in this operon is homologous to HCF106, whereas the gene affected in the study of Weiner *et al.* (1998) is a separate gene in the operon which would encode a protein unrelated to HCF106. In this study, we describe the effects of disrupting both the authentic *Hcf106* homologue in this operon and a second, unlinked, homologous gene that previously was designated *ybeC*. We show that the export of several cofactor-containing periplasmic proteins is adversely affected when the genes are disrupted singly, and is completely blocked in the double mutant. The data indicate a central role for the gene products in Sec-independent export, and suggest overlapping functions for the two proteins.

## Results

### *Organization of the genes encoding HCF106 homologues in E.coli*

The maize *Hcf106* gene encodes a protein of 243 amino acids with a predicted N-terminal transmembrane helix followed by an amphipathic α-helix and an acidic stromal domain (Settles *et al*., 1997). *Escherichia coli* encodes two proteins with sequence similarity to the predicted transmembrane and amphipathic helices of HCF106. One homologue is coded by the previously unassigned *ybeC* gene located in the 14 minute region of the chromosome. The *ybeC* gene appears to comprise a monocistronic transcriptional unit. The second HCF106 homologue is encoded in the *yigT* region at 86 minutes on the chromosome. The *yigT* locus has seen substantial sequence revisions in successive database releases (Daniels *et al*., 1992; Blattner *et al*., 1997), yet, surprisingly, the open reading frame (ORF) assignment still varies from that of apparently analogous gene clusters found in the closely related organisms *Haemophilus influenzae* and *Azotobacter chroococcum*. To clarify the gene assignments at the

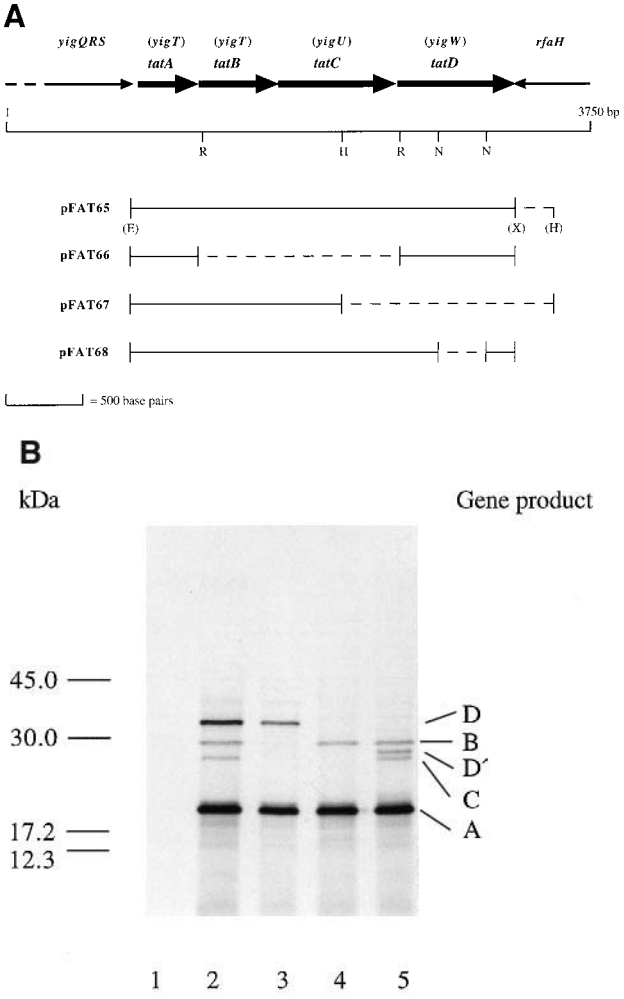


**Fig. 1.** The *tat* locus and *in vivo* synthesis of TatA–D. (**A**) A partial restriction map of the *tat* locus, showing the positions of the *tat* genes after DNA sequence correction. Plasmid pFAT65 carries 2.5 kb covering *tatABCD* cloned distal to the phage T7 ϕ10 promoter of plasmid pT7.5 as described in the text. The regions of deleted DNA in plasmids pFAT66–pFAT68 are indicated in the figure. Restriction site abbreviations are: E, *Eco*RI; H, *Hin*dIII; N, *Nru*I; R, *Eco*RV; X, *Xba*I. (**B**) *In vivo* synthesis of the *tat* gene products from plasmids pFAT65–pFAT68. Autoradiograph of a 12.5% gel after SDS–PAGE. Total proteins were prepared from K38/pGP1-2 containing the following plasmids: lane 1, pT7.5; lane 2, pFAT65; lane 3, pFAT66; lane 4, pFAT67; lane 5, pFAT68.

*E.coli* locus, we resequenced the functionally unassigned chromosomal region around the *Hcf106* homologue. This 2.5 kb area, originally assigned to the ORFs *yigTUVW*, is bounded by *o546* (formerly *yigQRS*), the final ORF in a cluster of genes involved in ubiquinone biosynthesis (Macinga *et al*., 1998), and *rfaH* which encodes a transcriptional regulator (Bailey *et al*., 1997). Our modified sequence has been submitted to the DDBJ/EMBL/GenBank database and will appear under the accession No. AJ005830. A revised ORF map for this region is given in Figure 1A.

The major area of sequence discrepancy between the results obtained here and those reported by the *E.coli* genome project and by Weiner *et al*. (1998) lies in the area encompassing the HCF106 homologue-coding sequence and the region directly 3′ to this. The resequencing carried out here results in the assignment of two

ORFs to this region. The first ORF codes for the HCF106 homologue and, on the basis of results presented below, is designated *tatA* for twin arginine translocation. Assuming the second possible methionine codon is the initiation site (because of favourable adjacent Shine–Dalgarno and promoter sequences), TatA is an 89 amino acid protein with 53% identity to YbeC, which is hereafter redesignated TatE. The second ORF, designated *tatB*, initiates 3 bp after the stop codon of *tatA* and terminates 2 bp before the start of *tatC*. The *tatB* gene is predicted to encode a protein of 171 amino acids. TatB is homologous to the products of possible *H.influenzae* and *A.chroococcum* ORFs that lie between genes coding for TatA and TatC homologues. The TatB proteins are predicted to have a structure in which an N-terminal transmembrane helix is followed by a cytoplasmic domain including a possible amphipathic helical region. The current *E.coli* genome project database release (DDBJ/EMBL/GenBank accession No. AE000459), while identifying *tatA* (*o103*), substitutes two fragmentary, overlapping ORFs (*o113* and *o145*) for *tatB*. In Weiner *et al*. (1998), *tatA* and *tatB* are assigned to a single ORF (*mttA*). As the mutation identified by Weiner and co-workers (Weiner *et al*., 1998) lies in the *tatB* coding region, their paper describes the phenotype of a *tatB* strain rather than a strain with a defect in the HCF106 homologue TatA.

The third gene at the locus, *tatC* (formerly *o258*, formerly *yigU* and *yigV*), encodes a protein of 258 amino acids. TatC is predicted to be an integral membrane protein composed of six transmembrane helices with the polypeptide N-terminus in the cytoplasm. TatC homologues (Prosite UPF0032 family) are found in the genomes of many prokaryotes and of some chloroplasts and mitochondria.

The *tatD* gene (formerly *o206* and *o113*, formerly *yigW*) encodes a protein of 264 amino acids which has homologues in many organisms (Prosite UPF0006 family) including the products of the two further functionally unassigned *E.coli* ORFs, *yjjV* and *ycfH*. TatD is predicted to be a water-soluble, cytoplasmically-located protein. The *tatCD* intergenic gap is 29 bp and *tatD* shares a 2 bp overlap with the *rfaH* gene transcribed in the opposite direction. The *tatD* gene has been assigned experimentally to the *tat* transcriptional unit by Weiner *et al*. (1998).

It was important to corroborate the revised ORF assignment experimentally given the conflict with previously published interpretations. Expression studies were carried out to analyse the polypeptides coded by this region. DNA covering the *tatABCD* region was placed under the control of the phage T7 promoter on plasmid pT7.5 (Tabor and Richardson, 1985) (Figure 1A). The resultant construct directed expression of the expected four polypeptides (Figure 1B) of estimated molecular masses 17.7, 26.4, 29.6 and 33.8 kDa, respectively (Table I). Deletion between the *Eco*RV sites shown in Figure 1A removes DNA from the fourth codon of *tatB* to the seventh codon of *tatD*. From Figure 1B, lane 2, it is apparent that only two peptides are now synthesized, one corresponding to *tatA* and the second to a protein fusion comprising three amino acids of TatB, and all but seven amino acids of TatD. Clearly the protein with an apparent mass of 17.7 kDa is the product of the *tatA* gene, since in-frame deletion of *tatD* between the *Nru*I sites leaves the 17.7 kDa band untouched (lane 5), but removes ~7 kDa from the 33.8 kDa band. The *tatB* gene product migrates with an apparent mol. wt of 29.6 kDa, and the *tatC* product with an apparent mass of 26.4 kDa. These results confirm the veracity of our sequence-derived reading frame assignments. The TatA and -B proteins fail to migrate close to their expected molecular masses (Table I). This is probably a reflection of the integral membrane character of these proteins as highly hydrophobic polypeptides may migrate anomalously in SDS–PAGE. It is notable that *tatA* expression in these experiments is considerably higher than that of the other *tat* genes (Figure 1B).

**Table I.** Apparent and predicted molecular masses of *tatA–D* gene products

| Gene | Predicted size from DNA sequence (kDa) | Estimated mass[a] of *tat* product (kDa) |
|---|---|---|
| *tatA* | 9.6 or 11.3[b] | 17.7 |
| *tatB* | 18.4 | 29.6 |
| *tatC* | 28.9 | 26.4 |
| *tatD* | 29.5 | 33.8 |

[a]Estimated by SDS–PAGE.
[b]Two possible translation initiation sites for *tatA*.

### Construction and growth phenotype of mutants

In order to test the function of the *E.coli* HCF106 homologues, chromosomal in-frame deletion mutations were constructed in the *tatA* and *tatE* genes, both singularly and in combination. The mutant strains grew as well as the parental strain under aerobic respiratory or anaerobic fermentative growth conditions. In *E.coli*, most periplasmic proteins with twin arginine signal sequences are involved in anaerobic respiratory pathways. The Δ*tatA* and Δ*tatE* single mutant strains grew on the non-fermentable carbon source glycerol with any of nitrate, fumarate, TMAO or DMSO as the sole electron acceptor. However, TMAO or DMSO would not support growth of the Δ*tatA*Δ*tatE* double mutant, indicating a defect in respiration involving these oxidants.

### TMAO reductase activity is mis-localized in the Δ*tatA*Δ*tatE* mutant strain

TMAO reductase is a soluble periplasmic enzyme which catalyses the reduction of the terminal electron acceptor TMAO to trimethylamine (Barrett and Kwan, 1985). The enzyme, which has *bis*(molybdopterin guanine dinucleotide)molybdenum (MGD) at its active site, is synthesized with a twin arginine signal peptide (Méjean *et al*., 1994) and exported by a Sec-independent mechanism (Santini *et al*., 1998). The cellular localization of TMAO reductase activity was assessed in cells cultured on a fermentable carbon source supplemented with TMAO to induce expression of the reductase (Pascal *et al*., 1984). In the parental strain, TMAO reductase activity is located almost exclusively in the periplasmic compartment (Table II). Strains deleted for either the *tatA* or *tatE* gene show an ~5-fold decrease in periplasmic TMAO reductase activity, and an accumulation of activity within the cytoplasm, suggesting that export of the enzyme is impaired. Deletion of both genes results in an almost complete absence of TMAO reductase activity in the periplasmic fraction. The very

**Table II.** A comparison of enzyme activities in *tat* mutants

| Enzyme | Fraction | Activity (U/g of cells) | | | |
|---|---|---|---|---|---|
| | | MC4100 parent strain | ELV15 Δ*tatA* | J1M1 Δ*tatE* | JARV15 Δ*tatA*Δ*tatE* |
| TMAO reductase[a] | periplasm | 115 | 28 | 20 | 1.1 |
| | membrane | 4.4 | 1.0 | 1.5 | <0.01 |
| | cytoplasm | 14 | 44 | 39 | 45 |
| DMSO reductase[b] | periplasm | 0.05 | 0.02 | 0.07 | <0.01 |
| | membrane | 1.5 | 0.18 | 0.90 | 0.14 |
| | cytoplasm | <0.01 | 1.9 | 0.54 | 3.1 |
| Nitrate reductase[c] | periplasm | 3.0 | 3.5 | 4.1 | 2.0 |
| | membrane | 36 | 59 | 56 | 55 |
| | cytoplasm | 1.2 | 2.9 | 2.0 | 2.8 |
| Fumarate reductase[d] | periplasm | <0.01 | <0.01 | <0.01 | <0.01 |
| | membrane | 2.3 | 7.6 | 5.8 | 7.3 |
| | cytoplasm | 0.18 | 1.2 | 2.0 | 1.4 |
| β-Lactamase[e] | total exported[f] | 0.99 | 0.99 | nd | 0.93 |
| (pBR322) | cytoplasm | 0.04 | 0.02 | nd | <0.01 |
| Acid phosphatase[a] | periplasm | 5.8 | 5.9 | 5.0 | 4.3 |

Oxidoreductase activities are expressed as substrate-dependent benzyl viologen oxidations (units are μmol of benzyl viologen oxidized per min). β-Lactamase activities were determined as μmol of 7-(thienyl-2-acetamido)-3-[2-(4-*N*,*N*-dimethylaminophenylazo)pyridinium-methyl]-3-cephem-4 carboxylic acid (PADAC) hydrolysed per min. Acid phosphatase activities were determined as μmol of *p*-nitrophenyl phosphate hydrolysed per min. Cells were grown anaerobically with 0.5% glycerol and 0.4% [a]TMAO, [b]DMSO, [c]nitrate, [d]fumarate or [e]0.2% glucose with 0.4% nitrate.
[f]'Total exported' combines activities from both periplasm and culture media.
nd = not determined.

low level of residual activity is probably due to low-level cytoplasmic contamination of the periplasmic fraction. TMAO reductase activity again accumulates in the cytoplasm. Thus, disruption of the genes coding for both HCF106 homologues abolishes periplasmic targeting of TMAO reductase.

DMSO reductase is a membrane-bound multisubunit enzyme which supports anaerobic growth on DMSO (Weiner *et al*., 1992). The MGD-binding α-subunit is synthesized with a twin arginine signal peptide (Bilous *et al*., 1988; Berks, 1996). In agreement with the results obtained for the TMAO reductase, a combined deletion of the *tatA* and *tatE* genes leads to an almost total loss of DMSO reductase activity in the cytoplasmic membrane and the single mutations have an incomplete phenotype, though in this case the Δ*tatA* mutation has a more severe effect on localization than the Δ*tatE* mutation (Table II). As with TMAO reductase, the mutant strains accumulate enzymatic activity in the cytoplasm. No DMSO reductase activity could be detected in the periplasm of any of the strains.

### *Formate dehydrogenase-N accumulates as an inactive precursor in the cytoplasm of the mutant strains*

*Escherichia coli* synthesizes three distinct membrane-bound formate dehydrogenase isozymes (Sawers, 1994). The formate dehydrogenase-N (FDH-N) isoenzyme is highly expressed during anoxic growth with nitrate as terminal electron acceptor. The MGD- and iron–sulfur cluster-binding FdnG subunit is located on the periplasmic side of the cell membrane and is synthesized as a precursor with a twin arginine signal peptide (Berg *et al*., 1991; Sawers, 1994; Berks, 1996; F.Sargent, N.R.Stanley, B.C.Berks, T.Palmer, J.Shi and V.Stewart, unpublished observations).

The subcellular localization of FDH-N in the parental and *tat* mutant strains was assessed by immunoelectrophoresis (Figure 2). In this technique, FDH-N is specifically immunoprecipitated by an antiserum raised against whole FDH-N. Enzymatically active immunoprecipitates are detected by formate dehydrogenase activity staining. In the parental strain, FDH-N activity is found almost exclusively in the membrane fraction (Figure 2B, lane 1), with only trace activity in the cytoplasmic fraction (Figure 2B, lane 5). Deletion of *tatA* results in a greatly lowered level of active enzyme in the membrane fraction (Figure 2B, lane 2). While there is no detectable formate dehydrogenase activity in the cytoplasmic fraction, protein staining shows that FDH-N-immunoreactive protein has accumulated (Figure 2A and B, lane 6). Thus, in contrast to the situation with DMSO and TMAO reductases, the FDH-N protein that accumulates in the mutant strain is enzymatically inactive. Deletion of *tatE* has a less severe effect on FDH-N assembly than the Δ*tatA* mutation, with only a 2-fold reduction in the level of membrane-bound formate dehydrogenase activity (Figure 2B, lane 3). Again, an accumulation of inactive soluble enzyme is observed (Figure 2A and B, lane 7). The Δ*tatA*Δ*tatE* strain completely lacks membrane-associated formate dehydrogenase activity (Figure 2B, lane 4). All of the FDH-N protein in this strain is cytoplasmically located and enzymatically inactive (Figure A and B, lanes 8). Note that the total amount of FDH-N antigen present in each strain (active membrane-bound plus enzymatically inactive cytoplasmic) was approximately equivalent, suggesting that the *tat* mutations have no gross effect on FDH-N polypeptide expression.

### *Hydrogenase 1 is found as an active soluble precursor in the mutant backgrounds*

Three hydrogenase isoenzymes are synthesized by *E.coli* (Sawers, 1994) and recent genetic evidence suggests

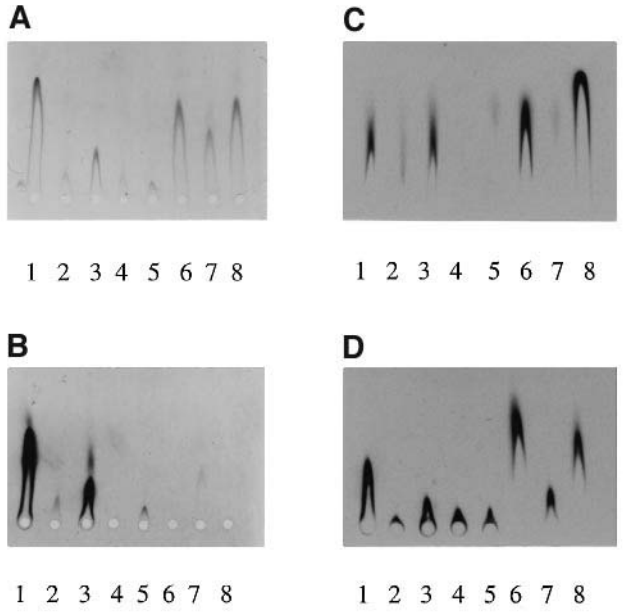


**Fig. 2.** FDH-N and hydrogenase 1 and 2 accumulate in the cytosol in strains deleted for *tatA* and *tatE*. Cells were grown on either glycerol plus nitrate to induce expression of FDH-N, or glycerol plus fumarate for maximal expression of the uptake hydrogenases (Sawers *et al*., 1985; Berg and Stewart, 1990; Brondsted and Atlung, 1994). Cells were fractionated, membrane proteins solubilized and rocket immunoelectrophoresis performed as described in Materials and methods. All plates are as follows: lane 1, MC4100 (parental strain), membrane fraction; lane 2, ELV15 (Δ*tatA*), membrane fraction; lane 3, J1M1 (Δ*tatE*), membrane fraction; lane 4, JARV15 (Δ*tatA* Δ*tatE*), membrane fraction; lane 5, MC4100, cytosolic fraction; lane 6, ELV15, cytosolic fraction; lane 7, J1M1, cytosolic fraction; lane 8, JARV15, cytosolic fraction. All the samples represent the same proportion (0.2%) of total protein present in the two fractions. (**A**) FDH-N, protein stained; (**B**) FDH-N, activity stained; (**C**) hydrogenase 1, activity stained; and (**D**) hydrogenase 2, activity stained.

there may be a fourth (Andrews *et al*., 1997). The core components of each enzyme are a large subunit that contains the Ni–Fe active site and a small subunit which binds iron–sulfur clusters. The active sites of hydrogenases 1 and 2 are at the periplasmic face of the membrane, where they catalyse respiratory hydrogen oxidation (Graham, 1981; Rodrigue *et al*., 1996; Sargent *et al*., 1998). The hydrogenase 1 and 2 small subunit precursors have twin arginine signal peptides (Menon *et al*., 1990; Sargent *et al*., 1998), while the large subunits, despite their periplasmic location, lack export signals. Insertion of Ni into the large subunit occurs in the cytoplasm and is accompanied by C-terminal processing of the subunit by a specific protease (Menon *et al*., 1993; Rossmann *et al*., 1995). Export of the small subunit is prevented if Ni insertion into the large subunit is blocked (Bernhard *et al*., 1997; Sargent *et al*., 1998), suggesting that the large and small subunits are assembled in a co-ordinated manner and that the twin arginine signal sequence of the small subunit directs export of both proteins.

A polyclonal antiserum raised against purified hydrogenase 1 readily identifies the protein as an activity-staining band in the membranes of the parental *E.coli* strain after rocket immunoelectrophoresis (Figure 2C, lane 1). The functional membrane localization of hydrogenase 1 is severely compromised in the Δ*tatA* strain (Figure 2C, lane 2) but is apparently unaffected by an isolated Δ*tatE* mutation (Figure 2C, lane 3). An involvement of TatE in hydrogenase 1 assembly can, however, be inferred from the observation that combining the Δ*tatE* and Δ*tatA* mutations completely eliminates membrane-bound hydrogenase 1 activity (Figure 2C, lane 4). In the mutant strains where hydrogenase 1 activity was lost from the membranes, there was a corresponding accumulation of active enzyme within the cytoplasm (Figure 2C, lanes 6 and 8). However, since activity staining is not quantitative, these experiments cannot determine how the specific activity of the cytoplasmic enzyme compares with the mature membrane-bound enzyme. We therefore compared the hydrogenase activities of hydrogenase 1 directly immunoprecipitated from Triton X-100-solublilized cell-free extracts. The hydrogen:benzyl viologen oxidoreductase activity of the parental and Δ*tatA*Δ*tatE* strains was 3.7 and 1.1 μmol of benzyl viologen reduced/min/g cells, respectively, indicating that although the level of hydrogenase 1 antigen is similar in the two strains (Figure 2C), the enzyme in the mutant is only ~30% as active.

### *The large and small subunits of hydrogenase 2 are processed in strains deleted for tatA and tatE*

The cellular localization of hydrogenase 2 was also investigated using an antiserum raised against the purified enzyme (Figure 2D). The Δ*tatE* mutation alone showed an ~2-fold reduction in the level of membrane-bound hydrogenase 2 (Figure 2D, lane 3). The Δ*tatA* mutation led to a dramatic reduction in the amount of active hydrogenase 2 in the membrane. A strain deleted for both *tatA* and *tatE* displayed a similar phenotype to that of the single *tatA* mutation. However, even in strains deleted for both genes, significant amounts of active enzyme were located in the membrane (Figure 2D, lanes 2 and 4). Enzymatically active immunoreactive protein accumulated in the cytoplasmic fraction in the strains exhibiting a hydrogenase 2 localization defect.

The processing of the hydrogenase 2 subunits in the mutant strains was investigated by immunoblotting following SDS–PAGE. In agreement with the results of the rocket immunoelectrophoresis experiments, analysis of subcellular fractions of the Δ*tatA*Δ*tatE* mutant showed that the majority of the large subunit of the enzyme was located in the cytoplasmic fraction (Figure 3A, lanes 5 and 6). The mature C-terminally processed form of the hydrogenase 2 large subunit can be distinguished electrophoretically from the precursor protein. This is demonstrated in Figure 3A, lanes 1 and 2, where the electrophoretic mobility of the mature subunit is compared with that of the precursor protein found in strain DHP-B (Δ*hypB*) which has a defect in Ni insertion into the large subunit (Jacobi *et al*., 1992; Sargent *et al*., 1998). Analysis of the cytoplasmically located hydrogenase 2 found in the Δ*tatA*Δ*tatE* mutant (Figure 3B, lane 6) clearly shows that the subunit has been processed, suggesting that, consistent with the positive activity stain of the corresponding immunoprecipitant rocket (Figure 2D, lane 8), the Ni–Fe active site has been inserted. The small subunit of hydrogenase 2 is unstable in cell extracts and it proved impossible to analyse this subunit by immunoblotting following subcellular fractionation. However, the subunit can be analysed if freshly harvested whole cells are added directly to the

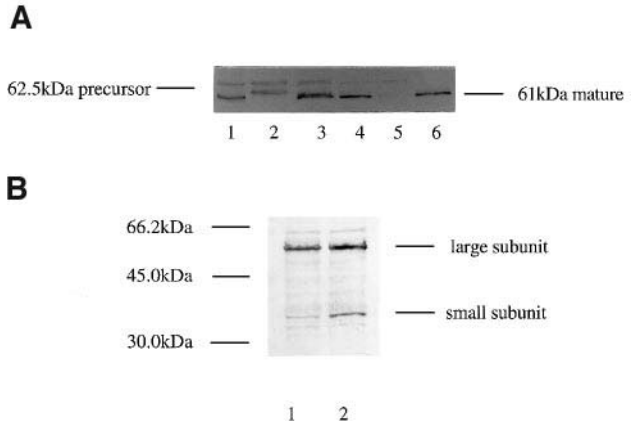


**Fig. 3.** The hydrogenase 2 large and small subunits are processed in JARV15. (**A**) Western blot analysis of the large subunit. Lanes 1–3 are total cell protein from ~25 μg of cells in each case: lane 1, MC4100 (parental strain); lane 2, DHP-B (Δ*hypB*); lane 3, JARV15 (Δ*tatA*Δ*tatE*). Lanes 4–6 are cell fractions of JARV15, and each represents the same fraction (0.25%) of the total protein present. Lane 4, crude cell extract; lane 5, membrane fraction; lane 6, cytoplasmic fraction. Precursor and mature forms of the large subunit are indicated. (**B**) Western blot analysis of the large and small subunits of total cell protein of: lane 1, JARV15; lane 2, MC4100. Each lane represents ~25 μg of cells in each case. Cross-reacting bands corresponding to the large and small subunits are indicated.

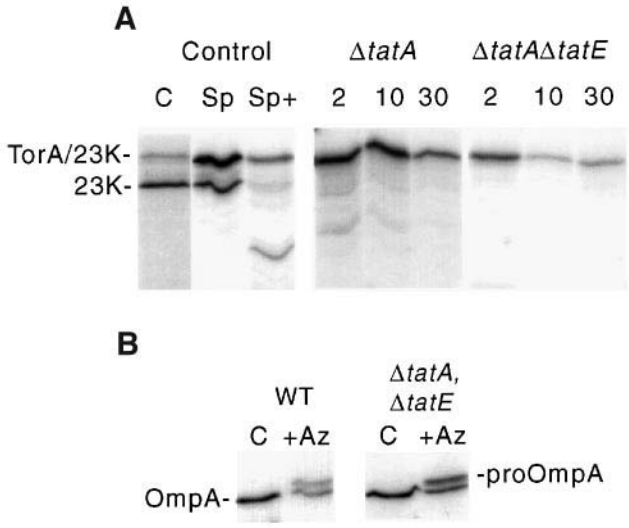


**Fig. 4.** A TorA–23K fusion protein is not exported in *tat* mutant strains. (**A**) Control panel: wild-type cells expressing a TorA–23K fusion protein were pulsed with [$^{35}$S]methionine for 1 min and chased with cold methionine for 10 min, after which samples were immunoprecipitated with 23K antiserum. Samples were analysed of cells (C), sphaeroplasts (Sp) and sphaeroplasts that were incubated with proteinase K (Sp+). Lanes Sp and Sp+ received a loading equivalent to twice that of lane C in order to enhance the signal of the precursor band. The remaining panels show pulse–chase data using the indicated mutant strains in which total cell contents were analysed after chase times of 2, 10 and 30 min as indicated above the lanes. (**B**) Wild-type (WT) and Δ*tatA*Δ*tatE* strains were pulsed as in (A), after which samples were immediately immunoprecipitated using antiserum to OmpA. Pulse–chases were carried out under control conditions (C) or in the presence of 1 mM sodium azide (+Az). The mobilities of mature and pro-OmpA are indicated.

electrophoretic sample buffer (Figure 3B). The small subunit is present in the Δ*tatA*Δ*tatE* mutant (Figure 3B, lane 1) although at lower levels than in the parental strain (Figure 3B, lane 2). This indicates that the small subunit might be more unstable in the mutant background. The small subunit has the same electrophoretic mobility in the two strains, suggesting that the 5 kDa signal sequence has been removed in the mutant. Attempts to confirm the site-specific processing of the hydrogenase 2 small subunit in the Δ*tatA*Δ*tatE* mutant by N-terminal sequence analysis of immunoprecipitated isoenzyme were unsuccessful.

### *Export of Sec-dependent proteins and assembly of cytoplasmically located cofactor-containing proteins are unaffected in the mutant strains*

The observations above show that TatA and TatE are required for correct localization of precursor proteins bearing twin arginine transfer peptides. The involvement of TatA and TatE in the export of proteins by the Sec pathway was tested by determining the effect of the *tat* mutations on localization of periplasmic acid phosphatase and β-lactamase activities. The overall activities and subcellular location of these enzymes, which are synthesized with standard Sec pathway signal peptides, were unaffected in the mutants (Table II). Thus, the *tat* mutations do not disrupt export via the Sec pathway.

It was conceivable that the export defect observed in the *tat* mutants was due to a general lesion in cofactor biosynthesis or insertion rather than to protein export. This possibility was tested by measuring the activities of cytoplasmically-located proteins with the same cofactor types as the twin arginine transfer peptide-dependent periplasmic enzymes analysed above. Membrane-bound nitrate reductase A has cytoplasmically-located subunits containing iron–sulfur clusters and the catalytic site MGD cofactor. The subunits of fumarate reductase that lie on the cytoplasmic side of the cell membrane bind FAD and iron–sulfur clusters. The activities of these two enzymes are not impaired in the *tat* mutants (Table II). The activities of the membrane-bound, but cytoplasmically-oriented, FDH-H, which binds MGD and a [4Fe–4S] cluster, and hydrogenase 3, which has the Ni–Fe hydrogenase active site cluster, are also unaffected in the *tat* mutants (data not shown). Further, with the exception of FDH-N, the substrate proteins that accumulate in the cytoplasm of the *tat* mutants with defects in export are enzymatically active and therefore must contain bound cofactors. Taken together, these measurements indicate that the *tat* mutants have no generalized defect in redox protein assembly.

### *Kinetics of a pre-TorA–23K fusion protein export in mutants with defects in TatA and/or TatE*

The observations so far presented have shown that Δ*tatA* and Δ*tatE* mutations affect correct localization of multiple enzymes whose precursors bear twin arginine transfer peptides. In order to demonstrate directly that the mutations affect export of such precursor proteins, and to assess the effects of the mutations on the kinetics of export and processing, we undertook the pulse–chase experiments shown in Figure 4. The twin arginine transfer peptide-associated proteins analysed above are not highly suitable for such experiments as they are either membrane-bound, which complicates the analysis of export, or are relatively large proteins, for which it is difficult to detect signal peptide processing by changes in electrophoretic mobility. Further, the kinetics of cofactor insertion would be difficult to separate from the kinetics of the transport process. To

overcome these difficulties, the pulse–chase experiments utilized a simplified model precursor in which the twin arginine transfer peptide of *E.coli* TMAO reductase (TorA) is fused in-frame to the mature 23 kDa protein (23K) of the plant thylakoid photosystem II oxygen-evolving complex. 23K is a relatively small, water-soluble protein that does not bind a cofactor. 23K is imported into thylakoids exclusively by the ΔpH pathway (Creighton *et al*., 1995; Bogsch *et al*., 1997) so was thought likely to be tolerated as a passenger protein by the apparently analogous bacterial export system.

When expressed in *E.coli* from an inducible promoter, the TorA–23K fusion protein undergoes time-dependent processing. Thus, after pulse-labelling followed by a 10 min chase, ~90% of the TorA–23K is processed to a smaller form of a size consistent with transfer peptide removal (Figure 4A). This mature form could be shown to be located in the periplasm, as the mature form but not the precursor is sensitive to externally added protease following sphaeroplast formation. Thus, the TorA transfer peptide is capable of mediating export of the plant 23K protein from the bacterial cytoplasm.

The kinetics of the 23K export were next examined in the Δ*tatA* and Δ*tatA*Δ*tatE* mutants. In contrast to the parental strain, no detectable processing of TorA–23K occurred in the mutant strains even after a 30 min chase. Protease accessibility experiments (data not shown) confirmed the cytoplasmic location of the precursor in these strains. Thus, both the Δ*tatA* and Δ*tatA*Δ*tatE* mutations severely retard the rate of export of a protein with a twin arginine transfer peptide. As 23K does not bind a cofactor, these experiments also indicate that the Δ*tatA* and Δ*tatE* mutations do not act solely at the level of cofactor insertion. We have also attempted to analyse the Δ*tatE* strain, but have failed to detect the labelled TorA–23K protein using immunoprecipitation (data not shown); the protein may thus be unstable in this strain.

The export of the Sec pathway substrate pro-OmpA in the mutant strains was also analysed by pulse–chase experiments. In each of the mutant strains, we observe correct export of OmpA, even after very short chase periods; the control lanes (C) in Figure 4B show only mature size OmpA in both wild-type cells and the Δ*tatA*Δ*tatE* double mutant, and identical results were obtained using the other strains (data not shown). Mature OmpA was sensitive to protease digestion following sphaeroplast formation (data not shown), which confirmed that the protein had been exported. The bands are confirmed as mature size by the inclusion of azide during the pulse–chase experiments analysed on the adjacent lanes (+Az); azide inhibits SecA-dependent ATP hydrolysis (Oliver *et al*., 1990) and leads to the accumulation of pro-OmpA in both the wild-type and mutant strains. These experiments provide additional evidence that Sec-dependent export is not impaired to any detectable extent by loss of TatA or TatE function.

## Discussion

We have shown here that two homologues of the thylakoidal ΔpH pathway component HCF106 are involved in the export of bacterial proteins bearing twin arginine signal peptides. An *E.coli* mutant lacking both TatA and TatE is completely unable to export five different proteins that between them bind a variety of redox cofactors, clearly indicating a central role for TatA/TatE in Sec-independent protein export in bacteria. Furthermore, the data provide strong support for the proposition that chloroplasts and bacteria contain Sec-independent protein translocases that are related in both structural and mechanistic terms. It is of interest, however, that neither protein is by itself essential, because neither the Δ*tatA* nor the Δ*tatE* single mutant is completely defective in the export of any of the proteins analysed in this study. It is also of interest that the severity of the export defect in the single knockouts appears to vary depending on the enzyme studied. Some enzymes are apparently unaffected in a given strain, whereas other enzymes (e.g. TMAO reductase) are almost equally affected in each of the single mutants. However, we should stress that the enzyme activity measurements may be misleading in the sense that they provide no information on export kinetics, and it remains possible that all of these proteins may be exported at a significantly slower rate in both the Δ*tatA* and Δ*tatE* single mutants.

Although comprehensive kinetic studies remain to be performed, the enzyme activity/localization data do indicate quite clearly that TatA plays a more important role than TatE in this Sec-independent export process, at least in *E.coli*. The Δ*tatE* strain is significantly less compromised than the Δ*tatA* strain in the localization of DMSO reductase, FDH-N, and hydrogenases 1 and 2, all of which are severely mis-localized in the Δ*tatA* strain. Nevertheless, TatE is clearly involved in the translocation process because some enzymes (e.g. TMAO reductase) are mis-localized to a notable extent in the Δ*tatE* strain, and the residual levels of export of TMAO reductase, DMSO reductase, FDH-N and hydrogenase 1 in the Δ*tatA* strain are completely abolished in the double mutant.

These findings raise an important question concerning the structure and substrate specificity of the twin arginine translocase: does this translocation system contain copies of both TatA and TatE, or do variants exist in which individual translocase complexes contain either TatA or TatE? This point remains to be clarified; however, the evidence favours the former explanation at present, for the following reason. Each of the individual mutations has a very severe effect on the export of some proteins, an example being TMAO reductase, most of which (~80–85%) is found predominantly in the cytoplasm in both the Δ*tatA* and Δ*tatE* strains. This 5-fold mis-localization must reflect an even greater level of inhibition in terms of export kinetics, and this is indeed found in pulse–chase experiments using the TorA–23K construct, where essentially no export is apparent in the Δ*tatA* mutant over a 30 min chase period. The logical interpretation of these data is that the rate of export is lowered by a considerable factor in the single mutants (probably several orders of magnitude), but that the block is not complete and a small proportion of protein is therefore exported during the extended time scale of the growth period. These findings are not consistent with a model in which translocase molecules contain either TatA or TatE, because removal of even one component has such a drastic effect on the export of TMAO reductase. Instead, the data suggest an involvement of both TatA and TatE in the translocation

apparatus, and our results suggest instead that TatA is simply more important than TatE in the translocation process. Our present hypothesis is that these proteins have overlapping functions such that each can partially substitute for the other.

It is notable that for one enzyme system, hydrogenase 2, the Δ*tatA*Δ*tatE* double mutation does not completely abolish membrane targeting. Several interpretations are possible. If this membrane-associated activity represents exported protein this would imply that TatA and TatE do not form the transporter itself, but regulate access to the transporter in some manner. Alternatively, the membranous activity might represent enzyme bound to a receptor site on the cytoplasmic side of the membrane but that cannot be exported because the TatA- and TatE-containing transporter is defective. In either case, the implication is that the twin arginine export system involves components other than TatA and TatE, prime candidates for which are the products of the *tatBCD* genes that, together with *tatA*, appear to constitute an operon. Weiner *et al*. (1998) have shown that the gene immediately following *tatA* (termed *mttA* in that study, but which we designate *tatB*) is required for export by this system, and future studies should help to define the remaining elements of the system.

It is also noteworthy that for all the enzyme systems tested, a localization defect leads to a corresponding accumulation of the substrate protein in the cytoplasmic compartment and that, in all but one case (FDH-N), this cytoplasmic form of the enzyme has enzymatic activity demonstrating that cofactor has been inserted. These observations support the view that proteins with twin arginine transit peptides receive their cofactors in the cytoplasm prior to export (reviewed by Berks, 1996; Santini *et al*., 1998). The lack of activity of the cytoplasmic form of FDH-N could be taken as an indication that this enzyme has not received cofactor. However, it was observed that the activity of the cytoplasmic form of hydrogenase 2 was far more labile than that of the mature membrane-bound enzyme (data not shown), and it is therefore possible that the failure to detect enzymatic activity in the cytoplasmic FDH-N could be explained by a gross instability of this species.

Weiner *et al*. (1998) previously have assigned a single reading frame, *mttA*, to approximately the chromosomal region here assigned to the two reading frames, *tatA* and *tatB*. The assignment in the current work is indisputably the correct one. The presence of a stop codon at the end of *tatA* has been confirmed by resequencing and is in agreement with the current *E.coli* genome project database entry for this region. The protein expression experiments described here definitively demonstrate that this region of the chromosome encodes two, rather than one polypeptide. The mutation described by Weiner *et al*. (1998) would lie in the reading frame *tatB* rather than in the gene encoding the HCF106-homologous TatA protein. A direct comparison of the phenotype of the *mttA* mutant (Weiner *et al*., 1998) with the *tatA* mutant described here is difficult because, in the main, different enzymes systems were investigated. Where a comparison is possible, the Δ*tatA* and *mttA* mutant strains are quite distinct, and this is most consistent with the lesions being in different genes. Specifically, the *mttA* mutant strain (Weiner *et al.*, 1998) fails to grow on the non-fermentable carbon source glycerol with either TMAO or DMSO as electron acceptor while the Δ*tatA* strain described here has no significant growth defect on such media. Further, while Weiner *et al*. (1998) report no active TMAO reductase in the periplasm of their mutant strain, TMAO reductase activity is still present, though at a low level, in the Δ*tatA* strain. The work of Weiner *et al*. (1998) is thus a demonstration that *E.coli tatB* is a third component of the twin arginine export pathway.

It was suggested originally that the thylakoid ΔpH pathway may have originated in photosynthetic eukaryotes, on the basis that the substrates for the pathway are not found in the cyanobacterial ancestors of the chloroplast. However, the possibility that an analogous pathway was present in prokaryotes was raised by the observations that a subset of the precursors of bacterial extracytoplasmic proteins have signal peptides resembling the thylakoid twin arginine transfer peptides (Berks, 1996) and that homologues of the HCF106 component of the plant ΔpH pathway are coded by many prokaryote genomes (Settles *et al*., 1997). In this work, it has been confirmed experimentally that the HCF106 homologues of *E.coli* are involved in export of the bacterial twin arginine transfer peptide-dependent proteins. This suggests that an export system structurally, and presumably mechanistically, similar to the thylakoid ΔpH pathway is very widely distributed in, and therefore probably originated in, prokaryotes. It is not advisable at this stage to assign the designation 'ΔpH' to the prokaryotic export pathway, as the energetics of the system have not been defined and the ΔpH component of the transmembrane protonmotive force in bacteria is usually negligible. Instead, the designation 'twin arginine', reflecting the distinctive feature of the substrate, is preferred.

Weiner and co-workers (1998) have proposed the alternative designation *mtt* (membrane targeting and transport) for genes in the *tat* operon. The *tat* (twin arginine translocation) designation is to be preferred because the membrane targeting function assigned by Weiner *et al*. (1998) is based on an unusual interpretation of their data. Essentially, Weiner *et al*. (1998) believe that DmsA, the peripheral membrane catalytic subunit of DMSO reductase, despite having a processed twin arginine transfer peptide, is located at the cytoplasmic face of the membrane. If this is the case, then the cytoplasmic accumulation of DmsA in the *tatB* mutation must indicate that the protein is, uniquely, not subject to an export event and that a twin arginine transfer peptide and TatB can have the sole physiological function of targeting a substrate protein to the membrane. The evidence for a cytoplasmic rather than periplasmic location for DmsA is not unequivocal, reflecting the difficulty in determining the topological organization of a membrane protein if the applicability of marker fusion approaches is suspect [for a more detailed discussion of the DmsA topology problem, see Berks (1996) and references therein]. Berks (1996) proposed that the topological organization of DmsA could be assigned definitively if the membrane anchor subunit of the enzyme was genetically removed and the location of the now water-soluble DmsA assessed following subcellular fractionation. Weiner and co-workers (1998) have carried out this experiment and find that the soluble DmsA is located in the periplasm. This observation is rationalized

most readily by proposing that DmsA normally is transported to the periplasm and binds to the portion of the integral membrane subunit exposed to the periplasmic face of the membrane. Weiner and co-workers (1998), however, suggest that, in the absence of the DMSO reductase integral membrane subunit, the transfer peptide functions to direct export of DmsA, but when the membrane subunit is present this binds DmsA on the cytoplasmic face of the membrane and export is prevented. The application of 'Occam's razor', as well as consistency with the behaviour of other 'twin arginine' enzyme systems, suggest that the explanation of Weiner *et al*. (1998) is highly unlikely and that it would be inadvisable to name the twin arginine transporter system on the basis of such a model.

Fincher *et al*. (1998) have proposed that the thylakoid HCF106 protein functions to target twin arginine transfer peptide-bearing precursors to the Sec translocon. Although the experiments described here do not exclude the possibility that Sec components are involved in the twin arginine transport pathway, Santini *et al*. (1998) have shown that export of the twin arginine transfer peptide-dependent enzyme *E.coli* TMAO reductase is independent of the core SecY, SecE and SecA components. Thus, the twin arginine transport system is completely distinct from the Sec pathway. It is proposed that the designation '(twin arginine) transfer peptide' be used for the signal peptides of the twin arginine pathway substrate proteins to emphasize the functional distinction from Sec pathway signal peptides and to be consistent with the nomenclature already adopted for substrates of the plant twin arginine pathway.

## Materials and methods

### *Bacterial strains, plasmids and growth conditions*

*Escherichia coli* strain MC4100 (F–Δ*lacU169 araD139 rpsL150 relA1 ptsF rbs flbB5301*; Casadaban and Cohen, 1979), a K-12 derivative, was the parental strain used. ELV15 (MC4100 Δ*tatA*) was constructed as follows: a 534 bp fragment covering the upstream region and the second possible start codon of *tatA* (see text) was amplified by PCR using primers TATA1 (5′-GCGCTCTAGAGGAAGTGCAGCCGCAACTGG-3′) and TATA2 (5′-GCGCGGATCCCATACATGTTCCTCTGTGG-3′) with MC4100 chromosomal DNA as template. The resulting product was digested with *Xba*I and *Bam*HI and cloned into the polylinker of pBluescript (Stratagene) to give plasmid pFAT4. A 520 bp fragment covering the last two codons of *tatA* and downstream DNA was amplified using primers TATA3 (5′-GCGCATCGATGTGTAATCCGTGTTTG-ATATCG-3′) and TATA4 (5′-GCGCGGTACCCTTCTACAGACA-TGTTTACGG-3′), digested with *Cla*I and *Kpn*I and cloned into pFAT4 to give plasmid pFAT8. The deletion construct, pFAT8, would therefore encode a protein of 16 amino acids, of which the N- and C-terminal residues are derived from TatA, and the remainder specified by pBluescript polylinker DNA. The DNA covering the in-frame deletion of *tatA* was excised by digestion with *Xba*I and *Kpn*I and cloned into the polylinker of pMAK705 (Hamilton *et al*., 1989), to give the construct pFAT12. The mutant allele of *tatA* was transferred to the chromosome as described (Hamilton *et al*., 1989). The mutant strain, ELV15, obtained from this procedure was verified by PCR using primers TATA1 and TATA4, and the chromosomal PCR product was sequenced to ensure that no mismatched bases had been introduced.

The *tatE* gene was deleted in a similar manner. A 715 bp fragment covering the upstream region and first three codons of *tatE* was amplified using the primers TATE1 (5′-GCGCTCTAGAGAAAGTAAAC-GTAACATGATGACG-3′) and TATE2 (5′-GCCGGATCCCTCAC-CCATAGATACCTTCTTGAC-3′), digested with *Xba*I and *Bam*HI and cloned into the polylinker of pBluescript to give plasmid pFAT42. A 545 bp fragment covering the last two codons of *tatE* and downstream DNA was likewise amplified using primers TATE3 (5′-GCGCATCG-ATGAGTGACGTGGCGAGCAGGACGCTC-3′) and TATE4 (5′-GCG-CGGTACCGGATCGTCGTCTGGATATTCTGAC-3′), digested with *Cla*I and *Kpn*I and cloned into pFAT42 to give plasmid pFAT43. The DNA covering the in-frame deletion of *tatE* was excised by digestion with *Xba*I and *Kpn*I and again cloned into the polylinker of pMAK705, to give the construct pFAT44. The mutant allele of *tatE* was transferred to the chromosome of MC4100 to give the Δ*tatE* strain J1M1, and to the chromosome of ELV15 to give the Δ*tatA* Δ*tatE* strain JARV15 as described above. Mutant strains were verified by PCR using the primers TATE1 and TATE4, and chromosomal PCR products were sequenced as before.

For expression of the *tatABCD* genes, chromosomal DNA was amplified from 13 nucleotides upstream of the first potential start site of *tatA* to 14 nucleotides downstream of the *tatD* stop codon using the primers TATA5 (5′-GCGCGAATTCCGTGTAACGTATAATGCG-3′) and TATD1 (5′-GCGCTCTAGACGATGGTGAGGCTCGCTCC-3′), digested with *Eco*RI and *Xba*I and cloned into the polylinker of pT7.5 (Tabor and Richardson, 1985) to give plasmid pFAT65. Plasmid pFAT66 was constructed by digestion of pFAT65 with *Eco*RV, followed by gel purification of the 3.6 kb fragment and religation. Plasmid pFAT67 was constructed by digestion of pFAT65 with *Hin*dIII, gel purification of the 3.8 kb fragment and religation. Plasmid pFAT68 was constructed by digestion of pFAT65 with *Nru*I, isolation of the 4.7 kb band and religation.

The *tatABCD* sequence was obtained using the pFAT65 insert as the sequencing template. Where the sequence differed from that deposited by the *E.coli* genome project, the sequence changes were confirmed by sequencing-independent PCR amplifications of the region of interest. The revised DNA sequence for this region has been deposited in the DDBJ/EMBL/GenBank database under the accession No. AJO05830.

Construction of the TorA signal sequence–23K fusion was as follows: a 176 bp fragment of chromosomal DNA was amplified with primers TorASS1 (5′-GCGGAATTCAAGAAGGAAGAAAAATAATG-3′) and TorASS2 (5′-GCGGAATTCGGTACCGTCAGTCGCCGCTTG-3′). This covered DNA from 17 bases upstream of the TorA start codon to the sixth codon of the mature TorA sequence. The product was digested with *Eco*RI and cloned into the polylinker of pBluescript (Stratagene). A clone with the insert in the correct orientation was determined by digestion with *Kpn*I, and designated pMW11. The gene encoding the mature region of the spinach 23 kDa oxygen-evolving complex was excised from plasmid pOEC23mp (Clausmeyer *et al*., 1993) by digestion with *Avi*II and *Sal*I, end-filled with Klenow DNA polymerase, and cloned into *Eco*RV-digested pMW11 to give plasmid pMW18. The DNA covering the TorA–23K fusion was excised by digestion with *Sac*I and *Sal*I and cloned into pDHB5700 (kindly provided by G.von Heijne) to give pMW23. All clones constructed from PCR-amplified DNA were sequenced to ensure that no mismatches had been introduced during amplification.

During all genetic manipulations, *E.coli* strains were grown aerobically in Luria–Bertani medium. Antibiotics were added as required to final concentrations: ampicillin, 125 μg/ml; chloramphenicol, 15 μg/ml; rifampicin, 200 μg/ml; kanamycin, 40 μg/ml. Growth phenotypes were determined by growth on M9 minimal medium (Sambrook *et al*., 1989). Biochemical characterizations were carried out on cultures grown routinely in a medium based on that of Cohen and Rickenberg (1956) containing bactopeptone (0.5% w/v), casamino acids (0.1% w/v), thiamine (0.001% w/v), 1 μM ammonium molybdate and 1 μM potassium selenite. For aerobic growth and anaerobic fermentations, glucose was added to a final concentration of 0.4% (w/v). For anaerobic respiration, glycerol (0.5% w/v) or glucose (0.2% w/v) were added as carbon source, and TMAO, DMSO, sodium nitrate or sodium fumarate (all 0.4% w/v) were added as electron acceptors where indicated.

### *Preparation of subcellular fractions*

Cultures were harvested by sedimentation at 7000 *g* for 15 min at 4°C, and washed twice in cold 50 mM Tris–HCl pH 7.6. Cells were resuspended in a volume of 1 ml of the above buffer per gram (wet weight) of cells, and the periplasmic fraction was prepared by the lysozyme/EDTA method of Osborn *et al*. (1972). The resulting sphaeroplasts were lysed by passage through a French pressure cell and separated into membrane and cytosolic fractions as described (Ballantine and Boxer, 1985). Fractionation efficiency was monitored by acid phosphatase and β-lactamase as periplasmic marker enzymes (Atlung *et al*., 1989; Nivière *et al*., 1992). Glucose-6-phosphate dehydrogenase was assayed as a cytoplasmic marker by monitoring the glucose-6-phosphate-dependent reduction of $NADP^+$ at 340 nm. The assay was performed at 37°C in a buffer of 50 mM Tris–HCl pH 7.5, containing 250 μM $NADP^+$. The reaction was initiated by addition of 12.5 mM glucose-6-phosphate. Membrane vesicles were dispersed by the method of Ballantine and

Boxer (1985). Protein concentration was determined by the Lowry method (Lowry *et al*., 1951).

### Enzyme assays and activity stains

Nitrate reductase, TMAO reductase, DMSO reductase and fumarate reductase were assayed spectrophotometrically at 600 nm as the substrate-dependent oxidation of reduced benzyl viologen, as previously described (Jones and Garland, 1977; Bilous and Weiner, 1985; Silvestro *et al*., 1988; Kalman and Gunsalus, 1989). Activity stains of immunoprecipitates of FDH-N and hydrogenases 1 and 2 after rocket immunoelectrophoresis were based on those described (Enoch and Lester, 1975; Ballantine and Boxer, 1985).

### Expression of the tatABCD gene products

Expression of the *tatABCD* gene products from plasmids pFAT65–pFAT68 was under control of the phage T7 ϕ10 promoter. Plasmids were transformed into the strain K38 (HfrC, *phoA4*, *pit-10*, *tonA22*, *ompF627*, *relA1*, $\lambda^+$) (Lyons and Zinder, 1972) which carries the compatible plasmid pGP1-2 coding for the T7 RNA polymerase (Tabor and Richardson, 1985). Synthesis of plasmid-encoded gene products was induced by a temperature shift from 30 to 42°C and followed by labelling with [$^{35}$S]methionine according to the procedure described by Tabor and Richardson (1985).

### Pulse–chase experiments

*Escherichia coli* strains freshly transformed with pMW23 were grown aerobically overnight in Luria–Bertani medium. The cultures were then diluted 1:75 in CR medium supplemented with 0.4% glucose and 40 mM sodium nitrate in 7 ml bijou bottles filled to the top, and grown anaerobically at 37°C until they reached mid-log phase. Cells were then harvested and resuspended in CR medium lacking peptone and casamino acids but supplemented with a methionine-free amino acid mixture (0.2 mg/ml each amino acid). Cultures were grown anaerobically for 1 h, after which expression of TorA–23K was induced by the addition of isopropyl-β-D-galactopyranoside (IPTG; 0.04 mM) for 30 min. Aliquots (5 ml) of the culture were then incubated with 30 μCi of [$^{35}$S]methionine for 1 min, after which cold methionine was added to a concentration of 0.5 mg/ml. Where appropriate, sphaeroplasts were formed after centrifugation at 14 000 r.p.m. for 2 min, resuspension in ice-cold buffer (40% w/v sucrose, 33 mM Tris pH 8.0) and incubation with lysozyme (5 μg/ml in 1 mM EDTA) for 15 min on ice. Aliquots of the sphaeroplasts were incubated on ice for 1 h in either the presence or absence of 0.3 mg/ml proteinase K. At the end of this period, phenylmethylsulfonyl fluoride (PMSF) was added (final concentration 0.33 mg/ml) and the samples were precipitated with trichloroacetic acid (5%). The precipitate was pelleted, resuspended in 10 mM Tris pH 7.5, 2% SDS and immunoprecipitated with antisera to OmpA (kindly provided by G.von Heijne) or 23K.

### Electrophoretic and immunological procedures

Proteins were separated by SDS–PAGE, as previously described (Laemmli, 1970). Antibodies to hydrogenase 1 were raised in New Zealand white rabbits using the protocol of Harboe and Ingild (1973). Western immunoblot was performed as previously described (Towbin *et al*., 1979) and antigen detected using the Promega Protoblot detection system (Promega, UK). Rocket immunoelectrophoresis was performed on soluble cell fractions and Triton X-100-dispersed membranes as described (Graham *et al*., 1980). Immunoprecipitation of hydrogenase 1 was performed as described (Sawers *et al*., 1985).

## Acknowledgements

We would like to thank Dr Gary Sawers for helpful discussion and advice, and for the gift of purified hydrogenase 1. We acknowledge Professor David Boxer for providing antibodies to hydrogenase 2 and FDH-N. We are grateful to G.von Heijne and J.-W.de Gier for advice on the pulse–chase experiments. This research was supported by a Royal Society Research Fellowship and BBSRC grant to T.P., and a project grant from the Leverhulme Trust (T.P. and B.C.B.). E.G.B. and N.R.S. are the recipients of, respectively, a BBSRC Research Studentship and a Norwich Research Park Studentship.

## References

Andrews,S.C., Berks,B.C., McClay,J., Ambler,A., Quail,M.A., Golby,P. and Guest,J.R. (1997) A 12-cistron *Escherichia coli* operon (*hyf*) encoding a putative proton-translocating formate hydrogenlyase system. *Microbiology*, **143**, 3633–3647.

Atlung,T., Nielson,A. and Hansen,F.G. (1989) Isolation, characterisation and nucleotide sequence of *appY*, a regulatory gene for growth-phase dependent gene expression in *Escherichia coli*. *J. Bacteriol.*, **171**, 1683–1691.

Bailey,M.J.A., Hughes,C. and Koronakis,V. (1997) RfaH and the *ops* element, components of a novel system controlling bacterial transcriptional elongation. *Mol. Microbiol.*, **26**, 845–851.

Ballantine,S.P. and Boxer,D.H. (1985) Nickel-containing hydrogenase isoenzymes from anaerobically grown *Escherichia coli* K-12. *J. Bacteriol.*, **163**, 454–459.

Barrett,E.L. and Kwan,H.S. (1985) Bacterial reduction of trimethylamine oxide. *Annu. Rev. Biochem*., **39**, 131–149.

Berg,B.L. and Stewart,V. (1990) Structural genes for nitrate-inducible formate dehydrogenase in *Escherichia coli* K-12. *Genetics*, **125**, 691–702.

Berg,B.L., Li,J., Heider,J. and Stewart,V. (1991) Nitrate-inducible formate dehydrogenase in *Escherichia coli* K-12. I. Nucleotide sequence of the *fdnGHI* operon and evidence that opal (UGA) encodes selenocysteine. *J. Biol. Chem.*, **266**, 22380–22385.

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

Bernhard,M., Benelli,B., Hochkoeppler,A., Zannoni,D. and Friedrich,B. (1997) Functional and structural role of the cytochrome *b* subunit of the membrane-bound hydrogenase complex of *Alcaligenes eutrophus* H16. *Eur. J. Biochem.*, **248**, 179–186.

Bilous,P.T., Cole,S.T., Anderson,W.F. and Weiner,J.H. (1988) Nucleotide sequence of the *dmsABC* operon encoding the anaerobic dimethyl sulfoxide reductase of *Escherichia coli*. *Mol. Microbiol*., **2**, 785–795.

Bilous,P.T. and Weiner,J.H. (1985) Dimethyl sulfoxide reductase activity by anaerobically grown *Escherichia coli* HB101. *J. Bacteriol*., **162**, 1151–1155.

Blattner,F.R. *et al*. (1997) The complete genome sequence of *Escherichia coli* K-12. *Science*, **277**, 1453–1462.

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.

Brondsted,L. and Atlung,T. (1994) Anaerobic regulation of the hydrogenase 1 (*hya*) operon of *Escherichia coli*. *J. Bacteriol.*, **176**, 5423–5428.

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

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

Clausmeyer,S., Klösgen,R.B. and Herrmann,R.G. (1993) Protein import into chloroplasts. The hydrophilic lumenal proteins exhibit unexpected import and sorting specificities in spite of structurally conserved transit peptides. *J. Biol. Chem.*, **268**, 13869–13876.

Chaddock,A.M., Mant,A., Karnauchov,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.

Cohen,G.N. and Rickenberg,H.W. (1956) Concentration specifique reversible des amino acids chez *Escherichia coli*. *Ann. Inst. Pasteur (Paris)*, **91**, 693–720.

Creighton,A.M., Hulford,A, Mant,A., Robinson,D. and Robinson,C. (1995) A monomeric, tightly folded stromal intermediate on the ΔpH-dependent thylakoidal protein transport pathway. *J. Biol. Chem.*, **270**, 1663–1669.

Daniels,D.L., Plunkett,G.,III, Buland,V. and Blattner,F.R. (1992) Analysis of the *Escherichia coli* genome: DNA sequence of the region from 84.5 to 86.5 minutes. *Science*, **257**, 771–778.

Dreusch,A., Bürgisser,D.M., Heizmann,C.W. and Zumft,W.G. (1997) Lack of copper insertion into unprocessed cytoplasmic nitrous oxide reductase generated by an R20D substitution in the arginine consensus motif of the signal peptide. *Biochim. Biophys. Acta*, **1319**, 311–318.

Enoch,H.G. and Lester,R.L. (1975) The purification and properties of formate dehydrogenase and nitrate reductase from *Escherichia coli*. *J. Biol. Chem.*, **250**, 6693–6705.

Fincher,V., McCaffery,M. and Cline,K. (1998) Evidence for a loop mechanism of protein transport by the thylakoid Delta pH pathway. *FEBS Lett.*, **423**, 66–70.

Glockner,A.B. and Zumft,W.G. (1996) Sequence analysis of an internal 9.72-kb segment from the 30-kb denitrification gene cluster of *Pseudomonas stutzeri*. *Biochim. Biophys. Acta*, **1277**, 6–12.

Graham,A. (1981) The organisation of hydrogenase in the cytoplasmic membrane of *Escherichia coli*. *Biochem. J.*, **197**, 283–291.

Graham,A., Jenkins,H.E., Smith,N.H., Mandrand-Berthelot,M.-A., Haddock,B.A. and Boxer,D.H. (1980) The synthesis of formate dehydrogenase and nitrate reductase proteins in various *fdh* and *chl* mutants of *Escherichia coli*. *FEMS Microbiol. Lett.*, **7**, 145–151.

Hamilton,C.M., Aldea,M., Washburn,B.K., Babitzke,P. and Kushner,S.R. (1989) New method for generating deletions and gene replacements in *Escherichia coli*. *J. Bacteriol.*, **171**, 4617–4622.

Harboe,N. and Ingild,A. (1973) Immunisation, isolation of immunoglobulins and estimation of antibody titre. *Scand. J. Immunol. Suppl. 1*, **2**, 161–164.

Izard,J.W. and Kendall,D.A. (1994) Signal peptides: exquisitely designed transport promoters. *Mol. Microbiol.*, **13**, 765–773.

Jacobi,A., Rossmann,R. and Böck,A. (1992) The *hyp* operon gene products are required for the maturation of catalytically active hydrogenase isoenzymes in *Escherichia coli*. *Arch. Microbiol.*, **158**, 444–451.

Jones,R.W. and Garland,P.B. (1977) Sites and specificity of the reaction of bipyridilium compounds with anaerobic respiratory enzymes of *Escherichia coli*. *Biochem. J.*, **164**, 199–211.

Kalman,L.V. and Gunsalus,R.P. (1989) Identification of a second gene involved in global regulation of fumarate reductase and other nitrate-controlled genes for anaerobic respiration in *Escherichia coli*. *J. Bacteriol.*, **171**, 3810–3816.

Laemmli,U.K. (1970) Cleavage of structural head proteins during assembly of the head of bacteriophage T4. *Nature*, **227**, 280–285.

Lowry,O.H., Rosebrough,N.J., Farr,A.L. and Randall,R.J. (1951) Protein measurements with the Folin phenol reagent. *J. Biol. Chem.*, **193**, 265–275.

Lyons,L.B. and Zinder,N.D. (1972) The genetic map of the filamentous bacteriophage f1. *Virology*, **49**, 45–60.

Macinga,D.R., Cook,G.M., Poole,R.K. and Rather,P.N. (1998) Identification and characterization of *aarF*, a locus required for production of ubiquinone in *Providencia stuartii* and *Escherichia coli* and for expression of 2′-*N*-acetyltransferase in *P.stuartii*. *J. Bacteriol.*, **180**, 128–135.

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.

Menon,N.K., Robbins,J., Peck.,H.D.,Jr, Chatelus,C.Y., Choi,E.-S. and Przybyla,A.E. (1990) Cloning and sequencing of a putative *Escherichia coli* [NiFe] hydrogenase-1 operon containing six open reading frames. *J. Bacteriol.*, **172**, 1969–1977.

Menon,N.K., Robbins,J., Dervartanian,M., Patil,D., Peck,H.D.,Jr, Menon,A.L., Robson,R.L. and Przybyla,A.E. (1993) Carboxy-terminal processing of the large subunit of [NiFe] hydrogenases. *FEBS Lett.*, **331**, 91–95.

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

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.

Osborn,M.J., Gander,J.E. and Parisi,E. (1972) Mechanism of assembly of the outer membrane of *Salmonella typhimurium*. *J. Biol. Chem.*, **247**, 3973–3986.

Pascal,M.C., Burini,J.F. and Chippaux,M. (1984) Regulation of trimethylamine *N*-oxide (TMAO) reductase of *Escherichia coli* K-12: analysis of *tor*::MudI fusion. *Mol. Gen. Genet.*, **195**, 351–355.

Pugsley,A.G. (1993) The complete general secretory pathway in Gram-negative bacteria. *Microbiol. Rev.*, **57**, 50–108.

Robinson,C. and Mant,A. (1997) Targeting of proteins into and across the thylakoid membrane. *Trends Plant Sci.*, **2**, 431–437.

Rodrigue,A., Boxer,D.H., Mandrand-Berthelot,M.A. and Wu,L.-F. (1996) Requirement for nickel of the transmembrane translocation of NiFe-hydrogenase 2 in *Escherichia coli*. *FEBS Lett.*, **392**, 81–86.

Roffey,R.A. and Theg,S.M. (1996) Analysis of the import of carboxyl-terminal truncations fo the 23-kilodalton subunit of the oxygen-evolving complex suggests that its structure is an important determinant for thylakoid transport. *Plant Physiol.*, **111**, 1329–1338.

Rossmann,R., Maier,T., Lottspeich,F. and Böck,A. (1995) Characterisation of a protease from *Escherichia coli* involved in hydrogenase maturation. *Eur. J. Biochem.*, **227**, 545–550.

Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

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

Sargent,F., Ballantine,S.P., Rugman,P.A., Palmer,T. and Boxer,D.H. (1998) Reassignment of the gene encoding the *Escherichia coli* hydrogenase 2 small subunit: identification of a soluble precursor of the small subunit in a *hypB* mutant. *Eur. J. Biochem.*, **255**, in press.

Sawers,G. (1994) The hydrogenases and formate dehydrogenases of *Escherichia coli*. *Antonie Van Leeuwenhoek*, **66**, 57–88.

Sawers,R.G., Ballantine,S.P. and Boxer,D.H. (1985) Differential expression of hydrogenase isoenzymes in *Escherichia coli* K-12: evidence for a third isoenzyme. *J. Bacteriol.*, **164**, 1324–1331.

Schatz,G. and Dobberstein,B. (1996) Common principles of protein translocation across membranes. *Science*, **271**, 1519–1526.

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

Silvestro,A., Pommier,J. and Giordano,G. (1988) The inducible trimethylamine-*N*-oxide reductase of *Escherichia coli* K12: biochemical and immunological studies. *Biochim. Biophys. Acta*, **954**, 1–13.

Tabor,S. and Richardson,C.C. (1985) A bacteriophage T7 polymerase/promoter system for controlled exclusive expression of specific genes. *Proc. Natl Acad. Sci. USA*, **82**, 1074–1078.

Towbin,H., Staehelin,T. and Gordon,J. (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets; procedure and some applications. *Proc. Natl Acad. Sci. USA*, **76**, 4350–4357.

Voelker,R. and Barkan,A. (1995) Two nuclear mutations disrupt distinct pathways for targeting proteins to the chloroplast thylakoid. *EMBO J.*, **14**, 3905–3914.

von Heijne,G. (1985) Signal sequences. The limits of variation. *J. Mol. Biol.*, **184**, 99–105.

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

Weiner,J.H., Bilous,P.T., Shaw,G.M., Lubitz,S.P., Frost,L. Thomas,G.H., Cole,J.A. and Turner,R.J. (1998) A novel and ubiquitous system for membrane targeting and secretion of cofactor-containing proteins. *Cell*, **93**, 93–101.

*Received April 14, 1998; accepted April 29, 1998*

## Note added in proof

To exclude the possibility that the differences between the *tat* sequences reported in this work and those reported by Weiner *et al*. (1998) was a consequence of the different *E.coli* parental strains employed in the two studies, DNA covering the *tatAB* region of *E.coli* K-12 strain HB101, the parental strain employed by Weiner *et al*. (1998), was amplified and sequenced. The HB101 sequence was identical to that reported here for strain MC4100.