Structure analysis of the protein translocating channel TatA in membranes using a multi-construct approach

Christian Lange [a], Sonja D. Müller [a], Torsten H. Walther [b], Jochen Bürck [a], Anne S. Ulrich [a,b,*]

[a] *Forschungszentrum Karlsruhe, POB 3640, D-76021 Karlsruhe, Germany*
[b] *Universität Karlsruhe (TH), Institut für Organische Chemie, Lehrstuhl für Biochemie, Fritz-Haber-Weg 6, D-76131 Karlsruhe, Germany*

Abstract

The *twin-arginine-t*ranslocase (Tat) can transport proteins in their folded state across bacterial or thylakoid membranes. In *Bacillus subtilis* the Tat-machinery consists of only two integral (inner) membrane proteins, TatA and TatC. Multiple copies of TatA are supposed to form the transmembrane channel, but little structural data is available on this 70-residue component. We used a multi-construct approach for expressing several characteristic fragments of TatA$_d$, to determine their individual structures and to cross-validate them comprehensively within the architecture of the full-length protein. Here, we report the design, high-yield expression, detergent-aided purification and lipid-reconstitution of five constructs of TatA$_d$, overcoming difficulties associated with the very different hydrophobicities and sizes of these membrane protein fragments. Circular dichroism (CD) and oriented CD (OCD) were used to determine their respective conformations and alignments in suitable, negatively charged phospholipid bilayers. CD spectroscopy showed an N-terminal α-helix, a central helical stretch, and an unstructured C-terminus, thus proving the existence of these secondary structures in TatA$_d$ for the first time. The OCD spectra demonstrated a transmembrane orientation of the N-terminal α-helix and a surface alignment of the central amphiphilic helix in lipid bilayers, thus supporting the postulated topology model and function of TatA as a transmembrane channel.
© 2007 Elsevier B.V. All rights reserved.

Keywords: Twin-arginine-translocation; TatA; Membrane protein; Circular dichroism; Oriented circular dichroism; *Bacillus subtilis*

1. Introduction

The translocation of proteins across cellular membranes follows distinct routes. Besides the well-known Sec-pathway for translocation of proteins in their unfolded state, a pathway for folded proteins was more recently discovered. Proteins destined for the latter pathway are synthesised with an N-terminal signal peptide containing two consecutive arginine residues. This so-called Tat (twin-arginine-translocation) pathway was first described for protein import into chloroplasts ([1–3]) and was then found also in bacteria and archaea [4–6]). In contrast to the Sec-pathway that is driven by ATP hydrolysis, Tat-mediated transport requires a transmembrane proton electrochemical gradient for protein translocation. The Tat-machinery has been extensively studied in *Escherichia coli*, for which a set of membrane proteins, TatA, TatB, TatC and TatE, has been identified by genetic analysis [5]. The small sequence-related components TatA, TatB and TatE are predicted to have a single transmembrane segment, while TatC is expected to contain six transmembrane helices [7]. In the currently accepted model for translocation in *E. coli*, the protein destined for transport has to bind via its signal sequence to a preformed TatBC complex, which is integrated in the membrane. Then, a TatA homo-oligomer is recruited to the complex and acts as the translocation channel, driven by the protonmotive force [8].

In *Bacillus subtilis* the translocase system is less complicated, as it contains only TatA and TatC proteins [9]. Namely, TatA$_y$ and TatC$_y$ are used for translocation of the protein YwbN,

---

```
                    TMS                      APH
       1        10        20          30        40           50        60
BACSU  MFSNIG.IPGLILIFVIAIIIFGPSKLPEIGRAAGRTLLEFKSATKSLVS......GDEKEEKSAELTAV
ECOLI  .MGGIS.IWQLLIIAVIVVLLFGTKKLGSIGSDLGASIKGFKKAMSDDE..PKQ..DKTSQDADFTAKTI
HAEIN  .MFGLS.PAQLIILLVVILLIFGTKKLRNAGSDLGAAVKGFKKAMKEDEKVKDAEFKSIDNETASAKKGK
PSEAE  .MGIFD.WKHWIVILIVVVLVFGTKRLKNLGSDVGEAIKGFRKAVNTEED......DKKDQPAAQPAQPL
MYCTU  .MGSLS.PWHWAILAVVVIVLFGAKKLPDAARSLGKSLRIFKSEVRELQ......NENKAEASIETPTP
AGRTU  .MGSFS.VWHWLIVLVIVLVLFGRGKIPELMGDVAKGIKSFKKGMADEDQ......TPPPADANANAKTV
HELPY  .MGGFTSIWHWVIVLLVIVLLFGAKKIPELAKGLGSGIKNFKKAVKDDEEEAKN..EPKTLDAQATQTKV
VIBCO  .MGGIS.IWQLLIIAVIVVLLFGTKKLRGIGSDLGSAVKGFKKAMSEEE..SNS..AANQKDADFETKNL
YERPE  .MGSIIG.WAQLLIIAVIVVLLFGTNKLRTLGSDLGASIKGFKKAMGDDSQTPPTNVDKTSNDADF.AKSI

              70
BACSU  KQDK......NAG...............
ECOLI  ADKQADTNQEQAKTEDAKRHDKEQV...
HAEIN  YKRE....RNRLNPCLILVFQNLFY...
PSEAE  NQPH......TIDAQAQKVEEPARKD..
MYCTU  VQSQ......RVDPSAASGQDSTEARPA
AGRTU  DHKA.....DEIK...............
HELPY  .HES.....SEIKSKQES..........
VIBCO  EQAK.....TNASAEVKK..DKEQA...
YERPE  TEKQ....QPVAKAEESKSHEKEQG...
```
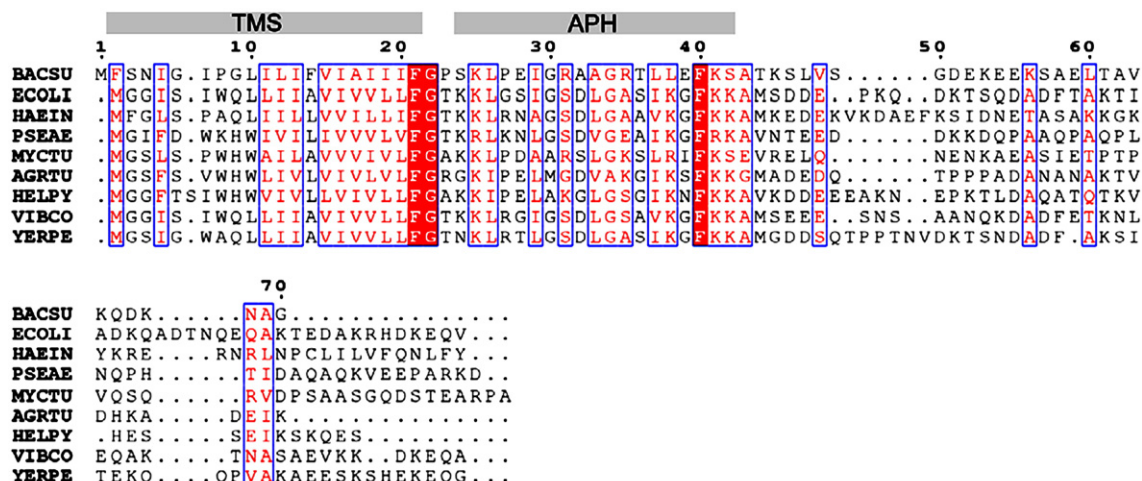
Fig. 1. Sequence alignment of TatA proteins from different bacterial species: *Bacillus subtilis* TatA$_d$ (BACSU), *Escherichia coli* (ECOLI), *Haemophilus influenza* (HAEIN), *Pseudomonas aeruginosa* (PSEAE), *Mycobacterium tuberculosum* (MYCTU), *Agrobacterium tumefaciens* (AGRTU), *Helicobacter pylorii* (HELPY), *Vibrio cholerae* (VIBCO) and *Yersinia pestis* (YERPE). The alignment was generated using the software CLUSTALW [29]. Sequence numbering is according to *Bacillus subtilis* TatA$_d$. Conserved residues are shown in boxes. The regions of the transmembrane segment (TMS) and the amphiphilic helix (APH) are marked with grey bars.

while TatA$_d$ and TatC$_d$ form a specific translocation unit for the enzyme phosphodiesterase PhoD [10]. TatA$_d$ from *B. subtilis* is reported to occur in the membrane and in the cytosol as high molecular weight homo-multimeric assemblies of 150–250 kDa size [11]. These assemblies have a micelle-like structure, as seen by electron microscopy. Furthermore, TatA$_d$ shows specific binding to the twin-arginine signal peptide of PhoD [12] and to TatC$_d$ [13]. These findings suggest that an oligomeric TatA$_d$ complex could recognize prePhoD in the cytosol and guide it to the membrane-embedded TatC$_d$ for translocation through a pore consisting of a variable number of TatA$_d$ units.

Based on its primary structure (see Fig. 1), TatA is predicted to consist of an N-terminal transmembrane segment (TMS), followed by an amphipathic helix (APH) and a highly charged, possibly unstructured C-terminus [14]. Electron microscopy studies have shown that TatA forms homo-multimeric assem-

blies of variable diameter [15]. An initial CD analysis of TatA from *E. coli* confirmed the helical character of this protein as a multimeric complex in detergent solution ($C_{12}E_9$) and in liposomes, though the extramembraneous portion did not retain much structure even in the presence of liposomes [16]. Based on topology studies, conformational changes have been proposed for TatA in connection with the translocation process [17]. The oligomerization behaviour and conformational flexibility of this membrane protein may thus explain why the molecular structure of TatA has so far evaded a more detailed analysis.

Here, we present a strategy using *E. coli* to express and purify large quantities of TatA$_d$ from *B. subtilis* for structure analysis. The three different regions of TatA were constructed and examined separately, in pairwise combinations, and finally compared with the full-length protein TatA$_{2-70}$. We have thus prepared the predicted transmembrane segment (TatA$_{1-27}$), as well
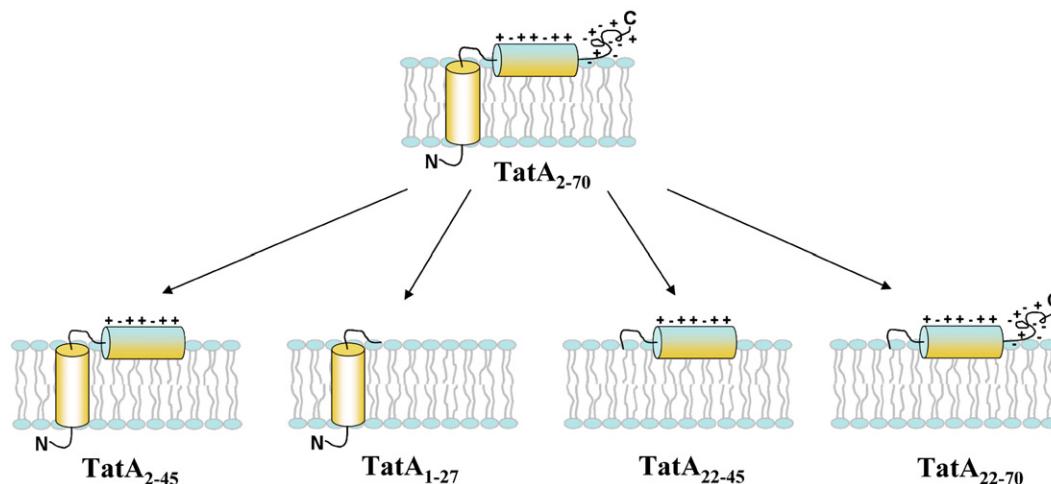

Fig. 2. Schematic illustration of the different TatA constructs investigated here, and their proposed alignment in a lipid bilayer. Cylinders are used to represent the predicted helical regions of the protein, and the N- and C-termini are indicated.

as the extramembraneous portion containing the putative amphiphilic APH region with the charged C-terminus (TatA$_{22-70}$). Furthermore, by removal of the potentially unstructured C-terminus, we have produced a fragment containing TMS plus APH (TatA$_{2-45}$), and another construct consisting of only APH (TatA$_{22-45}$). An overview of the fragments studied here is shown in Fig. 2. In view of their different hydrophobic/amphiphilic/charged characters, it was challenging to reconstitute and study them in one and the same representative bacterial model membrane system. Having selected DMPG:DMPC (70:30) as a suitable lipid bilayer composition in which the constructs were well folded, CD and OCD spectroscopy yielded a comprehensive and self-consistent picture of the structure and membrane alignment of TatA$_d$. CD spectra were measured in aqueous suspension of sonicated lipid vesicles to address the secondary structures of the respective fragments, and to obtain further qualitative information by generating difference spectra. The OCD data were acquired in macroscopically oriented membranes and revealed for the first time the alignment of the individual α-helical segments with respect to the membrane surface. These findings confirm the predicted topology model for TatA proteins, consisting of a transmembrane helix, a surface-bound amphiphilic α-helix, and an unstructured C-terminus.

2. Materials and methods

2.1. Materials

Oligonucleotides were purchased from MWGBiotech (Martinsried). Plasmid pET28a(+) and E. coli strain BL21(DE3) were from Novagen and mutagenesis was carried out using the QuikChange® Site-Directed Mutagenesis Kit from Stratagene. The TatA$_d$ expression strain TG1(pREP4/pQE9) was a gift from Dr. Jörg Müller, Friedrich Schiller Universität Jena, Germany. Ni-affinity chromatography was performed on HisTrap FF 5 ml columns from GE Healthcare. TPCK trypsin and NLS were from Fluka (SIGMA-Aldrich). HPLC was carried out using a Jasco system (Tokyo, Japan) equipped with a diode-array detector, and C18 columns (Grace-Vydac, Hesperia, CA, USA). Lipids were purchased from Avanti Polar Lipids (Alabaster, USA). Biobeads^{TM} SM-2 were purchased from BIORAD. CD and OCD spectra were recorded on a Jasco J-810 spectropolarimeter (Jasco Co., Tokyo, Japan).

2.2. Cloning and expression

Competent cells were prepared following the protocol of Chung [18]. Transformation was performed using standard protocols. DNA sequences of all constructs were verified by sequencing (MWGBiotech, Martinsried, Germany). All five constructs of TatA were grown at 37 °C and 250 rpm shaking speed in LB medium supplemented with the respective antibiotic. Expression was induced at an OD$_{600}$ of 0.6–0.8 by addition of 0.5 mM IPTG. All constructs were expressed as N-terminal His$_6$-tag fusion proteins. SDS-PAGE analyses were performed using 16% Tris-Tricine polyacrylamide gels [19].

2.2.1. TatA$_{2-70}$

Full-length TatA was expressed using E. coli strain TG1 containing the two plasmids pQE9TatA and pREP4 [12] for 20 h. To generate another TatA expression construct in a more convenient single-plasmid system, the TatA gene was amplified by PCR using primers TatA_1 (5′-GTG GAA TTC CAT ATG TTT TCA AAC ATT GG-3′) and TatA_2 (5′-GCG CCT CTC GAG TCA GCCCGCGTTT TTGTCC-3′), and the resulting product was cut with NdeI and XhoI and cloned into vector pET28a(+), yielding plasmid pET28TatA$_{1-70}$, which was then used to transform E. coli strain BL21 (DE3). Note that CNBr cleavage removed the first amino acid in the protein, thus yielding TatA$_{2-70}$.

2.2.2. TatA$_{2-45}$

To construct a TatA fragment lacking the highly charged C-terminus, a TAA stop codon was introduced in plasmid pET28TatA$_{1-70}$ at the position of Ser46 by site-directed mutagenesis, using the primers TatA$_{1-45}$_1 (5′-GCGCCACAAAA-TAACTTGTGTCTGG-3′) and TatA$_{1-45}$_2 (5′-CCAGACACAAGT-TATTTTGTGGCGC-3′). The resulting plasmid pET28TatA$_{1-45}$ was used to transform E. coli strain BL21(DE3), and the cells were allowed to express for 20 h.

2.2.3. TatA$_{1-27}$

For expression of the transmembrane segment TatA$_{1-27}$, a stop codon was introduced in plasmid pET28TatA$_{1-70}$ at position of Glu28 by site-directed mutagenesis, using the primers TMS_1 (5′-CCTTCCAAGCTGCCGTA-GATCGGGCGTGCCGCC-3′) and TMS_2 (5′-GGAAGGTTCGACGGCATC-TAGCCCGCACGGCGG-3′). To eliminate the vector-encoded tripeptide GSH between the His-tag and TatA$_{1-27}$, a second mutagenesis using the primers TMS_3 (5′-CCTGGTGCCGCGC ATGTTTTCAAACATTGG-3′) and TMS_4 (5′-CCAATGTTTGAAAACATGCGCGGCACCAGG-3′) followed. Expression of TatA$_{1-27}$ was performed in E. coli strains BL21(DE3) and C41(DE3), and expression was carried out for 16–20 h.

2.2.4. TatA$_{22-70}$

To generate a TatA construct without the hydrophobic N-terminal segment, the TatA gene was amplified by PCR using the primers TatA1 (5′-GTG GAA TTC CAT ATG TTT TCA AAC ATT GG -3′) and TatA3 (5′-GTG GAA TTC CAT ATG GGC CCT TCC AAG CTG C -3′). The PCR product was cut by NdeI and XhoI and ligated into vector pET28a(+). The resulting plasmid pET28TatA$_{22-70}$ was used to transform E. coli strain BL21(DE3), and expression of this construct was performed for 4–5 h.

2.2.5. TatA$_{22-45}$

The amphiphilic central region of TatA was constructed by introducing a Met codon (ATG) at the position of Phe21 in pET28TatA$_{1-45}$, using the primers a-TatA_1 (5′-CGCCCTCATTATTATGGGCCCTTCCAAGCTGCCGG-3′) and a-TatA_2 (5′-CCGGCAGCTTGGAAGGGCCCATAATAATGAGGGCG-3′). The protein variant TatA$_{1-45}$F21M was expressed in E. coli strain BL21(DE3) for 20 hrs, and the N-terminal segment was removed by CNBr cleavage.

2.3. Cell breakage

After expression, cells were harvested by centrifugation and the resulting cell pellet was resuspended in HisTrap binding buffer (20 ml of buffer for cells from a 1-l expression). The cells were broken using a French pressure cell (SLM Aminco) at 1500 psi, or with a sonifier (BANDELIN sonopuls). Benzamindine, Pefabloc SC and phenylmethylsulfonylfluorid were added as protease inhibitors, and benzonase for degradation of RNA and DNA. The cell debris was pelleted by centrifugation at 4 °C. To obtain the cell membranes, ultracentrifugation was performed for 1 h at 4 °C using a 50Ti rotor (Beckman Coulter). Cell membranes, cell debris and the supernatant, respectively were stored at −80 °C.

2.4. Protein purification

Ni-affinity chromatography binding buffer: 0.5 M NaCl, 0.02 M imidazole, 0.02 M Na-phosphate, pH 7.4; elution buffer: 0.5 M NaCl, 0.5 M imidazole, 0.02 M Na-phosphate, pH 7.4; Gel filtration buffer: 0.15 M NaCl, 0.02 M Na-phosphate, pH 7.0. All protein constructs were characterized by mass spectrometry using MALDI-TOF (Bruker Daltonics).

2.4.1. TatA$_{2-70}$ and TatA$_{2-45}$

The bulk of the His-tagged TatA$_{2-70}$ was found soluble in the cytosol and could be purified directly from the supernatant. TatA$_{2-45}$ was solubilized from the cell debris and membrane fractions by incubation in binding buffer supplemented with 2% (w/v) NLS overnight and slight sonication. The solubilizate and the supernatant, respectively, were then purified by Ni-affinity chromatography with 0.2% (w/v) NLS in all purification buffers to keep the proteins in solution. For cleavage of the His-tag by CNBr, the proteins had to be dialysed against water using dialysis tubings with 3.5 kDa cut-off at 4 °C overnight. The protein was then lyophilized and re-dissolved in 50% TFA

supplemented with some crystals of CNBr. This mixture was allowed to incubate at RT in the dark for 20 h. After reducing TFA concentration by adding 20 ml of H_2O, the solution was lyophilized and the resulting powder was re-dissolved in binding buffer containing 1–2% (w/v) NLS. A subtractive Ni-affinity chromatography step followed to remove free His-tag and uncut protein. When necessary, gel filtration on a HighLoad Superdex 200 pg (26/60) column was carried out as the final purification step. The pooled fractions were dialysed against water and lyophilized to obtain the product.

2.4.2. TatA$_{1-27}$

The construct TatA$_{1-27}$ was solubilized from the membrane fraction as described above and purified by Ni-affinity chromatography with 0.2% (w/v) NLS in all purification buffers. Removal of the His-tag was achieved by incubating the fusion protein in elution buffer with TPCK-treated trypsin (Fluka). Completeness of the tryptic digest was monitored by SDS-PAGE. After addition of protease inhibitors the protein was purified further using a HighLoad Superdex 200 pg (26/60) (GE Healthcare) column. Purified fractions were pooled and could be stored in detergent-containing binding buffer at 4 °C.

2.4.3. TatA$_{22-70}$

The water-soluble construct TatA$_{22-70}$ was purified from the supernatant by Ni- affinity chromatography without using detergent. To remove the His-tag, the fusion protein was treated with CNBr in 50% TFA as described above. TFA was removed by lyophilisation and the protein was re-dissolved using binding buffer. A subtractive Ni-affinity chromatography step was used to separate uncut fusion protein and free His-tag. The cleaved protein was purified further on a gel filtration column (Sephacryl S100 HR, GE Healthcare) equilibrated with gel filtration buffer without detergent. For desalting the protein was passed through a PD-10 column (GE-Healthcare), followed by lyophilization.

2.4.4. TatA$_{22-45}$

The variant TatA$_{2-45}$F21M was solubilized from the membrane and cell debris, purified by affinity chromatography and dialysed from detergent as described for TatA$_{2-45}$ above. To generate TatA$_{22-45}$, the His-tag fusion protein was treated with CNBr in 50% TFA for 20 h. To this mixture 3 volumes of DMF were added, and the precipitated N-terminal part of TatA$_{2-45}$F21M was centrifuged. The supernatant, containing TatA$_{22-45}$, was purified further by reverse phase HPLC. Analytical HPLC was performed using a C18 analytical column (4.6 × 250 mm), employing a linear water/acetonitrile gradient (separation part of the gradient: 5–20% B in 6 min) at 40 °C and a flow rate of 1.2 ml/min. Solvent A: 10% acetonitrile in 5 mM HCl; B: 90% acetonitrile in 5 mM HCl. Preparative HPLC was carried out using same solvents on a C18 column (22 × 250 mm) with a linear water/acetonitrile gradient (separation part of the gradient: 5–55% B in 8 min) at 40 °C and a flow rate of 16 ml/min. The injection volume for the analytical column was 20 µl and for the preparative column 4.8 ml.

2.5. Reconstitution into phospholipid bilayers

TatA$_{2-70}$ and TatA$_{1-27}$ were reconstituted into lipid vesicles (DMPG:DMPC molar ration of 70:30) at a molar protein-to-lipid ratio of approximately 1:100. Protein and lipids were co-solubilized with the aid of 0.5% (w/v) NLS in 10 mM Na-phosphate buffer (pH 7.0), and the detergent was subsequently removed using Biobeads™. The Biobeads™ were pre-washed with methanol and water and equilibrated either in buffer to prepare CD samples, or in pure water to prepare OCD samples. In either case, the protein–lipid–detergent mixed micelles were incubated for several minutes at 33 °C. Then, a spoon-full of Biobeads™ was added to the solution, and the mixture was stirred at 33 °C for 1 h. This step was repeated two times. The solution was finally separated from the Biobeads™ by filtration through glass wool.

TatA$_{2-45}$ and TatA$_{22-45}$ were reconstituted into lipid vesicles (DMPG: DMPC = 70:30) at a molar protein-to-lipid ratio of approximately 1:100 by co-solubilizing the protein and lipids in chloroform/methanol (v:v 1:1). The solvent was then evaporated with a gentle stream of nitrogen, and the resulting lipid film was placed under vacuum overnight to remove residual solvent. To prepare samples for CD spectroscopy, the lipid film was hydrated by adding 250 µl of H_2O, the suspension was incubated in a water bath at 37 °C for 1 h, and finally sonified to obtain small unilamellar vesicles. To prepare OCD samples of TatA$_{2-45}$, the dried lipid film was directly hydrated as described in the OCD

section. (In the case of TatA$_{22-45}$, however, we did not obtain meaningful spectra when reconstituting this way from $CHCl_3$/MeOH.) Therefore, for OCD analysis the protein and the lipids were co-solubilized at a protein-to-lipid ratio of 1:50 with the aid of 0.5% (w/v) NLS in water and further processed as described above for TatA$_{2-70}$.

TatA$_{22-70}$ is water-soluble and could be added directly to the lipid vesicles at a protein-to-lipid ratio of 1:100 in 10 mM Na-phosphate buffer (pH 7.0). Pure DMPC vesicles, as well as mixed DMPG/DMPC vesicles with a molar lipid ratio of either 20:80 or 70:30, were prepared from multilamellar lipid suspensions by repeated extrusion to create large unilamellar vesicles for CD measurements, following standard protocols (http://www.avantilipids.com). To measure TatA$_{22-70}$ in detergent micelles, the protein was added to SDS or DPC solution at a molar ratio of approximately 1:150. Difficulties were encountered, however, in the OCD analysis of TatA$_{22-70}$, as the spectra showed a high tendency for this construct to aggregate under many different conditions. Besides adding the protein to pre-formed vesicles (as was successful in the CD analysis), reconstitution was attempted via co-solubilization of protein and lipids in 80% TFE, as well as reconstitution with the help of detergent and Biobeads™ as described for TatA$_{2-70}$. Nevertheless, all methods yielded the same spectra.

2.6. CD spectroscopy

CD spectra of the vesicle suspensions and aqueous solutions (with detergent or TFE) were recorded in quartz glass cells (1 mm path length) between 260 and 185 nm at 0.1-nm intervals. Three repeat scans at a scan-rate of 10 nm min^{-1}, 4-s response time and 1-nm bandwidth were averaged for each sample and its respective baseline of the respective protein-free sample. All spectra were recorded at 30 °C, using a water-thermostated rectangular cell holder. In some samples the high hydrophobicity of the proteins appears to have resulted in incomplete dissolution/reconstitution yields. Therefore, no mean residue ellipticity values were calculated, given that the protein concentration of the samples could not be determined with the necessary accuracy by standard colorimetric assays. UV concentration determination at 280 nm could not be performed because TatA and the relevant fragments do not contain any tryptophan or tyrosine residues. Instead, the data are normalized as indicated in the respective figure legends, with the intention to emphasize changes in the line shape rather than showing inaccurate mean residue ellipticities.

2.7. OCD spectroscopy

For oriented CD spectroscopy, proteins were reconstituted into mechanically aligned phospholipid bilayers on a quartz plate and subsequently hydrated in the OCD-spectrometer. Samples were deposited on one of two circular quartz glass plates with 20-mm diameter (Suprasil QS, Hellma Optik, Jena), which served as UV transparent windows in the OCD cell. Before use, the quartz glass windows were cleaned in concentrated nitric acid and rinsed thoroughly and repeatedly with distilled water and finally with ethanol. Afterwards, the windows were dried in an oven at 60 °C. A 50- to 100-µl aliquot of the corresponding protein/lipid vesicle suspension was deposited on the quartz glass window with a pipet, and water (or solvent) was allowed to evaporate in a gentle stream of N_2 until the sample appeared dry. Afterwards, the window with the dried sample was assembled into the OCD cell and re-hydrated for ~15 h at 30 °C and 97.0% relative humidity using a saturated K_2SO_4 solution (see below). During hydration the lipid spontaneously aligns as multi-layers parallel to the quartz glass surface. The home-made cell for the OCD measurements is described in detail in [20], based on the set-up by [21]. The optical path of the CD spectropolarimeter is parallel to the cylindrical axis of the cell and normal to the quartz glass window with the oriented sample. The temperature of the cell can be controlled by a thermostat, and a small volume of saturated K_2SO_4 salt solution (~300–500 µl) was filled into the bottom of the cell to maintain 97% humidity. The OCD cell was mounted into the sample compartment of the J-810 spectropolarimeter on a rotation stage with a computer-controlled stepping motor (SKIDS-60YAW (θz), Sigma Koki Co. Ltd., Tokyo). To reduce possible spectral artefacts due to the linear dichroism due to imperfections in the sample, strain in the quartz glass windows, or imperfect alignment of the windows [22,23], the OCD spectra were recorded every 45.0 deg of rotation of the cell, as described by [21], as averages of three scans using the same data acquisition

parameters as in the normal CD measurements above. The eight spectra were subsequently averaged, and background spectra of lipid bilayers without protein were subtracted.

3. Results

3.1. Expression and purification of TatA constructs

TatA$_d$ from *B. subtilis* is a 70-amino acid protein, predicted to consist of an N-terminal helical transmembrane segment (TMS), followed by an amphiphilic helix (APH), and a possibly unstructured, highly charged C-terminus [14]. The sequence comparison for different bacteria in Fig. 1 shows that TMS and APH are conserved parts of the protein. Based on these characteristic regions, we have constructed four different TatA fragments to allow structural studies of the putative TMS and APH in more detail, as illustrated in Fig. 2. The full-length protein is termed here TatA$_{2-70}$, as the first methionine is cut off upon His-tag removal by CNBr cleavage. The construct representing mainly TMS is denoted TatA$_{1-27}$. The highly charged C-terminal part of TatA is less conserved and was shown in *E. coli* not to be important for function [14]. Therefore, the last 25 amino acids of TatA were deleted to generate the fragment TatA$_{2-45}$. To study the extramembraneous part further fragments were generated, namely the soluble TatA$_{22-70}$ lacking the first 21 amino acids, and TatA$_{22-45}$ corresponding to the APH alone.

All constructs were designed with an N-terminal His-tag to ease purification by Ni-affinity chromatography, followed by gel filtration when necessary. Only the small peptide TatA$_{22-45}$ was more conveniently purified by HPLC. For all constructs containing the TMS, purification buffers were supplemented with the anionic detergent N-lauroylsarcosine (NLS), which has been used before to solubilize proteins expressed in *E. coli* [24]. Amongst many other detergent tried, we found this one to be the best for solubilizing the hydrophobic TatA constructs from the cell membrane and stabilizing them in solution. NLS also helped to overcome binding problems of TatA$_{2-70}$ to the affinity matrix, given the tendency of soluble TatA$_{2-70}$ to self-assembly (results not shown). This detergent does not interfere with Ni-affinity chromatography and can be removed quickly by dialysis due to its relatively high CMC. After removing the detergent by dialysis, the N-terminal His-tag could be cleaved off using CNBr. In the case of the hydrophobic TatA$_{1-27}$, CNBr treatment was possible but resulted in an insoluble product after lyophilisation. Therefore, the His-tag was cleaved off with trypsin (hence this sequence begins with Met1) while keeping the protein soluble in NLS buffer.

TatA$_{2-70}$ was expressed in *E. coli* strain TG1 and found to be present in both the cytosolic fraction and the membrane fractions, but it was no longer water-soluble after CNBr removal of the His-tag. Then final yield after purification was about 25 mg of protein per liter expression medium.

TatA$_{2-45}$ was expressed in BL21(DE3) cells, using the TatA plasmid in which the codon of Ser46 had been exchanged for a stop codon. The protein was only present in the membrane and cell debris fractions. In contrast to full-length TatA$_{2-70}$, this deletion construct easily precipitates when NLS is removed by dialysis, suggesting that the C-terminus of TatA helps to stabilize the protein in the cytosol. The yield of pure TatA$_{2-45}$ was more than 20 mg/l.

TatA$_{1-27}$ was constructed to contain the transmembrane segment, and it could be expressed using *E. coli* strain C41 (DE3) [25]. Expression failed, however, in BL21(DE3) due to toxicity problems (data not shown), which were not encountered with any of the other constructs. Approximately 3–4 mg/l pure TatA$_{1-27}$ were obtained after trypsin cleavage.

TatA$_{22-70}$ corresponds to the water-soluble extramembraneous part. Expression was restricted to 5 h, since prolonged expression showed degradation of the fragment, and the yield was about 20 mg/l.

TatA$_{22-45}$ was first tried to be expressed in BL21(DE3) cells by replacing the codon of Ser46 by a stop codon in the TatA$_{22-70}$ construct (data not shown). Because expression of this construct failed, possibly due to degradation of the small peptide produced, the codon for Phe21 in construct TatA$_{2-45}$ was replaced by an ATG codon for Met. CNBr cleavage of the variant TatA$_{2-45}$F21M then yielded 15 mg/l after HPLC.

3.2. CD spectroscopy

To examine the predicted secondary structure elements of TatA by circular dichroism, the five TatA$_d$ constructs were reconstituted with phospholipid vesicles and other membrane mimicking environments. Fig. 3A shows the water-soluble extramembraneous fragment TatA$_{22-70}$ in different environments, namely in dodecylphosphocholine (DPC) and sodium dodecyl sulphate (SDS) micelles, in lipid vesicles made of DMPC and DMPG with different ratios, and in 30% TFE. The spectral lineshapes are clearly seen to fall into two distinct groups. TatA$_{22-70}$ reconstituted in zwitterionic detergent micelles (DPC), vesicles (DMPC) or in mixed lipid vesicles with a low percentage of the negatively charged DMPG (DMPG:DMPC=20:80) show a negative band at 198 nm and no positive intensity, indicating that the protein is largely unstructured. However, TatA$_{22-70}$ in a negatively charged environment, such as SDS micelles or mixed lipid vesicles with a high ratio of DMPG (DMPG:DMPC molar ratio of 70:30), show a strong positive band at 190 nm and two negative bands at 203 nm and 222 nm. Under these conditions the protein assumes a predominantly α-helical structure, similar to the situation in 30% TFE that is generally known to stabilize such conformation.

Having found that DMPG:DMPC (70:30) vesicles represent the most suitable environment to study TatA, we now compare in Fig. 3B the lineshapes of TatA$_{22-70}$ and TatA$_{22-45}$ under the same conditions. In the latter truncated fragment the highly charged C-terminus has been removed, hence we calculated the difference spectrum of (TatA$_{22-70}$)-(TatA$_{22-45}$). The resulting lineshape exhibits a distinct minimum at 200 nm, suggesting that the C-terminal region (positions 45 to 70) is indeed unstructured. In Fig. 3C, the CD lineshapes of all TatA fragments containing the putative transmembrane segment are compared (in DMPG/DMPC vesicles). As expected, they all
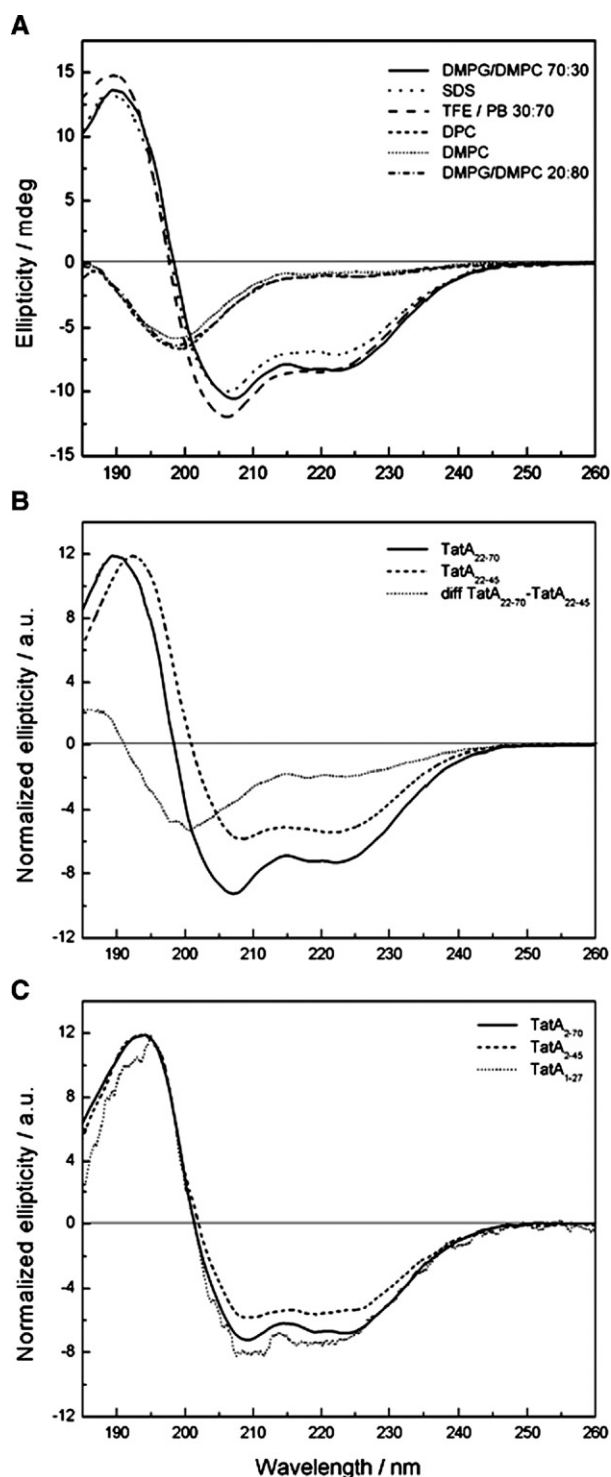
Fig. 3. CD spectra of the different TatA constructs in various membrane-mimicking environments. (A) Extramembraneous fragment TatA$_{2-70}$ in DMPC and DMPG/DMPC vesicle suspensions of different composition, TFE/buffer (30/70 v/v), and SDS and DPC micelles. (B) Extramembraneous fragment with (TatA$_{22-70}$) and without (TatA$_{22-45}$) the C-terminal region, and their calculated difference spectrum in DMPG/DMPC (70/30 w/w) vesicles. (C) Comparison of the TMS-containing constructs TatA$_{1-27}$, TatA$_{2-45}$ and TatA$_{2-70}$ in DMPG/DMPC (70/30 w/w) vesicles. (The positive bands of the experimental spectra in panels B and C are normalized to the same intensity in order to illustrate the characteristic spectral features, since the absolute protein concentrations could not be accurately determined).

show a high degree of α-helical structure, and in the full-length protein TatA$_{2-70}$ the slight shift of to lower wavelengths is indeed consistent with a contribution from the unstructured C-terminal portion.

3.3. OCD spectroscopy

Oriented CD spectroscopy allows the structure analysis of membrane-embedded proteins in samples that are macroscopically aligned with respect to the light beam. OCD is a powerful tool to discriminate between a transmembrane orientation and surface alignment of well-folded α-helical segments. Based on previous studies of membrane-active peptides [21,23] the characteristic lineshapes are illustrated in Fig. 4A for the transmembrane alignment (often referred to as the inserted state or I-state) and for the surface state (or S-state) of an α-helix. The most informative feature is the presence of absence of a negative band around 208 nm, being indicative of a surface or transmembrane helix alignment, respectively. (Note that for other secondary structure elements the OCD analysis is not yet that straightforward). To study the different TatA constructs by OCD, they were reconstituted at a protein-to-lipid ratio of about 1:100 (except for TatA$_{22-45}$ which had a ratio of 1:50) in DMPG: DMPC (70:30) bilayers, and spread onto the quartz windows to give macroscopically oriented membrane samples. First, let us consider the membrane alignment of the shortest helical fragments of TatA, namely the N-terminal and the central segment, as their OCD lineshapes should be straightforward to interpret. The spectrum of TatA$_{1-27}$ in Fig. 4B has a broad minimum around 225–230 nm, while the negative band at 209 nm is completely absent. These features are clearly indicative of a transmembrane alignment, thus confirming the predicted N-terminal TMS. In contrast, the lineshape of fragment TatA$_{22-45}$ with a double-minimum closely resembles the example of Fig. 4A for a surface alignment. Hence this result proves the planar orientation of the predicted amphiphilic helix in lipid membranes.

Next, consider fragment TatA$_{2-45}$ whose sequence contains approximately equal contributions from TMS and APH. Its lineshape is compared in Fig. 4B with the arithmetic average of the two separate fragments TatA$_{1-27}$ and TatA$_{22-45}$. A good correspondence is seen for the experimental and calculated lineshapes in the relevant spectral region of the negative bands above 200 nm. This observation demonstrates that both individual segments TMS and APH are not only well folded but they are also aligned in the membrane virtually the same way as they are in the combined construct. The qualitative information obtained from the OCD spectra of the various TatA constructs is thus in full accordance with the predicted topology of the TatA protein. Only in one case we found that the OCD spectrum of a fragment could not be cross-validated with the data from the rest of the protein. Namely, the lineshape of fragment TatA$_{22-70}$ shows a broad positive band in the range from 185 to 205 nm and a broad negative band around 218 nm, which closely resembles the OCD spectra of β-sheet proteins in the literature [26,27]. We repeatedly obtained such OCD lineshapes for TatA$_{22-70}$
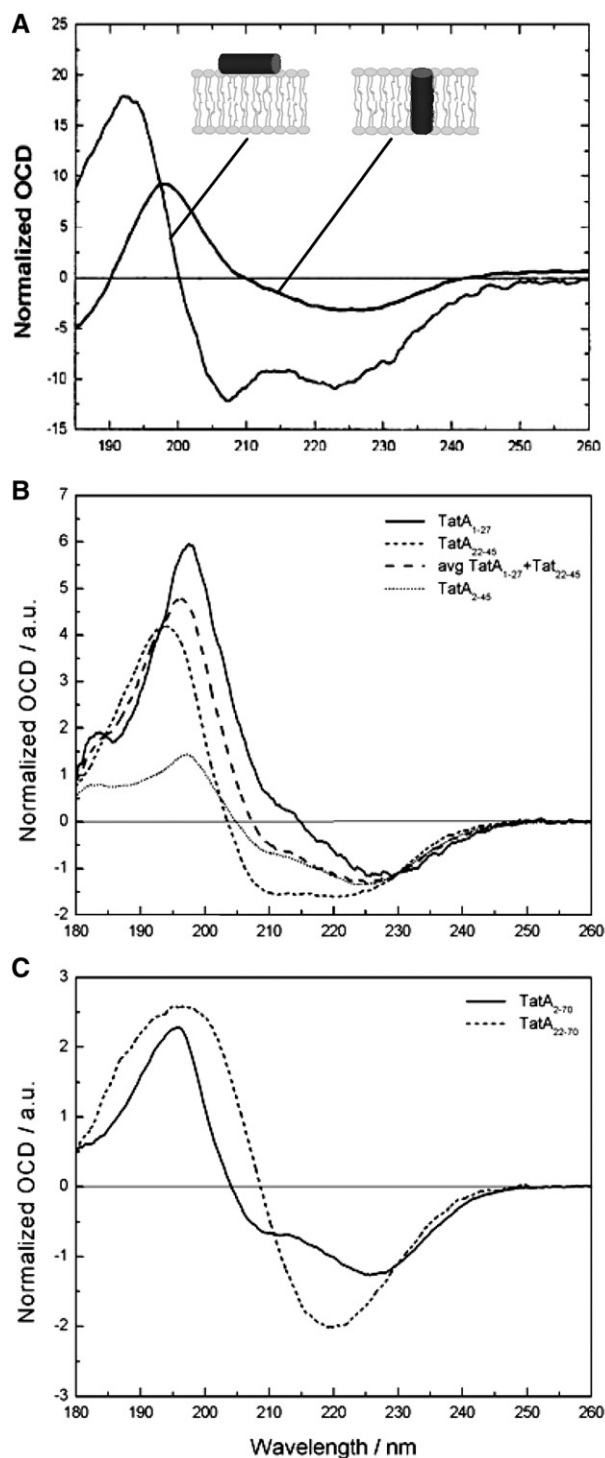
Fig. 4. OCD spectra of the TatA constructs measured in oriented DMPG/DMPC (70/30 w/w) bilayers for a peptide/lipid ratio of about 1:100 (for TatA$_{22-45}$ 1:50) at 30 °C and 97.0% relative humidity. (A) Illustration of the characteristic OCD lineshapes for the transmembrane and surface alignment of the antimicrobial peptide alamethicin in lipid bilayers (redrawn from Chen, Lee and Huang 2002). (B) OCD spectra of the helical fragments TatA$_{1-27}$, TatA$_{22-45}$, their calculated average, and of TatA$_{2-45}$. (C) OCD spectra of TatA$_{2-70}$ and TatA$_{22-70}$. (The spectra in panels B and C are normalized to give the same intensity at 230 nm, in order to illustrate the relative differences in the lineshapes, since absolute protein concentrations could not be determined accurately).

samples also with other protein-to-lipid ratios, prepared in different ways from different solvents, and reconstituted in different kinds of lipid mixtures (data not shown). Therefore, it appears that the isolated extramembraneous TatA$_{22-70}$ fragment has a strong tendency to form β-pleated aggregates. Such structures are frequently observed for peptides and proteins that do not fold correctly into their physiologically relevant conformation.

4. Discussion

TatA$_d$ is the main component of the *B. subtilis* translocation channel used for exporting folded proteins across the inner membrane, but little is known about the structure of this conformationally flexible protein, which is situated in membranes or forms large homooligomeric aggregates. TatA has been predicted to consist of three regions, namely a transmembrane segment, an amphiphilic α-helix, and an unstructured, highly charged C-terminus. These individual fragments were expressed separately (TatA$_{1-27}$, TatA$_{22-45}$) or pairwise (TatA$_{2-45}$, TatA$_{22-70}$) in *E. coli* for structure analysis, and to be compared with the full-length protein (TatA$_{2-70}$). It was challenging to tackle the different hydrophobicities, charges and sizes of these constructs, and to reconstitute all of them in one and the same representative model membrane. From a practical point of view, the detergent N-lauroylsarcosine was found to be critical for solubilising of the TatA homooligomers and for stabilizing those constructs containing the transmembrane segment. NLS was well suited for solubilizing the proteins from the bacterial membrane, for purification by Ni-affinity chromatography and gel filtration, and for reconstitution into lipid bilayers. All constructs were treated with CNBr to cleave the N-terminal His-tag, except for TatA$_{1-27}$ which required a more gentle method using trypsin. The preparation of the transmembrane segment TatA$_{1-27}$ was difficult due to its pronounced hydrophobicity, but the outlined expression and purification strategy might also be applicable to other transmembrane segments.

Having prepared all TatA fragments with high yield, they were successfully reconstituted into phospholipid bilayers, which is a suitable environment for circular dichroism and for future solid-state NMR studies. CD analysis in lipid vesicles showed that all fragments, except for the highly charged and disordered C-terminus, take on an α-helical conformation in negatively charged lipid membranes. In the presence of uncharged lipids or detergents, on the other hand, the water-soluble extramembraneous portion of TatA (with a net charge of +1) does not fold. This observation is in good agreement with previous CD studies by Porcelli et al. [16], who had focussed on a similar TatA fragment from *E. coli* in which the TMS-segment had been removed. That construct was also found to be unstructured in buffer solution and to show slightly more, but less than expected, helical structure in liposomes from *E. coli* total polar lipids. It had thus been suggested that the N-terminal transmembrane segment of TatA might be necessary for correct folding of the protein. Here, we demonstrate that for the amphiphilic segment TatA$_{22-70}$ from *B. subtilis* (with a net

charge of +2) a helical structure can be induced merely by a negatively charged environment. This finding is consistent with studies demonstrating that the efficiency of the Tat transport system is highly dependent on the presence of anionic phospholipids [28].

To elucidate the alignment of the different TatA segments in the lipid bilayer we have employed oriented circular dichroism spectroscopy. This technique is a valuable tool to study the orientation of helical segments of membrane-active peptides and membrane proteins in their native lipid environment under physiologically relevant conditions. By analyzing and cross-validating the individual fragments of TatA, we present here for the first time a comprehensive picture of the TatA structure based on experimental evidence. The OCD data demonstrate that the N-terminal region is indeed folded as membrane-spanning α-helix, and that the predicted amphiphilic helix is aligned parallel to the bilayer surface. Taken together, these results fully support the predicted topology of TatA, and support the idea that multiple copies of this protein may assemble to form a genuine transmembrane channel with additional amphiphilic and charged functionalities, across which proteins can be exported.

Acknowledgements

References

[1] R.M. Mould, C.A. Robinson, A proton gradient is required for the transport of two lumenal oxygen-evolving proteins across the thylakoid membrane, J. Biol. Chem. 266 (1991) 12189–12193.

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

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

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

[5] F. Sargent, E. Bogsch, M. Stanley, M. Wexler, C. Robinson, B. Berks, T. Palmer, Overlapping functions of components of a bacterial Sec-independent protein export pathway, EMBO J. 17 (1998) 3640–3650.

[6] E.G. Bogsch, F. Sargent, N.R. Stanley, B.C. Berks, C. Robinson, T. Palmer, An essential component of a novel bacterial protein export system with homologues in plastids and mitochondria, J. Biol. Chem. 273 (1998) 18003–18006.

[7] C. Robinson, A. Bolhuis, Tat-dependent protein targeting in prokaryotes and chloroplasts, Biochim. Biophys. Acta 1694 (2004) 135–147.

[8] B. Berks, T. Palmer, F. Sargent, Protein targeting by the bacterial twin-arginine translocation (Tat) pathway, Curr. Opin. Microbiol. 8 (2005) 174–181.

[9] J.D. Jongbloed, U. Grieger, H. Antelmann, M. Hecker, R. Nijland, S. Bron, J.M. van Dijl, Two minimal Tat translocases in Bacillus, Mol. Microbiol. 54 (2004) 1319–1325.

[10] O. Pop, U. Martin, C. Abel, J.P. Müller, The twin-arginine signal peptide of PhoD and the TatAd/Cd proteins of *Bacillus subtilis* form an autonomous Tat translocation system, J. Biol. Chem. 277 (2002) 3268–3273.

[11] M. Westermann, O.I. Pop, R. Gerlach, T.R. Appel, W. Schlormann, S. Schreiber, J.P. Müller, The TatAd component of the *Bacillus subtilis* twin-arginine protein transport system forms homo-multimeric complexes in its cytosolic and membrane embedded localisation, Biochim. Biophys. Acta 1758 (2006) 443–451.

[12] O.I. Pop, M. Westermann, R. Volkmer-Engert, D. Schulz, C. Lemke, S. Schreiber, R. Gerlach, R. Wetzker, J.P. Müller, Sequence-specific binding of prePhoD to soluble TatAd indicates protein-mediated targeting of the Tat export in *Bacillus subtilis*, J. Biol. Chem. 278 (2003) 38428–38436.

[13] S. Schreiber, R. Stengel, M. Westermann, R. Volkmer-Engert, O.I. Pop, J.P. Müller, Affinity of TatC$_d$ for TatA$_d$ elucidates its receptor function in the *Bacillus subtilis* twin arginine translocation (Tat) translocase system, J. Biol. Chem. 281 (2006) 19977–19984.

[14] P.A. Lee, G. Buchanan, N.R. Stanley, B.C. Berks, T. Palmer, Truncation analysis of TatA and TatB defines the minimal functional units required for protein translocation, J. Bacteriol. 184 (2002) 5871–5879.

[15] U. Golke, L. Pullan, C.A. McDevitt, I. Porcelli, E. deLeeuw, T. Palmer, H.R. Saibil, B.C. Berks, The TatA component of the twin-arginine protein transport system forms channel complexes of variable diameter, Proc. Natl. Acad. Sci. U. S. A. 102 (2006) 10482–10486.

[16] I. Porcelli, E. de Leeuw, R. Wallis, E. van den Brink-van der Laan, B. de Kruijff, B.A. Wallace, T. Palmer, B.C. Berks, Characterization and membrane assembly of the TatA component of the *Escherichia coli* twin-arginine protein transport system, Biochemistry 41 (2002) 13690–13697.

[17] K. Gouffi, F. Gérard, C.L. Santini, L.F. Wu, Dual topology of the *Escherichia coli* TatA protein, J. Biol. Chem. 279 (2004) 11608–11615.

[18] C.T. Chung, S.L. Niemela, R.H. Miller, One-step preparation of competent *Escherichia coli*: transformation and storage of bacterial cells in the same solution, Proc. Natl. Acad. Sci. U. S. A. 86 (1989) 2172–2175.

[19] H. Schägger, G. von Jagow, Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa, Anal. Biochem. 166 (1987) 368–379.

[20] J. Bürck, P. Wadhwani, S. Afonin, E. Strandberg, A.S. Ulrich, FInal title: Characterization of conformation and re-alignment of membrane-active peptides PGLa, MSI-103 and MAP in DMPC bilayers by oriented cicular dichroism, manuscript to be submitted to Biophys. J.

[21] F.Y. Chen, M.T. Lee, H.W. Huang, Sigmoidal concentration dependence of antimicrobial peptide activities: a case study on alamethicin, Biophys. J. 82 (2002) 908–914.

[22] G.A. Olah, H.W. Huang, Circular dichroism of oriented α-helices: I. Proof of the exciton theory, J. Chem. Phys. 89 (1988) 2531–2538.

[23] Y. Wu, H.W. Huang, G.A. Olah, Method of oriented circular dichroism, Biophys. J. 57 (1990) 797–806.

[24] S. Frankel, R. Sohn, L. Leinwand, The use of sarkosyl in generating soluble protein after bacterial expression, Proc. Natl. Acad. Sci. U. S. A. 88 (1991) 1192–1196.

[25] B. Miroux, J.E. Walker, Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels, J. Mol. Biol. 260 (1996) 289–298.

[26] M.D. Bazzi, R.W. Woddy, Interaction of amphipathic polypeptides with phospholipids: characterization of conformations and the CD spectra of oriented β-sheets, Biopolymers 26 (1987) 1115–1124.

[27] T.M. Weiss, L. Yang, L. Ding, A.J. Waring, R.I. Lehrer, H.W. Huang, Two states of cyclic antimicrobial peptide RTD-1 in lipid bilayers, Biochemistry 41 (2002) 10070–10076.

[28] M.I. Mikhaleva, C.L. Santini, G. Giordano, M.A. Nesmeyanova, L.F. Wu, Requirement for phospholipids of the translocation of the trimethylamine N-oxide reductase through the Tat pathway in *Escherichia coli*, FEBS Lett. 463 (1999) 331–335.

[29] R. Chenna, H. Sugawara, T. Koike, R. Lopez, T.J. Gibson, D.G. Higgins, D.J. Thompson, Multiple sequence alignment with the Clustal series of programs, Nucleic Acids Res. 31 (2003) 3497–3500.