Kent Academic Repository
Full text document (pdf)

Citation for published version

Sutherland, George A. and Grayson, Katie J. and Adams, Nathan B.P. and Mermans, Daphne M.J. and Jones, Alexander S. and Robertson, Angus J. and Auman, Dirk B. and Brindley, Amanda A. and Sterpone, Fabio and Tuffery, Pierre and Derreumaux, Philippe and Dutton, P. Leslie and Robinson, Colin and Hitchcock, Andrew and Hunter, C Neil (2018) Probing the quality control

DOI

https://doi.org/10.1074/jbc.RA117.000880

Link to record in KAR

http://kar.kent.ac.uk/66556/

Document Version

Publisher pdf

Probing the quality control mechanism of the *Escherichia coli* twin-arginine translocase using folding variants of a *de novo*-designed heme protein

Author Affiliation:

George Sutherland (University of Sheffield, United Kingdom), Katie Grayson (University of Sheffield, United Kingdom), Nathan Adams (University of Sheffield, United Kingdom), Daphne Mermans (University of Kent, United Kingdom), Alexander Jones (University of Kent, United Kingdom), Angus Robertson (University of Sheffield, United Kingdom), Dirk Auman (University of Pennsylvania, United States), Amanda Brindley (The University of Sheffield,), Fabio Sterpone (Université Paris Diderot, France), Pierre Tuffery (Université Paris Diderot, France), Philippe Derreumaux (Université Paris,), P. Leslie Dutton (University of Pennsylvania,), Colin Robinson (University of Kent, United Kingdom), Andrew Hitchcock (University of Sheffield, United Kingdom), C Hunter (University of Sheffield, United Kingdom)

Abstract:

Protein transport across the cytoplasmic membrane of bacterial cells is mediated by either the general secretion (Sec) system or the twin arginine translocase (Tat). The Tat machinery exports folded and cofactor containing proteins from the cytoplasm to the periplasm using the transmembrane proton motive force as a source of energy. The Tat apparatus apparently senses the folded state of protein substrates but the basis of this quality control mechanism, which prevents premature export of nascent unfolded or misfolded polypeptides, has not been determined. Here, we investigated the innate ability of the model *Escherichia coli* Tat system to recognize and translocate *de novo*-designed protein substrates with experimentally determined differences in the extent of folding. Water-soluble, four-helix bundle maquette proteins were engineered to bind 2, 1 or 0 heme *b* cofactors which results in a concomitant reduction in the extent of their folding, assessed using temperature-dependent CD spectroscopy and one-dimensional ^{1}H NMR spectroscopy. Fusion of the archetypal N-terminal Tat signal peptide of the *E. coli* trimethylamine-N-oxide (TMAO) reductase (TorA) to the N-terminus of the protein maquettes was sufficient to identify them as Tat substrates. The clear correlation between the level of Tat-dependent export and the degree of heme *b* induced folding of the maquette suggests that the membrane-bound Tat machinery is able to sense the extent of folding and conformational flexibility of its substrates. These artificial proteins are ideal substrates for future investigations of the Tat system's quality control mechanism.

Competing interests:

Author contributions:

George Sutherland: Data curation; Formal analysis; Validation; Visualization; Methodology; Writing-original draft; Writing-review and editing Katie Grayson: Data curation; Formal analysis; Investigation; Methodology; Writing-original draft Nathan Adams: Data curation; Software; Formal analysis; Investigation; Methodology; Writing-review and editing Daphne Mermans: Data curation; Formal analysis; Investigation; Methodology Alexander Jones: Data curation; Formal analysis; Methodology Angus Robertson: Data curation; Formal analysis; Investigation; Methodology Dirk Auman: Data curation; Formal analysis; Investigation; Methodology Amanda Brindley: Data curation; Formal analysis; Methodology Fabio Sterpone: Conceptualization; Formal analysis; Methodology Pierre Tuffery: Resources; Software; Formal analysis; Methodology Philippe Derreumaux: Resources; Software; Formal analysis; Visualization; Methodology P. Leslie Dutton: Conceptualization; Formal analysis; Methodology Colin Robinson: Conceptualization; Data curation; Formal analysis; Methodology Andrew Hitchcock: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Writing-original draft; Writing-review and editing C Hunter: Conceptualization; Resources; Data curation; Formal analysis; Funding acquisition; Methodology; Writing-original draft; Project administration; Writing-review and editing

Reviewer Instructions

Probing the quality control mechanism of the Escherichia coli twin-arginine translocase using folding variants of a de novo-designed heme protein

George A. Sutherland[a], Katie J. Grayson[a, 1], Nathan B. P. Adams[a], Daphne M. J. Mermans[b], Alexander S. Jones[b], Angus J. Robertson[a], Dirk Auman[c], Amanda A. Brindley[a], Fabio Sterpone[d], Pierre Tuffery[e], Philippe Derreumaux[d], P. Leslie Dutton[c], Colin Robinson[b], Andrew Hitchcock[a] and C. Neil Hunter[a, *]

[a] Department of Molecular Biology and Biotechnology, University of Sheffield, Sheffield, UK.

[b] School of Biosciences, University of Kent, Canterbury, UK.

[c] Department of Biochemistry and Biophysics, University of Pennsylvania, Philadelphia, USA.

[d] Laboratoire de Biochimie Théorique, UPR 9080 CNRS, Université Paris Diderot, Sorbonne Paris Cité, 75005 Paris, France.

[e] INSERM U973, Université Paris Diderot, Sorbonne Paris Cité, 75013 Paris, France.

[1] Present Address: School of Biochemistry, University of Bristol, Bristol, UK.

[*]**Author for correspondence**: Professor C. Neil Hunter FRS, Department of Molecular Biology and Biotechnology, University of Sheffield, Sheffield S10 2TN, United Kingdom. Email: c.n.hunter@sheffield.ac.uk. Phone: +44 (0)114 222 4191. Fax: +44 (0)114 222 2711.

Abstract

Protein transport across the cytoplasmic membrane of bacterial cells is mediated by either the general secretion (Sec) system or the twin arginine translocase (Tat). The Tat machinery exports folded and cofactor containing proteins from the cytoplasm to the periplasm using the transmembrane proton motive force as a source of energy. The Tat apparatus apparently senses the folded state of protein substrates but the basis of this quality control mechanism, which prevents premature export of nascent unfolded or misfolded polypeptides, has not been determined. Here, we investigated the innate ability of the model Escherichia coli Tat system to recognize and translocate de novo-designed protein substrates with experimentally determined differences in the extent of folding. Water-soluble, four-helix bundle maquette proteins were engineered to bind 2, 1 or 0 heme b cofactors which results in a concomitant reduction in the extent of their folding, assessed using temperature-dependent CD spectroscopy and one-dimensional ^1H NMR spectroscopy. Fusion of the archetypal N-terminal Tat signal peptide of the E. coli trimethylamine-N-oxide (TMAO) reductase (TorA) to the N-terminus of the protein maquettes was sufficient to identify them as Tat substrates. The clear correlation between the level of Tat-dependent export and the degree of heme b induced folding of the maquette suggests that the membrane-bound Tat machinery is able to sense the extent of folding and conformational flexibility of its substrates. These artificial proteins are ideal substrates for future investigations of the Tat system's quality control mechanism.

Introduction

The transport of proteins across membranes is one of the great challenges faced by the cell. In prokaryotes, two major pathways are used to achieve protein translocation across the cytoplasmic (inner) membrane. The general Secretion (Sec) pathway transports proteins in an unfolded configuration using energy provided by both ATP hydrolysis and the transmembrane proton gradient (1). In contrast, the Twin arginine translocase (Tat) system transports fully folded proteins (2, 3) and is energized solely by the transmembrane proton gradient.

The majority of Tat substrates are cofactor-containing proteins that require assembly in the cytoplasm (4-6) including those that fold too quickly for Sec transport (7) and those that assemble into oligomeric complexes (8). Proteins translocated by the Tat pathway have an N-terminal signal sequence characterized by a twin arginine (RR) motif (9); the signal sequence is cleaved from the precursor protein during or immediately after translocation, releasing the mature protein into the periplasm.

How components of the Tat machinery assess the folding state of a protein substrate remains poorly understood (6, 10, 11). Many studies have attempted to prove the presence of a quality control or proofreading mechanism that exists to prevent futile export of misfolded or misassembled proteins (12-14). In Escherichia coli the Tat apparatus comprises TatA/B family proteins and the TatC protein. Single point mutations in either TatA or TatC were identified that allow translocation of unfolded substrates, and point mutations in both TatB and TatC enabled export of a broader range of unfolded substrates, suggesting that the TatABC subunits cooperatively assess the folding state of proteins independently of protein translocation (15).

Richter et al. (16) showed that small, unstructured hydrophilic FG repeat proteins could be exported by the Tat system, and that the presence of hydrophobic surface patches was sufficient to abort transport, raising the possibility that the Tat system screens proteins based on their surface hydrophobicity. It has been reported that the length of the unstructured FG repeat polypeptide dramatically affects Tat export, with longer regions abolishing Tat export altogether (17). Conversely, Jones et al. (18) recently reported that the Tat system was surprisingly tolerant of hydrophobic patches on the surface of structured single-chain variable fragment proteins, and export efficiency was increased with greater structural rigidity. Chaperones may also prevent export of a protein until cofactor insertion has taken place (19-21), and mutants incapable of cofactor binding are rapidly degraded once in contact with the Tat machinery (22).

In order to further investigate the Tat quality control mechanism, we used maquettes, which are simple, repetitive protein structures designed de novo from first principles with minimal reference to natural protein structures (23-26). As natural protein structures contain unnecessary complexity, accumulated from perpetual rounds of blind natural selection (27), altering protein residues can have unpredictable effects on protein structure and dynamics. In contrast, the role of each amino acid in the simple maquette structure has been rationalized at the outset, so changes to structure and function become more predictable.

The maquettes used in this study are based on the BT6 maquette developed by Farid et al. (28), and consist of loops linking four largely identical α-helices enclosing a water-excluding cavity that can accommodate cofactors (Fig. 1A-C). Histidine residues within the maquette cavity ligate hemes, producing artificial proteins resembling b-type cytochromes and myoglobin (29, 30).

Here, three different maquette variants were utilized, each with a different heme b binding capacity (Fig. 1). Using nuclear magnetic resonance (NMR) and circular dichroism (CD) spectroscopy we show that binding 2, 1 or 0 hemes (Fig. 1A, B, C, respectively) imparts changes in the extent of folding of the maquette variants. The archetypal trimethylamine N-oxide reductase (TorA) Tat signal peptide was fused to the N-terminus of the three proteins and the E. coli Tat machinery was challenged to differentiate between the folding variants. We show that the Tat apparatus is able to sense the conformational flexibility of the different maquette substrates, and that increasingly well-folded maquettes are exported with enhanced efficiency.

Results

Production of heme-reconstituted maquettes–The aim of this study was to test whether the E. coli Tat system could recognize and export a de novo designed di-heme protein, and then to use variants of this protein with experimentally confirmed differences in conformational flexibility to test if the Tat apparatus selectively processes the more folded proteins.

As described above and depicted in the molecular dynamics simulated structures

presented in Fig. 1, three variants of the BT6 maquette were produced, facilitating the incorporation of 2 (BT6), 1 (BT6M1) or 0 (BT6M0) heme b cofactors. The requirement for bis-histidine ligation of heme b in the artificial constructs enabled the generation of the 1 and 0 heme binding variants through H53A, and H53A/H88A (histidine to alanine) point mutations, respectively (Table 1).

Maquettes are typically overproduced in E. coli in large amounts, with bound heme largely absent following purification (see Materials and Methods) because the native tetrapyrrole biosynthesis pathway cannot keep pace with the induced synthesis of maquette (Fig. S1A). Apo-proteins were reconstituted in vitro with an excess of heme and unbound pigment was removed by anion exchange chromatography. The Soret absorption bands for the BT6, BT6M1 and BT6M0 maquettes were normalized for protein concentration (absorbance at 280 nm) and had maxima at 413 nm, 412 nm and 396 nm, respectively (Fig. 2A), with the amplitudes corresponding to binding 2, 1, and 0 hemes, respectively. The weak, blue-shifted absorption for the BT6M0 maquette suggested a low level of adventitiously bound heme. We compared heme absorption in the maquette-bound, and solvated (buffer or dimethyl sulfoxide, DMSO) states which had respective maxima at 413 nm, 404 nm and 384 nm (Fig. S1B). The blue-shifted absorption maxima for solvated heme are consistent with the weak, blue-shifted absorption for BT6M0 (Fig. 2A) arising from residual, weakly bound heme; ligation into the internal cavity of BT6 and BT6M1 causes a red-shift absorption of the heme.

Thermostability and folding of maquette variants–Temperature-dependent CD spectroscopy was performed to assess the effect of heme ligation on the thermal stability of each maquette scaffold. At 15 °C the CD spectra of all three maquette variants, with or without bound heme, were typical of α-helical structures (Fig. S2, black lines). Likewise, irrespective of heme binding heating to 80 °C resulted in spectra typical of that of a denatured protein (31) (Fig. S2 from black to red to grey lines).

Ellipticity at 222 nm was monitored during temperature cycling from 15 to 80 to 15 °C to observe denaturing and subsequent re-folding of protein (Fig. 2B-D). For BT6, heme incorporation displaced the midpoint of the melting curve from 22 °C to 56 °C, the 34 °C difference indicating increased stabilization following heme binding. The same procedure with BT6M1 increased stabilization by only 16 °C (25 °C to 41 °C), and had very little effect on the BT6M0 scaffold (32 °C to 35 °C) (Table 2). Thus, bis-histidine ligation of two hemes within the maquette cavity of BT6 significantly stabilizes the four-helix bundle structure, and there is a smaller effect with only one bis-histidine ligation in BT6M1. In the case of BT6M0, where heme ligation is not possible, there is correspondingly no significant stabilization of maquette structure in the presence of heme.

One-dimensional proton (^{1}H) NMR spectroscopy was used to assess conformational changes in tertiary structure upon heme binding to maquettes. A ^{1}H NMR spectrum characteristic of a protein with limited tertiary structure was observed for all maquette scaffolds in the absence of heme (Fig. 3A, S3). Addition of heme to BT6 showed greater dispersion of resonances in the amide proton region, with a notable increase in the number of peaks at around 10 ppm, suggesting that the heme is binding to the scaffold and stabilizing the protein tertiary structure (Fig. 3B). BT6M1 showed a small increase in resonance dispersion following heme addition (Fig. 3C), whilst BT6M0 showed no significant change in amide proton dispersion (Fig. 3D). The increase in the chemical shift dispersion observed in the methyl proton region (at around 1.0-0.0 ppm) for BT6, BT6M1, and BT6M0 mirrored the behavior observed for the amide proton resonances (Fig. S4). Together, the increase in ^{1}H NMR dispersion in both backbone amide and methyl regions, confirm the ligation of heme into the BT6 and BT6M1 variants and indicate an increase in protein folding upon ligation. These results are consistent with a previous NMR study that showed poor dispersion for a related apo maquette, and progressive structuring of the maquette iwhen one then two equivalents of heme were added (28).

In vivo Tat export assays–The E. coli TorA signal peptide is sufficient to direct green

fluorescent protein (GFP) to the E. coli periplasm via the Tat system (13). Constructs in which the sequence encoding the TorA signal peptide (residues 1-39) and first four amino acids of the mature TorA protein (residues 40-43) were added in-frame to the 5'terminus of the genes encoding the BT6 maquette variants followed by a C-terminal 6xHis tag (Fig. S5) were synthesized and cloned into the pEXT22 vector (Table S1). In the TorA-BT6 variants residues H95 and H130 correspond to H53 and H88 in the maquette scaffolds presented in Fig. 1.

The resulting plasmids were co-transformed into E. coli BL21 (DE3) along with empty pET-21a(+). The pEXT22 vector was used as it allows isopropyl β-D-1-thiogalactopyranoside (IPTG) inducible expression of the synthetic gene under the control of a tightly regulated tac promoter. The low copy number of pEXT22 (R100 origin of replication, 1-1.5 copies per chromosome) (32, 33) limits production of recombinant protein, to avoid overwhelming the E. coli Tat apparatus with substrate.

Following growth and protein production, periplasmic fractions were prepared from harvested cells as described in Materials and Methods and analyzed alongside cell-free extracts prepared from identically grown cells. The C-terminal 6xHis tag allows immunodetection of the recombinant protein; the predicted molecular weight of the unprocessed precursor protein is ~21 kDa while the Tat-processed mature protein is ~17 kDa. Anti-GroEL was used in order to probe the degree of cytoplasmic contamination in our periplasmic preparations (Fig. 4A); for all samples only a faint signal was detectable in the periplasm compared to that detected in cell free extracts, indicating the level of contamination was very low. The Tat export assays were performed in cells which also maintained the pET-21a(+) vector, allowing immunoblot signals to be normalized by immunodetection of β-lactamase (Fig. 4B), allowing direct comparison of the degree of Tat transport between samples.

For TorA-BT6, a 17 kDa signal was present for both periplasmic and cell free extract (Fig. 4C) preparations with no signal detected at 21 kDa, indicating complete processing and export of TorA-BT6. To confirm the periplasmic localization was Tat-dependent, a R12K/R13K (KK) variant of the TorA signal peptide was generated and fused to BT6 (Table 1); it has previously been demonstrated that the KK change completely abolishes Tat-transport (12, 34). When this construct was tested, a 21 kDa signal corresponding to the unprocessed apo protein is observed in blots of total lysates and no processed protein was detectable in the periplasm (Fig. 4C). Together these data indicate that the BT6 scaffold is able to acquire heme from the native biosynthetic pathway and fold in vivo, and that this folding event can be detected by the Tat system, with the TorA signal peptide sufficient for Tat-mediated recognition and transport of heme-loaded BT6 to the E. coli periplasm.

To probe the proofreading ability of the Tat machinery to recognize and efficiently export only folded proteins, similar constructs were generated in which the TorA signal peptide was fused to the BT6M1 and BT6M0 maquettes, which, in the presence of heme, are either partially folded (BT6M1) or unfolded (BT6M0) relative to BT6 (Figs. 2 and 3). Compared to BT6, considerably less protein was found in the periplasm for BT6M1, and almost no signal for BT6M0 (Fig. 4C). Conversely, blots of cell free extracts of the same strains reveal a significant signal for the unprocessed maquettes for BT6M0, and to a lesser extent BT6M1, as observed for the to the TorA-KK-BT6 control, showing that the pre-protein accumulates in the cytoplasm when it is less well folded (Fig. 4B).

The cellular levels of maquette variants should be very similar as all the proteins are identical apart from 1 or 2 point mutations and were expressed from the same plasmid in the same cell line under the same conditions. Indeed, the data presented in Fig. 4C shows that the total signal (preprotein plus exported protein where relevant) for all three maquettes is similar. However, to further confirm that the differential periplasmic targeting of the three BT6 variants by the Tat system was due to differences in their heme loading and associated folding and not in expression, synthesis or stability, we performed two additional control experiments.

4

First we produced each maquette under the same conditions as for the Tat export assays but without the TorA signal peptide, and immunoblotting confirmed approximately equal levels of each (Fig. S6A), ruling out that BT6M1 and BT6M0 are produced at lower levels or are more rapidly degraded in the cytoplasm than BT6.

We also performed Sec-system mediated export assays of BT6, BT6M1 and BT6M0. As the Sec system transports unfolded proteins across the cytoplasmic membrane, differences in folding, mediated in the case of BT6 by heme binding, will not affect transport. Maquette genes were cloned into the pET-22b(+) vector (Table S1) in frame with an N-terminal PelB (Erwinia carotovora pectate lyase B) Sec leader peptide (22 amino acids) (35, 36) (Table 1). Stop codons were omitted so proteins had a C-terminal 6xHis tag for immunodetection; the unprocessed proteins have a predicted molecular weight of approximately 19 kDa and the mature maquettes would be approximately 17 kDa following cleavage of the PelB signal sequence. Fig. S6B shows that when normalized amounts of cell lysates were loaded that single, similarly intense signals were detected for BT6, BT6M1 and BT6M0, indicating all three apo-maquettes are equally synthesized and transported to the periplasm when directed through the Sec system.

Purification of maquettes from the periplasm–To demonstrate that heme ligation-into BT6 was responsible for its Tat-dependent localization in the periplasm, production of His-tagged maquettes was performed on a larger scale using the recently described E. coli W3110-TatExpress cell line (37). Cells were fractionated and maquette proteins were purified from the periplasmic fraction by immobilized metal affinity chromatography (IMAC). IMAC eluates were concentrated and striking differences in pigmentation between the maquette variants were observed, with a deep red color for BT6, weaker pigmentation for BT6M1, and no visible color for BT6M0 (Fig. 5 inset panels). SDS-PAGE of the concentrated eluates revealed the same pattern as the Tat exports presented in Fig. 4, with more BT6 than BT6M1 and only a trace amount of BT6M0 (Fig. S7). This also ruled out any leakage of

unprocessed protein across the cytoplasmic membrane as the larger precursor species were not purified (Fig. S7). The absorbance spectra of Tat-exported BT6, BT6M1 and BT6M0 samples normalized at 280 nm are shown in Fig. 5, along with an in vitro heme reconstituted maquette. There is a close match between the spectra of the Tat-exported heme-loaded BT6 purified from the periplasm and the reconstituted BT6 sample (Fig. 5 solid red versus dashed red lines). The high affinity of BT6 for heme (28) ensures full occupancy of the two heme binding sites, so the Tat-transported BT6 purified from the periplasm is also fully heme bound.

These data, alongside the export assays shown in Fig. 4, reveal that the artificial TorA-BT6 maquette is able to sequester heme b from the native E. coli biosynthetic pathway, correctly fold around its substrate, and retain the pigment following export by the Tat-apparatus.

Discussion

The Tat pathway is present in the cytoplasmic membranes of most prokaryotic organisms and is evolutionarily conserved in the thylakoids of plant chloroplasts and some mitochondrial membranes (reviewed in Refs. 3 and 7). As well as a biosynthetic role in numerous important cellular processes, the Tat pathway is required for colonization and virulence of globally significant human pathogens (38, 39). The Tat system differs from the universally conserved Sec pathway, which translocates unstructured polypeptides, as it transports folded, and typically cofactor containing, proteins. How the Tat apparatus is able to determine that a protein is folded, co-factor loaded, and suitable for export is not understood.

Here, we have used artificial heme 'maquette' proteins to investigate the ability of the archetypal E. coli Tat machinery to discriminate between folding states of protein substrates. The maquettes used in this study are four-helix bundle proteins enclosing a hydrophobic cavity that can accommodate heme cofactors ligated by histidine residues. Point mutations to the BT6 maquette, affecting only single histidine ligands, generated the BT6M1 and BT6M0 maquettes, and were sufficient to reduce the number of bound hemes, verified by

absorption spectroscopy (Fig. 2a). The number of bound hemes therefore provides a simple, well-defined method to alter the stability of a protein substrate. The consequent effects of heme content on the thermostability and folding of each maquette were measured using temperature-dependent CD spectroscopy (Fig. 2b, Table 2) and ^{1}H NMR (Fig. 3).

When directed for export through the Tat export machinery of E. coli there were clear differences in the extent of translocation for BT6, BT6M1 and BT6M0 maquettes (Fig. 4) that correlate with the in vitro stability and folding experiments for heme-bound constructs. Export efficiency decreased with increasing conformational flexibility; the unstructured BT6M0 maquette is largely rejected for export, while limited export was observed in the intermediate case of the single heme-binding BT6M1 maquette.

The export data shown would only be obtained if the BT6 maquettes were able to ligate heme, in vivo, produced by the native biosynthetic pathway of E. coli, as heme coordination is a pre-requisite for protein stability and folding. Evidence for in vivo heme ligation was shown with the purification of heme-containing maquettes from the periplasmic fraction of E. coli (Fig. 5), the spectra of which strongly resemble those of heme proteins generated in vitro (Fig. 2a).

Although the native E. coli Tat translocase will never have encountered the artificial protein substrates described here, our data shows that it is able to recognize and distinguish between them, even to the extent of processing the intermediate state BT6M1 differently from the folded BT6 and the least structured BT6M0. Thus, the Tat proofreading process must involve a generic form of discrimination; the correlation between the structural flexibility of the Tat substrate and its suitability for export suggests that there are initial encounters between Tat components and the substrate at the membrane surface. Such interactions might sense flexible motions of the substrate that are transmitted to other components of the Tat machinery, preventing transport across the membrane.

This proof-of-concept study shows that an artificial protein, engineered to bind heme b

from a native biosynthetic pathway, can be exported from the cell in its correctly folded, functional state. Development of this concept may contribute a significant advance in biotechnology, where the principle could be applied to other organisms and to important biomolecules and protein-cofactor complexes, particularly those that may cause toxicity to the host.

Experimental procedures

Molecular dynamics and structure prediction–To generate a structural model for BT6, the amino acid sequence was run through the PEP-FOLD structure prediction algorithm (40) for a total of 100 simulations, from which the five best-scoring structures were selected for further analysis (all 4-helix bundles resembling the anticipated fold of the protein). From these, a structure with both pairs of histidines positioned closest to their respective heme-ligating positions was selected as a starting structure. From this structure, mutants BT6M1 and BT6M0 were generated using PyMol v1.7 (PyMol Molecular Graphics System Version 1.7, Schrödinger, LLC). Hemes (0, 1 or 2) were manually docked into the bis-his sites of the structures.

To relax the starting structure into an energy minimized conformation, molecular dynamics (MD) preprocessing, production runs, and post-processing were performed with Gromacs v4.6 (41). MD simulations were performed using the CHARMM27 Force Field for proteins and TIP3P-CHARMM model for water. For the heme cofactor, parameters were used for reduced, deprotonated, bis-his ligated heme included in the Charmm27 force field. Covalent bonds were explicitly specified between the heme iron and the epsilon nitrogen of the relevant histidine residues. NaCl was added to a total concentration of 150 mM. Each structure was equilibrated for 100 ps in the NVT ensemble, followed by 100 ps in the NPT ensemble. Temperature and pressure/density plots (following NVT and NPT equilibrations, respectively) were checked for convergence before proceeding to production MD runs of 50 ns. Following production runs, trajectories were corrected for periodicity and centered on the protein. Snapshots of the trajectories at 50 ns

were exported to pdb files. Electrostatics and solvent accessible surface calculations were performed using APBS (42). Visualization of calculations and structures was performed with the PyMol Molecular Graphics System Version 1.7, Schrödinger, LLC.

DNA manipulation–Plasmids and primers used in this study are provided in Tables S1 and S2, respectively. Synthetic genes encoding maquette protein variants codon optimized for expression in E. coli were purchased from DNA2.0 (now ATUM) or Integrated DNA Technologies. Point mutations were generated using the QuikChange II Site-Directed Mutagenesis Kit (Agilent). The TorA Tat signal peptide was amplified from E. coli genomic DNA and joined in frame to maquette constructs by overlap extension (OLE)-PCR using Q5 High-Fidelity DNA Polymerase (NEB). All plasmids were sequence verified by automated DNA sequencing (GATC Biotech). Competent E. coli JM109 (Promega) was used for cloning and was grown in Luria-Bertani (LB) broth/agar supplemented with the appropriate antibiotic(s) (Table S1).

Production and purification of untagged maquettes–The high copy number pJexpress414 plasmid was used for recombinant protein production under the control of an IPTG inducible T7 promoter. Escherichia coli BL21(DE3) containing the desired plasmid was grown with shaking (230 rpm) at 37 °C in LB broth with 100 µg ml^{-1} ampicillin to an optical density at 600 nm (OD$_{600}$) of ~0.6. At this point IPTG was added to a final concentration of 1 mM to induce expression and the cultures were incubated for a further 16 h at 37 °C. Cells were harvested by centrifugation (4,400 xg, 15 min, 4°C) and resuspended in buffer A (50 mM HEPES pH 7.4, 500 mM NaCl, 5 mM imidazole). Cells were lysed by sonication on ice and the lysate was clarified by centrifugation (53,000 xg, 30 mins, 4°C). The supernatant was filtered through a 0.45 µm filter and applied to a Chelating Sepharose Fast Flow column (GE Healthcare) pre-equilibrated with 10 mg mL^{-1} nickel sulfate. The column was washed with 5 column volumes of buffer A with the flow-through and wash collected and pooled. The pooled sample was buffer exchanged into buffer B (50 mM HEPES pH 7.4) and further purified

by ion exchange chromatography (IEC) on a Fast Flow Q-Sepharose column (GE Healthcare) with a linear gradient of 0-1 M NaCl in buffer B. Where required maquettes were further purified by size exclusion chromatography on a Superdex 200 increase column (GE Healthcare) in buffer C (50 mM HEPES pH7.4, 200 mM NaCl). Where necessary protein was concentrated using Vivaspin centrifugal concentrators (Sartorius).

Heme reconstitution into apo-maquettes–Hemin (Sigma-Aldrich) stocks (1 mg mL^{-1}) were prepared in 100% DMSO. Protein concentrations were calculated by absorbance at 280 nm following the method described by Gill and von Hippell (43), and using the experimentally determined (extinction coefficient of 32.6 mM^{-1} cm^{-1}. Reconstitutions were conducted with a 10-fold molar excess of hemin in buffer D (50 mM HEPES pH 7.4, 200 mM NaCl, 20 % (v/v) DMSO) and incubated for 45 minutes at 25 °C, before being transferred to ice and buffer exchanged into buffer B using Vivaspin centrifugal concentrators. Unbound cofactor was removed by IEC on a DEAE Sepharose (Sigma Aldrich) column.

UV-visible absorbance spectroscopy–Protein samples were buffer exchanged into buffer E (5 mM sodium phosphate buffer pH 7.4) and UV-visible absorbance was measured in a 1 cm path length polystyrene UV cuvette in a Cary 60 UV-Vis spectrophotometer (Agilent) at room temperature.

Circular dichroism spectroscopy–Mean residue ellipticity ([θ]$_{MRW}$) of protein samples was measured in a 1 mm path length quartz cuvette on a Jasco J-810 spectropolarimeter with a Jasco PFD-425S Peltier to enable temperature control. Spectra were obtained from 15 °C to 80 °C at 5 °C intervals. Spectra were recorded continuously at a scan speed of 100 nm min^{-1}, with 1 nm resolution and a 4 s response with 4 accumulations. Ellipticity ([θ]) at 222 nm was measured every 1 °C from 15 °C to 80 °C at 1 °C min^{-1} with a 4 s response. Melting temperatures (T$_m$) and refolding temperatures (T$_f$) are the temperature at which 50% of the protein is unfolded or folded respectively as determined by fitting melting data to a sigmoidal Boltzmann distribution according to the following equation where T is the temperature, T$_m$ is the melting (or

folding) temperature and T_0 is the initial temperature of the experiment.

$$[\theta]_{222\text{nm}} = [\theta]_{222\text{nm}}^{\text{base}} + \left(\frac{[\theta]_{222\text{nm}}^{\text{max}}}{1 + exp^{\left(\frac{T_m - T_0}{\frac{\partial T}{\partial [\theta]}}\right)}}\right)$$

Proton nuclear magnetic resonance spectroscopy–Spectra were recorded at 298 K on 0.2-0.5 mM protein samples in buffer E, with the addition of 10% D_2O (spectrometer lock), and 1 mM trimethylsilylpropanate (TSP) (reference standard). ^{1}H NMR spectra were recorded using a Bruker Avance 800 MHz spectrometer fitted with a 5-mm QXI room temperature probe, equipped with z-axis gradients. One-dimensional experiments were acquired as accumulations of 4096 transients over a spectral width of 24.038 kHz, corresponding to a proton spectral width of 30.0 ppm. All data were processed using an EM window function and 5-Hz line broadening, without linear prediction in TopSpin (Bruker). Spectra were referenced to TSP at 0 ppm prior to overlay and analysis.

Escherichia coli fractionation–In order to avoid overloading the Tat system for in vivo transport assays, maquette genes were cloned into the KpnI and XbaI sites of the low copy number pEXT22 plasmid for expression from a tightly controlled tac promoter (33). The pEXT22 constructs (Table S1) were co-transformed into E. coli BL21(DE3) cells along with empty pET-21a(+). Cultures (50 ml) were grown at 37°C with shaking (230 rpm) in LB medium with 30 µg ml^{-1} kanamycin and 100 µg ml^{-1} ampicillin in 250 mL Erlenmeyer flasks. At an OD_{600} of ~0.6 expression was induced with 0.5 mM IPTG for 2 h. Cells were harvested by centrifugation (3,900 xg, 30 min, 4°C) and washed with buffer F (100 mM Tris-acetate pH 8.2, 500 mM sucrose, 5 mM ethylenediaminetetraacetic acid (EDTA)). To prepare cell free lysates, cells were re-suspended to OD_{600} ~2 in 2 mL chilled Buffer C and lysed by sonication on ice. The supernatant, following clarification by centrifugation (16,600 xg, 30 min 4°C), was collected as the cell-free lysate.

Periplasmic fractions were obtained using a procedure based on the EDTA/lysozyme/osmotic shock method described by Randall and Hardy (44). Briefly, cells were re-suspended to OD_{600} ~10 in 500 µL chilled Buffer F followed by the addition of 5 µl of 1 mg ml^{-1} lysozyme. 500 µl of chilled QH_2O was added and cells were incubated on ice for 5 min prior to the addition of 20 µl 1 M $MgSO_4$. Spheroplasts were pelleted by centrifugation at 16,600 xg for 30 min in a pre-chilled (4°C) microcentrifuge and the supernatant was collected as the periplasm.

Immunoblotting–Immunoblotting was performed essentially as described previously (45). Briefly, proteins were separated by SDS-PAGE on 12% Bis-Tris gels (Invitrogen), transferred to polyvinylidene difluoride membranes (Invitrogen), and incubated with anti-6-His (Bethyl Laboratories, Inc.), anti-β-Lactamase (Abcam) or anti-GroEL (Sigma-Aldrich) primary antibodies followed by an appropriate horseradish peroxidase conjugated secondary antibody (Sigma-Aldrich). Chemiluminescence was detected using the WESTAR ETA C 2.0 chemiluminescent substrate (Cyanagen) on an Amersham Imager 600 (GE Healthcare).

Purification of Tat-exported proteins from E. coli periplasm–The E. coli W3110 'TatExpress' cell line (37) was transformed with the pEXT22 vector containing either TorA-BT6, TorA-BT6M1 or TorA-BT6M0 (Table S1). 500 mL cultures were grown in 2 L Erlenmeyer flasks at 30 °C with 220 rpm agitation. At an OD_{600} of ~0.6 protein production was induced with 0.5 mM IPTG and cultures were incubated for 24 h, after which cells were harvested by centrifugation (3,900 x g, 30 min, 4°C). To obtain periplasmic fractions, cells were resuspended in 10 mL chilled buffer F. 10 mL chilled QH_2O was added followed by 800 µL 1 mg mL^{-1} lysozyme and samples were incubated on ice for 10 min. 800 µL 1 M $MgSO_4$ was added and the solution was centrifuged (16,600 xg, 30 min, 4°C) with the supernatant collected as periplasm.

Periplasmic fractions were applied to a Chelating Sepharose Fast Flow column (GE Healthcare) pre-equilibrated with 10 mg mL^{-1} nickel sulfate. The column was washed with 20

mL of buffer A and 20 mL buffer G (50 mM HEPES pH 7.4, 500 mM NaCl, 50 mM imidazole). Protein was eluted with 10 mL buffer H (50 mM HEPES pH 7.4, 100 mM NaCl, 400 mM imidazole) and the elution fractions were collected.

References

1. De Keyzer, J., Van Der Does, C., and Driessen, A. J. M. (2003) The bacterial translocase: A dynamic protein channel complex. Cell Mol. Life Sci. **60**, 2034–2052
2. 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
3. Berks, B. C. (2015) The Twin-Arginine Protein Translocation Pathway. Annu. Rev. Biochem. **84**, 843–864
4. Berks, B. C. (1996) A common export pathway for proteins binding complex redox cofactors? Mol. Microbiol. **22**, 393–404
5. Santini, C. L., Ize, B., Chanal, A., Müller, M., Giordano, G., and Wu, L. F. (1998) A novel Sec-independent periplasmic protein translocation pathway in Escherichia coli. EMBO J. **17**, 101–112
6. Natale, P., Brüser, T., and Driessen, A. J. M. (2008) Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane - Distinct translocases and mechanisms. Biochim. Biophys. Acta – Biomembr. **1778**, 1735–1756
7. Palmer, T., and Berks, B. C. (2012) The twin-arginine translocation (Tat) protein export pathway. Nat. Rev. Microbiol. **10**, 483–496
8. Rodrigue, A., Chanal, A., Beck, K., Müller, M., and Wu, L. (1999) Co-translocation of a periplasmic enzyme complex by a hitchhiker mechanism through the bacterial tat pathway. J. Biol. Chem. **274**, 13223–13228
9. 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
10. Sargent, F., Stanley, N. R., Berks, B. C., and Palmer, T. (1999) Sec-independent protein translocation in Escherichia coli. J. Biol. Chem. **274**, 36073–36082
11. Patel, R., Smith, S. M., and Robinson, C. (2014) Protein transport by the bacterial Tat pathway. Biochim. Biophys. Acta – Mol. Cell. Res. **1843**, 1620–1628
12. Halbig, D., Wiegert, T., Blaudeck, N., Freudl, R., and Sprenger, G. A. (1999) The efficient export of NADP-containing glucose-fructose oxidoreductase to the periplasm of Zymomonas mobilis depends both on an intact twin- arginine motif in the signal peptide and on the generation of a structural export signal induced by cofactor bind. Eur. J. Biochem. **263**, 543–551
13. Thomas, J. D., Daniel, R. A., Errington, J., and Robinson, C. (2001) Export of active green fluorescent protein to the periplasm by the twin-arginine translocase (Tat) pathway in Escherichia coli. Mol. Microbiol. **39**, 47–53
14. DeLisa, M. P., Tullman, D., and Georgiou, G. (2003) Folding quality control in the export of proteins by the bacterial twin-arginine translocation pathway. Proc. Natl. Acad. Sci. U.S.A. **100**, 6115–6120
15. Rocco, M. A., Waraho-Zhmayev, D., and DeLisa, M. P. (2012) Twin-arginine translocase mutations that suppress folding quality control and permit export of misfolded substrate proteins. Proc. Natl. Acad. Sci. U.S.A. **109**, 13392–13397
16. Richter, S., Lindenstrauß, U., Lücke, C., Bayliss, R., and Brüser, T. (2007) Functional tat transport of unstructured, small, hydrophilic proteins. J. Biol. Chem. **282**, 33257–33264
17. Lindenstrauß, U., and Brüser, T. (2009) Tat transport of linker-containing proteins in Escherichia coli. FEMS Microbiol. Lett. **295**, 135–140
18. Jones, A. S., Austerberry, J. I., Dajani, R., Warwicker, J., Curtis, R., Derrick, J. P., and Robinson, C. (2016) Proofreading of substrate structure by the Twin-Arginine Translocase is highly dependent on substrate conformational flexibility but surprisingly tolerant of surface charge and hydrophobicity changes. Biochim. Biophys. Acta - Mol. Cell Res. **1863**, 3116–3124
19. Oresnik, I. J., Ladner, C. L., and Turner, R. J. (2001) Identification of a twin-arginine leader-binding protein. Mol. Microbiol. **40**, 323–331
20. Ray, N., Oates, J., Turner, R. J., and Robinson, C. (2003) DmsD is required for the biogenesis of DMSO reductase in Escherichia coli but not for the interaction of the DmsA signal peptide with the Tat apparatus. FEBS Lett. **534**, 156–160

21.   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

22.   Matos, C. F. R. O., Robinson, C., and Di Cola, A. (2008) The Tat system proofreads FeS protein substrates and directly initiates the disposal of rejected molecules. EMBO J. **27**, 2055–2063

23.   Degrado, W. F., Wasserman, Z. R., and Lear, J. D. (1989) Protein Design, a Minimalist Approach. Science **243**, 622–628

24.   Shifman, J. M., Moser, C. C., Kalsbeck, W. A., Bocian, D. F., and Dutton, P. L. (1998) Functionalized de novo designed proteins: mechanism of proton coupling to oxidation / reduction in heme protein maquettes. Biochemistry **37**, 16815–16827

25.   Grosset, A. M., Gibney, B. R., Rabanal, F., Moser, C. C., and Dutton, P. L. (2001) Proof of principle in a de novo designed protein maquette: An allosterically regulated, charge-activated conformational switch in a tetra-α-helix bundle. Biochemistry **40**, 5474–5487

26.   Lichtenstein, B. R., Farid, T. A., Kodali, G., Solomon, L. A., Anderson, J. L. R., Sheehan, M. M., Ennist, N. M., Fry, B. A., Chobot, S. E., Bialas, C., Mancini, J. A., Armstrong, C. T., Zhao, Z., Esipova, T. V., Snell, D., Vinogradov, S. A., Discher, B. M., Moser, C. C., and Dutton, P. L. (2012) Engineering oxidoreductases: maquette proteins designed from scratch. Biochem. Soc. Trans. **40**, 561–566

27.   Edelman, G. M., and Gally, J. A. (2001) Degeneracy and complexity in biological systems. Proc. Natl. Acad. Sci. U.S.A. **98**, 13763–13768

28.   Farid, T. A., Kodali, G., Solomon, L. A., Lichtenstein, B. R., Sheehan, M. M., Fry, B. A., Bialas, C., Ennist, N. M., Siedlecki, J. A., Zhao, Z., Stetz, M. A., Valentine, K. G., Anderson, J. L. R., Wand, A. J., Discher, B. M., Moser, C. C., and Dutton, P. L. (2013) Elementary tetrahelical protein design for diverse oxidoreductase functions. Nat. Chem. Biol. **9**, 826–833

29.   Choma, C. T., Lear, J. D., Nelson, M. J., Dutton, P. L., Robertson, D. E., and DeGrado, W. F. (1994) Design of a Heme-Binding Four-Helix Bundle. J. Am. Chem. Soc. **116**, 856–865

30.   Koder, R. L., Anderson, J. L. R., Solomon, L. A., Reddy, K. S., Moser, C. C., and Dutton, P. L. (2009) Design and engineering of an O2 transport protein. Nature **458**, 305–309

31.   Greenfield, N. J. (2006) Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. Nat. protoc. **1**, 2527–2535

32.   Liu, C., Churchward, G., and Caro, L. (1983) The repA2 gene of the plasmid R100.1 encodes a repressor of plasmid replication. Plasmid **10**, 148–155

33.   Dykxhoorn, D. M., St. Pierre, R., and Linn, T. (1996) A set of compatible tac promoter expression vectors. Gene **177**, 133–136

34.   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

35.   Lei, S. P., Lin, H. C., Wang, S. S., Callaway, J., and Wilcox, G. (1987) Characterization of the Erwinia-carotovora pelB gene and its product pectate lyase. J. Bacteriol. **169**, 4379–4383

36.   Yoon, S., Kim, S., and Kim, J. (2010) Secretory production of recombinant proteins in Escherichia coli. Recent Pat. Biotechnol. **4**, 23–29

37.   Browning, D. F., Richards, K. L., Peswani, A. R., Roobol, J., Busby, S. J. W., and Robinson, C. (2017) Escherichia coli "TatExpress" strains super-secrete human growth hormone into the bacterial periplasm by the Tat pathway. Biotechnol. Bioeng. 10.1002/bit.26434

38.   Hitchcock, A., Hall, S. J., Myers, J. D., Mulholland, F., Jones, M. A., and Kelly, D. J. (2010) Roles of the twin-arginine translocase and associated chaperones in the biogenesis of the electron transport chains of the human pathogen Campylobacter jejuni. Microbiology **156**, 2994–3010

39.   De Buck, E., Lammertyn, E., and Anne, J. (2008) The importance of the twin-arginine translocation pathway for bacterial virulence. Trends Microbiol. **16**, 442–453

40.   Shen, Y., Maupetit, J., Derreumaux, P., and Tufféry, P. (2014) Improved PEP-FOLD approach for peptide and miniprotein structure prediction. J. Chem. Theory Comput. **10**, 4745–4758

41. Hess, B., Kutzner, C., van der Spoel, D., and Lindahl, E. (2008) GROMACS 4: algorithms for highly efficient, load balanced, and scalable molecular simulations. J. Chem. Theory Comput. **4**, 435–447

42. Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. Proc. Natl. Acad. Sci U.S.A. **98**, 10037–10041

43. Gill, S. C., and von Hippel, P. H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. Anal. Biochem. **182**, 319–326

44. Randall, L. L., and Hardy, S. J. S. (1986) Correlation of competence for export with lack of tertiary structure of the mature species: A study in vivo of maltose-binding protein in E. coli. Cell **46**, 921–928

45. Grayson, K. J., Faries, K. M., Huang, X., Qian, P., Dilbeck, P., Martin, E. C., Hitchcock, A., Vasilev, C., Yuen, J. M., Niedzwiedzki, D. M., Leggett, G. J., Holten, D., Kirmaier, C., and Hunter, C.N. (2017) Augmenting light coverage for photosynthesis through YFP-enhanced charge separation at the Rhodobacter sphaeroides reaction centre. Nat. Commun. **8**, 10.1038/ncomms13972

Footnotes

Tables

Table 1. Amino acid sequences of maquettes and signal peptides used in this study.

Name	Sequence*	Details
BT6	MGGDGENLYFQG EIWKQ**H**EDALQKFEEALNQFEDLKQLGGSGSGSGG EIWKQ**H**EDALQKFEEALNQFEDLKQLGGSGSGSGG EIWKQ**H**EDALQKFEEALNQFEDLKQLGGSGSGSGG EIWKQ**H**EDALQKFEEALNQFEDLK	Di-heme binding
BT6M1	MGGDGENLYFQG EIWKQ**H**EDALQKFEEALNQFEDLKQLGGSGSGSGG EIWKQAEDALQKFEEALNQFEDLKQLGGSGSGSGG EIWKQ**H**EDALQKFEEALNQFEDLKQLGGSGSGSGG EIWKQ**H**EDALQKFEEALNQFEDLK	Single heme binding due to removal of 1 coordinating histidine residues
BT6M0	MGGDGENLYFQG EIWKQ**H**EDALQKFEEALNQFEDLKQLGGSGSGSGG EIWKQAEDALQKFEEALNQFEDLKQLGGSGSGSGG EIWKQAEDALQKFEEALNQFEDLKQLGGSGSGSGG EIWKQ**H**EDALQKFEEALNQFEDLK	No heme binding due to removal of 2 coordinating histidine residues
TorA	MNNNDLFQAS*RR*RFLAQLGGLTVAGMLGPSLLTPRRATAAQAA	TorA signal peptide for periplasmic localization by the Tat system
TorA R12/13K	MNNNDLFQAS*KK*RFLAQLGGLTVAGMLGPSLLTPRRATAAQAA	R12K/R13K TorA signal peptide
PelB	MKYLLPTAAAGLLLLAAQPAMA	PelB signal peptide for periplasmic localization by Sec system

*Heme coordinating histidine residues are shown in bold. The twin arginine motif of the Tat signal peptide is shown in bold italics.

Table 2. Melting temperatures (T_m) and folding temperatures (T_f) of maquettes in the absence (apo) or presence of heme

Maquette	Apo		+Heme	
	T_m (°C)	T_f (°C)	T_m (°C)	T_f (°C)
BT6	22	17	56	54
BT6M1	25	24	41	39
BT6M0	32	32	35	36

† The T_m and T_f (temperature at which 50% of the protein is unfolded or folded respectively) values are determined from Boltzman distribution fits as shown in Fig. 2B.

Figures
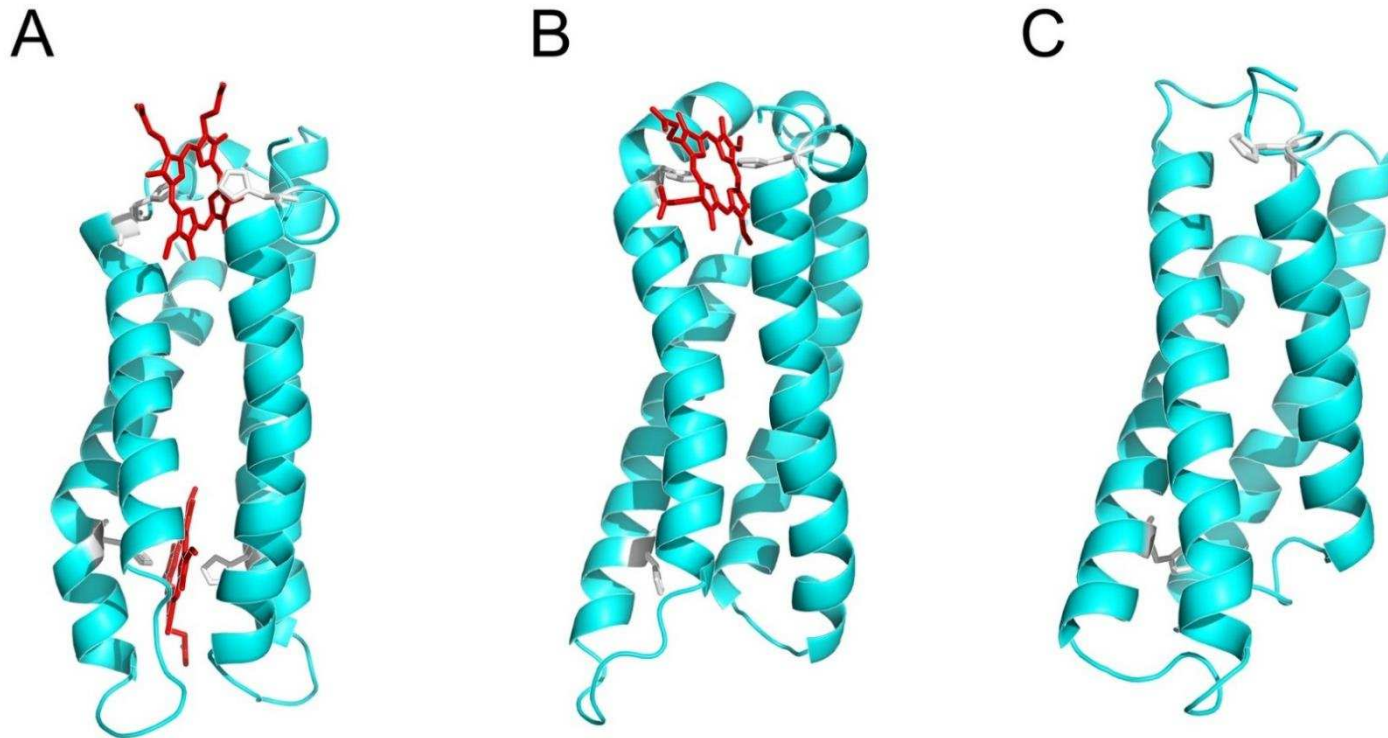
Figure 1. Simulated molecular dynamics trajectory structures of the BT6 maquette proteins used in this study. **(A)** BT6 (see ref. 28) coordinates two heme b molecules (red) using four histidine ligands (white). **(B)** In BT6M1 the H53A substitution means the protein can only coordinate 1 heme b. **(C)** In BT6M0 the double H53A/H88A substitution prevents heme binding. All images taken from 50 ns trajectories.
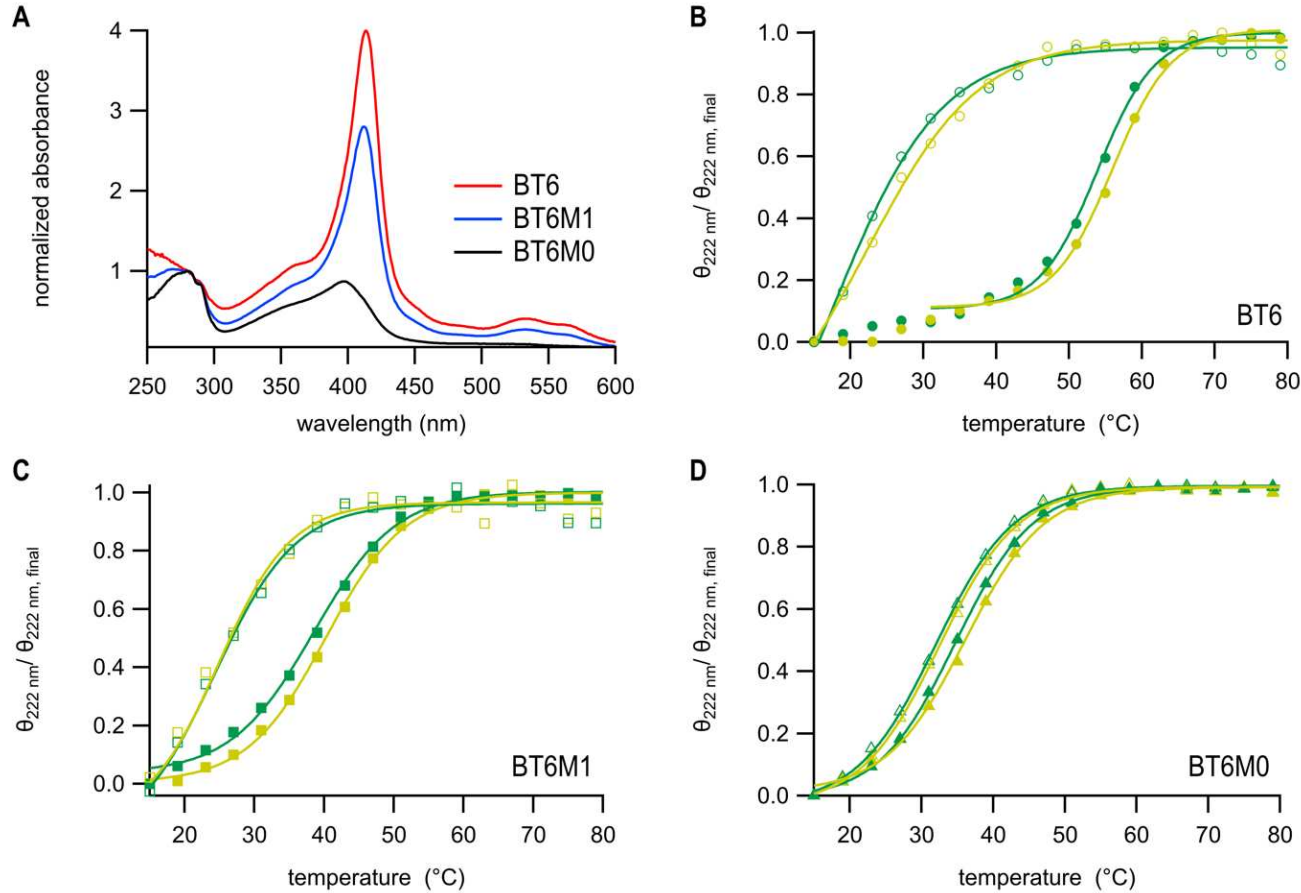
Figure 2. Spectroscopic analysis of apo- and heme-reconstituted maquette proteins. (A) UV-visible absorption spectra of heme-reconstituted maquette samples normalized to absorbance at 280 nm. **(B) – (D)** Normalized melting (yellow) and refolding (green) of secondary structure measured as the reduction in circular dichroism ellipticity at 222 nm across a temperature gradient in the absence (open markers) or presence (solid markers) of heme. Elipticity was recorded ever 1 °C, but only ever 3 data points are shown for clarity. Lines are theoretical and described by a Boltzmann distribution (see Materials and methods). Melting (Tm) and refolding (Tf) temperatures are reported in Table 2.
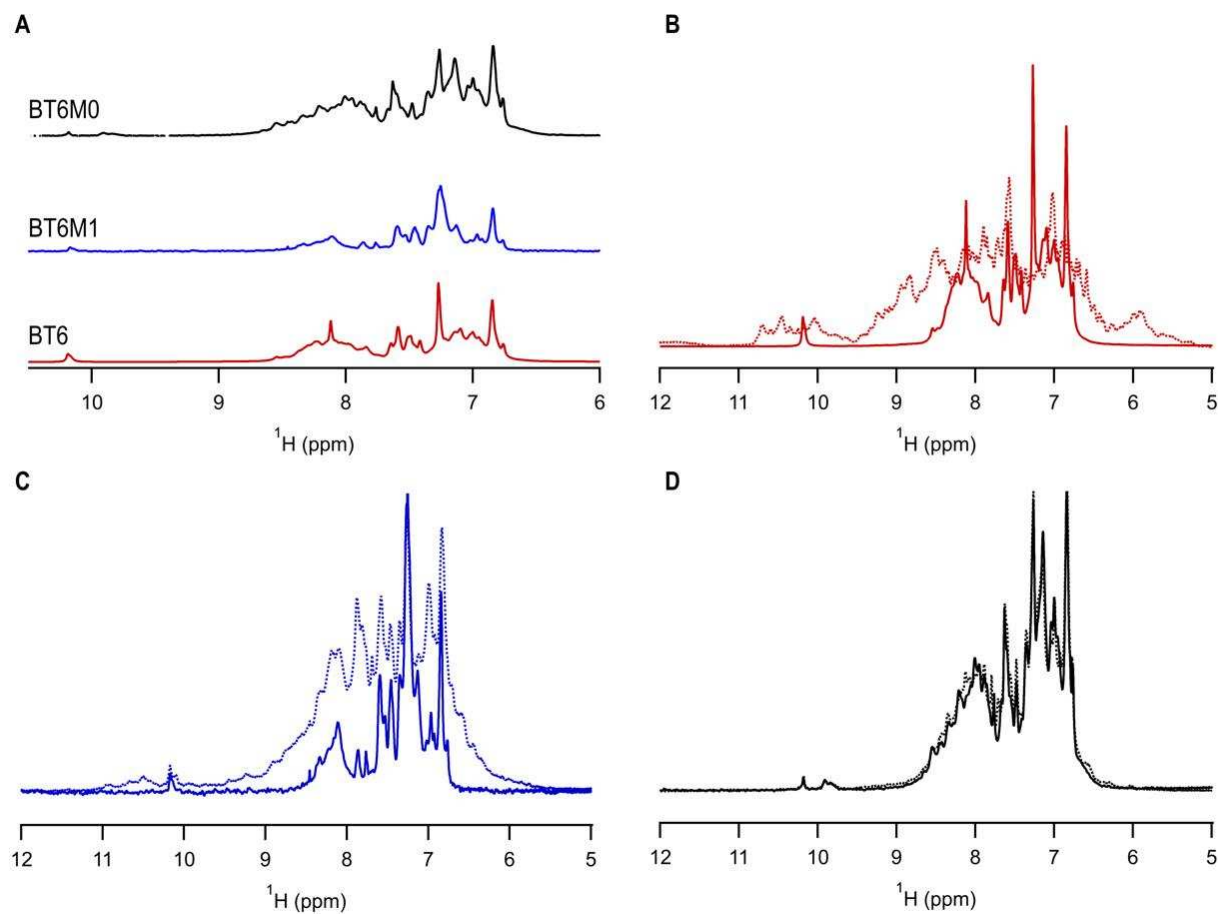
Figure 3. Proton NMR resonances of maquettes with and without heme cofactor. The amide proton region **(A)** for apo-BT6 (red), apo-BT6M1 (blue), and apo-BT6M0 (black), and comparisons in the absence (solid line) and presence (dashed line) of heme for BT6 **(B)**, BT6M1 **(C)** and BT6M0 **(D)**. For BT6 and BT6M1, the presence of heme induces changes in the amide proton resonance dispersion, whereas for BT6M0 negligible changes are observed. These chemical shift changes are consistent with a heme binding event coupled with a change in protein conformation
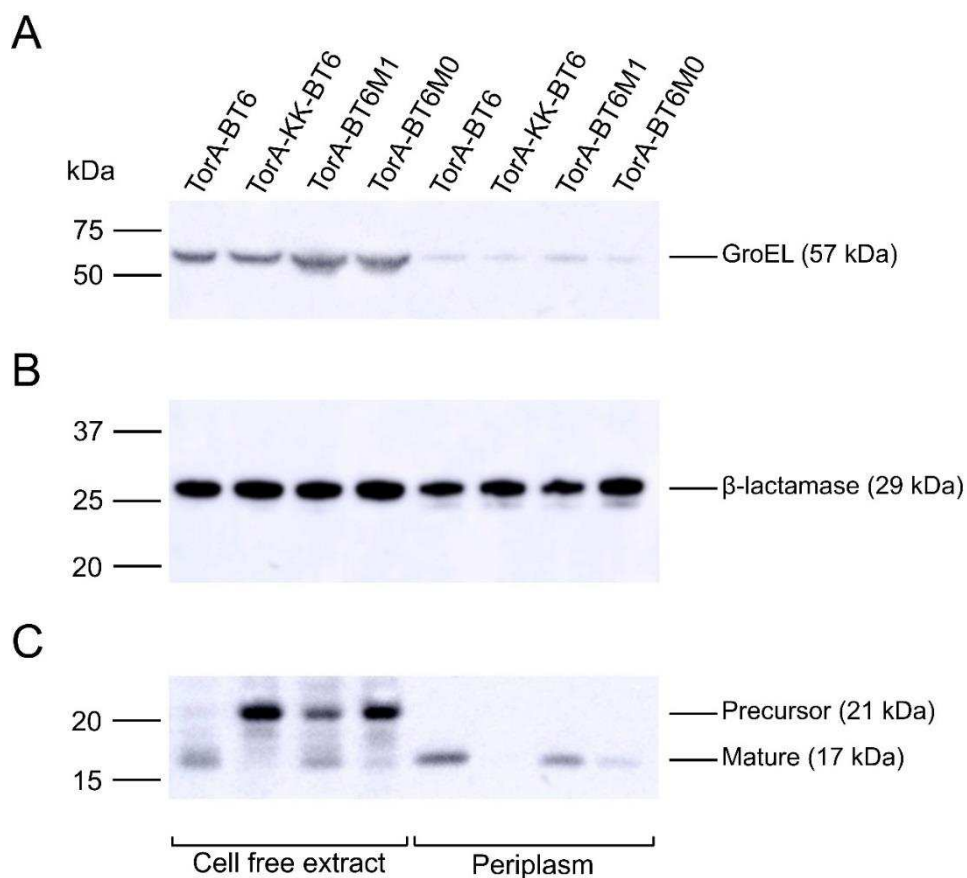
Figure 4. Differential export of BT6 variants by the Tat system determined by immunoblotting. (A) GroEL was used as a cytoplasmic marker to confirm only very minor cytoplasmic contamination of periplasmic preparations. **(B)** β-lactamase was used to confirm equal loadings of cell free extracts or periplasm samples. **(C)** The C-terminal His-tag on the BT6 maquettes was used to determine the degree of Tat-dependent periplasmic localization and the level of the unprocessed precursor proteins in cell free extracts. For all panels the positions of molecular weight markers and the expected size of proteins are indicated alongside the blots. Each blot is representative of at least 3 independent experiments.
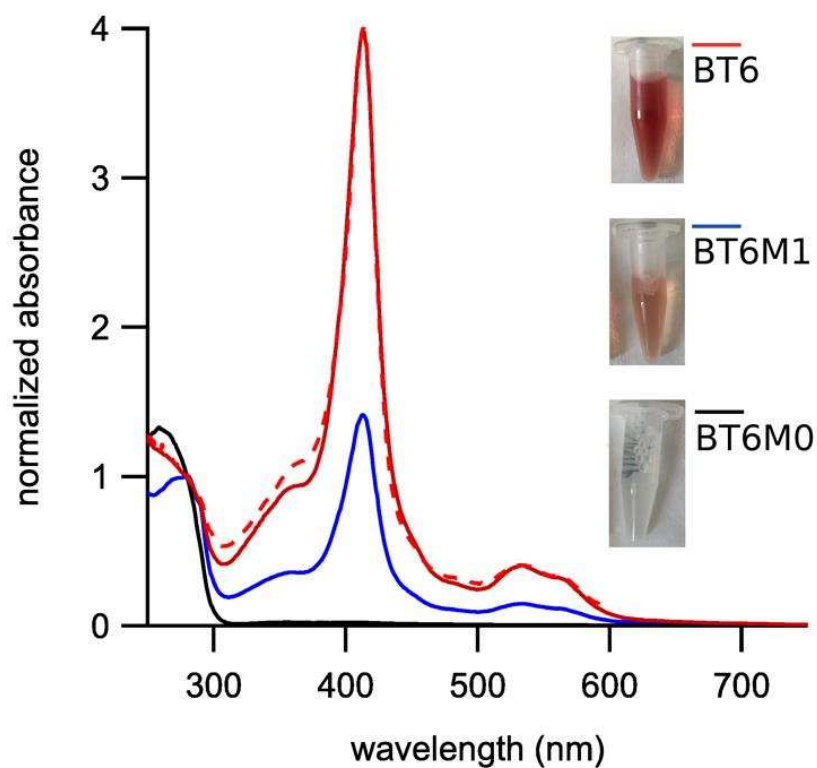
Figure 5. Spectroscopic analysis of purified maquette proteins. UV-visible absorbance spectra of maquettes purified from the periplasm of E. coli (BT6, red line; BT6M1, blue line; BT6M0, black line) compared to the BT6 reconstituted with hemin in vitro (red dashed line). Spectra are normalized to absorbance at 280 nm. Inset panels show the purified proteins. See Figure S7 for SDS-PAGE analysis of the purified maquettes.