Escherichia coli TatA and TatB Proteins Have N-out, C-in Topology in Intact Cells

Koch, Sabrina; Fritsch, Maximilian J.; Buchanan, Grant; Palmer, Tracy

Published in:
Journal of Biological Chemistry

DOI:
10.1074/jbc.M112.354555

Publication date:
2012

Document Version
Peer reviewed version

The *Escherichia coli* TatA and TatB proteins have an N-out C-in topology in intact cells*.

Sabrina Koch[†], Maximilian J. Fritsch[†], Grant Buchanan and Tracy Palmer[1]

Division of Molecular Microbiology, College of Life Sciences, University of Dundee, Dundee DD1 5EH, UK

*Running title: TatA and TatB have an N-out C-in topology

To whom correspondence should be addressed: Tracy Palmer, Division of Molecular Microbiology, College of Life Sciences, University of Dundee, Dundee DD1 5EH, UK telephone +44 (0)1382 386464, fax +44 (0)1382 388216, e-mail t.palmer@dundee.ac.uk

[†]These authors contributed equally to this study.

Keywords: Twin arginine protein transport, *Escherichia coli*, TatA, TatB, topology, thiol labelling.

Background: The Tat pathway transports folded proteins across energy coupling membranes.
Results: TatA has a fixed N-out, C-in topology in intact cells that is not altered by the absence of other Tat components or over-production of a Tat substrate.
Conclusion: The TatA amphipathic helix does not re-orient during protein translocation.
Significance: Topological inversion of TatA does not accompany protein transport by the Tat pathway.

SUMMARY

The twin-arginine protein transport (Tat) system translocates folded proteins across the cytoplasmic membrane of prokaryotes and the thylakoid membrane of chloroplasts. In *Escherichia coli* TatA, TatB and TatC, are essential components of the machinery. A complex of TatB and TatC acts as the substrate receptor, whilst TatA is proposed to form the Tat transport channel. TatA and TatB are related proteins that comprise an N-terminal transmembrane helix and an adjacent amphipathic helix. Previous studies addressing the topological organization of TatA have given conflicting results. In this study we have addressed the topological arrangement of TatA and TatB in intact cells by labelling of engineered cysteine residues with the membrane impermeable thiol reagent methoxypolyethylene glycol maleimide. Our results show that TatA and TatB share an N-out, C-in topology, with no evidence that the amphipathic helices of either protein are exposed at the periplasmic side of the membrane. We further show that the N-out, C-in topology of TatA is fixed and is not affected by the absence of other Tat components or by the overproduction of a Tat substrate. These data indicate that topological reorganization of TatA is unlikely to accompany Tat-dependent protein transport.

The Sec and Tat pathways operate in parallel to transport proteins across the cytoplasmic membrane of bacteria and archaea and the thylakoid membrane of plants. Whilst the Sec machinery can only export unfolded proteins, substrates of the Tat pathway are folded prior to transport. Transport by the Tat pathway is driven solely by the transmembrane proton electrochemical gradient, Δp (reviewed in (1,2). Substrates are targeted to each of these transport machineries by means of N-terminal cleavable signal peptides. One of the major distinguishing features of Tat targeting signals is that they contain a conserved twin arginine motif, which harbors consecutive and usually invariant arginine residues that are essential for efficient transport by the Tat machinery (3,4).

In Gram negative bacteria and in plant thylakoids, the Tat machinery is made up of three membrane proteins which in *Escherichia coli* are termed TatA, TatB and TatC (5-8). In *E. coli* a fourth protein, TatE, is a minor component of the Tat pathway and has an identical function to TatA (7,9). The TatA and TatB proteins share some primary sequence homology and evolved from a common ancestor, but have functionally distinct roles during Tat transport (8,10). TatB is found almost exclusively as part of the TatBC complex (11). This complex, which contains multiple copies of each protein, interacts with twin arginine

signal peptides and acts as the receptor for Tat substrates (11-16).

TatA can be purified as an array of large homo-oligomeric complexes. Analysis of these complexes by negative stain electron microscopy reveals that they form a series of related channel-like structures of different sizes, with internal cavities big enough to accommodate folded proteins, consistent with the idea that TatA forms the protein conducting channel (17,18). Large assemblies of fluorophore-tagged TatA have also been observed *in vivo*, and homo-oligomers containing at least 16 copies of the plant ortholog of TatA, Tha4, have been detected by cross-linking during Tat transport in thylakoid membranes (19,20). In *E. coli* cells the formation of TatA assemblies is dependent upon the presence of the TatBC complex. In the absence of TatB or TatC, TatA is arranged as much smaller units, possibly tetramers, suggesting that interaction with TatBC is required to drive the polymerisation of smaller units of TatA into larger assemblages (20). In resting thylakoid membranes, cross-linking studies of Tha4 are also consistent with this protein existing as a tetrameric unit (19). Transient interactions of Tha4 with the thylakoid equivalent of the TatBC complex have detected by cross-linking, dependent upon the presence of substrate and a Δp (21).

The TatA and TatB proteins have a common structural arrangement, comprising a single transmembrane helix (TMH), followed by an amphipathic helix (APH) and an unstructured C-terminal tail (18,22; Fig 1A,B). Application of the positive inside rule (23) would suggest that the N-termini of both proteins are located at the periplasmic side of the membrane with the C-termini in the cytoplasm, and this is consistent with protease accessibility experiments which indicated that the C-termini of TatA and TatB are only accessible to protease digestion in inside-out and not right-side-out membrane vesicles (18,22). A similar topology was inferred from protease mapping of Tha4 in isolated thylakoids (24). The N-out topology of TatA is also supported by the observation that the TatA protein from *Providencia stuartii*, which is synthesized with an inactivating N-terminal extension of eight amino acids, is processed by a membrane-embedded rhomboid protease which has its active site close to the periplasmic face of the membrane (25-27). However, a recent study probing the accessibility of cysteine substituted TatA to sulfhydryl labelling reagents in whole cells appeared to show that the N-terminus of TatA was localized at the cytoplasmic side of the membrane (28).

Some models for Tat transport assume that the APH of TatA may re-orient during transport (Fig 1C), for example by folding into a channel assembled from TatA transmembrane helices like a trap-door in response to a pulling force on the substrate (1,29,30). Early support for this model was provided by Gouffi *et al*. (31) who used compartment-sensitive marker proteins fused to the end of the APH of TatA to infer that this region of TatA was exposed at both sides of the membrane. Similar dual topology was also seen when a much smaller fusion, that of a TEV protease cleavage sequence, was inserted between residues 53 and 54 of TatA since this site was also shown to be protease accessible from either side of the membrane. This led the authors to conclude that the TatA APH has a dual topology and that topology changes of this region of TatA are associated with protein transport (31). Support for a helical hairpin arrangement of TatA was also provided by Chan *et al*. (28) who showed that in whole cells cysteine residues in the APH or C-terminus of TatA were not labelled by a membrane-impermeable thiol reagent. They further showed that in the presence of an uncoupler the labelling pattern of a cysteine present in the APH of TatA was altered, suggesting that topological changes in the APH were dependent upon Δp.

In this study we have re-visited the topological organization of TatA and TatB in whole cells by direct labelling of engineered cysteine residues. Our results clearly show that TatA has a fixed N-out, C-in topology that is not altered by the absence of other Tat components or by the overproduction of a Tat substrate.

EXPERIMENTAL PROCEDURES

E. coli Strains and Plasmids. Strains and plasmids used in the present work are shown in Tables 1 and 2, respectively. All strains used in this study are derivatives of MC4100 (32).

Plasmids in strain DADE-P (22) were used to overexpress the *E. coli tatABC* operon with alanine substitutions of all four cysteine codons in *tatC* and single cysteine codon substitutions in *tatA* or *tatB*. Cysteine substitutions were introduced by QuikChange™ site-directed mutagenesis (Stratagene) in plasmid pUNITATA (30) or UNITATB (22) resulting in plasmid series pUNITATAX#C or pUNITATBX#C, respectively, where X corresponds to the single letter amino acid code and # to the position of the substituted

codon. Each *tatA* mutation was sub-cloned into the *Eco*RI and *Pml*I sites of plasmid pUNITATCC4 (22) giving rise to plasmid series pUNITATCC4AX#C. Each *tatB* mutation was sub-cloned into the *Pml*I and *Afl*II sites of plasmid pUNITATCC4 (22) giving rise to plasmid series pUNITATCC4BX#C. Primer sequences used for QuikChange™ mutagenesis are available on request.

For the cysteine insertion between residues 1 and 2 of TatA, DNA covering the *tatA* start codon and upstream DNA was amplified with primers TatPromXcaBamrev (5'-GCGCGGATCCGTATA CATGTTCCTCTGTGGTAGATG-3') and TATA5 (7) and cloned into the *Eco*RI and *Bam*HI sites of pBluescript KS^+ (Stratagene) to give plasmid pBSTatAPromXcaI. Plasmid pBSTatAins2C was constructed following amplification of *tatA* using primers TatAins2C (5'-TGTGTGGTGGTATCAGTATT-3') and TatAEcPmlBam (5'-GCGCGGATCCCACGTGT TACACCTGCTCTTTATCG-3') and subsequent cloning of the PCR product into the *Bam*HI and *Xca*I (blunt-end) sites of plasmid pBSTatAPromXcaI. The *tatA* gene with the cysteine insertion was subsequently excised by digestion with *Eco*RI and *Pml*I and cloned into similarly digested pUNITATCC4 to give pUNITATCC4Ains2C.

A cysteine insertion between codons G21 and T22 was constructed as follows. DNA covering *tatA* up to codon 21 was amplified using primers TatATMH1 (5'-GCGCGGATCCTGGCCAAAAA GCAGTACAACGATGA-3') and TATA5 and cloned into pBluescript as an *Eco*RI-*Bam*HI fragment. DNA covering *tatA* from codon 22 onwards was amplified using TatAGly1 (5'-CGGCACCAAAAAGCTCGGCTCCATCGG-3') and UNIA1 (30), digested with *Bam*HI and cloned into the above plasmid that had been previously digested with *Msc*I (blunt end) and *Bam*HI, to give plasmid pTatAGly1. This construct codes for TatA with a glycine insertion between G21 and T22. The extra gly codon was subsequently changed to a cys codon by QuikChange™ site-directed mutagenesis to give plasmid pBSTatAins21C. The *tatA* genes with the cysteine insertion was subsequently excised by digestion with *Eco*RI and *Pml*I and cloned into similarly digested pUNITATCC4 to give plasmid pUNITATCC4Ains21C.

The *tatABC* operon in single copy harboring single cysteine codon substitutions in *tatA* and alanine substitutions of all four cysteine codons of *tatC*, was expressed from the chromosomal lambda phage attachment site, *attB*, of strain DADE (33). The *tatA* gene and approximately 100 bp of upstream promoter region were amplified from *E. coli* chromosomal DNA with primers UNIREP1 and UNIA1 and cloned into the *Eco*RI and *Bam*HI sites of pBluescript KS^+ generating plasmid pKSuniA. Single cysteine substitutions in *tatA* were introduced by QuikChange™ site-directed mutagenesis in pKSuniA giving plasmid series pKSuniAX#C. The *tatABC* operons under control of the *tatA* promoter in plasmids pUNICC–AX#C were sub-cloned into the *Eco*RI and *Bam*HI sites of plasmid pRS552 (34) resulting in plasmid series pRSUNICC–AX#C. The *tatABC* operon under control of the *tatA* promoter from pRSUNICC–AX#C was subsequently integrated into the *attB* site of strain DADE as described (34) resulting in strains MF1–13 (Listed in Table 1). A series of strains expressing only *tatA* were also constructed. In this case, the *tatA* variants with single cysteine substitutions under control of the *tat* promoter in the pUNITATAX#C plasmid series were sub-cloned into the *Eco*RI and *Bam*HI site of pRS552 and subsequently integrated into the *attB* site of strain DADE resulting in strains SK1–13 (Table 1).

The *Paracoccus panthotrophus soxYZ* genes were amplified from plasmid pVS005 (35) using primers BamHAsoxY (5-GCGCGGATCCATGTA TCCGTACGATGTGCCGGACTATGCGAGCAC CGTTGACGAGTTG-3') and soxZHind (5'-GCGCAAGCTTTTAGGCGACTGCG). This also introduces an N-terminal hemagglutinin-tag (HA-tag) in place of the SoxY signal sequence. The *soxYZ* PCR product (as a *Bam*HI-*Hind*III fragment) and the *tatA* promoter (released as an *Eco*RI-*Bam*HI fragment from pSUPROM (36) were cloned by three-way ligationn into *Eco*RI-*Hind*III-digested pSU20 (37) to give plasmid pHASoxYZ (38). The *tatA* promoter and *soxYZ* from pHASoxYZ were subsequently cloned into the *Eco*RI and *Hind*III sites of the low-copy vector pTH19cr (39) resulting in plasmid pTH19SoxYZ.

Culture Conditions, Fractionation and Protein Methods. Unless indicated otherwise, liquid cultures were inoculated with 1/100 volume of an overnight culture in LB medium supplemented with appropriate antibiotics and grown with vigorous shaking at 37° (40). Growth assays to test resistance to SDS were performed as described (41). Growth assays to test Tat-dependent growth with trimethylamine-*N*-oxide (TMAO) as sole electron acceptor were carried out by growing strains anaerobically at 37°C for up to four days on

solid M9 minimal medium supplemented with 0.4% (w/v) TMAO and 0.5% (v/v) glycerol (41). Cultures for TMAO reductase assays were grown anaerobically in 50 ml LB medium supplemented with 0.4% (w/v) TMAO and 0.5% (v/v) glycerol at 37°C overnight (8). Cells were separated into spheroplast and periplasmic fractions using lysozyme/EDTA treatment in a high sucrose buffer followed by centrifugation (42). TMAO:benzylviologen oxidoreductase activity was measured in periplasmic fractions of TMAO-grown cells as described previously (43).

SDS-PAGE and immunoblotting analysis was performed as described (44,45) and immunoreactive bands were detected using a chemiluminescent horseradish peroxidase (HRP) substrate kit (Millipore). Antisera against *E. coli* TatA and TatB (46) were used to detect the proteins, with an anti-rabbit IgG HRP conjugate (Bio-Rad) used as secondary antibody. SufI was detected using a polyclonal anti-SufI antiserum (41). Hemagglutinin-tagged SoxY was detected with anti-hemagglutinin tag (HA tag) HRP conjugate (Sigma-Aldrich).

Sulfhydryl Labelling. For sulfhydryl labelling with crude membrane fractions, cells were grown in 25 ml LB medium at 37°C overnight, harvested by centrifugation at 5000 × *g* and washed in Buffer K (50 mM triethanolamine-HCl, pH 7.5, 250 mM sucrose, 1 mM Na_2EDTA). Cell pellets were resuspended in 1 ml Buffer K supplemented with protease inhibitors (Complete Mini, EDTA-free protease inhibitor mixture tablets; Roche Applied Science). Cells were disrupted on ice by five cycles of sonication with 15 second pulses and 15 second intervals in between. Cell debris was pelleted by centrifugation at 16,000 × *g* for 10 minutes and the crude membrane fraction was pelleted from the supernatant by ultra-centrifugation at 278,000 × *g* for 30 minutes. The crude membrane fraction was resuspended in 1 ml HEPES/NaCl buffer (50 mM HEPES, pH 6.8, 50 mM NaCl) and diluted to 10–15 μg μl^{-1} of total protein. 10 μl of the membrane fraction were labelled with 5 mM Methoxypolyethylene glycol maleimide (MW 5000 Da, Sigma-Aldrich) in a final volume of 50 μl HEPES/NaCl buffer at room temperature for one hour. Control samples were incubated with buffer alone or were treated with 1% Triton-X 100 or 1% sodium dodecyl sulfate prior to labelling. The sulfhydryl labelling reaction was stopped by addition of 5 μl 0.5 M dithiothreitol (DTT) and mixed 1:1 with 2× Laemmli sample buffer (Bio-Rad).

For sulfhydryl labelling with intact cells, cultures were grown at 37°C to mid-logarithmic phase and normalised to an OD_{600} of 0.3. Cells were harvested by centrifugation at 5000 × *g*, washed with HEPES/$MgCl_2$ buffer (50 mM HEPES, pH 6.8, 5 mM $MgCl_2$) and cell pellets were resuspended in 1 ml of HEPES/$MgCl_2$ buffer. 80 μl of cell suspension were incubated with 5 mM Methoxypolyethylene glycol maleimide (MAL-PEG) at room temperature for one hour in presence of 5 mM EDTA and in a final volume of 100 μl HEPES buffer. The sulfhydryl labelling reaction was stopped with 25 μl 0.5 mM DTT and proteins were precipitated as described (47). Protein precipitates were resolubilized in 70 or 100 μl 2× Laemmli sample buffer depending if TatA variants of samples were expressed from the chromosome or from plasmid, respectively.

RESULTS

Cysteine-substituted TatA proteins can be labelled with methoxypolyethylene glycol maleimide (MAL-PEG). Previous studies looking at cysteine accessibility of *E. coli* TatA and TatC proteins have used an indirect method to detect sulfhydryl labelled proteins (28,48). N^{α}-(3-maleimidylpropionyl)-biocytin or MPB is a membrane permeable maleimide reagent with a molecular mass of around 500 Da that can subsequently be detected by binding of streptavidin. 4-acetamido-4'-maleimidylstilbene-2,2'-disulfonic acid (AMS) is a membrane impermeable maleimide of similar molecular mass to MPB. Periplasmically-located cysteines in TatA and TatC were identified by pre-treating cells with AMS, which will only react with external cysteines, followed by MPB, which will react with any cysteines that have not been previously labelled with AMS (28,48). Since these reagents effectively do not alter the apparent mass of the protein with which they react, labelling can only be detected once the protein has been purified and incubated with streptavidin. According to this method periplasmic cysteines are identified because they show a lower level of MPB labelling than cytoplasmic cysteines.

Because this approach is rather time-consuming, and requires significant amounts of material (due to losses at the purification step), we employed a more direct approach to assess cysteine accessibility. Methoxypolyethylene glycol maleimide (MAL-PEG) is a membrane-impermeable maleimide of much larger mass (approx. 5,000 Da) that results in a clear size-shift of the labelled protein which can be detected by

western blotting (38,49). To assess whether cysteine substitutions in TatA that might be expected to lie close to the membrane were accessible to labelling by the relatively large MAL-PEG reagent, we firstly isolated membrane fractions from *E. coli* strains producing plasmid encoded TatA variants along with TatB and a cysteine-less variant of TatC. Cysteine residues close to the N-terminus of TatA (i.e. a G2C substitution or a variant with a cys insertion between M_1 and G_2 termed Ins2C) clearly reacted with MAL-PEG to give a species that migrated more slowly than unlabelled TatA (Fig 2A). Also as expected a TatA protein lacking cys residues did not label with MAL-PEG, and a cys residue that is predicted to be buried in the TMH (I12C) did not label with MAL-PEG unless detergent was also included in the labelling reaction.

Cysteine residues in the hinge region between the TMH and APH showed poor labelling in the absence of detergents (Fig 2B). Thus cys residues at positions 18, 19 and 20 did not react with MAL-PEG in isolated membranes, nor did a cys residue that had been inserted between G21 and T22. The first residue in this hinge region to become accessible to PEGylation was a cys substitution of position 22. This suggests that the preceding residues are buried within the membrane, which is in accordance with the findings of a solid state NMR study on the *Bacillus subtilis* $TatA_d$ protein (50). All of the cys substitutions that were tested along the APH were freely accessible to MAL-PEG labelling in isolated membranes (Fig 1C), indicating that for at least a proportion of the TatA molecules the APH must lay along the membrane rather than being in a transmembrane orientation.

Labelling of accessible cys residues with MAL-PEG in intact cells. The outer membranes of Gram negative bacteria are generally permeable to hydrophilic molecules < 600 Da due to the presence of porins, but do not allow passage of molecules larger than this. Therefore to test whether any of the membrane-extrinsic cys substitutions of TatA were accessible to MAL-PEG in whole cells, it was necessary to devise a method which would allow the compound to permeate across the outer membrane. EDTA chelates divalent metal ions that are associated with the lipopolysaccharide, offsetting their stabilizing effect and increasing permeability (51). We therefore added 5mM EDTA to cell suspensions prior to addition of MAL-PEG. Control experiments showed that the EDTA treatment did not affect the viability of cells (Fig S1).

MAL-PEG accessibility was tested for TatA variants harboring cys substitutions at position 2, Ins2 and 22 to determine the topology of the TMH, and 33, 35, 39, 41, 45, 47, 49, 60, 78 and 89 to determine the localization and topology of the APH and C-terminal tail. Some of these cys substitutions have been described previously (30) and it was shown that the G2C, S35C and K41C variants supported a high level of Tat transport activity, T22C supported a low but detectable level of Tat transport whilst G33C and F39C inactivated the function of TatA. When we assessed the activity of all of the new substitutions constructed here (Ins2C, D45C, E47C, K49C, T60C, T78C and V89C) each of them supported high Tat transport activity (Fig S2A) indicating that TatA function was preserved.

As a control for each of the labelling reactions the *E. coli* strains co-produced a cytoplasmic form of the *Paracoccus pantotrophus* SoxY protein carrying an N-terminal hemagglutanin epitope tag. This protein contains a highly exposed single cys residue on a flexible arm at the C-terminus of the protein (35) and any MAL-PEG labelling of this indicated that the integrity of the cytoplasmic membrane was compromised.

TatA variants containing cys substitutions at the N-terminus were clearly accessible to PEGylation in intact cells, whilst the T22C variant was not (Fig 3A). The T22C variant was, however, PEGylated in the presence of SDS. Analysis of SoxY labelling indicated that the cytoplasmic membrane was not breached during the procedure. These data show that the TatA TMH has an N-out, C-in topology. We noted that the detectability of SoxY and TatA appears to increase dramatically after PEGylation. This may be due to better transfer and/or increased binding of the PEGylated proteins to the nitrocellulose membrane during electroblotting.

Fig 3B and C show the labelling patterns for cys residues located in the APH (Fig 3B) and C-tail (Fig 3C). Each of the positions gives a labelling pattern that is identical to that for T22C, with the single exception of the V89C substitution, which showed some labelling in the absence of SDS. However, in this case we noticed that the integrity of the inner membrane was compromised resulting in PEGylation of SoxY. We routinely observed that the plasmid-produced V89C substitution induced leakage of the inner membrane towards MAL-PEG, for reasons that are not clear. However, taken together the results clearly indicate that the APH and C-tail of TatA have a cytoplasmic localization, with no evidence

that the APH can flip across the membrane.

TatA shows the same topology when produced at native levels and in the absence of TatBC. All of the results obtained above, and in the experiments undertaken by Chan *et al*. (28) studied TatA variants that were overproduced from a multicopy plasmid (along with TatB and TatC). It has been shown that even a relatively low level of *tatABC* overexpression can result in an aberrant localization of TatA, which accumulates in the cytoplasm as tube-like structures (52). The presence of cytoplasmic forms of TatA might be expected to interfere with labelling experiments because a proportion of the TatA protein will have its N-terminus located in the cytoplasm rather than in a transmembrane orientation. We therefore repeated our PEGylation experiments with TatA proteins produced at chromosomal levels.

DNA encoding TatA variants G2C, ins2C, T22C, G33C, S35C, F39C, K41C, D45C, E47C, T60C, T78C, V89C along with TatB and cysteine-less variant of TatC, under the control of the *tatA* promoter were placed into the lambda phage attachment site, *attB*, on the chromosome of strain DADE (which is deleted for all natively-encoded *tat* genes). The activity of the Tat system producing these variants at chromosomal level was assessed by determining the activity of the Tat substrate, trimethylamine-*N*-oxide (TMAO) reductase, TorA, in the periplasmic fraction. As shown in Fig 4, the S35C, D45C, E47C, T78C and V89C variants had levels of periplasmic TorA activity that was indistinguishable from the strain carrying wild type *tatABC* at the *attB* site. The other strains had very low levels of periplasmic TorA activity, close to that seen for the negative control. These strains were also examined for Tat activity using two phenotypic growth tests, growth on media containing 2% SDS (which requires the export of the Tat-dependent cell wall amidases, AmiA and AmiC; 53,54) and growth on minimal medium containing TMAO as sole electron acceptor (which assesses the combined export of TorA and the membrane-bound dimethylsulfoxide reductase, which can also reduce TMAO). All strains were able to grow on 2% SDS with the exception of those producing TatA variants G33C and T60C. Likewise all strains could grow with TMAO as sole electron acceptor except for those producing the ins2C, T22C, G33C, F39C and K41C TatA variants. Taken together it would seem that at native expression level the G33C substitution of TatA completely inactivates Tat function, whereas the remaining substitutions allow detectable transport of at least one Tat substrate protein.

We next confirmed that all of the variant TatA proteins were stably produced (Fig S3) and then we repeated the MAL-PEG accessibility assays in EDTA permeabilised whole cells. It is clear from the data shown in Fig 5 that cys residues close to the N-terminus of TatA were accessible to reaction with MAL-PEG in the absence of SDS (Fig 5A), and thus reside at the periplasmic side of the membrane. By contrast cys substitutions in the hinge region, the APH or C-tail of TatA could only be PEGylated in the presence of SDS (Fig 5A-C), consistent with cytoplasmic localization of these regions. Thus it is clear that the topology of TatA produced at close to native levels, is similar to that of overproduced TatA.

To ascertain whether the topology of TatA was affected by the presence of other Tat components we constructed a series of analogous strains where mutant *tatA* genes alone were expressed from the *attB* site of strain DADE. These cysteine-subsituted TatA variants, which were again shown to be stably produced (Fig S3), gave MAL-PEG labelling results that were largely indistinguishable from the labelling results obtained in the presence of TatBC (compare Fig 5 and Fig 6). We conclude that TatA has a similar N-out, C-in topology in the presence or absence of other Tat components.

Overproduction of a Tat substrate does not alter the topology of the TatA APH. The results presented above support an N-out, C-in topology for TatA under all conditions tested, and provide no evidence for dual topology of the APH suggested by others (28,31). Whilst the cells clearly remained viable and presumably also transport active during the EDTA treatment and PEGylation procedure (Fig S1), we maximized the likelihood that the Tat machinery was operational during the labelling experiments by saturating the Tat pathway through the high-level overproduction of the plasmid-encoded Tat substrate SufI. As shown in Fig 7A, SufI was massively overproduced from this construct when compared to native level, and the Tat machinery appeared to be saturated since there was significant SufI present in the spheroplast fraction. It can also be seen that some SufI was detected in the periplasmic fractions of strains producing each of the G33C, S35C, F39C, K41C, D45C and E47C TatA proteins, confirming functionality of the variant Tat machineries, even for those TatA proteins that supported only low levels of Tat transport of TorA. Differential effects of *tat* point mutations have previously been noted, depending

upon the nature of reporter proteins used to assess functionality of the Tat system and on whether Tat components or substrates proteins are overproduced (2). Control experiments (Fig 7A) showed that there was little disruption to the spheroplasts during the procedure since no HA-tagged SoxY could be detected in the periplasmic fractions. It should be noted that the SufI protein present in the spheroplast fractions had approximately the same mobility as that in the periplasmic fraction. It has been shown previously that SufI overproduced from the same construct used here is subjected to proteolytic processing of the N-terminal 17 amino acids, presumably by a cytoplasmic protease (55). Alternatively these observations may suggest that there is incomplete spheroplast formation in these experiments.

When the one hour PEGylation reaction was repeated in the presence of overproduced SufI, again none of the cys residues tested reacted with the MAL-PEG unless SDS was also included in the samples (Fig 7B). Thus we conclude that it is extremely unlikely that the TatA APH becomes exposed at the periplasmic side of the membrane during substrate transport.

TatB has an N-out, C-in topology assessed by MAL-PEG accessibility. To confirm that TatA and TatB show similar overall topology we applied the same MAL-PEG labelling approach to cysteine-substituted TatB proteins. Cysteine substitutions at position 2 and 41 of TatB were already available (22) and we constructed a further three substitutions, two more in the extended APH of TatB at positions 45 and 50, and a substitution at the very C-terminus of the protein (P171C). Previous studies have shown that TatB is tolerant to mutagenesis (22,56) and indeed each of the cys-substituted TatB proteins supported a good level of Tat activity (Fig 8A).

In these experiments the TatB variants were produced from a plasmid along with TatA and a cysteine-less variant of TatC, in a strain that was devoid of chromosomally-encoded Tat components. MAL-PEG labelling in whole cells demonstrated that the only cysteine accessible to PEGylation in the absence of SDS was at position 2 (Fig 8B), confirming that the N-terminus of TatB resides at the periplasmic side of the membrane. Cys residues in the APH and C-terminus of TatB were not PEGylated unless SDS was also included in the reaction, indicating that these residues most likely reside at the cytoplasmic side of the membrane. We therefore conclude that TatB has the same topological organization as TatA.

DISCUSSION

The development of even the most basic models for Tat pathway function requires knowledge about the location and topology of the essential Tat components. Whilst the membrane localization of the *E. coli* TatABC components is not in doubt (e.g. 57), the orientation and number of membrane-spanning domains present in TatA has been a source of contention. Previously the only experimental evidence directly addressing the location of the N-terminus of TatA led to the conclusion that it resided at the cytoplasmic side of the membrane, contrary to the positive inside rule (28). Furthermore, two separate studies indicated that the APH of TatA showed dual topology, being exposed at both the cytoplasmic and periplasmic side of the membrane. This supported models where the topological re-orientation of the TatA APH accompanied substrate movement across the membrane (1,29,30).

In this study we re-examined the topology of *E. coli* TatA, as well as addressing for the first time the topological orientation of the related protein TatB. We developed a method which would allow direct detection of labelled cysteine residues introduced at strategic positions throughout each protein by a size shift, whilst at the same time ensuring that the cytoplasmic membrane remained intact. Our results clearly showed that the N-termini of TatA and TatB are exposed at the periplasmic side of the membrane whilst residues along the APH and within the C-termini of both proteins were not labelled unless the inner membrane was dispersed with detergent. Topological changes of the TatA APH have been linked to substrate transport (31), however even when we grossly over-produced the Tat substrate SufI we did not detect any trace of exposure of any cysteines in the APH at the periplasmic side of the membrane. Since the cells clearly remained viable during these experiments, which were carried out at room temperature, it is reasonable to suppose that the Tat system was still functional during this time. Coupled with the observations that at least some of the TatA variants we tested were not impaired for Tat function, we conclude that Tat transport is not accompanied by exposure of the APH to the periplasm.

We also addressed the accessibility of cys residues located around the hinge region of TatA, between the TMH and APH. The results indicate that most of the hinge region is buried within the membrane and not accessible to PEGylation. This is in agreement with a solid-state NMR

spectroscopy study of a truncated *Bacillus subtilis* $TatA_d$ construct comprising just the TMH and APH reconstituted into lipid bicelles (50). It was shown that the N-terminal part of the APH is pulled into the membrane by virtue of the fact that the TMH is unusually short. A solution NMR study of the same protein also showed that the hinge region was unexpectedly rigid rather than flexible with extensive contacts between residues F_{14}, L_{18}, P_{23} and L_{26} (58). The rigidity of this region results in quite a steep upward tilt of the APH (50) and would be expected to constrain movement of the APH relative to the TMH. It is not known whether the hinge region of *E. coli* TatA is similarly rigid, particularly since most of the residues in $TatA_d$ involved in this interaction network are not conserved between the two proteins. However whilst it would appear that the TMH-APH hinge region plays an essential role in TatA function, it is unlikely to act as a pivot that allows these two domains of TatA to form a helical hairpin. Future models for Tat transport should consider alternative functions for the TatA APH.

REFERENCES

1. Cline, K., Theg, S.M. (2007) The Sec and Tat protein translocation pathways in chloroplasts. in *Molecular Machines Involved in Protein Transport across Cellular Membranes* (Dalbey, R. E., Koehler, C.M., Tamanoi, F. ed.), Elsevier, London. pp 463-492
2. Palmer, T., Sargent, F. and Berks, B.C. (2010) The Tat protein export pathway. in *EcoSal—Escherichia coli and Salmonella: Cellular and Molecular Biology. http://www.ecosal.org.* (Böck, A., Curtiss III, R., Kaper, J.B., Karp, P.D., Neidhardt, F.C., Nyström, T., Slauch, J.M., Squires, C.L. and Ussery, D. ed.), ASM Press, Washington, D.C. pp
3. Berks, B. C. (1996) A common export pathway for proteins binding complex redox cofactors? *Mol Microbiol* **22**, 393-404
4. Stanley, N. R., Palmer, T., and Berks, B. C. (2000) The twin arginine consensus motif of Tat signal peptides is involved in Sec-independent protein targeting in *Escherichia coli*. *J Biol Chem* **275**, 11591-11596
5. 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
6. 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
7. 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
8. Sargent, F., Stanley, N. R., Berks, B. C., and Palmer, T. (1999) Sec-independent protein translocation in *Escherichia coli*. A distinct and pivotal role for the TatB protein. *J Biol Chem* **274**, 36073-36082
9. Jack, R. L., Sargent, F., Berks, B. C., Sawers, G., and Palmer, T. (2001) Constitutive expression of *Escherichia coli tat* genes indicates an important role for the twin-arginine translocase during aerobic and anaerobic growth. *J Bacteriol* **183**, 1801-1804
10. Yen, M. R., Tseng, Y. H., Nguyen, E. H., Wu, L. F., and Saier, M. H., Jr. (2002) Sequence and phylogenetic analyses of the twin-arginine targeting (Tat) protein export system. *Arch Microbiol* **177**, 441-450
11. Bolhuis, A., Mathers, J. E., Thomas, J. D., Barrett, C. M., and Robinson, C. (2001) TatB and TatC form a functional and structural unit of the twin-arginine translocase from *Escherichia coli*. *J Biol Chem* **276**, 20213-20219
12. Richter, S., and Bruser, T. (2005) Targeting of unfolded PhoA to the TAT translocon of *Escherichia coli*. *J Biol Chem* **280**, 42723-42730
13. Tarry, M. J., Schafer, E., Chen, S., Buchanan, G., Greene, N. P., Lea, S. M., Palmer, T., Saibil, H. R., and Berks, B. C. (2009) Structural analysis of substrate binding by the TatBC component of the twin-arginine protein transport system. *Proc Natl Acad Sci U S A* **106**, 13284-13289
14. Cline, K., and Mori, H. (2001) Thylakoid DeltapH-dependent precursor proteins bind to a cpTatC-Hcf106 complex before Tha4-dependent transport. *J Cell Biol* **154**, 719-729

15. de Leeuw, E., Granjon, T., Porcelli, I., Alami, M., Carr, S. B., Muller, M., Sargent, F., Palmer, T., and Berks, B. C. (2002) Oligomeric properties and signal peptide binding by *Escherichia coli* Tat protein transport complexes. *J Mol Biol* **322**, 1135-1146
16. Alami, M., Luke, I., Deitermann, S., Eisner, G., Koch, H. G., Brunner, J., and Muller, M. (2003) Differential interactions between a twin-arginine signal peptide and its translocase in *Escherichia coli*. *Mol cell* **12**, 937-946
17. Gohlke, U., Pullan, L., McDevitt, C. A., Porcelli, I., de Leeuw, E., Palmer, T., Saibil, H. R., and Berks, B. C. (2005) The TatA component of the twin-arginine protein transport system forms channel complexes of variable diameter. *Proc Natl Acad Sci U S A* **102**,10482-10486
18. Porcelli, I., de Leeuw, E., Wallis, R., van den Brink-van der Laan, E., de Kruijff, B., Wallace, B. A., Palmer, T., and Berks, B. C. (2002) Characterization and membrane assembly of the TatA component of the *Escherichia coli* twin-arginine protein transport system. *Biochemistry* **41**, 13690-13697
19. Dabney-Smith, C., and Cline, K. (2009) Clustering of C-Terminal Stromal Domains of Tha4 Homo-Oligomers during Translocation by the Tat Protein Transport System. *Mol Biol Cell* **20,** 2060-2069
20. Leake, M. C., Greene, N. P., Godun, R. M., Granjon, T., Buchanan, G., Chen, S., Berry, R. M., Palmer, T., and Berks, B. C. (2008) Variable stoichiometry of the TatA component of the twin-arginine protein transport system observed by *in vivo* single-molecule imaging. *Proc Natl Acad Sci U S A* **105**, 15376-15381
21. Mori, H., and Cline, K. (2002) A twin arginine signal peptide and the pH gradient trigger reversible assembly of the thylakoid [Delta]pH/Tat translocase. *J Cell Biol* **157**, 205-210
22. Lee, P. A., Orriss, G. L., Buchanan, G., Greene, N. P., Bond, P. J., Punginelli, C., Jack, R. L., Sansom, M. S., Berks, B. C., and Palmer, T. (2006) Cysteine-scanning mutagenesis and disulfide mapping studies of the conserved domain of the twin-arginine translocase TatB component. *J Biol Chem* **281**, 34072-34085
23. Heijne, G. (1986) The distribution of positively charged residues in bacterial inner membrane proteins correlates with the trans-membrane topology. *EMBO J* **5**, 3021-3027
24. Mori, H., Summer, E. J., Ma, X., and Cline, K. (1999) Component specificity for the thylakoidal Sec and Delta pH-dependent protein transport pathways. *J Cell Biol* **146**, 45-56
25. Stevenson, L. G., Strisovsky, K., Clemmer, K. M., Bhatt, S., Freeman, M., and Rather, P. N. (2007) Rhomboid protease AarA mediates quorum-sensing in *Providencia stuartii* by activating TatA of the twin-arginine translocase. *Proc Natl Acad Sci U S A* **104**, 1003-1008
26. Maegawa, S., Koide, K., Ito, K., and Akiyama, Y. (2007) The intramembrane active site of GlpG, an *E. coli* rhomboid protease, is accessible to water and hydrolyses an extramembrane peptide bond of substrates. *Mol Microbiol* **64**, 435-447
27. Wang, Y., Zhang, Y., and Ha, Y. (2006) Crystal structure of a rhomboid family intramembrane protease. *Nature* **444**, 179-180
28. Chan, C. S., Zlomislic, M. R., Tieleman, D. P., and Turner, R. J. (2007) The TatA subunit of *Escherichia coli* twin-arginine translocase has an N-in topology. *Biochemistry* **46**, 7396-7404
29. Dabney-Smith, C., Mori, H., and Cline, K. (2006) Oligomers of Tha4 organize at the thylakoid Tat translocase during protein transport. *J Biol Chem* **281**, 5476-5483
30. Greene, N. P., Porcelli, I., Buchanan, G., Hicks, M. G., Schermann, S. M., Palmer, T., and Berks, B. C. (2007) Cysteine scanning mutagenesis and disulfide mapping studies of the TatA component of the bacterial twin arginine translocase. *J Biol Chem* **282**, 23937-23945
31. Gouffi, K., Gerard, F., Santini, C. L., and Wu, L. F. (2004) Dual topology of the *Escherichia coli* TatA protein. *J Biol Chem* **279**, 11608-11615
32. 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 U S A* **76**, 4530-4533
33. Wexler, M., Sargent, F., Jack, R. L., Stanley, N. R., Bogsch, E. G., Robinson, C., Berks, B. C., and Palmer, T. (2000) TatD is a cytoplasmic protein with DNase activity. No requirement for TatD family proteins in sec-independent protein export. *J Biol Chem* **275**, 16717-16722
34. Simons, R. W., Houman, F., and Kleckner, N. (1987) Improved single and multicopy *lac*-based cloning vectors for protein and operon fusions. *Gene* **53**, 85-96
35. Sauve, V., Bruno, S., Berks, B. C., and Hemmings, A. M. (2007) The SoxYZ complex carries sulfur cycle intermediates on a peptide swinging arm. *J Biol Chem* **282**, 23194-23204

36. Jack, R. L., Buchanan, G., Dubini, A., Hatzixanthis, K., Palmer, T., and Sargent, F. (2004) Coordinating assembly and export of complex bacterial proteins. *EMBO J* **23**, 3962-3972
37. Bartolome, B., Jubete, Y., Martinez, E., and de la Cruz, F. (1991) Construction and properties of a family of pACYC184-derived cloning vectors compatible with pBR322 and its derivatives. *Gene* **102**, 75-78
38. Bauer, J., Fritsch, M. J., Palmer, T., and Unden, G. (2011) Topology and accessibility of the transmembrane helices and the sensory site in the bifunctional transporter DcuB of *Escherichia coli*. *Biochemistry* **50**, 5925-5938
39. Hashimoto-Gotoh, T., Yamaguchi, M., Yasojima, K., Tsujimura, A., Wakabayashi, Y., and Watanabe, Y. (2000) A set of temperature sensitive-replication/-segregation and temperature resistant plasmid vectors with different copy numbers and in an isogenic background (chloramphenicol, kanamycin, *lacZ*, *repA*, *par*, *polA*). *Gene* **241**, 185-191
40. Sambrook, J., and Russell, D. W. (2001) *Molecular cloning : a laboratory manual*, 3rd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
41. Buchanan, G., de Leeuw, E., Stanley, N. R., Wexler, M., Berks, B. C., Sargent, F., and Palmer, T. (2002) Functional complexity of the twin-arginine translocase TatC component revealed by site-directed mutagenesis. *Mol Microbiol* **43**, 1457-1470
42. Palmer, T., Berks, B. C., and Sargent, F. (2010) Analysis of Tat targeting function and twin-arginine signal peptide activity in *Escherichia coli*. *Methods Mol Biol* **619**, 191-216
43. 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
44. Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685
45. 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 U S A* **76**, 4350-4354
46. Sargent, F., Gohlke, U., De Leeuw, E., Stanley, N. R., Palmer, T., Saibil, H. R., and Berks, B. C. (2001) Purified components of the *Escherichia coli* Tat protein transport system form a double-layered ring structure. *Eur J Biochem* **268**, 3361-3367
47. Wessel, D., and Flugge, U. I. (1984) A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Anal Biochem* **138**, 141-143
48. Punginelli, C., Maldonado, B., Grahl, S., Jack, R., Alami, M., Schroder, J., Berks, B. C., and Palmer, T. (2007) Cysteine scanning mutagenesis and topological mapping of the *Escherichia coli* twin-arginine translocase TatC Component. *J Bacteriol* **189**, 5482-5494
49. Lu, J., and Deutsch, C. (2001) Pegylation: a method for assessing topological accessibilities in Kv1.3. *Biochemistry* **40**, 13288-13301
50. Walther, T. H., Grage, S. L., Roth, N., and Ulrich, A. S. (2010) Membrane alignment of the pore-forming component TatA(d) of the twin-arginine translocase from *Bacillus subtilis* resolved by solid-state NMR spectroscopy. *J Amer Chem Soc* **132**, 15945-15956
51. Leive, L. (1968) Studies on the permeability change produced in coliform bacteria by ethylenediaminetetraacetate. *J Biol Chem* **243**, 2373-2380
52. Berthelmann, F., Mehner, D., Richter, S., Lindenstrauss, U., Lunsdorf, H., Hause, G., and Bruser, T. (2008) Recombinant expression of *tatABC* and *tatAC* results in the formation of interacting cytoplasmic TatA tubes in *Escherichia coli*. *J Biol Chem* **283**, 25281-25289
53. Ize, B., Stanley, N. R., Buchanan, G., and Palmer, T. (2003) Role of the *Escherichia coli* Tat pathway in outer membrane integrity. *Mol Microbiol* **48**, 1183-1193
54. Bernhardt, T. G., and de Boer, P. A. (2003) The *Escherichia coli* amidase AmiC is a periplasmic septal ring component exported via the twin-arginine transport pathway. *Mol Microbiol* **48**, 1171-1182
55. Tarry, M., Arends, S. J., Roversi, P., Piette, E., Sargent, F., Berks, B. C., Weiss, D. S., and Lea, S. M. (2009) The *Escherichia coli* cell division protein and model Tat substrate SufI (FtsP) localizes to the septal ring and has a multicopper oxidase-like structure. *J Mol Biol* **386**, 504-519
56. Maldonado, B., Kneuper, H., Buchanan, G., Hatzixanthis, K., Sargent, F., Berks, B. C., and Palmer, T. (2011) Characterisation of the membrane-extrinsic domain of the TatB component of the twin arginine protein translocase. *FEBS letts* **585**, 478-484

57. De Leeuw, E., Porcelli, I., Sargent, F., Palmer, T., and Berks, B. C. (2001) Membrane interactions and self-association of the TatA and TatB components of the twin-arginine translocation pathway. *FEBS letts* **506**, 143-148
58. Hu, Y., Zhao, E., Li, H., Xia, B., and Jin, C. (2010) Solution NMR structure of the TatA component of the twin-arginine protein transport system from gram-positive bacterium *Bacillus subtilis*. *J Amer Chem Soc* **132**, 15942-15944
59. Bryson, K., McGuffin, L. J., Marsden, R. L., Ward, J. J., Sodhi, J. S., and Jones, D. T. (2005) Protein structure prediction servers at University College London. *Nucleic acids res* **33**, W36-38

Acknowledgements - We thank Ben Berks for advice and helpful discussion and for providing us with construct pVS005.

FOOTNOTES

* This work was supported by the Biotechnology and Biological Sciences Research Council through a PhD studentship (to M.F.).

[1]To whom correspondence should be addressed: Tracy Palmer, Division of Molecular Microbiology, College of Life Sciences, University of Dundee, Dundee DD1 5EH, UK telephone +44 (0)1382 386464, fax +44 (0)1382 388216, e-mail t.palmer@dundee.ac.uk

[2]The abbreviations used are: AMS, 4-acetamido-4'-maleimidylstilbene-2,2'-disulfonic acid; APH, amphipathic helix; Δ*p*, transmembrane electrochemical proton potential; DTT dithiothreitol; MAL-PEG methoxypolyethylene glycol maleimide; MPB, N^{α}-(3-maleimidylpropionyl)-biocytin; Tat, twin-arginine translocation; TEV, tobacco etch virus; TMAO, trimethylamine *N*-oxide; TMH, transmembrane helix.

FIGURE LEGENDS

FIGURE 1. Secondary structure predictions of the *E. coli* TatA and TatB proteins and possible topological arrangements of TatA. Primary amino acid sequence of A. *E. coli* TatA; B. *E. coli* TatB. The position of the invariant glycine found throughout TatA and TatB family proteins is boxed. Predicted secondary structure elements are shown above the amino acid sequences, with the transmembrane helix in dark grey and the amphipathic helix in light grey. Secondary structure predicted using PSIPRED 3.0 (59) and predicted helical regions are shown as cylinders; Note only the first 100 amino acids of TatB are shown. C. Possible topological organizations for *E. coli* TatA based on previous experimental observations.

FIGURE 2. MAL-PEG labelling of cysteine-substituted TatA in crude membrane fractions of *E. coli*. Cells of strain DADE-P (Δ*tatABCD*, Δ*tatE, pcnB*) producing plasmid encoded TatB and cysteine-less TatC along with either wild type TatA (WT) or the indicated TatA cys-substituted variants were disrupted by sonication. The crude membrane fraction was isolated by ultracentrifugation and divided into four aliquots (each of 100 – 150 μg protein). One aliquot was incubated with buffer alone, whilst the remaining three aliquots were incubated with MAL-PEG, MAL-PEG plus 1 % Triton-X 100 or MAL-PEG plus 1 % SDS. The reaction was quenched by addition of with 45 mM DTT and a sample of each aliquot (4.5 – 7 μg protein) was separated by SDS-PAGE on 15% Tris-glycine gels. Proteins were transferred to nitrocellulose membrane and TatA was detected with anti-TatA antiserum. The positions of the molecular weight markers are indicated to the left of the panels and the positions of unlabelled and PEGylated TatA to the right of the panels. A possible TatA-lipid cross-link seen for some of the TatA variants and reported previously (30) is indicated with an asterisk. Labelling of TatA cys substitutions shown for A. the N-terminus and within the TMH; B. around the hinge region and C. in the APH.

FIGURE 3. MAL-PEG labelling of plasmid produced, cysteine-substituted TatA in intact cells shows an N-out, C-in topology. Cells of strain DADE-P (Δ*tatABCD*, Δ*tatE, pcnB*) producing plasmid encoded TatB and cysteine-less TatC along with either wild type TatA (WT) or the indicated TatA cys-substituted variants in 50 ml culture were grown to mid-exponential phase (OD600 of approx. 0.4), harvested and resuspended in 1 ml buffer. 80 μl aliquots of cell suspension were incubated with buffer alone, or 5 mM MAL-PEG in the presence or absence of 1% SDS for 1 hour at room temperature. Reactions were quenched with 45 mM DTT and proteins were precipitated with chloroform and methanol. Resolubilised samples were separated by SDS-PAGE, electroblotted and immunoreactive bands were revealed by incubation with either anti-TatA antiserum

or an anti-hemagglutinin HRP conjugate (to detect SoxY). The positions of the molecular weight markers are indicated to the left and the positions of PEGylated and non-PEGylated protein to the right. Labelling of TatA cys substitutions shown for A. either side of the TMH; B. the APH and C. the C-terminal tail.

FIGURE 4. Periplasmic TMAO reductase activities of *E. coli* strains producing chromosomally-encoded cysteine-substituted TatA variants. Relative TMAO reductase activities were determined in the periplasmic fractions of strain DADE (Δ*tatABCD*, Δ*tatE*; annotated as Δ in the figure) or DADE producing TatB and cysteine-less TatC along with either wild type TatA (strain MF1, labelled WT) or the indicated TatA variants. The TMAO reductase activity of strain MF1 (DADE *attB::*P_{tatA} (*tatABC*) is classed as 100% activity and represents a specific activity of approximately 2.6 μmol benzyl viologen oxidised/min/mg protein. The error bars represent the standard error of the mean, *n*=3.

FIGURE 5. MAL-PEG labelling of chromosomally-encoded cysteine-substituted TatA in intact cells shows an N-out, C-in topology. Cells of strains MF2 – MF13 producing TatB, cysteine-less TatC and cysteine-substituted TatA at chromosomal levels were grown in 50 ml culture until mid-exponential phase was reached. Cells were harvested and resuspended in 1 ml HEPES/$MgCl_2$ buffer. 80 μl aliquots of cell suspension were incubated with buffer alone, or 5 mM MAL-PEG in the presence or absence of 1% SDS for 1 hour at room temperature. Reactions were quenched with 45 mM DTT, proteins were precipitated with chloroform and methanol and separated by SDS PAGE. After electroblotting, TatA was detected with anti-TatA antiserum and SoxY using an anti-hemagglutinin HRP conjugate. The positions of the molecular weight markers are indicated to the left and the positions of PEGylated and non-PEGylated protein to the right. Labelling of TatA cys substitutions shown for A. either side of the TMH; B. the APH and C. the C-terminal tail.

FIGURE 6. TatA has an N-out, C-in topology in the absence of other Tat components. Cells of strains SK2 – SK13 producing cysteine-substituted TatA at chromosomal levels in an otherwise *tat*$^-$ background were grown in 50 ml culture until mid-exponential phase was reached. Cells were harvested and resuspended in 1 ml HEPES/$MgCl_2$ buffer. 80 μl aliquots of cell suspension were incubated with buffer alone, or 5 mM MAL-PEG in the presence or absence of 1% SDS for 1 hour at room temperature. Reactions were quenched with 45 mM DTT, proteins were precipitated with chloroform and methanol and separated by SDS PAGE. After electroblotting, TatA was detected with anti-TatA antiserum and SoxY using an anti-hemagglutinin HRP conjugate. The positions of the molecular weight markers are indicated to the left and the positions of PEGylated and non-PEGylated protein to the right. Labelling of TatA cys substitutions shown for A. either side of the TMH; B. the APH and C. the C-terminal tail.

FIGURE 7. Topology of the TatA APH does not change when the Tat substrate SufI is overproduced. A. Overproduction and Tat-dependent transport of SufI mediated by variant TatA proteins. Cultures of strains MF1 (WT), MF5 (G33C), MF6 (S35C), MF7 (F39C), MF9 (D45C) and MF10 (E47C) harboring pQE-SufI and pTH19SoxYZ were cultured aerobically in 50 ml LB medium until an OD_{600} of approximately 0.6 was reached. Cells were harvested, washed with 50 mM HEPES, pH 7.0, 250 mM NaCl, resuspended in 1 ml HEPES/sucrose buffer (50 mM HEPES, pH 7.0, 0.5 M sucrose) and fractionated to give a final volume of 1ml periplasm (P) and 5 ml spheroplasts (S). 5 μl samples of each fraction were separated by SDS-PAGE, electroblotted and immunoreactive protein bands of SufI or SoxY were detected as indicated. Note that expression of *sufI* from pQE60 is constitutive in these strain backgrounds because the chromosomal copy of *lacI* is deleted. B. Cells of strains MF1, MF5, MF6, MF7, MF9 and MF10 harboring pQE-SufI and pTH19SoxYZ were grown in 50 ml culture until mid-exponential phase was reached. Cells were harvested and resuspended in 1 ml HEPES/$MgCl_2$ buffer. 80 μl aliquots of cell suspension were incubated with buffer alone, or 5 mM MAL-PEG in the presence or absence of 1% SDS for 1 hour at room temperature. Reactions were quenched with 45 mM DTT, Proteins were precipitated with chloroform and methanol and separated by SDS PAGE. After electroblotting, TatA was detected with anti-TatA antiserum and SoxY using an anti-hemagglutinin HRP conjugate.

FIGURE 8. TatB has an N-out, C-in topology. A. Relative TMAO reductase activities were determined in the periplasmic fractions of strain DADE-P (Δ*tatABCD*, Δ*tatE*, *pcnB*) harboring either pQE60, or pQE60 encoding TatA and cysteine-less TatC along with TatB variants containing a single cysteine substitution as

indicated. The TMAO reductase activity of strain DADE-P transformed with pUNITATCC4, encoding wild type TatA, TatB and cysteine-less TatC was defined as one hundred percent activity and corresponds to 1.9 μM benzyl viologen oxidised per minute and per mg of protein. The error bars represent standard error of the mean, $n = 3$. B. Cells of strain DADE-P producing plasmid encoded TatA and cysteine-less TatC along with either wild type TatB (WT) or the indicated TatB cys-substituted variants in 50 ml culture were grown to mid-exponential phase (OD600 of approx. 0.4), harvested and resuspended in 1 ml buffer. 80 μl aliquots of cell suspension were incubated with buffer alone, or 5 mM MAL-PEG in the presence or absence of 1% SDS for 1 hour at room temperature. Reactions were quenched with 45 mM DTT and proteins were precipitated with chloroform and methanol. Resolubilised samples were separated by SDS-PAGE, electroblotted and immunoreactive bands were revealed by incubation with either anti-TatB antiserum or an anti-hemagglutinin HRP conjugate (to detect SoxY). The positions of the molecular weight markers are indicated to the left and the positions of PEGylated and non-PEGylated protein to the right.

	Description	Reference
MC4100	F^-, $[araD139]_{B/r}$, Δ(*argF-lac*)*U169*, λ^-, *e14*$^-$-, *flhD5301*, Δ(*fruK-yeiR*)*725*(*fruA25*), *relA*1, *rpsL150*(Str^R), *rbsR22*, Δ(*fimB-fimE*)*632*(::IS1), *deoC1*	(32)
JARV16-P	MC4100 Δ*tatA* Δ*tatE pcnB1 zad-981*::*Tn*10d	(8)
DADE	MC4100 Δ*tatABCD* Δ*tatE*	(33)
DADE-P	MC4100 Δ*tatABCD* Δ*tatE pcnB1 zad-981*::*Tn*10d	(22)
MF1	DADE *attB::*P_{tatA} (*tatABC*)	This study
MF2	DADE *attB::*P_{tatA} ($tatA_{ins2C}BC$)	This study
MF3	DADE *attB::*P_{tatA} ($tatA_{G2C}BC$)	This study
MF4	DADE *attB::*P_{tatA} ($tatA_{T22C}BC$)	This study
MF5	DADE *attB::*P_{tatA} ($tatA_{G33C}BC$)	This study
MF6	DADE *attB::*P_{tatA} ($tatA_{S35C}BC$)	This study
MF7	DADE *attB::*P_{tatA} ($tatA_{F39C}BC$)	This study
MF8	DADE *attB::*P_{tatA} ($tatA_{K41C}BC$)	This study
MF9	DADE *attB::*P_{tatA} ($tatA_{D45C}BC$)	This study
MF10	DADE *attB::*P_{tatA} ($tatA_{E47C}BC$)	This study
MF11	DADE *attB::*P_{tatA} ($tatA_{T60C}BC$)	This study
MF12	DADE *attB::*P_{tatA} ($tatA_{T78C}BC$)	This study
MF13	DADE *attB::*P_{tatA} ($tatA_{V89C}BC$)	This study
SK1	DADE *attB::*P_{tatA} (*tatA*)	This study
SK2	DADE *attB::*P_{tatA} ($tatA_{ins2C}$)	This study
SK3	DADE *attB::*P_{tatA} ($tatA_{G2C}$)	This study
SK4	DADE *attB::*P_{tatA} ($tatA_{T22C}$)	This study
SK5	DADE *attB::*P_{tatA} ($tatA_{G33C}$)	This study
SK6	DADE *attB::*P_{tatA} ($tatA_{S35C}$)	This study
SK7	DADE *attB::*P_{tatA} ($tatA_{F39C}$)	This study
SK8	DADE *attB::*P_{tatA} ($tatA_{K41C}$)	This study
SK9	DADE *attB::*P_{tatA} ($tatA_{D45C}$)	This study
SK10	DADE *attB::*P_{tatA} ($tatA_{E47C}$)	This study
SK11	DADE *attB::*P_{tatA} ($tatA_{T60C}$)	This study
SK12	DADE *attB::*P_{tatA} ($tatA_{T78C}$)	This study
SK13	DADE *attB::*P_{tatA} ($tatA_{V89C}$)	This study

Table 1. Strains used in this study

plasmid	Description	Reference
pUNITATA	Wild type *tatA* in pBluescript KS^+	(30)
pUNITATB	Wild type *tatB* in pBluescript KS^+	(22)
pUNITATAX#C	As pUNITATA, *tatA* harboring single cysteine codon substitution as indicated*	This study and (30)
pUNITATAins2C	As pUNITATA, *tatA* harboring a cysteine codon insertion between codons 1 and 2.	This study
pUNITATAins21C	As pUNITATA, *tatA* harboring a cysteine codon insertion between codons 21 and 22.	This study
pUNITATBX#C	As pUNITATB, *tatB* harboring single cysteine codon substitution as indicated*	This study and (22)
pUNITATCC4	*tatABC* operon in pQE60, all 4 cysteine codons in *tatC* substituted for alanine codons	(22)
pUNITATCC4AX#C	As pUNITATCC4 harboring single cysteine codon substitutions in *tatA* as indicated*	This study and (30)
pUNITATCC4Ains2C	As pUNITATCC4 *tatA* harboring a cysteine codon insertion between codons 1 and 2	This study
pUNITATCC4Ains21C	As pUNITATCC4 *tatA* harboring a cysteine codon insertion between codons 21 and 22	This study
pUNITATCC4BX#C	As pUNITATCC4 harboring single cysteine codon substitutions in *tatB* as indicated*	This study and (22)
pKSuniA	WT *tatA* under control of the *tat* promoter in pBluescript KS^+	This study
pKSuniAx#C	As pKSuniA harboring single cysteine codon substitutions in *tatA* as indicated*	This study
pKSuniAins2C	As KSuniA, *tatA* harboring a cysteine codon insertion between codons 1 and 2.	This study
pUNICC–AX#C	*tatA* promoter and *tatABC* operon in pQE60, no cysteine codons in *tatC*, single cysteine substitutions in *tatA* as indicated*	This study
pUNICCins2C	*tatA* promoter and *tatABC* operon in pQE60, no cysteine codons in *tatC*, *tatA* harboring a cysteine codon insertion between codons 1 and 2.	This study
pRSUNICC–AX#C	*tatA* promoter and *tatABC* operon in pRS552, no cysteine codons in *tatC*, single cysteine substitutions in *tatA* as indicated*	This study
pRSUNICCins2C	*tatA* promoter and *tatABC* operon in pRS552, no cysteine codons in *tatC*, *tatA* harboring a cysteine codon insertion between codons 1 and 2.	This study
pRSUNIAX#C	*tatA* promoter and *tatA* (without *tatBC*) in pRS552, single cysteine substitutions in *tatA* as indicated*	This study
pRSUNICCins2C	*tatA* promoter and *tatA* (without *tatBC*) in pRS552, *tatA* harboring a cysteine codon insertion between codons 1 and 2.	This study
pVS005	*P. panthotrophus soxYZ* expression plasmid	(35)
pHASoxYZ	*E. coli tatA* promoter controlling expression of *P. panthotrophus HA-soxY* and *soxZ* in pSU20	(38)
pTH19SoxYZ	*E. coli tatA* promoter controlling expression of *P. panthotrophus HA-soxY* and *soxZ* in pTH19cr	This study
pQE60-SufI	Full length *sufI* gene cloned in vector pQE60	(55)

*In constructs with single cysteine substitutions of *tatA* and *tatB*, X represents the single letter amino acid code and # represents the position of the substituted codon.

Table 2. Plasmids used in this study.

Figure 1

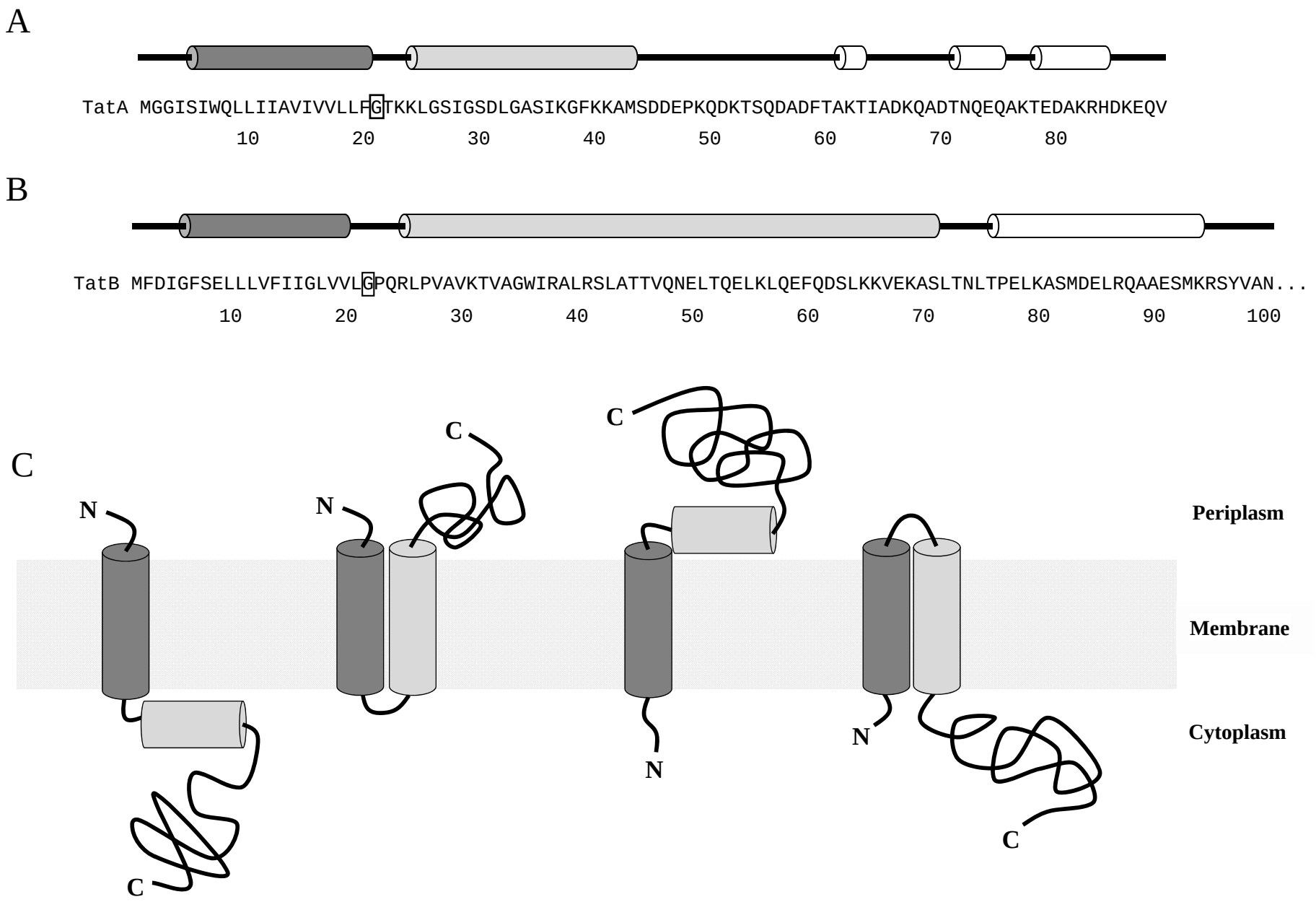

A
TatA MGGISIWQLLIIAVIVVLLFGTKKLGSIGSDLGASIKGFKKAMSDDEPKQDKTSQDADFTAKTIADKQADTNQEQAKTEDAKRHDKEQV
10 20 30 40 50 60 70 80
B
TatB MFDIGFSELLLVFIIGLVVLGPQRLPVAVKTVAGWIRALRSLATTVQNELTQELKLQEFQDSLKKVEKASLTNLTPELKASMDELRQAAESMKRSYVAN...
10 20 30 40 50 60 70 80 90 100
C
N
C
N
C
N
C
N
C
Periplasm
Membrane
Cytoplasm

Figure 2

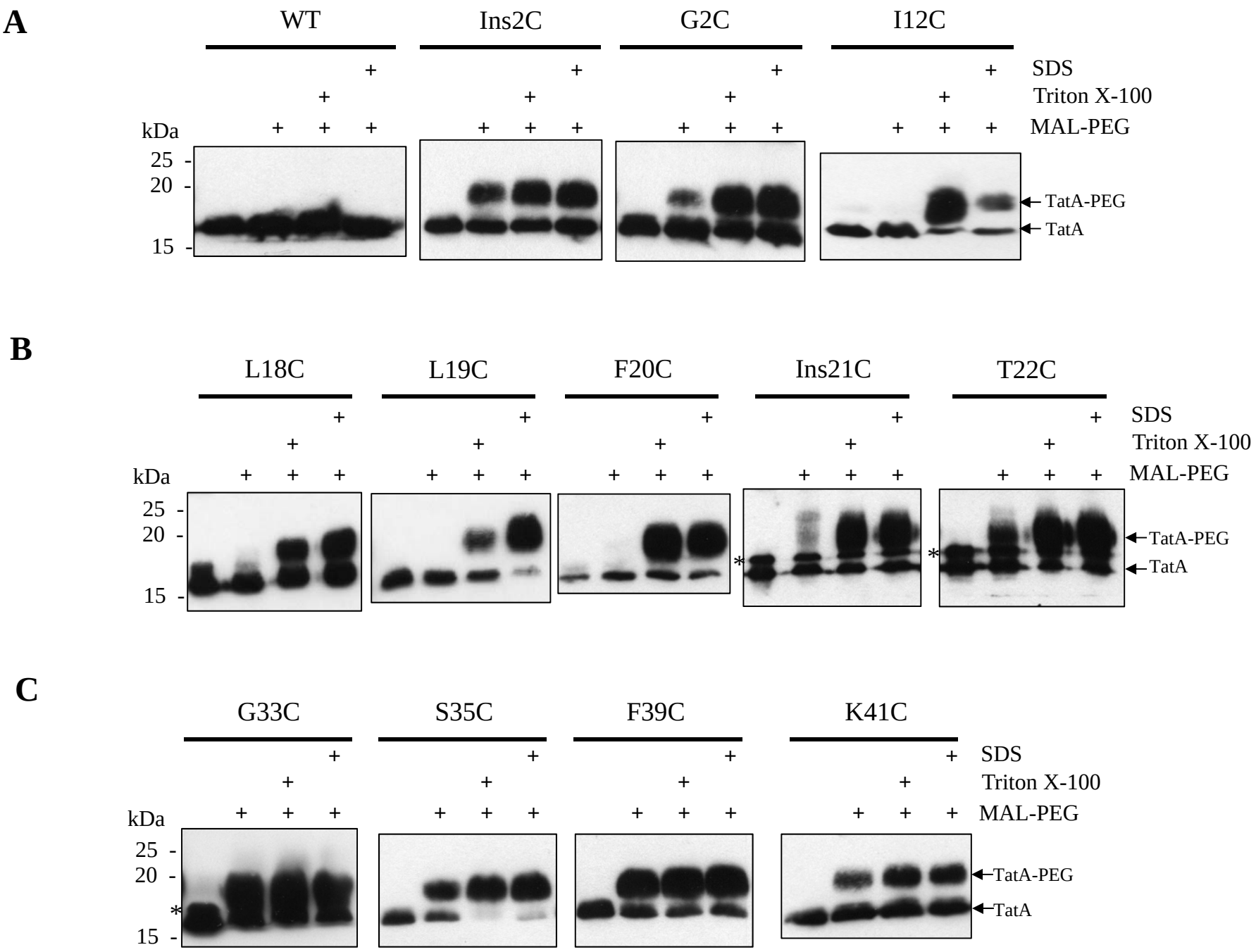

A
WT
Ins2C
G2C
I12C
SDS
Triton X-100
MAL-PEG
kDa
25
20
15
TatA-PEG
TatA
B
L18C
L19C
F20C
Ins21C
T22C
SDS
Triton X-100
MAL-PEG
kDa
25
20
15
*
*
TatA-PEG
TatA
C
G33C
S35C
F39C
K41C
SDS
Triton X-100
MAL-PEG
kDa
25
20
*
15
TatA-PEG
TatA

Figure 3

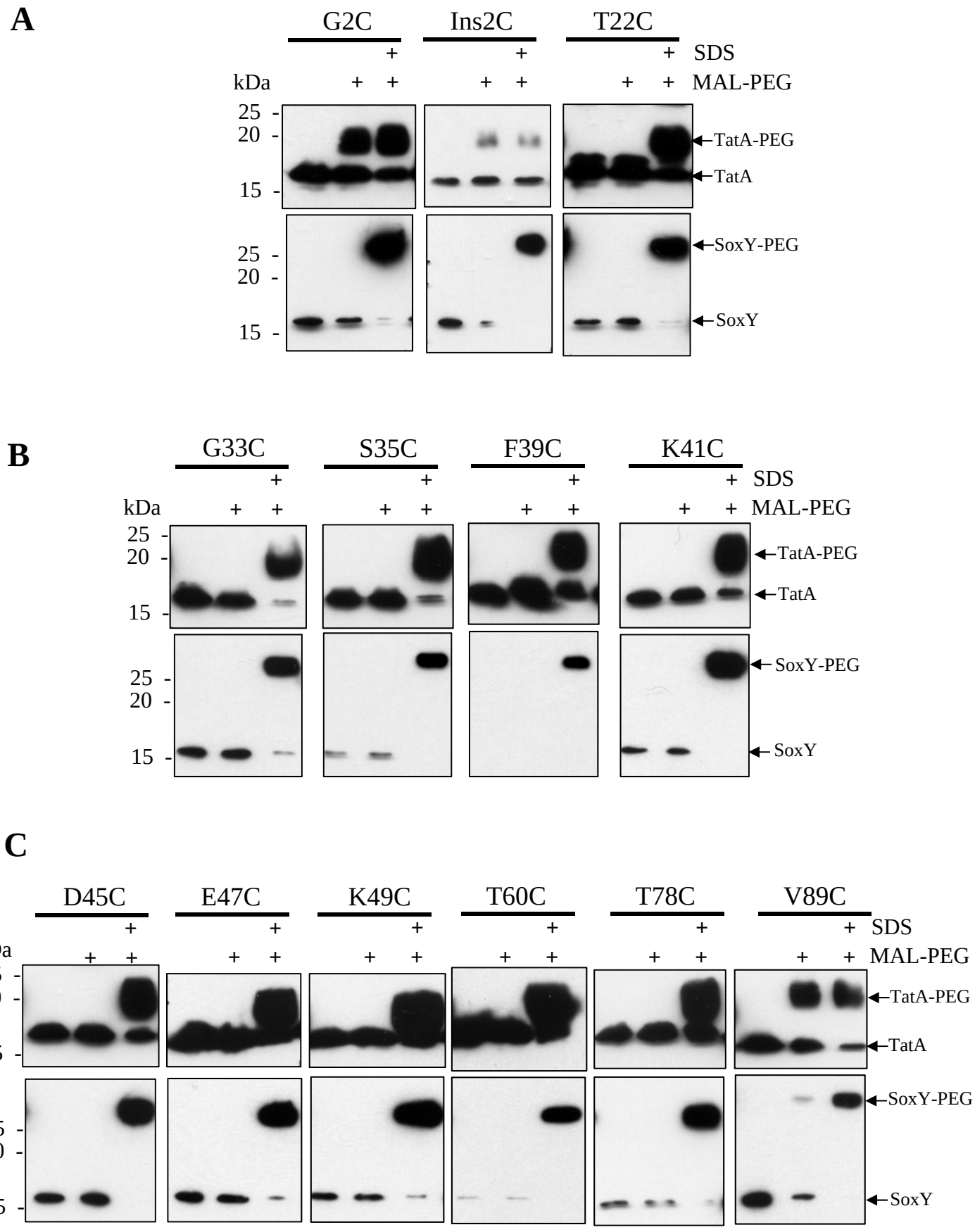

A
G2C
Ins2C
T22C
SDS
kDa
MAL-PEG
25
20
15
TatA-PEG
TatA
SoxY-PEG
SoxY
B
G33C
S35C
F39C
K41C
C
D45C
E47C
K49C
T60C
T78C
V89C

Figure 4

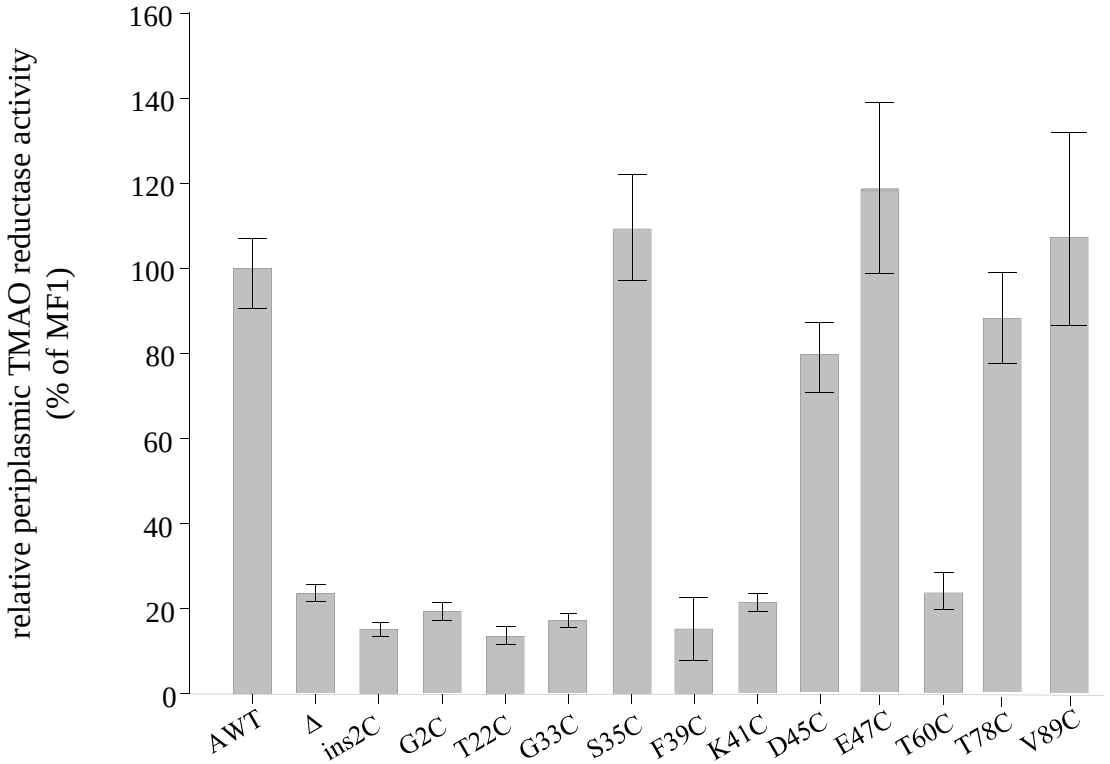

relative periplasmic TMAO reductase activity
(% of MF1)
160
140
120
100
80
60
40
20
0
AWT
Δ
ins2C
G2C
T22C
G33C
S35C
F39C
K41C
D45C
E47C
T60C
T78C
V89C

Figure 5

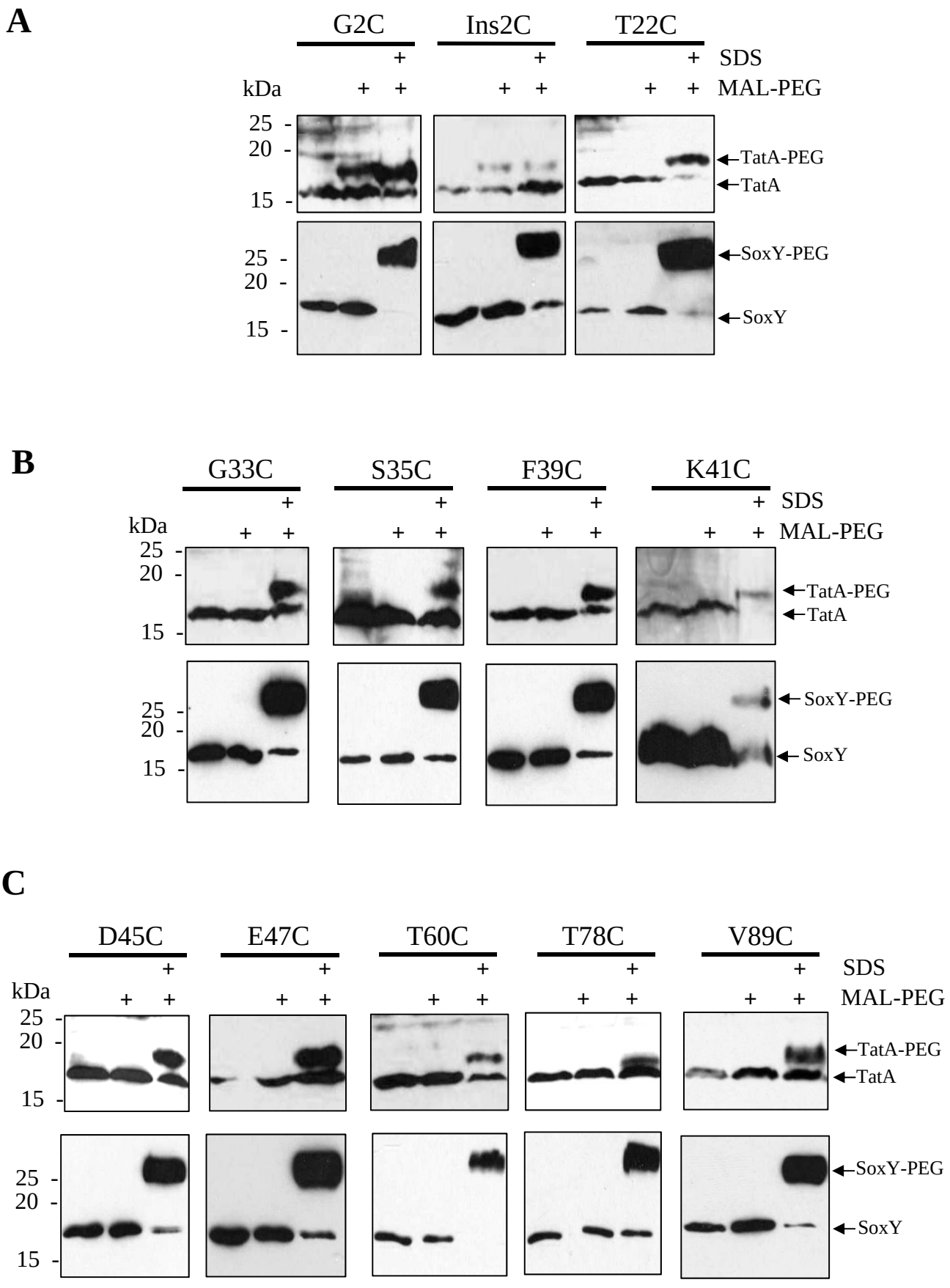

A
G2C
Ins2C
T22C
SDS
kDa
MAL-PEG
25
20
15
TatA-PEG
TatA
SoxY-PEG
SoxY
B
G33C
S35C
F39C
K41C
C
D45C
E47C
T60C
T78C
V89C

Figure 6

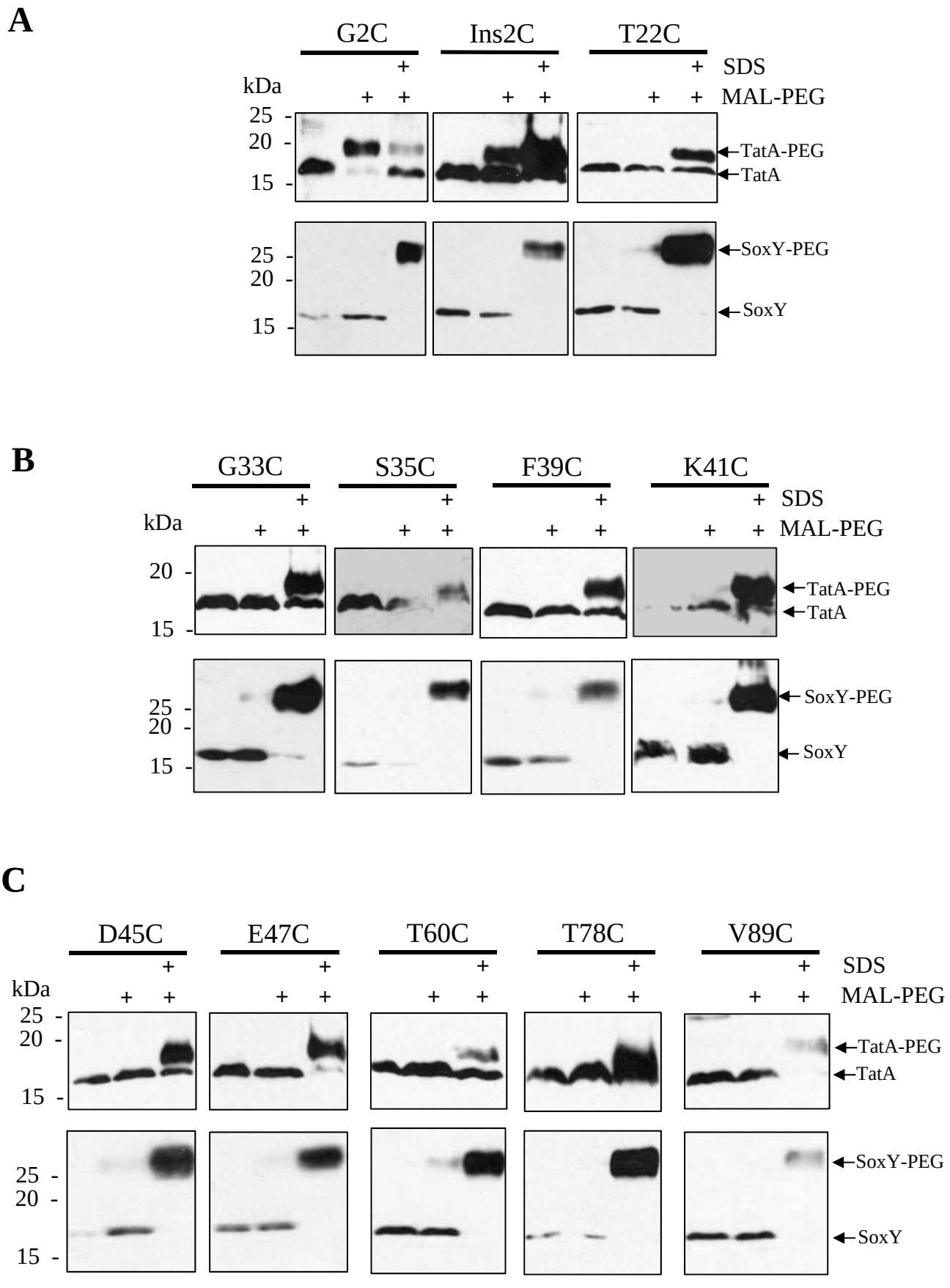

A
G2C
Ins2C
T22C
SDS
kDa
MAL-PEG
25
20
15
TatA-PEG
TatA
SoxY-PEG
SoxY
B
G33C
S35C
F39C
K41C
C
D45C
E47C
T60C
T78C
V89C

Figure 7

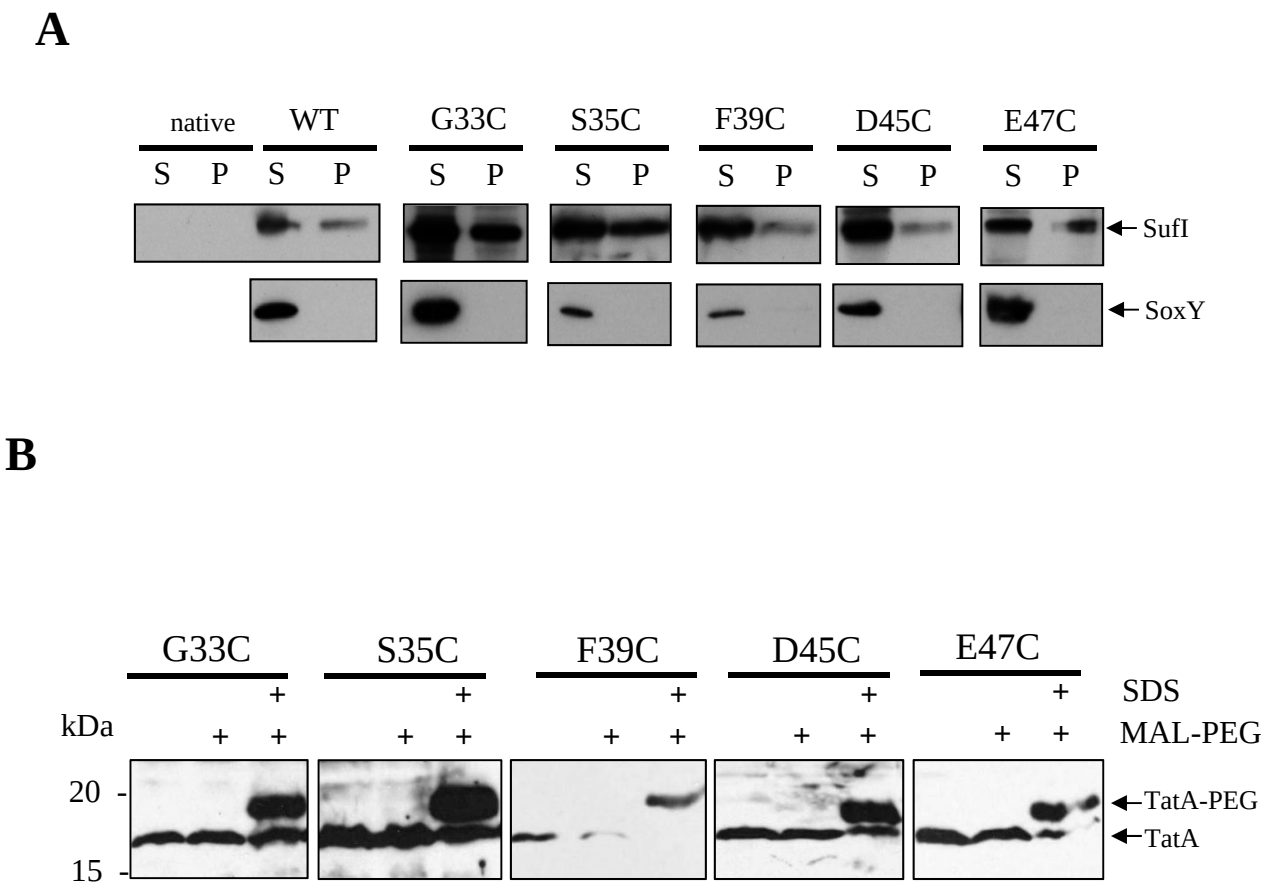
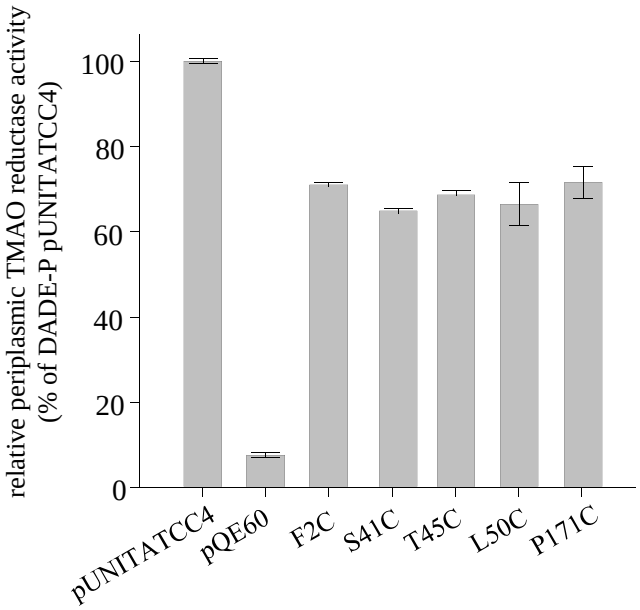

A
native
WT
G33C
S35C
F39C
D45C
E47C
S P S P S P S P S P S P S P
SufI
SoxY
B
G33C
S35C
F39C
D45C
E47C
SDS
MAL-PEG
kDa
20
15
TatA-PEG
TatA

Figure 8

A


relative periplasmic TMAO reductase activity
(% of DADE-P pUNITATCC4)
100
80
60
40
20
0
pUNITATCC4
pQE60
F2C
S41C
T45C
L50C
P171C

B

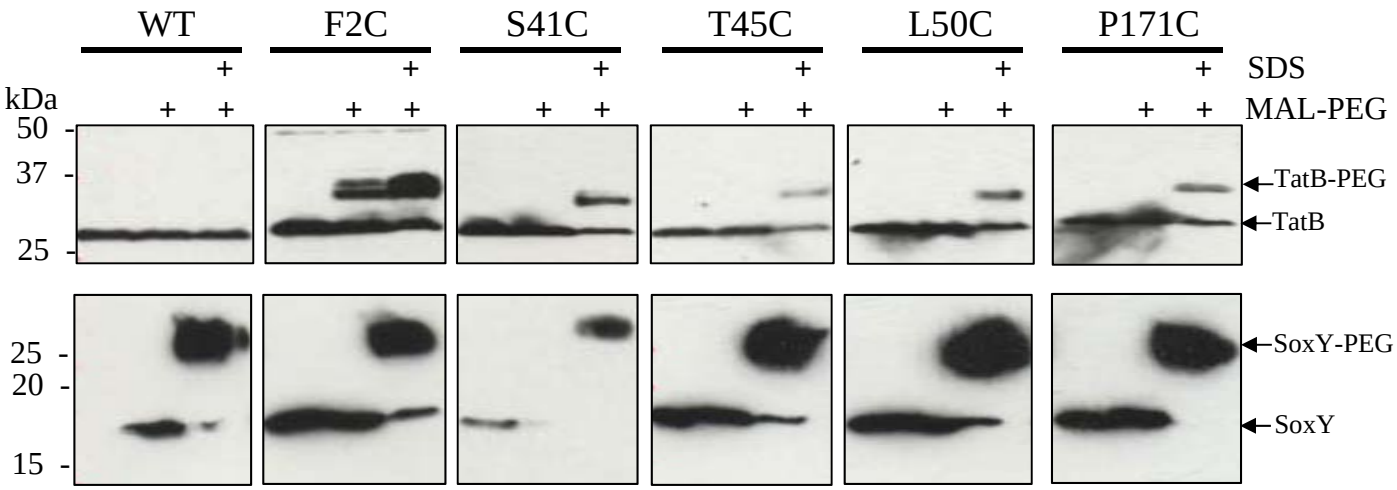

WT
F2C
S41C
T45C
L50C
P171C
SDS
MAL-PEG
kDa
50
37
25
TatB-PEG
TatB
25
20
15
SoxY-PEG
SoxY

FIGURE S1. *E. coli* remains viable after EDTA treatment and sulfhydryl labelling. Strains MF1 (WT) or MF2 (G2C) were grown in 50 ml LB medium harvested and resuspended in 1 ml buffer. 80 µl aliquots of cell suspension were incubated with buffer alone, or with 5 mM EDTA in the presence or absence of 5 mM MAL-PEG for 1 hour at room temperature. After this time, 3 µl aliquots of each sample were inoculated, in triplicate, into 300 µl LB medium and cultured with aeration in a 96 well plate reader and the optical density at 600 nm was monitored for ten hours. Data points and error bars correspond to the mean ± SD (n=3). Samples are buffer alone (WT open diamond; G2C closed diamond), 5 mM EDTA treatment (WT open circle; G2C closed circle) or 5 mM EDTA PLUS 5 mM MAL-PEG treatment (WT open triangle; G2C closed triangle).

FIGURE S2. Periplasmic TMAO reductase activities of *E. coli* strains producing plasmid-encoded TatA variants. Relative TMAO reductase activities were determined from the periplasmic fractions of strains DADE-P (Δ*tatABCD*, Δ*tatE, pcnB1*) harboring either pQE60, or pQE60 encoding TatB and cysteine-less TatC along with TatA variants containing a single cysteine substitution as indicated. The TMAO reductase activity of strain DADE-P transformed with pUNITATCC4, encoding wild type TatA (along with TatB and cysteine-less TatC), was defined as one hundred percent activity and corresponds to 9.5 µM benzyl viologen oxidised per minute and per mg of protein. The error bars represent standard error of the mean, n = 3.

FIGURE S3. Levels of TatA variants produced from single copy *tatA* genes present at the chromosomal *attB* site. Crude membrane fractions were prepared from 50 ml cultures of *E. coli* strains expressing either A. the *tatABC* operon (strains MF1 - 13) or B. *tatA* only (strains SK1 - 13) from the chromosomal lambda phage attachment site. Crude membrane fractions (20 µg) were separated by SDS-PAGE (15% acrylamide), electroblotted, and immunoreactive bands of TatA were detected using anti-TatA antiserum. Expression levels of strains producing the indicated cys variant are compared to strains MC4100 (WT), DADE (Δ) and either MF1 (DADE *attB::*P_{tatA} (*tatABC*); panel A) or SK1 (DADE *attB::*P_{tatA} (*tatA*); panel B)

Figure S1

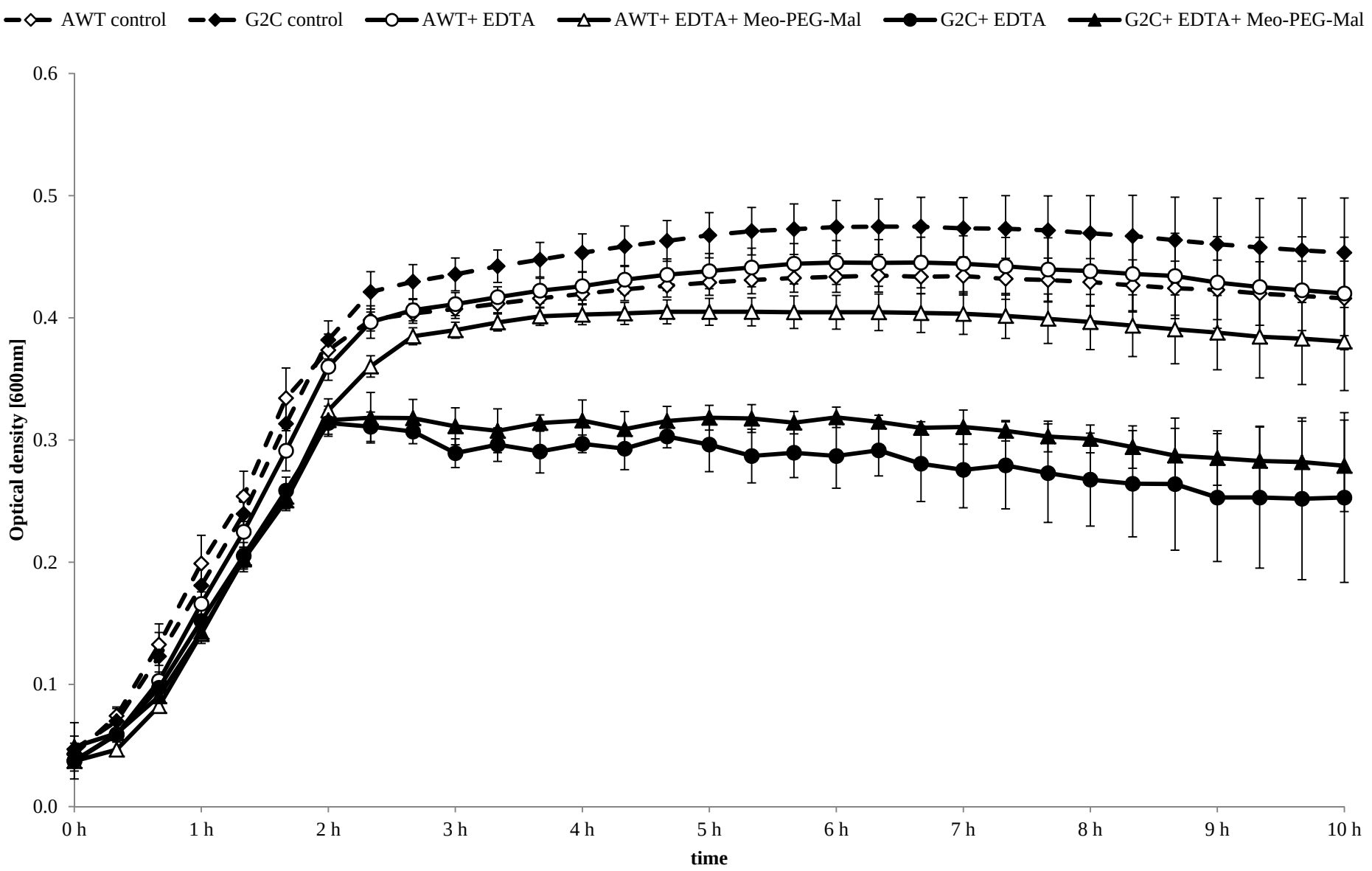

AWT control
G2C control
AWT+ EDTA
AWT+ EDTA+ Meo-PEG-Mal
G2C+ EDTA
G2C+ EDTA+ Meo-PEG-Mal
Optical density [600nm]
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0 h
1 h
2 h
3 h
4 h
5 h
6 h
7 h
8 h
9 h
10 h
time

Figure S2

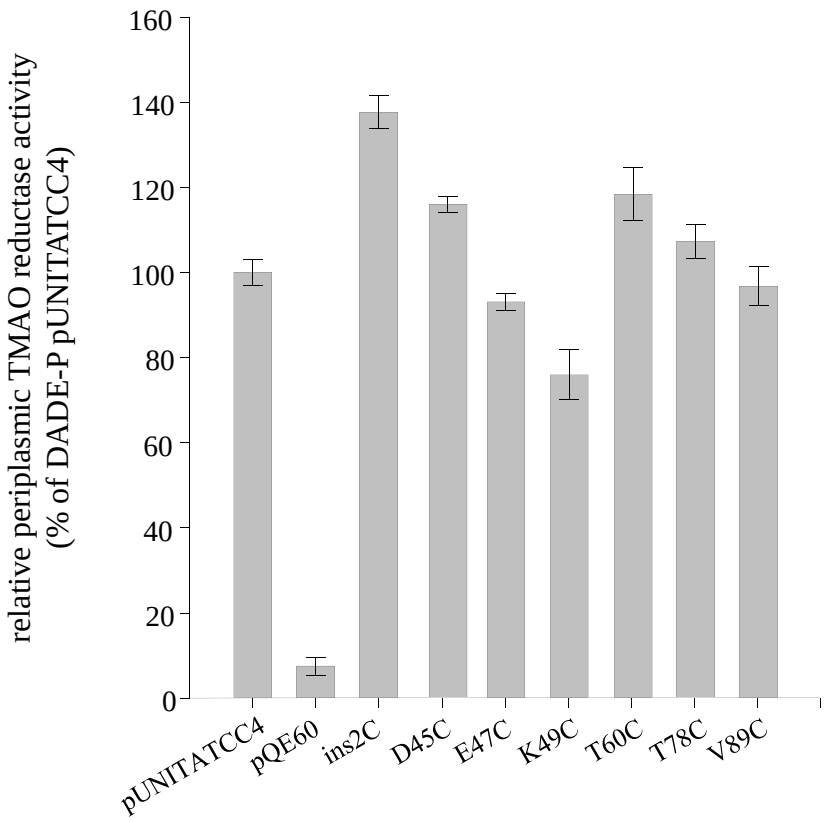
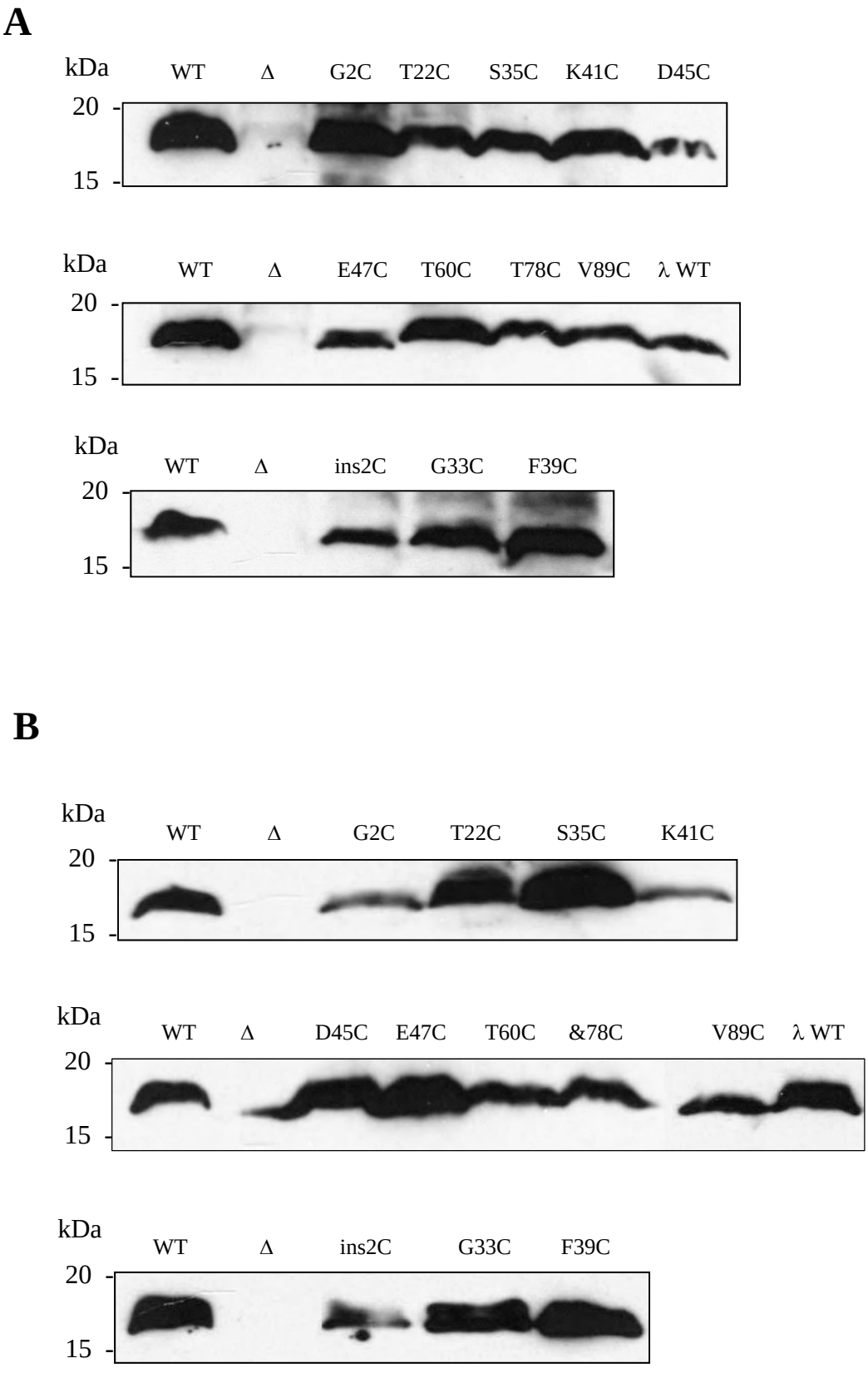

relative periplasmic TMAO reductase activity
(% of DADE-P pUNITATCC4)
160
140
120
100
80
60
40
20
0
pUNITATCC4
pQE60
ins2C
D45C
E47C
K49C
T60C
T78C
V89C

Figure S3


A
kDa
WT
Δ
G2C
T22C
S35C
K41C
D45C
20
15
kDa
WT
Δ
E47C
T60C
T78C
V89C
λ WT
20
15
kDa
WT
Δ
ins2C
G33C
F39C
20
15
B
kDa
WT
Δ
G2C
T22C
S35C
K41C
20
15
kDa
WT
Δ
D45C
E47C
T60C
&78C
V89C
λ WT
20
15
kDa
WT
Δ
ins2C
G33C
F39C
20
15