Characterization of a pre-export enzyme–chaperone complex on the twin-arginine transport pathway

Jennifer M. DOW*, Frank GABEL†, Frank SARGENT*† and Tracy PALMER*†

*Division of Molecular Microbiology, College of Life Sciences, University of Dundee, Dundee DD1 5EH, U.K., and †Institut de Biologie Structurale, Commissariat à l’Energie Atomique, CNRS, Université Joseph Fourier, 41 rue Jules Horowitz, 38027 Grenoble, France

INTRODUCTION

The Tat (twin-arginine translocation) system is a protein targeting pathway utilized by prokaryotes and chloroplasts. Tat substrates are produced with distinctive N-terminal signal peptides and are translocated as fully folded proteins. In Escherichia coli, Tat-dependent proteins often contain redox cofactors that must be loaded before translocation. Trimethylamine N-oxide reductase (TorA) is a model bacterial Tat substrate and is a molybdopterin cofactor-dependent enzyme. Co-ordination of cofactor loading and translocation of TorA is directed by the TorD protein, which is a cytoplasmic chaperone known to interact physically with the TorA signal peptide. In the present study, a pre-export TorAD complex has been characterized using biochemical and biophysical techniques, including SAXS (small-angle X-ray scattering). A stable, cofactor-free TorAD complex was isolated, which revealed a 1:1 binding stoichiometry. Surprisingly, a TorAD complex with similar architecture can be isolated in the complete absence of the 39-residue TorA signal peptide. The present study demonstrates that two high-affinity binding sites for TorD are present on TorA, and that a single TorD protein binds both of those simultaneously. Further characterization suggested that the C-terminal ‘Domain IV’ of TorA remained solvent-exposed in the cofactor-free pre-export TorAD complex. It is possible that correct folding of Domain IV upon cofactor loading is the trigger for TorD release and subsequent export of TorA.

Key words: bacterial protein transport, Escherichia coli, molybdoenzyme, twin-arginine translocation protein transport pathway (Tat protein transport pathway), twin-arginine translocation proofreading chaperone (Tat proofreading chaperone).

The Tat (twin-arginine translocation) protein transport pathway exports proteins across the cytoplasmic membranes of bacteria and archaea, and the thylakoid membranes of plant chloroplasts [1,2]. Protein substrates of the Tat pathway are folded before translocation and are targeted to the Tat machinery by cleavable N-terminal signal peptides containing an almost invariant pair of arginine residues [3]. Transport by the Tat pathway is energized solely by the protonmotive force [1,2].

The model bacterium Escherichia coli K-12 produces 28 proteins bearing twin-arginine signal peptides. Approximately two-thirds of these proteins are known or predicted to contain redox cofactors and some of them also form complexes with partner subunits that lack a signal peptide before export [4]. The complexity of these Tat substrates necessitates a pre-export ‘proofreading’ process to ensure that only folded and assembled proteins are presented for translocation. One hypothesis involves dedicated chaperone proteins that interact directly with twin-arginine signal peptides to prevent premature targeting of the substrate before cofactor insertion and binding of any protein partners [5].

The E. coli TMAO (trimethylamine N-oxide) reductase TorA is a soluble periplasmic enzyme containing the MGD (molybdopterin guanine dinucleotide) cofactor at its active site [6]. TorA is a well-characterized substrate of the Tat pathway; indeed, transport of TorA to the periplasm was found to be blocked in a mutant strain that was unable to synthesize MGD, which led to the original hypothesis that a transport pathway for folded proteins may exist in bacteria [7]. The TorA protein is encoded by the torCAD operon (Figure 1A) where TorC is a haem-containing quinol oxidase and TorD is a cytoplasmic protein [8]. Cellular levels of enzymatically active TorA are significantly reduced in a strain lacking the torD gene [9], and the TorD protein was found subsequently to interact with an unfolded (heat-denatured) form of the mature TorA enzyme [9]. This initial work therefore established that there was a clear binding site for TorD within the mature region of the protein and suggested a role for TorD in loading the MGD cofactor into the TorA apoenzyme [9,10]. In addition to this, a subsequent genetic screen identified TorD as a binding partner for the TorA twin-arginine signal peptide [11]. This was confirmed in vitro when recombinant TorD was shown to bind directly to the isolated TorA signal peptide with a 1:1 stoichiometry and an apparent $K_D$ of $\sim 59$ nM [12,13]. The binding epitope for TorD has been mapped to the C-terminal end of the TorA signal peptide, close to its junction with the mature portion of the enzyme [12,14]. Interestingly, binding of the TorA signal peptide by TorD seems to be an activity not strictly connected with cofactor loading, since co-expression of torD can enhance the export of synthetic signal peptide fusion proteins on the Tat pathway [11,15]. The TorD protein itself is 22.5 kDa and is known to exist in equilibrium between monomeric and dimeric forms [16–18], have a low level of GTPase activity associated with the dimeric form [18], and have the ability to bind the MGD cofactor in vitro [19]. However, although it is clear that the monomeric form of TorD can bind the isolated TorA signal peptide [12,13], the physiological role and significance of TorD dimerization is not clear.

Current models for TorA biosynthesis assume that TorD interacts with the TorA precursor by binding at two distinct sites: one being within the twin-arginine signal peptide with the other, uncharacterized, site lying elsewhere in the mature portion of the TorA protein.
of the protein [5,20]. These two binding events are thought to
not only delay the Tat transport event, but also actively facilitate
efficient insertion of the MGD cofactor into the apoenzyme. This
pre-export ‘Tat proofreading’ process would be completed by the
cellular release of the bound chaperones and subsequent
translocation of the folded active enzyme to the periplasm.
However, it is not known whether each of the TorD-binding
sites on the TorA apoenzyme are bound by separate TorD monomers
or a single TorD dimer, or whether the TorD monomer itself
contains two distinct TorA-binding sites. Moreover, the degree of
folding of the TorA precursor when bound by TorD has not been
explored, and it remains a mystery how release of the chaperones is
ultimately triggered.

In the present study, a stable and cofactor-free complex of
TorD with the TorA apoenzyme has been isolated, both with and
without the twin-arginine signal peptide. Using a combination of
biophysical and biochemical approaches, we have demonstrated
that a single TorD monomer binds directly to a single TorA apoprotein regardless of the presence of the signal peptide. This
suggests that a single TorD monomer recognizes two binding sites
on the TorA polypeptide and that these epitopes are nearby or
overlapping. Finally, we have demonstrated that binding of TorA
by TorD changes the conformation of apo-TorA in relation to the
mature active enzyme by inducing exposure of the C-terminal
domain of the TorA protein.

**EXPERIMENTAL**

**Plasmid and strain construction**

The torAD genes form a transcription unit on the E. coli chromosome with overlapping stop and start codons. To
overproduce TorA and TorDHis, this natural transgenic coupling
was retained. DNA encoding SPA (sequential peptide affinity)-
tagged TorD along with TorA was amplified from the chromosome
of E. coli strain DY330::torAhis [25] by using oligonucleotides
TorAMfelup (5’-GCGCCAATTTGCGATAAGAAAGGAAAGA-
AATAATG-3’) and SPAphidown (5’-GCGGAGCTGGCTTACAT-
GTCATGCATCC-3’). The resulting PCR product was
digested with MfeI and SphI. To replace DNA encoding the
SPA-tagged variant of TorD with a His-tagged variant, the
pQE80torADSPA plasmid was digested with NheI which
removes DNA covering the 3’-end of torD, the SPA-tag-coding
sequence and a small region of the vector. Plasmid pQE60torD
[13] which codes for a C-terminally His-tagged TorD was also
digested with NheI and the fragment covering the 3’-end of
torD along with the His-tag-coding sequence was ligated into the
digested pQE80torADSPA vector. Resultant clones were analysed
by sequencing to ensure that the NheI fragment had been cloned
into the correct orientation, and the construct was designated
pQE80torADhis (Figure 1B).

To overproduce TorDHis along with TorA lacking its N-
terminal signal peptide, the encoding DNA was amplified using
oligonucleotides TorAMfeldelSS (5’-GGCCCTATGTGGCGATAA-
GAAAGGAAAGAATAATTGCGCAAGCCGCGACTGAGG-
CTGTCATCTCG-3’) and QERreverse (5’-GTTCTGAGGTCTATT-
GCTGGAT-3’), with pQE80 TorADhis as a template. The resultant
PCR product was digested with MfeI and HindIII and
cloned into pQE80 that had been pre-digested with EcoRI and
HindIII, to give construct pQE80TorAdelSShis (Figure 1C).

To produce a C-terminally truncated TorA along with TorDHis,
a fragment of torA terminating at codon 675 and followed by a
stop codon was amplified with oligonucleotides TorAfor
(5’-GACCTCAATCCTGGCTTTGCG-3’) and TorArev
(5’-GCTATCGGAAGACGTGAGG-3’) which codes for a C-terminally His-tagged TorD was also
constructed as follows. A 582 bp fragment of
torA immediately before the ribosome-binding site and a fragment of the
torD gene was amplified with oligonucleotides TorAHRS
(5’-GGGACGATTGTCGTTTAAATTACGAGG-3’) and TorDrev
(5’-GTCATCGCGCAGGACG-3’) and digested with NruI and SacII. The resultant clone was
digested with MrEl and HindIII and cloned into similarly digested pFA210 [23].
DNA covering the torD ribosome-binding site and approximately
600 nucleotides of downstream DNA was amplified with
oligonucleotides TorAHIS1 (5’-CGCCCATGCGAGGGATTAAATT-
ACGTGCGCATGAC-3’) and TorAHIS2 (5’-CGCCGAGTG-
CTGAGGTTCAATGGAGG-3’), digested with NcoI and BglII and cloned into similarly digested
pFA210. DNA covering the torD ribosome-binding site and approximately
600 nucleotides of downstream DNA was amplified with
oligonucleotides TorAHIS3 (5’-CGCCGACGTTTCTTGCAGC-
GTGCGAAGGATTAAATTACGAGG-3’) and TorAHIS4 (5’-GTCATCGCGC-
AGCAGTTCATGCGTTTCTTGCAGC-3’), digested with XbaI and HindIII cloned into the above construct. DNA covering the
torD allele was subsequently excised by digestion with
XbaI and KpnI and cloned into similarly digested pMAK705
[24]. The torA allele was then moved on to the chromosome of strain DSS401 (as MC4100, ΔdmaABC::kanR) [25] to give strain
FHTh007.

All constructs were fully sequenced to ensure that no mistakes
had been introduced inadvertently during the cloning procedure.

**TMAO reductase assays**

TMAO reductase assays were performed using strain GB426 (as
MC4100, ΔtorCAD::Apr8, ΔdmaABC::kanR [22]), which
contained either pQE80, pQE80torADhis, pQE80torAdelSShis

© The Authors. Journal compilation © 2013 Biochemical Society
or pQE80TorAtruncDhis. Cells were cultured anaerobically overnight in LB (Luria–Bertani) medium supplemented with 0.5% glycerol and 0.4% TMAO, along with the antibiotics apramycin (50 μg/ml), kanamycin (25 μg/ml) and ampicillin (100 μg/ml). Cells were harvested, washed and lysed by passage through a French press pressure cell as described previously [26]. After removal of unbroken cells and debris by a short centrifugation step, the resultant crude cell extract was assayed for TMAO reductase activity essentially as described previously [6,27]. Protein concentration was determined using the Lowry assay [28].

Preparation of recombinant proteins

All plasmid-encoded TorA/TorDHis proteins were overproduced and purified in an identical manner. A single colony of freshly transformed E. coli BL21(DE3) [F-ompT hsdS (rB- mB-) gal dcm (DE3)] was cultured overnight in 5 ml of LB medium supplemented with 100 μg/ml ampicillin. This provided the inoculum for a 500 ml culture, which was grown aerobically at 37°C for 5 ml pre-culture, bottles were incubated for 16 h without shaking before induction of recombinant protein production by addition of 2 mM IPTG (isopropyl β-D-thiogalactopyranoside), followed by a temperature shift to 18°C. After 16 h, cells were harvested and resuspended (10 ml per g of cells) in buffer A [50 mM Tris/HCl (pH 7.5), 200 mM KCl, 1 mM DTT (dithiothreitol) and 25 mM imidazole]. Protease inhibitor [EDTA-free CompleteTM protease inhibitor cocktail (Roche)], lysozyme and DNase were added to the resuspension and cells were lysed using an Emulsiflex C3 high-pressure homogenizer. Cell debris was removed by centrifugation at 18 500 × g for 15 min, followed by removal of membranes by ultracentrifugation at 45 000 rev./min for 1 h using a Beckman Ti70 rotor. All centrifugation steps were conducted at 4°C. The supernatant obtained following ultracentrifugation was filtered through a 0.22-μm-pore-size membrane filter (Millipore), lysozyme and DNase were added to a concentration of 10 ml/g of cells in 50 mM Tris/HCl, 200 mM KCl and 1 mM DTT; 100 μl sample and buffer volumes were loaded in a flowthrough quartz capillary cell at 25°C. The sample volume exposed to the X-ray beam was approximately 10 μl. Samples were checked for radiation damage by using ten successive exposure times of 10 s each. Final exposure time was 100 s for all samples/buffers. The two-dimensional diffraction patterns were normalized to an absolute scale and azimuthally averaged to obtain the intensity profiles I(Q), within BSxCuBE (ESRF beamline data collection software). Solvent contributions (buffer backgrounds collected before and after every protein sample) were averaged and subtracted from the associated protein sample using the program PRIMUS [29].

CD analysis

CD analysis of protein samples was provided as a service by the Scottish Circular Dichroism Facility, University of Glasgow. Spectra of protein solutions were obtained using a JASCO J-810 spectropolarimeter. Far-UV CD measurements (185–260 nm) were collected in quartz cells of 0.02 cm pathlength at 25°C with a scan speed of 10 nm/min, bandwidth of 1 nm, response of 2 s and data pitch of 0.2 nm. Before collection of CD data, proteins were buffer-exchanged into CD-appropriate buffer (50 mM Tris/HCl, pH 7.5, 200 mM K2SO4, and 1 mM DTT). Protein concentrations were estimated using a NanoDrop spectrophotometer before the experiment. Near-UV CD measurements (250–320 nm) were obtained in a 0.2-cm-pathlength quartz cuvette using the following parameters: temperature, 25°C; scan speed, 10 nm/min; bandwidth, 1 nm; response, 2.0 s; and data pitch, 0.2 nm. Protein was analysed in appropriate buffer (50 mM Tris/HCl, pH 7.5, 200 mM KCl and 1 mM DTT). The spectra were corrected for protein concentration and cell pathlength before being analysed by DichroWeb (http://dichroweb.cryst.bbk.ac.uk), an online server which hosts the various algorithms used to estimate protein secondary structures.

Metal analysis

Metal content was analysed by ICP-AES (inductively coupled plasma–atomic emission spectrometry)/ICP-MS (inductively coupled plasma–MS) and was provided as a service by the School of Chemistry at the University of Edinburgh.

Limited trypsinolysis of TorA–TorDHis and TorAaspx–TorDHis complexes

Protein samples (1 mg/ml in 50 mM Tris/HCl, pH 7.5, 200 mM KCl and 1 mM DTT; 100 μl final volume) were incubated with 1% (w/v) trypsin (porcine pancreas, proteomics grade, Sigma–Aldrich) at 37°C. Aliquots of 5 μl were withdrawn at various intervals and the reaction stopped by addition of 20 μl of Laemmli buffer (Bio-Rad Laboratories) followed by immediate boiling for 10 min. Once the time course was completed, samples were analysed by SDS/PAGE (12% gels). Proteins were identified by tryptic peptide mass fingerprinting of gel slices performed by the FingerPrints Proteomic Service (College of Life Sciences, University of Dundee).
Figure 2  Isolation of a TorA–TorDHis and TorA_{AS}–TorDHis complex

(A and C) TorD_{AS}-containing fractions after metal chelate chromatography of cell extracts overproducing (A) TorA and TorD_{His} or (C) TorA_{AS} and TorD_{His} were pooled, concentrated and applied to a HiLoad 16/60 Superdex 200 Prep Grade size-exclusion column. Eluted protein was monitored by measuring absorbance at 280 nm. The column was calibrated with the standard proteins ribonuclease (14 kDa), carbonic anhydrase (29 kDa), ovomucoid (44 kDa), conalbumin (75 kDa), aldolase (158 kDa) and ferritin (440 kDa), and the linear regression analysis is shown as inset boxes. R² = 0.9976, y = −23.5x + 126.83. MW, molecular mass. (B and D) SDS/PAGE analysis (12 % gels) of (B) the concentrated fractions after metal chelate chromatography (Ni pool), and the non-concentrated and concentrated peak fractions from SEC of the TorA–TorD_{His} complex and (D) the concentrated peak fraction from SEC of the TorA_{AS}–TorD_{His} complex. Molecular masses are indicated in kDa.

Analysis of TorA–TorD_{His} complexes by SEC (size-exclusion chromatography)–MALLS (multi-angle laser light scattering)

Size estimates of TorA–TorD_{His} complexes were obtained by SEC–MALLS, which was provided as a service by the University of York Bioscience Technology Facility. Measurements were conducted on a system comprising a Wyatt HELEOS-II multi-angle light-scattering detector and a Wyatt rEX refractive index detector linked to a Shimadzu HPLC system (SPD-20A UV detector, LC20-AD isocratic pump system, DGU-20A3 degasser and SIL-20A autosampler). Work was conducted at room temperature (20 ± 2°C). Solvent was filtered through a 0.2-μm-pore-size filter before use and a further 0.1-μm-pore-size filter was present in the flow path. Filtered (0.2 μm pore size) TorA–TorD_{His} samples (100 μl, corresponding to 100 μg of protein) were applied to a Superdex 200 10/300 GL column (GE Healthcare) equilibrated with 50 mM Tris/HCl (pH 7.5), 200 mM KCl and 1 mM DTT at 0.5 ml/min; Shimadzu LC Solutions software was used to control the HPLC, and Astra V software was used for the HELEOS-II and rEX detectors. BSA was used for molecular mass calibration. The Astra data collection was 1 min shorter than the LC solutions run to maintain synchronization. Blank buffer injections were used as appropriate to check for carry-over between sample runs. Data were analysed using the Astra V software. MWs were estimated using the Zimm fit method [30] with degree 1. A value of 0.19 was used for protein refractive index increment (dn/dc).

RESULTS

Isolation of a TorA–TorD_{His} complex

Isolation of a complex between untagged TorA and affinity-tagged TorD was originally demonstrated on a small scale using a C-terminally SPA-tagged form of TorD [22]. In order to scale up the co-expression system to allow detailed analysis of the complex, the original torAD_{SPA} allele was cloned into expression vector pQE80 and the DNA coding for the SPA tag at the C-terminus of TorD replaced with a hexahistidine tag (Figure 1B). The soluble cell extract containing overproduced TorA and TorD_{His} was subjected to IMAC and fractions containing proteins of the expected size of TorA and TorD_{His} were identified in a single peak at 110 mM imidazole. The IMAC samples were pooled, concentrated and subjected to SEC using a HiLoad 16/60 Superdex 200 Prep Grade column. As shown in Figure 2(A), protein was eluted at 78.5 ml, corresponding to an estimated molecular mass of 114 kDa. Analysis of the peak fraction

© The Authors Journal compilation © 2013 Biochemical Society
by SDS/PAGE revealed the presence of two major proteins (Figure 2B), which were confirmed by tryptic peptide mass fingerprinting to be TorA and TorD\textsubscript{His}. The overall protein yield from this process was approximately 1.5 mg/g of bacterial cells.

It was notable during the purification of TorAD\textsubscript{His}, particularly following concentration of the samples, that more than one electrophoretic form of TorA was evident by SDS/PAGE. A low-percentage acrylamide gel was therefore used to separate these forms, which were excised and analysed by tryptic peptide mass fingerprinting. The analysis showed that, whereas the majority of the TorA was present as full-length intact polypeptide, there was a proportion of TorA that had undergone limited proteolysis and was lacking up to 150 amino acids from its C-terminal end (Supplementary Figure S1 at http://www.biochemj.org/bj/452/bj4520057add.htm). In addition, in some instances, up to 12 amino acid residues could be lost from the N-terminus of TorA. It is concluded that, whereas a complex of TorAD\textsubscript{His} is stable to purification, the extremities of TorA can be subjected to variable levels of proteolysis during isolation.

Removal of the entire Tat signal peptide from the TorA protein (TorA \(\Delta 2–39\)) still allows isolation of a stable complex with TorD\textsubscript{His}

As well as signal peptide binding, previous biochemical and genetic experiments pointed to a binding site for TorD within the mature region of the TorA enzyme (e.g. [9,10]). To investigate this further, a vector was designed that would co-overproduce a version of TorA lacking its entire 39-residue Tat signal peptide (TorA\textsubscript{ASp}) with TorD\textsubscript{His} (Figure 1C). When TorD\textsubscript{His} was isolated from the soluble cell fraction by IMAC, a protein of the expected size of TorA\textsubscript{ASp} was also co-purified. The TorD\textsubscript{His} containing fractions were pooled, concentrated and applied to a Superdex 200 size-exclusion column. The TorD\textsubscript{His} eluted at 78.8 ml, corresponding to an estimated molecular mass of 110 kDa (Figure 2C) and analysis of the concentrated peak fraction by SDS/PAGE revealed the presence of two major proteins (Figure 2D), which were confirmed by tryptic peptide mass fingerprinting to be TorA\textsubscript{ASp} and TorD\textsubscript{His}. Thus it is clear that a stable complex of TorD\textsubscript{His} and TorA can be isolated in the absence of the twin-arginine signal peptide.

TorD\textsubscript{His} binds the full-length TorA precursor, and the TorA\textsubscript{ASp} species lacking its entire signal peptide (\(\Delta 2–39\)), at a 1:1 stoichiometry

SEC indicates that the complex of TorD\textsubscript{His} with TorA elutes at a similar molecular mass regardless of whether the signal peptide was present on TorA (Figure 2). A more accurate technique is SEC–MALLS, which uses HPLC-linked SEC coupled with static laser light scattering to provide a direct measure of molecular mass. Using this technique, the TorA–TorD\textsubscript{His} complex was observed to exhibit a mass range of 100–120 kDa and a peak mass of 114.7 kDa (Table 1), which is very close to the calculated mass (117.9 kDa) for a 1:1 complex of the two proteins (Table 1). The TorA\textsubscript{ASp}–TorD\textsubscript{His} complex exhibited a mass range of 95–115 kDa and a peak mass of 108.2 kDa (Table 1), which is close to the calculated theoretical mass (113.8 kDa) for a 1:1 complex of the two proteins (Table 1). The relatively wide mass range seen for both samples reflects the observed proteolytic events leading to heterogeneity of the TorA sample observed by SDS/PAGE (Figure 2). Overall, the SEC–MALLS analyses point to complexes containing one TorA protein and one TorD protein, and suggest strongly that TorD can remain stably bound to TorA in the absence of the twin-arginine signal peptide.

Table 1 The apparent molecular masses of TorAD complexes

<table>
<thead>
<tr>
<th>Sample</th>
<th>Major peak elution range (min)</th>
<th>Molecular mass at peak (kDa)</th>
<th>Range across peak (kDa)</th>
<th>Predicted molecular mass of a 1:1 complex (kDa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TorA–TorD\textsubscript{His}</td>
<td>26.4–30.4</td>
<td>114.7</td>
<td>100–120</td>
<td>117.9</td>
</tr>
<tr>
<td>TorA\textsubscript{ASp}–TorD\textsubscript{His}</td>
<td>26.3–30.0</td>
<td>108.2</td>
<td>95–115</td>
<td>113.8</td>
</tr>
<tr>
<td>TorA\textsubscript{ACT}–TorD\textsubscript{His}</td>
<td>27.6–30.2</td>
<td>99.7</td>
<td>97–107</td>
<td>99.2</td>
</tr>
</tbody>
</table>

The Escherichia coli TorAD complex

To gain more insight into the architecture of the TorAD complexes under investigation in the present study, limited proteolysis experiments were undertaken. First, the purified TorA–TorD\textsubscript{His} complex was incubated with a low concentration of trypsin over a time course of 60 min. As shown in Figure 3(A), TorD\textsubscript{His} was almost completely resistant to trypsinolysis under these conditions. In contrast, TorA was cleaved to yield a stable proteolytic fragment of approximately 70 kDa that was apparently resistant to further digestion by trypsin. Peptide mass fingerprinting of the stable 70 kDa TorA proteolysis product revealed that the initial 675 amino acid residues from the N-terminus of the protein were intact and that the trypsin treatment had therefore removed a short section of polypeptide from the C-terminal end of TorA (Supplementary Figure S1).

© The Authors Journal compilation © 2013 Biochemical Society
Next, the TorA<sub>ASP</sub>–TorD<sup>His</sup> complex, where TorA had already had its 39-residue signal peptide deliberately removed, was incubated with trypsin in a similar manner. As shown in Figure 3(B), TorD<sup>His</sup> was almost completely resistant to protease digestion, whereas the TorA<sub>ASP</sub> was degraded to a smaller stable product. Analysis of this protease-resistant fragment by tryptic peptide mass fingerprinting showed that it was cleaved at Lys<sup>675</sup>, which was the identical position where the full-length TorA precursor was cleaved. As a control, a C-terminally His-tagged, but otherwise native, form of TorA was purified from the periplasm of E. coli strain FTH007. The fully folded mature form of TorA was completely resistant to digestion with trypsin, even after a 16-h incubation period (Figure 3C). These results are consistent with the suggestion that the TorA–TorD<sup>His</sup> and TorA<sub>ASP</sub>–TorD<sup>His</sup> complexes have a similar overall architecture, both being susceptible to limited trypsinolysis in an identical manner. In both cases, a C-terminal stretch of immature TorA is obviously exposed to the aqueous phase in a conformation that is clearly not adopted by the mature fully assembled enzyme.

The C-terminal domain of TorA is not involved in stable TorAD complex formation

TorA is a member DMSO reductase family of molybdoenzymes. Proteins in this family adopt a fold that can be divided into four structurally distinct domains [31], and, of these, Domain IV is formed by the extreme C-terminal part of the protein and is the only domain in this fold that comprises a clear contiguous stretch of polypeptide chain (Supplementary Figure S2 at http://www.biochemj.org/bj/452/bj4520057add.htm). Although the crystal structure of the E. coli TorA protein has not been solved, structure predictions based on the Shewanella massilia TorA crystal structure [32] suggest the position of the trypsin cleavage site in the E. coli TorA–TorD<sup>His</sup> complexes lies within an unstructured loop region that connects Domain III to Domain IV (Supplementary Figure S2). To investigate this further, a construct was designed that would co-overproduce TorD<sup>His</sup> together with the first 675 amino acids of TorA, including the signal peptide (designated TorA<sub>ACT</sub>). First, the construct was expressed in a strain lacking endogenous TMAO reductase activity under anaerobic conditions, which would induce expression of the MGD biosynthetic machinery. Then a crude cell extract was prepared and the TMAO reductase activity assayed. The resultant data showed that the C-terminally truncated form of TorA was completely devoid of enzymatic activity (Supplementary Table S1 at http://www.biochemj.org/bj/452/bj4520057add.htm), establishing that this domain of the protein is absolutely critical for function. Next, TorD<sup>His</sup> was isolated from the soluble cell fraction by IMAC, which resulted in a protein of the expected size of TorA<sub>ACT</sub> co-purifying (Figure 4). The two proteins remained fully associated during SEC through a Superdex 200 column (Figure 4A), with the complex eluting with an estimated molecular mass of 100 kDa (Figure 4A). Analysis of the peak fractions by SDS/PAGE confirmed that both proteins were present in the sample, but this time very little degradation of the deliberately truncated TorA was apparent (Figure 4B). SEC–MALLS analysis of the purified complex (Table 1) confirmed that the TorA<sub>ACT</sub>–TorD<sup>His</sup> complex had a narrow mass range of 97–107 kDa, reflecting the relative homogeneity of the sample, and a peak mass of 99.7 kDa, which is close to the calculated mass (99.2 kDa) for a 1:1 complex of the two proteins (Table 1). It can be concluded that, although the C-terminal 173 amino acids are essential for the ultimate enzymatic activity of TorA, they do not influence the formation or stability of the pre-export complex between TorA and TorD.

The TorAD complexes are devoid of MGD, but retain most of their secondary structure

Whereas it has generally been assumed that a Tat substrate protein is maintained in an ‘unfolded conformation’ when in complex with its pre-export chaperone, the degree of folding of such Tat precursors has not been examined in detail. In the present study, recombinant TorD<sup>His</sup>, active TorA<sub>His</sub>, and the TorA–TorD<sup>His</sup> and TorA<sub>ASP</sub>–TorD<sup>His</sup> complexes were analysed by CD spectroscopy. The CD spectra recorded for each sample are shown in Figure 5 and the calculated percentages of helix, strand, turn and disorder are given in Supplementary Table S2 (at http://www.biochemj.org/bj/452/bj4520057add.htm). Consistent with the highly α-helical structure of TorD<sup>His</sup> [17], the CD spectrum of purified TorD<sup>His</sup> shows a high proportion of helix (Figure 5 and Supplementary Table S2). Enzymes of the DMSO reductase family are mixed α/β proteins [31–34] and this is reflected in the CD spectrum of purified TorA<sub>His</sub> (Figure 5 and Supplementary Table S2). Interestingly, CD analysis of the TorA–TorD<sup>His</sup> and TorA<sub>ASP</sub>–TorD<sup>His</sup> complexes showed that both had a significant degree of secondary structure and
The observed percentage of disorder was no higher than that of isolated fully assembled TorA<sub>His</sub>. Using the parameters determined for TorA<sub>His</sub> and TorD<sub>His</sub> (Supplementary Table S2), the predicted percentages of helix, strand, turn and disorder that would be expected for a simple 1:1 mixture of these two proteins can be calculated (Supplementary Table S2). It is notable that these predicted values are a close match to those experimentally determined for the purified TorA–TorD<sub>His</sub> and TorA<sub>/Delta1</sub>SP–TorD<sub>His</sub> complexes, with the values particularly close for the TorA–TorD<sub>His</sub> complex (Supplementary Table S2). Taken together, these results indicate that the TorA component of the TorA–TorD<sub>His</sub> complex has probably adopted its native secondary structure. Moreover, whereas the TorA<sub>His</sub> protein was isolated from an anaerobic culture and exhibited a straw colour in solution, the TorA–TorD<sub>His</sub> and TorA<sub>/Delta1</sub>SP–TorD<sub>His</sub> complexes examined were produced from cultures grown under aerobic conditions without added molybdate. As a result, they would be expected to be largely devoid of cofactor and indeed were colourless and devoid of enzymatic activity. Metal content in the complexes by ICP-MS showed that the molar ratio of molybdenum to TorA/D complex was less than 1:1000. Taken together with the trypsinolysis experiments (Figure 3), these data suggest that the cofactor-free TorD-bound species of the TorA apoenzyme adopts a fold very similar to that of its fully assembled counterpart, save for an exposed C-terminal Domain IV.

**SAXS analysis of full-length TorA–TorD<sub>His</sub> and signal peptide-less TorA<sub>/Delta1</sub>SP–TorD<sub>His</sub> complexes reveals shared structural features**

SAXS is a powerful technique that can be used to compare the overall shapes and sizes of proteins in solution. We analysed the TorA–TorD<sub>His</sub> and TorA<sub>/Delta1</sub>SP–TorD<sub>His</sub> complexes, together with isolated TorD<sub>His</sub>, since the high-resolution structure of TorD from *S. massilia* suggested that the protomer could adopt a highly extended conformation [17]. Representative SAXS scattering curves from each of the three samples are shown in Figure 6. It is clear that the scattering curves for the TorA–TorD<sub>His</sub> and TorA<sub>/Delta1</sub>SP–TorD<sub>His</sub> complexes are almost indistinguishable, indicating that they have an identical overall shape. An overview of the general parameters extracted from the SAXS data including the radii of gyration ($R_g$; obtained from the Guinier approximation [35]), the distance distribution function, $P(r)$ and the maximum particle size ($D_{max}$; obtained from the GNOM [36] analysis) are given in Table 2.

A low-resolution model of TorD<sub>His</sub>, obtained using the program DAMMIN [37], is shown in Figure 7(A). The overall shape of TorD<sub>His</sub> is compact and elliptical with dimensions of approximately 36 Å×50 Å×34 Å. From the crystal structures of *S. massilia* TorD [17] and the TorD family protein DmsD...
from E. coli [38–40], scattering envelopes can be calculated [41]. It is clear from Supplementary Figure S3 (http://www.biochemj.org/bj/452/bj4520057add.htm) that neither the S. massilia TorD domain-swapped dimer nor one of the protomers involved in forming that dimer would give scattering curves similar to that measured for TorD\textsubscript{His} in solution. By contrast, the scattering curve predicted from the X-ray structure of DmsD is very similar in shape, and the electron density of monomeric DmsD, which can be readily docked into the TorD\textsubscript{His} SAXS envelope (Figure 7B).

Since the scattering curves indicated that the TorA–TorD\textsubscript{His} complexes were indistinguishable, a single SAXS-derived shape corresponding to either the TorA–TorD\textsubscript{His} or the TorA\textsubscript{ASp}–TorD\textsubscript{His} complex is shown in Figures 7(C) and 7(D). The complex is highly elongated, with an overall length of approximately 120 Å. There is a large central area of density with one relatively large lobe of density protruding from one side of the central region and a smaller lobe of density protruding from the opposite side. The large central area has approximate dimensions of 70 Å × 74 Å × 44 Å and is a reasonably close match to the dimensions of the crystal structure of S. massilia TorA (Figures 7C and 7D).

**DISCUSSION**

The process of co-ordinated assembly (‘proofreading’ or ‘quality control’) is an important consideration in the biosynthesis of all complex cofactor-containing Tat-dependent enzymes [5]. The simplest model system used to study this process is the E. coli TMAO reductase, which has a single MGD cofactor and a single biosynthetic chaperone, TorD. Whereas early research has seen analyses of the TorD–signal peptide interaction in isolation [12,42], the present study has focused on the characterization of the entire pre-export TorAD complex.

Design of co-overproduction vectors established that untagged TorA formed a tight and stable complex with hexahistidine-tagged TorD. Surprisingly, however, despite the strong evidence for two separate TorD-binding sites on TorA [9,11], we have clearly demonstrated using the accurate SEC–MALLS technique in conjunction with other experiments that TorD binds TorA with a 1:1 stoichiometry. Even more surprisingly, given the well-characterized interaction between TorD and the TorA signal peptide [11–15,42], a truncated TorA enzyme completely lacking its entire 39-residue signal peptide retained the ability to interact in a stable manner with TorD. Whereas the present study clearly establishes that there is a high-affinity TorD-binding site within the mature portion of TorA, it also raises questions about the role and importance of the signal peptide interaction.

These findings can be interpreted in two possible ways. First, that the signal peptide-binding activity of TorD is of a much lower affinity than that of the second binding site within the mature TorA sequence, so much so that there is a strong bias towards the higher-affinity species during the purification protocol. Although this scenario is not impossible, recent studies place the apparent \(K_d\) for TorD binding to the isolated TorA signal peptide at 59–330 nM [12,42], which is indicative of very tight binding such that it is likely that a detectable portion of TorA would not fail to have two TorD molecules bound. The second possibility is that the TorD-binding site within the mature portion of TorA is very close to the signal peptide-binding epitope such that a single TorD protein could interact with both simultaneously. This hypothesis would require there to be two different TorA-binding sites on a single TorD protein. Although the location of the TorA signal peptide-binding site on TorD has not yet been reported, there is a larger body of evidence available on the location of the signal peptide-binding site on another member of the TorD family, DmsD. The E. coli DmsD protein was the first twin-arginine signal peptide-binding chaperone to be described [43]. It is essential for the assembly of molybdenum-dependent TMAO reductase and selenate reductase enzymes in E. coli and binds to the signal peptides of those enzymes with apparent \(K_d\) values in the range 10–100 nM [43–45]. Genetic, biochemical and structural analyses predict the signal peptide-binding site to occupy only one face of DmsD (primarily involving residues of the N-terminal helices) and so suggests that a significant proportion the DmsD protein is not involved in signal peptide recognition [39,45,46]. This leaves the possibility, at least, that a second protein- or peptide-recognition site may be present on this type of chaperone. Indeed, variants of E. coli TorD have been described (C79R and L83P).
that retain signal peptide-binding activity, but are impaired in their binding to mature TorA [10].

Further evidence for overlapping TorD-binding sites near the N-terminus of TorA come from analysis of the SAXS data obtained in the present study. To investigate the possible arrangement of TorA and TorD in the TorAD complex, rigid body modelling was undertaken using the SASREF program [47] and the structures of S. massilia TorA and E. coli DmsD as building blocks. To facilitate this analysis, TorA was split into two parts to give the proteolytically stable N-terminal part (Domains I–III) and the proteolytically sensitive Domain IV identified in the present study. A constraint was then applied that the last amino acid of N-terminal part should be within 10 Å of the first amino acid of Domain IV, in order to incorporate flexibility into the TorA component of the complex, again consistent with the protease accessibility results of the present study. The results from the rigid body modelling show a similar overall shape to that obtained during ab initio analysis (the fit of the model with the experimental data is shown in Supplementary Figure S4 at http://www.biochemj.org/bj/452/bj4520057add.htm). The rigid body model is shown docked into the ab initio SAXS-derived envelope of TorAD in Figure 8. When docked into the ab initio model, the density for TorD fits very well into the C-terminal Domain IV of TorA, whereas the remainder of the ab initio envelope is of an appropriate volume to accommodate TorA. The C-terminal Domain IV of TorA distal from TorD fits into the smaller protruding lobe, which is consistent with the biochemical data of the present study that Domain IV is not required for interaction of TorD with TorA. Bearing in mind that the SAXS scattering curves for both signal peptide-containing and signal peptide-free TorAD complexes are identical, this must be taken as good evidence that the binding epitope for TorD on the mature portion of TorA must lie near the N-terminus and thus also near the second epitope within the signal peptide.

The CD and SAXS analyses suggest that there is a very high degree of folding by the TorA protein even in the absence of the MGD cofactor. However, it is clear that the TorA–TorD<sup>ΔII</sup> and TorA<sub>ΔN</sub>–TorD<sup>ΔII</sup> complexes are not in the native conformation and that the C-terminal Domain IV of the TorA protein is exposed. This could suggest that the locking of Domain IV into position represents the final act in cofactor insertion. One possible role of TorD could be to prevent premature closing of the Domain IV ‘flap’ before MGD has bound, or, alternatively, that closing of Domain IV after cofactor loading is the trigger that releases TorD from the now mature enzyme. The molecular mechanism of how TorD senses the events surrounding cofactor loading remains to be unearthed.

Acknowledgements

We acknowledge Dr Bérengère Ize, Dr Grant Buchanan and Dr Dave Guymer for useful discussions and preliminary studies. We also thank Dr Cyril Dian for technical assistance at the ID14-3 BioSAXS beamline at the ESRF.

Funding

This work was funded via a Wellcome Trust Ph.D. Studentship Award to J.M.D. (grant number WT089629/2092) and a Royal Society Wolfson Research Merit Award to T.P.
16 Tranier, S., Mortier-Barriere, I., Ibert, M., Bricz, C., Iobbii-Nivioz, C., Mejean, V. and Samama, J. P. (2002) Characterization and multiple molecular forms of TorD from Shewanella massiliensis, the putative chaperone of the molybdoenzyme TorA. Protein Sci. 11, 2148–2157
SUPPLEMENTARY ONLINE DATA
Characterization of a pre-export enzyme–chaperone complex on the twin-arginine transport pathway

Jennifer M. DOW*, Frank GABEL†, Frank SARGENT*1 and Tracy PALMER*1

*Division of Molecular Microbiology, College of Life Sciences, University of Dundee, Dundee DD1 5EH, U.K., and †Institut de Biologie Structurale, Commissariat à l‘Energie Atomique, CNRS, Université Joseph Fourier, 41 rue Jules Horowitz, 38027 Grenoble, France

Figure S1  The E. coli TorA precursor primary sequence

The points of natural proteolysis, and deliberate truncation, of the TorA precursor are highlighted. The 39-residue N-terminal twin-arginine signal peptide is highlighted in red. The position of the engineered initiation site for the construct expressing the signal-less TorA protein is indicated by the black arrow. The location of the experimentally determined trypsin-cleavage site for TorA within the TorA–TorDHis complex is indicated by the red arrow. This is also the point of truncation for the TorA_light protein. A small fraction of the TorA sample is subject to degradation during the purification procedure. Following IMAC, a small fraction of TorA is found to be degraded from the C-terminus and loses the purple-coloured stretch of polypeptide. Following the subsequent SEC step, a small fraction of TorA within the TorA–TorDHis complex becomes proteolysed at the N-terminus and further at the C-terminus (blue). This shows that the extremities of TorA are not shielded or protected by the tightly bound TorDHis protein, are exposed to solvent and are therefore susceptible to proteolysis.

1 Correspondence may be addressed to either of these authors (email f.sargent@dundee.ac.uk or t.palmer@dundee.ac.uk).
Figure S2  X-ray structure of S. massilia TorA

The C-terminal Domain IV is shaded in blue and the approximate position of the trypsin-cleavage site determined for E. coli TorA in the TorA–TorDHis and TorA Δ1SP–TorDHis complexes is indicated. The image was created using PyMOL (http://www.pymol.org) with PDB code 1TM0.

Figure S3  Comparison of SAXS scattering curves

Theoretical curves were calculated from S. massilia TorD (PDB code 1N1C) monomer (blue line) and dimer (red line) and E. coli DmsD (PDB code 3CWO) monomer (black line) using CRYSOL and compared with experimental scattering curves of TorDHis (green dots).

Figure S4  SAXS scattering curve of the TorA–TorDHis complex

Fitted curve of TorADHis rigid body modelling conducted using the TorA structure (PDB code 1TM0) (split into two components; Domains I–III and Domain IV) and the DmsD structure (PDB code 3CWO), performed using SASREF. The fitted curve is shown in comparison with the E. coli TorADHis scattering curve.

Table S1  TMAO reductase activity in strains producing variants of TorA

Crude cell extracts were prepared from anaerobically grown cells before TMAO-dependent Benzyl Viologen-oxidation assays were performed. Results are means ± S.D. of three measurements.

<table>
<thead>
<tr>
<th>Strain and plasmid</th>
<th>Complex produced</th>
<th>TMAO reductase activity (μmol of Benzyl Viologen oxidized/min per mg of protein)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GB426 (tor−, dms−) + pQE80 (control)</td>
<td>–</td>
<td>0.15 ± 0.01</td>
</tr>
<tr>
<td>GB426 + pQE80TorADHis</td>
<td>TorA/TorDHis</td>
<td>28.8 ± 4.0</td>
</tr>
<tr>
<td>GB426 + pQE80TorAdelSSDHis</td>
<td>TorA Δ1SP/TorDHis</td>
<td>14.8 ± 1.6</td>
</tr>
<tr>
<td>GB426 + pQE80TorAtruncDHis</td>
<td>TorA ΔCT/TorDHis</td>
<td>0.12 ± 0.01</td>
</tr>
</tbody>
</table>

Table S2  Characterization of the TorA and TorD proteins by CD

Proportions of secondary-structural elements were calculated from CD analysis for each of the indicated samples.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Total helix (%)</th>
<th>Total strand (%)</th>
<th>Turns (%)</th>
<th>Unordered (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TorADHis (experimental)</td>
<td>26</td>
<td>25</td>
<td>20</td>
<td>29</td>
</tr>
<tr>
<td>TorDHis (experimental)</td>
<td>56</td>
<td>7</td>
<td>17</td>
<td>20</td>
</tr>
<tr>
<td>TorADHis and TorDHis combined (theoretical)</td>
<td>31</td>
<td>22</td>
<td>19</td>
<td>27</td>
</tr>
<tr>
<td>TorA–TorDHis (experimental)</td>
<td>28</td>
<td>22</td>
<td>22</td>
<td>29</td>
</tr>
<tr>
<td>TorA ΔSP–TorDHis (experimental)</td>
<td>20</td>
<td>29</td>
<td>22</td>
<td>30</td>
</tr>
</tbody>
</table>