

Reinhard Albrecht,^a Monika Schütz,^b Philipp Oberhettinger,^b Michaela Faulstich,^c Ivan Bermejo,^d Thomas Rudel,^c Kay Diederichs^e and Kornelius Zeth^{a,d,f,*}

^aDepartment of Protein Evolution, Max-Planck-Institut für Entwicklungsbiologie, Spemannstrasse 35, 72076 Tübingen, Germany, ^bInstitut für Medizinische Mikrobiologie und Hygiene, Elfriede-Aulhorn-Strasse 6, 72076 Tübingen, Germany, ^cDepartment for Microbiology, Biocenter of the University of Würzburg, Am Hubland, 97074 Würzburg, Germany, ^dUnidad de Biofísica (CSIC-UPV/EHU), Barrio Sarriena s/n, 48940 Leioa, Vizcaya, Spain, ^eDepartment of Biology, University of Konstanz, 78457 Konstanz, Germany, and ^fIKERBASQUE, Basque Foundation for Science, Bilbao, Spain

Correspondence e-mail:
kornelius.zeth@gmail.com

Structure of BamA, an essential factor in outer membrane protein biogenesis

Outer membrane protein (OMP) biogenesis is an essential process for maintaining the bacterial cell envelope and involves the β -barrel assembly machinery (BAM) for OMP recognition, folding and assembly. In *Escherichia coli* this function is orchestrated by five proteins: the integral outer membrane protein BamA of the Omp85 superfamily and four associated lipoproteins. To unravel the mechanism underlying OMP folding and insertion, the structure of the *E. coli* BamA β -barrel and P5 domain was determined at 3 Å resolution. These data add information beyond that provided in the recently published crystal structures of BamA from *Haemophilus ducreyi* and *Neisseria gonorrhoeae* and are a valuable basis for the interpretation of pertinent functional studies. In an 'open' conformation, *E. coli* BamA displays a significant degree of flexibility between P5 and the barrel domain, which is indicative of a multi-state function in substrate transfer. *E. coli* BamA is characterized by a discontinuous β -barrel with impaired β 1– β 16 strand interactions denoted by only two connecting hydrogen bonds and a disordered C-terminus. The 16-stranded barrel surrounds a large cavity which implies a function in OMP substrate binding and partial folding. These findings strongly support a mechanism of OMP biogenesis in which substrates are partially folded inside the barrel cavity and are subsequently released laterally into the lipid bilayer.

Received 19 November 2013

Accepted 3 April 2014

PDB reference: BamA, 4c4v

1. Introduction

Outer membrane protein (OMP) biogenesis is essential in order to maintain the indispensable functions of the bacterial outer membrane (OM). Outer membranes act as a protective physical barrier and at the same time facilitate transmembrane trafficking of nutrients and signalling molecules. Thus, the proper targeting, folding and insertion of a variety of different integral β -barrel proteins are essential for the viability of Gram-negative bacteria (Hagan *et al.*, 2011; Knowles *et al.*, 2009; Ricci & Silhavy, 2012). Owing to their evolutionary relationship, the machinery facilitating OMP biogenesis exists in Gram-negative bacteria, mitochondria and chloroplasts. In bacteria, the complex in charge of substrate recognition and processing is referred to as the β -barrel assembly machinery (BAM; Noinaj *et al.*, 2013; Voulhoux *et al.*, 2003). The central component of this multiprotein complex, BamA, consists of an integral outer membrane protein and belongs to the conserved Omp85 superfamily. BamA is composed of a membrane-embedded C-terminal β -barrel domain and a variable number of N-terminal polypeptide transport-associated (POTRA) domains which locate to the periplasmic space (Gentle *et al.*, 2005; Schleiff & Becker, 2011; Zeth & Thein, 2010).

In *Escherichia coli* the Bam complex consists of five proteins, BamA–BamE, which appear to interact *via* various

protein–protein interactions (Knowles *et al.*, 2009; Malinverni *et al.*, 2006; Albrecht & Zeth, 2011). Two proteins of this

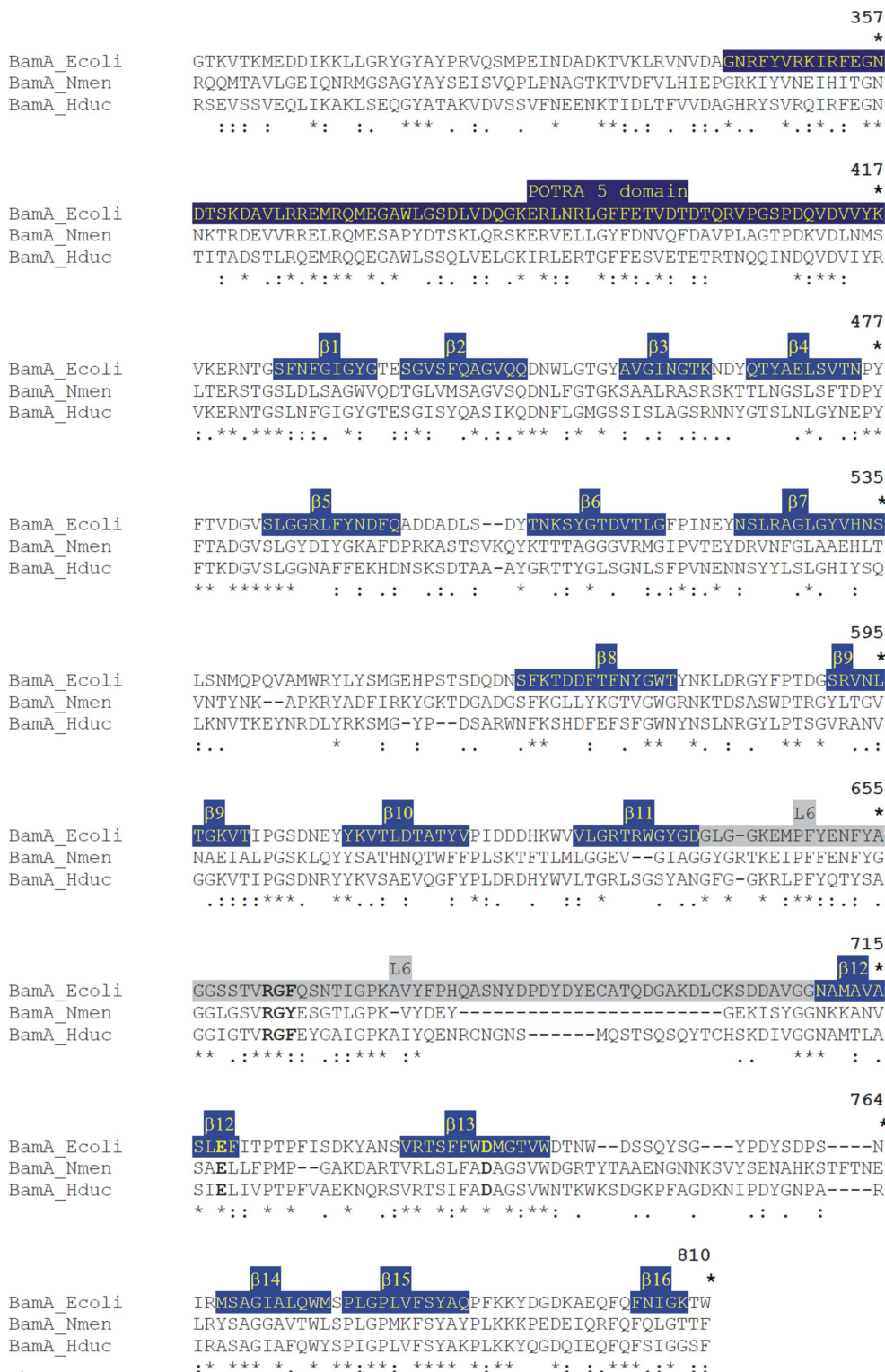


Figure 1

Sequence alignment of the P5 and OMP domains of the BamA proteins from *E. coli*, *N. meningitidis* and *H. ducreyi*. Structures of BamA from these different organisms have recently been determined. The alignment displays their variability in sequence. The alignment was performed using the *MUSCLE* server (<http://www.ebi.ac.uk/Tools/msa/muscle/>) and identical positions in the alignment are indicated by asterisks, while strongly homologous residues are marked by colons. Weakly homologous residues are marked by single dots. The POTRA5 domain sequence is colour-coded in dark blue, while β-strands of the barrel are highlighted in pale blue and numbered accordingly (β1–β16). The elongated loop L6 is depicted in grey.

complex, namely BamA and BamD, are essential for cell survival, while BamB, BamC and BamE modulate complex activity and stability (Kim *et al.*, 2007; Ruiz *et al.*, 2005; Sklar *et al.*, 2007; Wu *et al.*, 2005). In contrast to recent studies using *E. coli*, which demonstrated that all of the POTRA domains are required for normal growth (Browning *et al.*, 2013), it has previously been shown that in *Neisseria meningitidis* POTRA domains 1–4 are dispensable for protein function, while POTRA domain 5 (P5) is essential for cell survival (Bos *et al.*, 2007). This finding emphasizes that there are differences in BamA between organisms and indicates the need to determine BamA crystal structures from several model organisms. Additional mutational data obtained for *E. coli* BamA *in vivo* indicated that a conserved RGF motif in the longest loop L6 is important for the correct function of BamA (Clantin *et al.*, 2007; Rigel *et al.*, 2013; Browning *et al.*, 2013).

Crystal structures of the N-terminal part of BamA (Gatzeva-Topalova *et al.*, 2008, 2010; Kim *et al.*, 2007; Knowles *et al.*, 2007; Zhang *et al.*, 2011) and of BamB–BamE (Albrecht & Zeth, 2011; Dong, Yang *et al.*, 2012; Jansen *et al.*, 2012; Kim & Paetzel, 2011; Noinaj *et al.*, 2011; Dong, Hou *et al.*, 2012; Sandoval *et al.*, 2011; Endo *et al.*, 2011; Kim *et al.*, 2011; Knowles *et al.*, 2011) have been solved in recent years. POTRA-domain constructs of BamA from *E. coli* have been studied by NMR and SAXS (Gatzeva-Topalova *et al.*, 2008; Kim *et al.*, 2007; Zhang *et al.*, 2011). These studies together showed differing quaternary arrangements of the five domains, which indicate flexibility to be one important feature of the entire periplasmic portion of BamA. Crystal structures of BamA from *Haemophilus ducreyi*

(lacking three POTRA domains) and *N. gonorrhoeae* (full length; Noinaj *et al.*, 2013) as well as of the structurally related *E. coli* TamA, a protein involved in translocation of auto-transporters (Gruss *et al.*, 2013), have recently become available. The amino-acid sequence alignment in Fig. 1 displays the analogies and differences between these three organisms; when compared with the *E. coli* sequence, *H. ducreyi* BamA has 51% identity and *N. gonorrhoeae* BamA has 34% identity. The most striking differences are visible in loop 6 (L6; coloured grey) and in the terminal β -strand (β 16), two important structural features which have implications for the suggested mechanism of BamA, as we will discuss later here. Previous work on BamA has revealed important features in structure and function. (i) Two different conformations of the POTRA domains exist relative to the β -barrel. One conformation renders the barrel cavity closed, while the other leads to an open conformation in which the barrel pore is freely accessible from the periplasmic side. (ii) In *N. gonorrhoeae* BamA, the first and the last β -strands (which close the barrel) only interact by two hydrogen bonds, whereas in most other β -barrels this interaction is characterized by the presence of multiple hydrogen bonds. This weak interaction poses the possibility of a lateral barrel-opening event, for example in order to release OM substrates into the lipid bilayer or to expose β -strands to enable the folding of nascent OMPs by β -augmentation. Molecular-dynamics simulations have indeed demonstrated that a lateral opening of both *N. gonorrhoeae* and *H. ducreyi* BamA is likely to occur. (ii) The aromatic girdle is a characteristic trait of β -barrel proteins. It creates a hydrophobic surface anchoring the barrel at the lipid interface. Crystal structures of both *H. ducreyi* and *N. gonorrhoeae* BamA revealed a reduced hydrophobic surface on one side of the barrel. This area is localized in close proximity to β -strand 16. A disruption of the hydrophobic interface at this position might induce local distortions of the lipid bilayer and also a reduction in membrane thickness. Both effects might facilitate the efficient insertion of substrates into the OM.

Taken together, these findings support the following model of BamA-mediated OMP biogenesis, which represents a synopsis of several models that have recently been discussed: nascent OMPs interact with the POTRA domains and L6. Thereby, the OMP is threaded into the BamA barrel. Subsequent folding of the nascent OMP provides the energy to disrupt the hydrogen bonds closing the BamA barrel. Finally, the OMP is released laterally into the OM. Local distortions of the lipid bilayer may additionally support this release. Another conceivable mechanism is a direct folding of OMPs into the locally distorted lipid bilayer, which does not involve entry and lateral release by the BamA barrel.

As most biochemical and molecular-biological studies analysing the mode of action of BamA use *E. coli* as a model system, we set out to solve the crystal structure of *E. coli* BamA. To this end, we crystallized a construct comprising the C-terminal β -barrel domain and POTRA domain 5 (BamA_{P5}). During the preparation of our manuscript, Noinaj and coworkers reported the crystal structures of *N. gonorrhoeae* and *H. ducreyi* BamA (Noinaj *et al.*, 2013). However, sequence

alignment of *H. ducreyi*, *N. gonorrhoeae* and *E. coli* BamA reveals numerous differences which might be of critical importance for further analysis and will be described and discussed here. Taken together, the structure of *E. coli* BamA displays an overwhelming number of interesting details. This knowledge may serve as a basis for a consistent characterization of the final step in OMP biogenesis through lateral substrate release.

2. Experimental procedures

2.1. Cloning of BamA_{P5} from *E. coli*

The *bamA*_{P5} fragment consisting of the β -barrel domain and the fifth POTRA domain (residues 344–810) of *E. coli* BL21 (DE3) was amplified by PCR using chromosomal DNA as a template, adding a His tag at the C-terminal end. Primer sequences are available upon request. The PCR product was digested with *Nde*I and *Xho*I and cloned under the control of an IPTG-inducible promoter of the pET-30b expression vector (Novagen, Darmstadt, Germany). The construct was verified by DNA sequencing.

2.2. Purification of BamA_{P5} from inclusion bodies

The protein was prepared by expression in the bacterial cytoplasm and isolation and purification of the protein as inclusion bodies. *E. coli* BL21(DE3) cells harbouring the BamA_{P5} plasmid were grown at 310 K in lysogeny broth containing 50 mg kanamycin per litre of cell culture. Cells were harvested by centrifugation at 20 000g, resuspended in 150 mM NaCl, 20 mM Tris pH 8 (buffer *X*) containing 1 mg ml⁻¹ lysozyme and DNase and broken using a French press. The suspension was centrifuged at 75 000g for 20 min to remove soluble proteins. The pellet was washed twice with buffer *X* before resuspension in buffer *X* containing 5 mg ml⁻¹ lysozyme and 1% lauryldimethylamine *N*-oxide (LDAO). This solution was stirred for 3 h to allow entire degradation of the bacterial peptidoglycan and solubilization of the inner and outer membranes. After centrifugation of the solution at 75 000g for 20 min, this step was repeated one more time in the absence of lysozyme followed by another centrifugation step. Finally, the pellet was washed with buffer *X* to remove the detergent and pellet at 75 000g for 20 min. Homogenization of the pellets was carefully performed using a hand homogenisator.

BamA_{P5} inclusion bodies were solubilized in denaturing buffer (6 M Gua-HCl, 20 mM Tris pH 8.5). After centrifugation and filtration (pore size 0.22 μ m), the solution was loaded onto a 20 ml Ni-NTA column. Unbound material was removed by washing the column with denaturing buffer containing 20 mM imidazole. BamA_{P5} was eluted with denaturing buffer containing 300 mM imidazole and displayed a purity of ~98% as judged by SDS-PAGE. For refolding, the eluate was added dropwise to a tenfold volume of 20 mM Tris, 0.5% LDAO pH 8.4 at 281 K while stirring. After 12 h, the solution was further diluted using 20 mM Tris pH 8.4 until an LDAO concentration of 0.2% was reached, and the resulting

Table 1

Data-collection and refinement statistics for BamA_{P5}.

Values in parentheses are for the highest resolution shell.

Data collection	
Space group	$P2_1$
Unit-cell parameters (Å, °)	$a = 79.58, b = 67.35, c = 109.79,$ $\alpha = 90, \beta = 93.46, \gamma = 90$
Resolution (Å)	47–3.0 (3.18–3.00)
R_{merge}	0.15 (0.99)
$\langle I/\sigma(I) \rangle$	8.6 (1.3)
Completeness (%)	99.6 (99.0)
Multiplicity	4.6 (4.75)
Refinement	
Resolution (Å)	47–3.0 (3.12–3.00)
No. of reflections	28588
$R_{\text{work}}/R_{\text{free}}$	0.25/0.29 (0.26/0.31)
No. of atoms	
Protein	7026
Water	51
B factors (Å ²)	
Protein	98.1
Water	105.3
R.m.s. deviations	
Bond lengths (Å)	0.009

solution was slowly passed through a 20 ml Ni–NTA column. After loading, the column was washed with ten column volumes of 20 mM Tris, 0.5% octyl-polyoxyethylene (o-POE) to exchange the detergent. BamA_{P5} was eluted with 20 mM Tris, 300 mM imidazole, 0.5% o-POE pH 8.4. The protein eluted from the column was concentrated to ~20 mg ml⁻¹ using a 50 kDa molecular-weight cutoff membrane (Amicon Ultra, Millipore). The concentrated protein solution was divided into portions, which were separately dialyzed through a 25 kDa membrane against 20 mM Tris pH 8.5 containing either 0.2% LDAO, 0.1% dodecyl-maltoside (DDM) or 1% octyl-glucoside (OG), which was added prior to dialysis into the dialysis chamber and the surrounding buffer. The purity of the dialyzed proteins was again confirmed by SDS–PAGE (data not shown).

2.3. Crystallization of BamA_{P5}

Concentrated BamA_{P5} (16 mg ml⁻¹) with one of the three detergents LDAO, DDM or OG was screened against ten different crystallization screens comprising 96 conditions each (Qiagen). Protein-drop setup was performed by a Honeybee 961 robot (Genomic Solutions), which mixed 0.4 µl protein solution with 0.4 µl reservoir solution on 96:3 Intelli-Plate sitting-drop plates (Art Robbins). After 5 d, small crystals appeared with LDAO under two conditions from The PEGs II Suite [condition Nos. 4 (200 mM sodium citrate, 15% PEG 400, 100 mM Tris pH 8.5) and 12 (200 mM sodium citrate, 30% PEG 400, 100 mM Tris pH 8.5)]. The diffraction of these initial crystals did not extend beyond 5 Å resolution. However, condition No. 12 appeared to be promising for crystal quality improvement and was manually varied using hanging drops, yielding disc-shaped oval crystals with a significantly increased size but that were mostly embedded in precipitate. The best diffracting crystals were obtained with a reservoir solution consisting of 21% PEG 400, 0.5 M sodium citrate, 100 mM Tris pH 8.5. Crystals were briefly immersed in reservoir solution

supplemented with 20% glycerol/LDAO and were then flash-cooled in liquid nitrogen. A data set comprising 400 images of 0.5° rotation was collected on beamline X10SA at the Swiss Light Source, Villigen, Switzerland at 100 K using a Pilatus 6M detector (Dectris).

2.4. Structure determination of BamA_{P5}

Crystallographic images were integrated and scaled using the *XDS/XSCALE* package (Kabsch, 2010). Initial molecular-replacement trials using the FhaC structure as a search model failed. However, when the coordinates of BamA from *H. ducreyi* became available we succeeded in solving the structure by molecular replacement using *MOLREP* and *Phaser* and could subsequently place the different subdomains (Vagin & Teplyakov, 2010; Zwart *et al.*, 2008). Refinement of the structure was initially performed using the *PHENIX* package, but during refinement it turned out that *BUSTER* yielded a better geometry of the structure and refinement was completed with two protein models in the asymmetric unit to an R and R_{free} of 25 and 29%, respectively, at 3 Å resolution (Smart *et al.*, 2012). The geometry of the structure was analyzed by *MolProbity* (<http://molprobity.biochem.duke.edu>) and the refinement values are given in Table 1.

2.5. Normal-mode analysis

Normal-mode analysis of BamA_{P5} using the *elNémo* server was used to identify potential structural changes based on the B -factor distribution of the residues (<http://www.igs.cnrs-mrs.fr/elneemo/index.html>). The parameters used for the low-frequency mode calculations were $\text{DQMIN} = -300$, $\text{DQMAX} = 300$ and $\text{DQSTEP} = 20$. Three different conformations of the undisturbed protein model in orange ($\text{DQMIN} = 0$), the weakly disturbed model in green ($\text{DQMIN} = -140$) and a more strongly disturbed model in blue ($\text{DQMIN} = -300$) are shown in Fig. 2(b).

3. Results

3.1. Protein purification, crystallization and structure solution of BamA_{P5}

The BamA_{P5} protein was produced in *E. coli* as inclusion bodies, which were carefully purified from remaining cellular proteins before being refolded using LDAO as a detergent. Protein purification of the refolded protein was conducted in a single step using Ni–NTA affinity chromatography. Exchange into the detergent o-POE with a higher critical micellar concentration was performed on the same column. Further detergent substitution into LDAO, DDM or OG to screen for detergents suitable for crystallization was carried out by dialysis and the protein was successfully crystallized in the presence of 0.2% LDAO. Crystallization of the protein was achieved using 21% PEG 400, 500 mM citrate, 100 mM Tris–HCl pH 8.5. BamA_{P5} crystals were formed in space group $P2_1$ and diffracted to 3 Å resolution. The phase problem was solved by molecular replacement using the BamA structure from *H. ducreyi* (Noinaj *et al.*, 2013) and the P5 domain

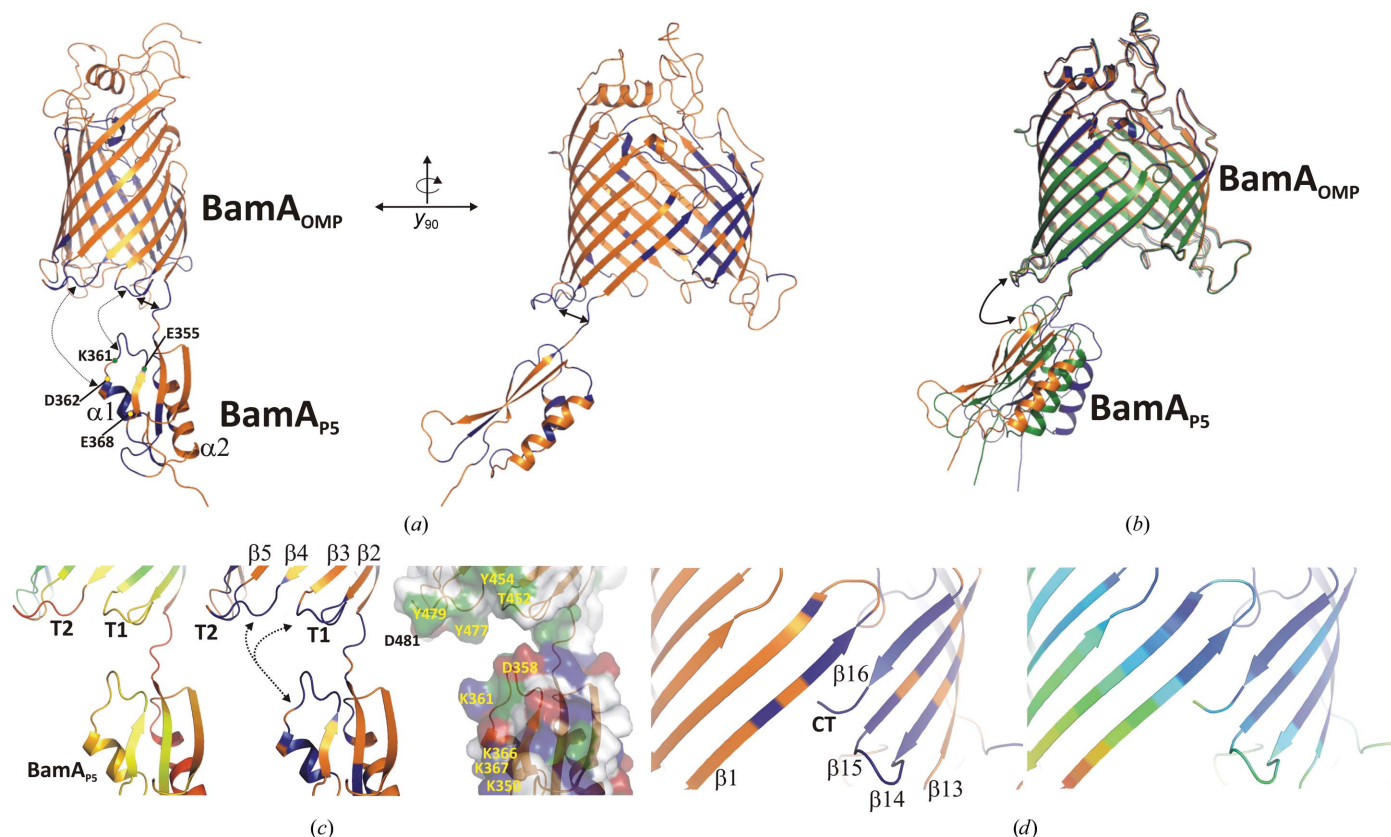


Figure 2

Analysis of conserved residue patterns and *B*-factor distribution in *E. coli* BamA_{P5}. Conserved residues colour-coded in dark blue were identified on the basis of the sequence alignment shown in Supplementary Fig. S2 and subsequently mapped onto the structure. (a) The structure in orange is shown from the narrow side of the barrel. The hinge region between the BamA_{P5} and BamA_{OMP} domains may allow flexible movements between these domains as indicated by arrows (left panel). The right panel depicts the same structure rotated around the *y* axis by 90°. (b) While there are minor interactions between BamA_{OMP} and BamA_{P5} in the undisturbed structure, the third elastic network model from the *eINémo* analysis in blue exhibits interactions between L1/L2 and the POTRA domain. (c) The BamA structure illustrated from essentially the same orientation as in (b). Conserved residues of the structure are marked in dark blue according to the conservation pattern presented in Supplementary Fig. S2. The left panel shows a close-up of the BamA_{P5} and BamA_{OMP} domain interfaces and indicates increased flexibility in turns 1 and 2 (T1 and T2) and helix 1 of BamA_{P5} (α 1) according to the analysis of *B* factors (*B* factors are represented using rainbow colours; small *B* factors are marked in blue, large *B* factors are represented in red). The middle panel shows the conserved residues (as in Fig. 3a) of BamA_{P5} as well as the residues of T1 and T2 colour-coded in dark blue in the same orientation. In the right panel selected conserved residues are marked by numbers and colours (green, amphipatic residues; red, negatively charged residues; blue, positively charged residues). In BamA_{P5} several conserved residues are predominantly positively charged, for example Lys350, Lys361, Lys366 and Lys367, while the conserved residues in BamA_{OMP} are mostly amphipatic. This distribution reflects the residues accounting for the interaction between FhaC_{OMP} and FhaC_{P1-2}, where positively charged residues in FhaC_{P1-2} contact hydrophilic residues in the barrel domain. (d) Analysis of conservation (left panel) and *B*-factor distribution (right panel) at the β 1/ β 16 zip-like interface of BamA_{OMP}. Residues of strands β 1 and β 12– β 16 are particularly conserved. The majority of the β -barrel displays lower *B* factors and consequently a decreased mobility in comparison to the POTRA domain, turns and loop structures.

structure from *E. coli* (Gatzeva-Topalova *et al.*, 2008) as models. Initially, the two OMP domains of *H. ducreyi* were placed in the asymmetric unit and refined before the P5 domain was localized in difference maps. The asymmetric unit contains two monomers which face each other from their wide side, which is stabilized mostly through hydrophobic contacts (Supplementary Fig. S1a¹). The two monomers of the asymmetric unit are structurally almost identical, with an r.m.s.d. of 0.7 Å for the C α positions of 442 aligned residues (Supplementary Fig. S1b). The monomer with the slightly higher structural completeness and order was used in the description and discussion.

3.2. The architecture of the POTRA-5–OMP interface differs between *E. coli*, *H. ducreyi* and *N. meningitidis* BamA

The extended orientation between POTRA domain 5 and the OMP module is only stabilized through weak intramolecular interactions between conserved residues of the inter-domain element (connecting P5 and the OMP domain) and turn 1 (T1) of the OMP domain which links strands β 2 and β 3 (Figs. 2a and 2c). Residues of the inter-domain element contributing hydrogen bonds are Asn448 (to Lys419 and Asn422) and Asn422 to Asp447. Further stabilization of POTRA domain 5 is achieved through intermolecular interactions with loop structures of symmetry-related molecules. This conformation is likely to resemble one of several transient conformations which might be adopted by the POTRA domains (Fig. 2b). High crystallographic temperature factors

¹ Supporting information has been deposited in the IUCr electronic archive (Reference: MH5121).

of the P5 domain and T1/T2 support this flexible domain arrangement and are also in line with normal-mode analysis of the protein domain movements indicating a mobile interface (Fig. 2c). This interface, including T1 and T2 and two small segments of P5 (Fig. 2c), is strongly conserved in BamA-like proteins and therefore indicates a functionally important interface (Supplementary Fig. S2). Further strongly conserved residues of BamA_{OMP} with implications for BamA function and mechanism are found in strands $\beta 1$ and $\beta 12$ – $\beta 16$ and the elongated loop L6 (Fig. 1 and Supplementary Fig. S2). Most of these conserved residue patches can be assigned to specific functions or to structural maintenance of the β -barrel fold (Supplementary Fig. S2). In *N. meningitidis* BamA the stability of the entire periplasmic POTRA domain is maintained by interactions between POTRA domain 1 and the loop structure folding inside the barrel and connecting $\beta 8$ and $\beta 9$. The elongated part of the POTRA domain is stabilized by inter-

actions with the adjacent POTRA and barrel domains of two symmetry-related proteins in the crystal packing. In the *H. ducreyi* BamA $\Delta 3$ structure the stability of the three POTRA domains is ensured by interactions between POTRA domain 1 and the small periplasmic loop structure connecting $\beta 1$ and $\beta 2$. Furthermore, POTRA domain 4 forms an extended network of interactions with the extraplasmic loop structure of a symmetry-related molecule (see Fig. 5a).

3.3. The OMP domain represents a weakly closed β -barrel forming a large cavity

The structure of BamA_{OMP} is formed by a barrel of 16 β -strands with variable length (5–13 residues). The ellipsoidal barrel form shows a strong similarity to the BamA structures from *N. meningitidis* and *H. ducreyi* (r.m.s.d.s of 2.7 Å for 353 C α positions and 1.6 Å for 367 C α positions, respectively;

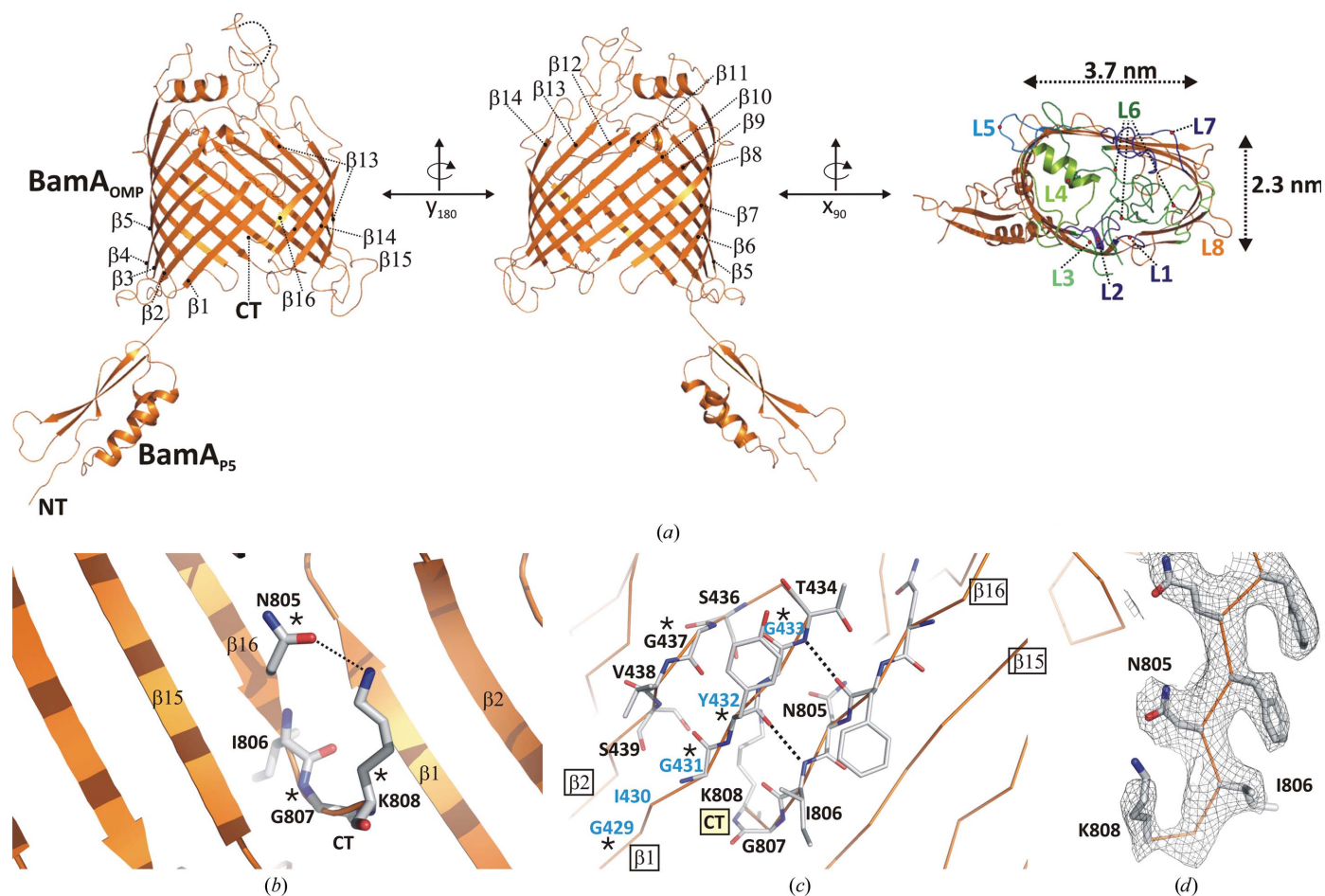


Figure 3

The crystal structure of BamA_{P5} displays a discontinuous β -barrel. (a) The structure of BamA_{P5} shown as a cartoon model in orange from opposite sides related by a rotation around the y axis of 180°. The POTRA and the OMP domains are termed BamA_{P5} and BamA_{OMP}, respectively; β -strands ($\beta 1$ – $\beta 16$) and the termini (NT and CT) are indicated. On the right the structure is shown as a top view from the extracellular side with the loops (defined as L1–L8) colour-coded in blue and green tones. (b) Enlarged view of the loosely closed barrel moiety at the transition between $\beta 1$ and $\beta 16$ demonstrating the unusual organization of the C-terminus folding inside the barrel interior. This conformation is stabilized by residue Lys808 shown in stick representation through the formation of a hydrogen bond to the $\beta 16$ residue Asn805 (marked by a dashed line). Conserved residues are marked with asterisks. (c) View onto the $\beta 1/\beta 2/\beta 16$ strands in line representation. Two hydrogen-bond interactions between $\beta 1$ and $\beta 16$ lead to the weak closure of the β -barrel. Residues are shown in stick representation and hydrogen bonds between interacting residues are shown as dashed lines. The GIGYG motif in $\beta 1$ is marked with blue labels. All figures were prepared using PyMOL (<http://www.pymol.org>). (d) $2F_o - F_c$ OMIT map of the C-terminus calculated at 1σ with residues 802–808 removed from the refinement and phasing. The backbone is shown as a ribbon while side chains are shown in stick representation.

Noinaj *et al.*, 2013) and a significantly lower structural similarity to the two-component protein-export machinery FhaC (r.m.s.d. of 3.1 Å for 259 C $^{\alpha}$ positions; Clantin *et al.*, 2007). The insertase structure appears with an open cavity laterally surrounded by hydrophobic barrel residues and vertically by the extended loop structures (L3 and L6–L8) which tightly close the barrel for the incoming substrate (Figs. 3*a*, 3*b* and Fig. 4). The very C-terminus of BamA_{OMP} comprising the

conserved aromatic residue Trp810 that is ordered in all OMP proteins is disordered and the terminus folds backwards into the barrel interior (Fig. 3*b*). In *E. coli* BamA this particular conformation is stabilized through hydrogen bonding between Lys808 (the last structured residue) and Asn805 (Figs. 3*b* and 3*c*). The interface between β 1 and the short β 16 strand is weak and is stabilized only by two main-chain hydrogen bonds (Fig. 3*c*). Thus, the C-terminus of BamA_{OMP} could not only

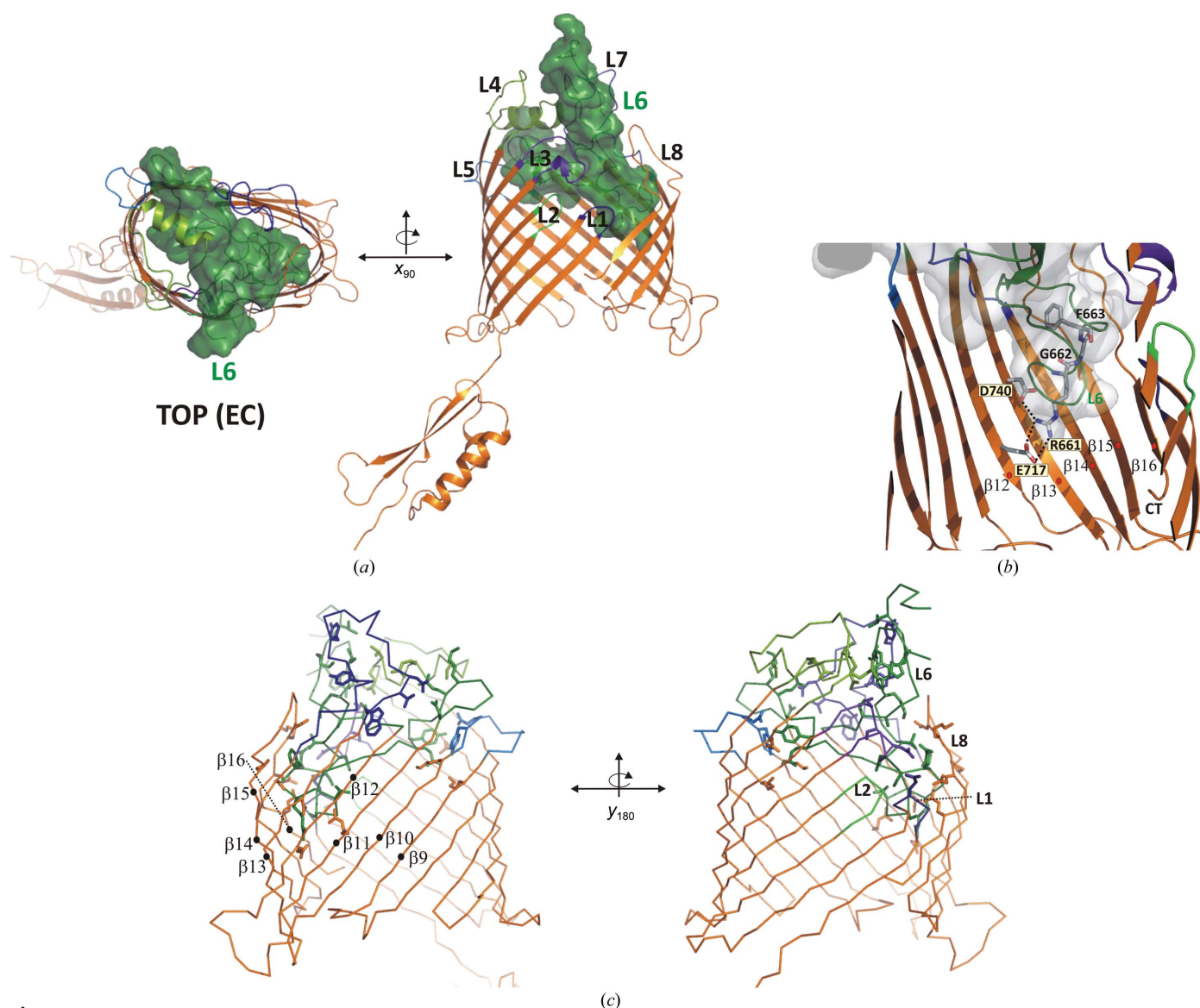


Figure 4

The L6 loop of BamA_{P5} shows an extended and very complex fold. (a) The two panels of this figure represent independent views of BamA from the top (facing the extracellular space; EC) and from the side. The views are related to each other by a rotation of the protein by 90° around the x axis. The loops L1–L5 and L7–L8 are colour-coded in blue and green tones, while the L6 loop is colour-coded in forest green and shown as a surface representation. The extension and volume of the loop leads to a complete closure of the barrel at the side exposed to the extracellular space. Mutational data of the highly conserved RGF motif of L6 have shown the influence of this loop in BamA_{P5} in supporting folding of outer membrane proteins and BamA biogenesis. (b) A conserved and functionally essential RGF motif in loop L6 attaches the loop tip to the barrel wall. The figure shows the BamA protein in an orange ribbon representation. All loops are shown in the same colours as in (a). The L6 loop is shown in forest green surrounded by a surface representation of the loop in grey. Residues involved in the loop–barrel interaction are marked with numbers (Glu717 and Asp740 on the barrel and Arg661 on the loop tip). All of these residues are strongly conserved as shown in the alignment in Supplementary Fig. S2. Besides a number of additional interactions, the L6 loop is strongly stabilized in this configuration by the two salt bridges marked by dashed lines. (c) The structure of BamA_{P5} in ribbon representation with residues of the loop L6 marked in stick representation. Side view of BamA_{P5} with the colour-coding used in (a). Residues on loop L6 and the interacting structures (loops and β -strands) are marked in stick representation. Strands with residues interacting with the L6 loop are marked (β 9– β 16). The structure is shown from two perspectives related by a rotation of 180° around the y axis.

function as a targeting signal for membrane insertion, but could also play a role in the lateral gating event. The stability of the $\beta 1$ strand is further decreased through a GIGYG motif increasing flexibility at this position together with a small β -bulge on $\beta 2$ (residues Ser436–Ser439) which does not allow saturated intra-strand $\beta 1/\beta 2$ hydrogen-bond pattern formation, similar to the porins from *Comamonas acidovorans* and *N. gonorrhoeae* (Zeth *et al.*, 2000, 2013). A representative OMIT map is shown in Fig. 3(d).

3.4. The loop L6 forms an extended structure in BamA_{P5} closing the extraplasmic side

Another functionally important motif of BamA is the extended loop structure of L6, which closes the extracellular

side of the β -barrel (Fig. 4a). This elongated loop of 68 residues (Asp639–Asn706) spans the entire width of the pore in a complex fashion and forms a variety of mostly conserved interactions with the barrel and interactions stabilizing the particular three-dimensional conformation (Fig. 4a and Supplementary Fig. S2). One loop motif comprising residues Arg661–Phe663 was shown to be functionally essential for OMP and BamA folding (Fig. 4b; Leonard-Rivera & Misra, 2012). Based on the structure, the strongly conserved Arg661 residue appears to be particularly important for the attachment of the loop to the barrel wall through interactions with the conserved residues Glu717 (on $\beta 13$) and Asp740 (on $\beta 14$) (Fig. 4b). In contrast to FhaC, in which the homologous loop folds back to the periplasmic space, this loop in BamA does not strongly contribute to a narrowing of the pore diameter.

Owing to the number of inter-loop and intra-loop interactions and the low *B* factors of the L6 residues embedded in the OMP barrel, it appears plausible that this loop has an important function in the formation of a plug which firmly closes the barrel structure (Fig. 4). Furthermore, based on the numerous interactions of L6 with β -sheets and surrounding loop structures, L6 has an important function in the overall architectural maintenance of the protein (Fig. 4c).

4. Discussion

Outer membrane protein biogenesis is the process which ultimately leads to the folding and insertion of outer membrane proteins (OMPs) as a result of multiple interactions and transfer steps of the preprotein with periplasmic chaperones and the BAM complex. Following these multiple interactions and the formation of a transient complex between the substrate and BamA, the OMP is released into the outer membrane.

Two structures of BamA from *N. gonorrhoeae* (the full-length protein) and *H. ducreyi* (BamA $\Delta 3$; a construct missing three POTRA domains) have recently been solved by X-ray crystallography and have allowed insights into the terminal mechanistic step by which OMP proteins could be assembled and

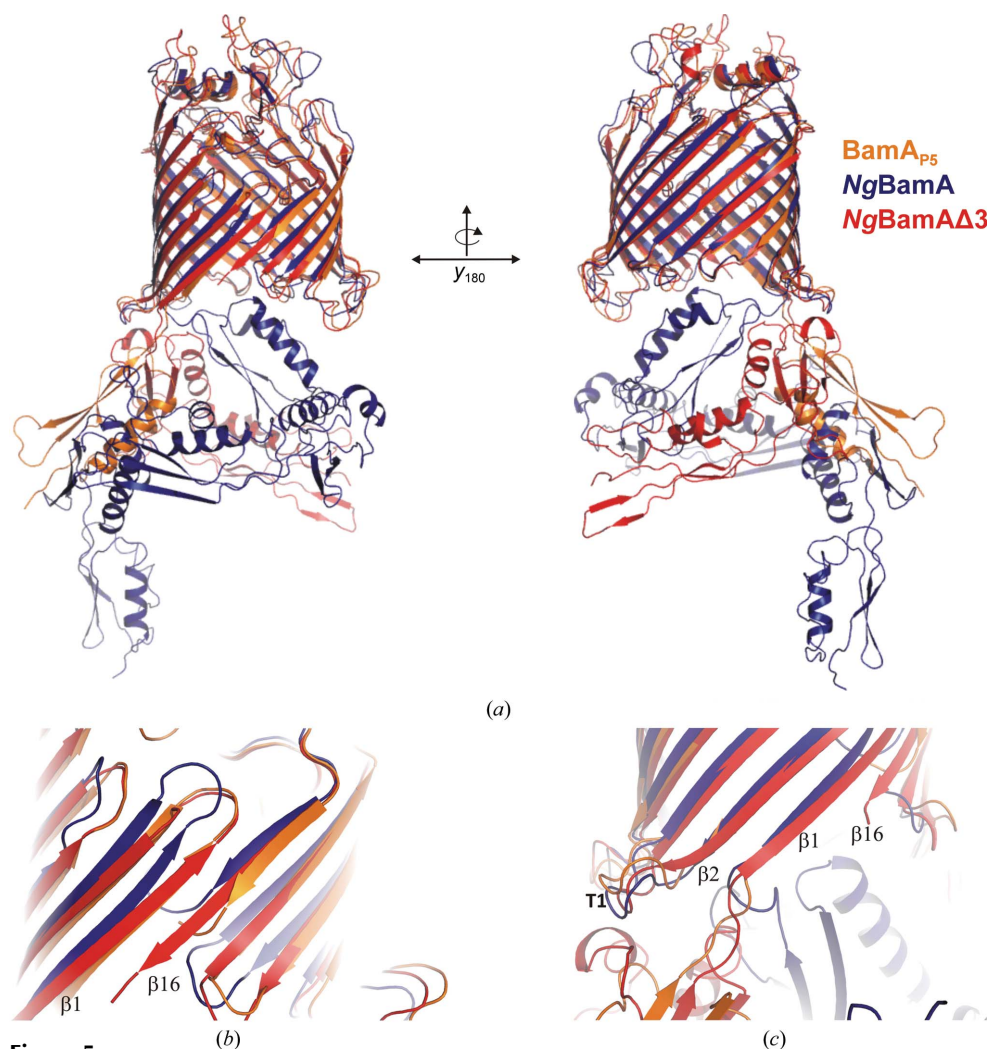


Figure 5 Structure comparison of BamA from *E. coli*, *N. gonorrhoeae* and *H. ducreyi*. (a) The structures are depicted in cartoon representation. The superposition of the structures resulted in r.m.s.d.s of 2.7 Å for the protein derived from *N. gonorrhoeae* (NgBamA; 353 residues were superimposed) and 1.6 Å for the protein derived from *H. ducreyi* (HdBamA $\Delta 3$; 367 residues were superimposed). These values coincide well with the sequence relationship of the proteins, with 34% identity between BamA_{P5} and *N. gonorrhoeae* BamA and 51% identity between BamA_{P5} and *H. ducreyi* BamA $\Delta 3$. (b) Close-up of the $\beta 1/\beta 16$ -strand barrel opening of the three structures in the same orientation as (a) using the same colour-coding as defined in (a). (c) Close-up of the domain boundaries between the OMP and P5 domains (colour-coding as above). The figure shows the different orientations of the POTRA domains relative to the OMP domains of the three BamA structures.

released into the OM (Noinaj *et al.*, 2013). Conformational changes in the POTRA-domain arrangements of the two systems and the weakly locked β -barrel in *H. ducreyi* BamA Δ 3 are indicative of a mechanism which involves the POTRA domain in substrate transfer followed by opening of the barrel domain. These experimental data were underpinned by molecular-dynamics studies, which show the transient opening of the *H. ducreyi* BamA Δ 3 barrel at the β 1/ β 16 transition.

In our complementary study, we present the structure of BamA from *E. coli* comprising the terminal POTRA domain 5 (for a structural comparison of all BamA proteins, see Fig. 5). In analogy to the BamAs from *N. gonorrhoeae* and *H. ducreyi*, *E. coli* BamA is also composed of 16-stranded β -barrel, as recently reinforced for *in vivo*-assembled chimeric BamA proteins (Volokhina *et al.*, 2013), and not of a 12-stranded barrel as predicted earlier (Voulhoux *et al.*, 2003). Our structure not only exhibits a number of distinct features but will also serve as the basis for further biochemical studies specifically obtained using the *E. coli* model system.

One significant difference between the published structures and BamA_{P5} is the orientation of P5 relative to the β -barrel domain, which is maintained by only two hydrogen bonds (see Fig. 3). Moreover, the P5 domain forms crystal contacts to crystallographically related molecules. Notably, the formation of crystal contacts mediated by the POTRA domain is a general feature of all BamA structures, and these contacts may have implications for the quaternary architecture of these structures. As a consequence, the conformational space of the POTRA domain relative to the barrel domains potentially reflects only a subpopulation of states which in solution may be much larger than represented by the three structures. Although the interface between the OMP domain and P5 is strongly conserved and stably maintained in FhaC and

H. ducreyi BamA Δ 3, it becomes obvious from the structures of BamA_{P5} and *N. gonorrhoeae* BamA that this interface can be broken based on weak mechanical forces. Another important feature for BamA function is the β -barrel architecture of BamA. This weakly connected barrel in BamA_{P5} shows similarity to the arrangement observed in *N. gonorrhoeae* BamA, with a weakly connected barrel wall interaction between β 1 and β 16, and supports the idea of lateral substrate release. In contrast to *N. gonorrhoeae* BamA, in which the POTRA architecture blocks access to the hydrophobic cavity while the β 1/ β 16 transition remains weakly formed, in BamA_{P5} the hydrophobic cavity is unlocked in the presence of a weak β 1/ β 16 contribution (Noinaj *et al.*, 2013).

Lateral opening of the barrel as a mechanism in OMP biogenesis has previously been proposed for Tom40, a component of the mitochondrial TOM machinery (translocation of the outer membrane; Rapaport, 2005) and more recently also for the TAM complex (Gruss *et al.*, 2013), which is involved in autotransporter biogenesis (Selkrig *et al.*, 2013), as well as for *N. gonorrhoeae* BamA and *H. ducreyi* BamA Δ 3. This opening may be accompanied by a transient fusion of the two barrels followed by OMP release into the membrane or by release through a narrow gate as demonstrated for the Sec system for the biogenesis of α -helical inner membrane proteins (Lycklama & Driessen, 2012). Such a transient fusion event of two β -barrels could be accomplished by displacing β -strands 1 and 16 of BamA and adding laterally β -strands from the substrate (β -augmentation). Beta-augmentation as an important interaction mechanism for initiation of substrate folding has previously been described for the BamA POTRA domains forming a hydrophobic groove for β -barrel substrates (Kim *et al.*, 2007; Gatzeva-Topalova *et al.*, 2008). In addition, β -augmentation of the BamA barrel with an OMP substrate barrel could also be important for passenger domain secretion

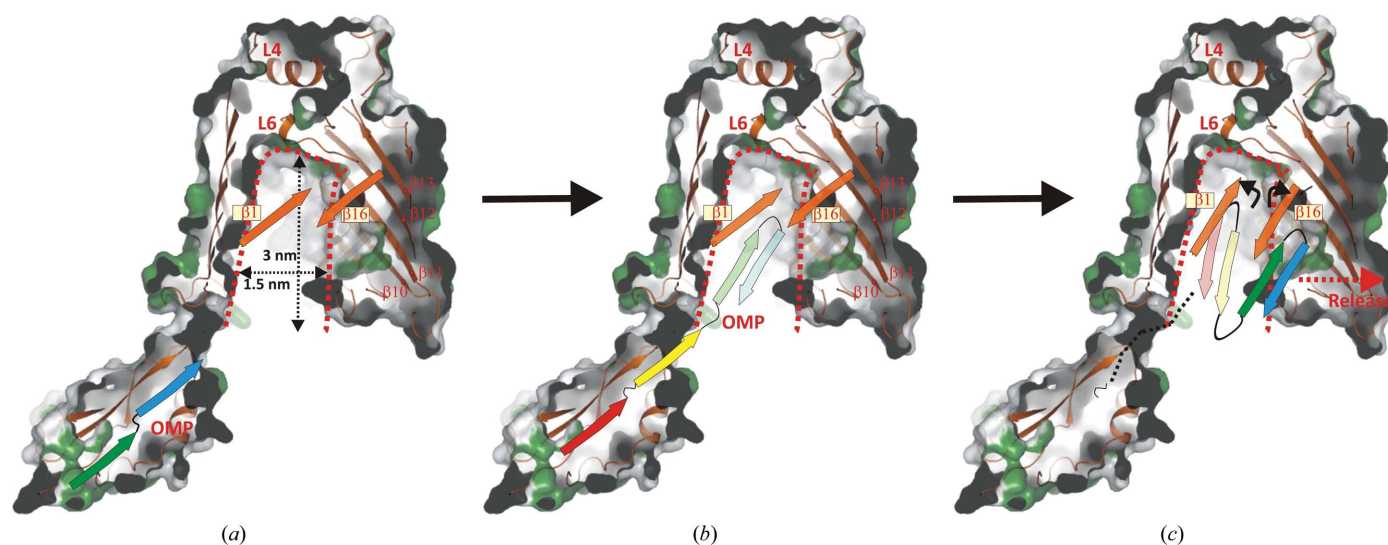


Figure 6

Model of the final step in OMP biogenesis catalyzed by BamA. (a) Unfolded but stabilized OMP protein is aligned at the POTRA interface. (b) Strand (pairs) subsequently entering the cavity provided by BamA with a diameter of $\sim 3 \times 1.5$ nm. Within the cavity, folding is enhanced and the stability of an initial hairpin contributes to the alignment of β -strands. (c) Further accessing β -hairpin elements may lead to additional folding events and laterally force the zip-like transition of the β 1/ β 16 fracture to open temporarily. While part of the protein is released into the membrane, additional parts of the OMP can access the cavity until the entire protein has passed through the folding chamber and can finally be released into the outer membrane.

of autotransporter (AT) proteins. A transient interaction with BamA or TamA, respectively, allows modulation of the flexibility and size of an AT β -barrel, which might also facilitate the secretion of larger or partially folded passenger domains (van Ulsen *et al.*, 2013; Sauri *et al.*, 2009; Soprova *et al.*, 2010; Gruss *et al.*, 2013). In agreement with barrel fusion is the experiment by Rigel *et al.* (2013) which suggested a cyclic mechanism for β -barrel assembly in *E. coli*. They showed two distinct conformations of the extended loop L6: a protease-sensitive surface-exposed conformation and a protease-resistant conformation in which the loop is folded back into the lumen of the BamA β -barrel. Although a surface-exposed state of the L6 loop was not observed for BamA of *E. coli* as well as for BamA of *N. gonorrhoeae* and *H. ducreyi*, such a conformation cannot be excluded as the suggested cyclic mechanism of β -barrel assembly is dependent on BamD and BamE (Rigel *et al.*, 2013).

Although the fusion of two barrels is a charming hypothesis, it remains difficult to explain how the strong interaction between the L6 loop and the surrounding structural topology is broken on a simple energetic basis. In agreement with the low *B* factors, it does not seem to be a plausible mechanism that this structural entity becomes destabilized during substrate processing, because this loop of *E. coli* BamA is in contact with eight β -strands (β 9– β 16) and four adjacent loop structures, and these contacts are mostly conserved. Hence, it is tempting to speculate that the loop primarily fulfils the function of structural architecture maintenance of an extra-cytoplasmic plug (Fig. 4). Moreover, in the particular case of BamA from *E. coli* and *H. ducreyi* BamA Δ 3 the number of hydrogen bonds which may be contributed to the barrel fusion is limited owing to the reduced length of the β 1 and β 16 strands.

BamA is a highly conserved protein of the Omp85 superfamily and is present in bacteria, mitochondria and chloroplasts (Webb *et al.*, 2012). However, the first insight into the molecular details of a member of this superfamily was obtained when the crystal structure of FhaC was determined. FhaC is a member of the two-partner secretion system which mediates the translocation of filamentous haemagglutinin (FHA) into the extraplasmic medium (Clantin *et al.*, 2007). It was shown that the FHA substrate initially interacts with POTRA domain 1 and is further translocated to the FhaC barrel domain which mediates export of the protein into the extracellular space. The export of substrate proteins, in comparison to lateral release as for the Omp85 superfamily, requires another protein architecture with lateral stability and extraplasmic access. This requirement is facilitated through removal of the N-terminal helix and extraplasmic opening, allowing the export of FHA substrate.

In summary, four features of the BamA architecture are important and are very likely to have functional implications: (i) the discontinuous barrel, (ii) the flexible POTRA/OMP arrangement, (iii) the strongly conserved and essential loop L6 and (iv) the flexible arrangement of POTRA domains as determined by NMR, SAXS and X-ray crystallography (Gatzeva-Topalova *et al.*, 2008; Kim *et al.*, 2007; Ward *et al.*,

2009). In addition, a conserved and elongated turn element T4 between β 8 and β 9 constricts access to this cavity from the opposite side of the barrel. Together, these elements allow the plausible mechanism deduced from the insertase structure, explaining OMP protein folding and biogenesis on the basis of the BamA scaffold. POTRA-domain flexibility within POTRA domains and relative to the barrel domain may drive the substrate movement forwards to the OMP barrel domain. The most fragile part of the barrel exists between β 1 and β 16, which may be easily ruptured by small mechanic interactions, for example through the initial formation of β -hairpins in the BamA cavity which can serve as a folding compartment for the incoming OMP barrel. The volume of this cavity in BamA_{P5} is 15 200 Å³ and its dimensions would allow the folding of at least two β -hairpin motifs, which is a spontaneous energetically driven process through simple hydrogen-bond formation within a chaperone-like environment (Fig. 6). Additional hairpins entering the cavity would further extend the protein volume and thereby force the newly folding subdomains to overcome the small energetic barrier of transient barrel opening. Subsequent β -hairpins will be released and diffuse into the membrane to assemble therein until the last β -hairpin has folded and is released from the cavity. The current models do not exclude partial or entire fusion of BamA with the OMP substrate protein and its subsequent release, as suggested by Gruss *et al.* (2013). However, this seems unlikely given the multiple interactions between the extended loop structures and additional interactions of the barrel domain.

This work was funded by the German Science Foundation (DFG; SFB766 to MS, single application ZE522/5-1 to KZ), UKT TÜFF 2160-0-0 to MS and UKT fortune F1433253 to PO. The authors cordially thank the staff of the Swiss Light Source (SLS), Villigen, Switzerland for excellent beamline maintenance and help during data collection. This work was supported by the Max Planck Society. The authors are grateful to Dr Susan Buchanan and her team at the NIH, USA for preliminary access to the PDB files for BamA from *N. gonorrhoeae* and *H. ducreyi*. We would like to thank Liam for critical reading.

References

- Albrecht, R. & Zeth, K. (2011). *J. Biol. Chem.* **286**, 27792–27803.
- Bos, M. P., Robert, V. & Tommassen, J. (2007). *EMBO Rep.* **8**, 1149–1154.
- Browning, D. F., Matthews, S. A., Rossiter, A. E., Sevastyanovich, Y. R., Jeeves, M., Mason, J. L., Wells, T. J., Wardius, C. A., Knowles, T. J., Cunningham, A. F., Bavro, V. N., Overduin, M. & Henderson, I. R. (2013). *PLoS One*, **8**, e84512.
- Clantin, B., Delattre, A.-S., Rucktooa, P., Saint, N., Méli, A. C., Loch, C., Jacob-Dubuisson, F. & Villeret, V. (2007). *Science*, **317**, 957–961.
- Dong, C., Hou, H.-F., Yang, X., Shen, Y.-Q. & Dong, Y.-H. (2012). *Acta Cryst. D* **68**, 95–101.
- Dong, C., Yang, X., Hou, H.-F., Shen, Y.-Q. & Dong, Y.-H. (2012). *Acta Cryst. D* **68**, 1134–1139.
- Endo, T., Kawano, S. & Yamano, K. (2011). *EMBO Rep.* **12**, 94–95.
- Gatzeva-Topalova, P. Z., Walton, T. A. & Sousa, M. C. (2008). *Structure*, **16**, 1873–1881.
- Gatzeva-Topalova, P. Z., Warner, L. R., Pardi, A. & Sousa, M. C. (2010). *Structure*, **18**, 1492–1501.

- Gentle, I. E., Burri, L. & Lithgow, T. (2005). *Mol. Microbiol.* **58**, 1216–1225.
- Gruss, F., Zähringer, F., Jakob, R. P., Burmann, B. M., Hiller, S. & Maier, T. (2013). *Nature Struct. Mol. Biol.* **20**, 1318–1320.
- Hagan, C. L., Silhavy, T. J. & Kahne, D. (2011). *Annu. Rev. Biochem.* **80**, 189–210.
- Jansen, K. B., Baker, S. L. & Sousa, M. C. (2012). *PLoS One*, **7**, e49749.
- Kabsch, W. (2010). *Acta Cryst.* **D66**, 125–132.
- Kim, K. H., Kang, H.-S., Okon, M., Escobar-Cabrera, E., McIntosh, L. P. & Paetzel, M. (2011). *Biochemistry*, **50**, 1081–1090.
- Kim, K. H. & Paetzel, M. (2011). *J. Mol. Biol.* **406**, 667–678.
- Kim, S., Malinverni, J. C., Sliz, P., Silhavy, T. J., Harrison, S. C. & Kahne, D. (2007). *Science*, **317**, 961–964.
- Knowles, T. J., Bobat, S., Jeeves, M., Henderson, I. R. & Overduin, M. (2007). *Biomol. NMR Assign.* **1**, 113–115.
- Knowles, T. J. *et al.* (2011). *EMBO Rep.* **12**, 123–128.
- Knowles, T. J., Scott-Tucker, A., Overduin, M. & Henderson, I. R. (2009). *Nature Rev. Microbiol.* **7**, 206–214.
- Leonard-Rivera, M. & Misra, R. (2012). *J. Bacteriol.* **194**, 4662–4668.
- Lycklama, A. N. J. A. & Driessen, A. J. (2012). *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **367**, 1016–1028.
- Malinverni, J. C., Werner, J., Kim, S., Sklar, J. G., Kahne, D., Misra, R. & Silhavy, T. J. (2006). *Mol. Microbiol.* **61**, 151–164.
- Noinaj, N., Fairman, J. W. & Buchanan, S. K. (2011). *J. Mol. Biol.* **407**, 248–260.
- Noinaj, N., Kuszak, A. J., Gumbart, J. C., Lukacik, P., Chang, H., Easley, N. C., Lithgow, T. & Buchanan, S. K. (2013). *Nature (London)*, **501**, 385–390.
- Rapaport, D. (2005). *J. Cell Biol.* **171**, 419–423.
- Ricci, D. P. & Silhavy, T. J. (2012). *Biochim. Biophys. Acta*, **1818**, 1067–1084.
- Rigel, N. W., Ricci, D. P. & Silhavy, T. J. (2013). *Proc. Natl Acad. Sci. USA*, **110**, 5151–5156.
- Ruiz, N., Falcone, B., Kahne, D. & Silhavy, T. J. (2005). *Cell*, **121**, 307–317.
- Sandoval, C. M., Baker, S. L., Jansen, K., Metzner, S. I. & Sousa, M. C. (2011). *J. Mol. Biol.* **409**, 348–357.
- Sauri, A., Soprova, Z., Wickström, D., de Gier, J. W., Van der Schors, R. C., Smit, A. B., Jong, W. S. & Luirink, J. (2009). *Microbiology*, **155**, 3982–3991.
- Schleiff, E. & Becker, T. (2011). *Nature Rev. Mol. Cell Biol.* **12**, 48–59.
- Selkrig, J., Leyton, D. L., Webb, C. T. & Lithgow, T. (2013). *Biochim. Biophys. Acta*, doi:10.1016/j.bbamcr.2013.10.009.
- Sklar, J. G., Wu, T., Gronenberg, L. S., Malinverni, J. C., Kahne, D. & Silhavy, T. J. (2007). *Proc. Natl Acad. Sci. USA*, **104**, 6400–6405.
- Smart, O. S., Womack, T. O., Flensburg, C., Keller, P., Paciorek, W., Sharff, A., Vonrhein, C. & Bricogne, G. (2012). *Acta Cryst.* **D68**, 368–380.
- Soprova, Z., Sauri, A., van Ulsen, P., Tame, J. R. H., den Blaauwen, T., Jong, W. S. P. & Luirink, J. (2010). *J. Biol. Chem.* **285**, 38224–38233.
- Ulsen, P. van, Rahman, S. U., Jong, W. S. P., Daleke-Schermerhorn, M. H. & Luirink, J. (2013). *Biochim. Biophys. Acta*, doi:10.1016/j.bbamcr.2013.11.006.
- Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* **D66**, 22–25.
- Volokhina, E. B., Grijpstra, J., Beckers, F., Lindh, E., Robert, V., Tommassen, J. & Bos, M. P. (2013). *PLoS One*, **8**, e85799.
- Voulhoux, R., Bos, M. P., Geurtsen, J., Mols, M. & Tommassen, J. (2003). *Science*, **299**, 262–265.
- Ward, R., Zoltner, M., Beer, L., El Mkami, H., Henderson, I. R., Palmer, T. & Norman, D. G. (2009). *Structure*, **17**, 1187–1194.
- Webb, C. T., Heinz, E. & Lithgow, T. (2012). *Trends Microbiol.* **20**, 612–620.
- Wu, T., Malinverni, J., Ruiz, N., Kim, S., Silhavy, T. J. & Kahne, D. (2005). *Cell*, **121**, 235–245.
- Zeth, K., Diederichs, K., Welte, W. & Engelhardt, H. (2000). *Structure*, **8**, 981–992.
- Zeth, K., Kozjak-Pavlovic, V., Faulstich, M., Fraunholz, M., Hurwitz, R., Kepp, O. & Rudel, T. (2013). *Biochem. J.* **449**, 631–642.
- Zeth, K. & Thein, M. (2010). *Biochem. J.* **431**, 13–22.
- Zhang, H., Gao, Z.-Q., Hou, H.-F., Xu, J.-H., Li, L.-F., Su, X.-D. & Dong, Y.-H. (2011). *Acta Cryst.* **F67**, 734–738.
- Zwart, P. H., Afonine, P. V., Grosse-Kunstleve, R. W., Hung, L.-W., Ioerger, T. R., McCoy, A. J., McKee, E., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K., Storoni, L. C., Terwilliger, T. C. & Adams, P. D. (2008). *Methods Mol. Biol.* **426**, 419–435.