doi:10.1016/j.jmb.2010.12.020 J. Mol. Biol. (2011) 406, 667–678

Contents lists available at www.sciencedirect.com

Journal of Molecular Biology

journal homepage: http://ees.elsevier.com.jmb

Crystal Structure of *Escherichia coli* BamB, a Lipoprotein Component of the β-Barrel Assembly Machinery Complex

Kelly H. Kim and Mark Paetzel*

Department of Molecular Biology and Biochemistry, Simon Fraser University, South Science Building, 8888 University Drive, Burnaby, British Columbia, Canada V5A 1S6

Received 19 October 2010; received in revised form 7 December 2010; accepted 9 December 2010
Available online 17 December 2010

Edited by R. Huber

Keywords:
outer membrane biogenesis; outer membrane protein; lipoprotein; protein structure; protein folding

In Gram-negative bacteria, the BAM (β-barrel assembly machinery) complex catalyzes the essential process of assembling outer membrane proteins. The BAM complex in *Escherichia coli* consists of five proteins: one β-barrel membrane protein, BamA, and four lipoproteins, BamB, BamC, BamD, and BamE. Despite their role in outer membrane protein biogenesis, there is currently a lack of functional and structural information on the lipoprotein components of the BAM complex. Here, we report the first crystal structure of BamB, the largest and most functionally characterized lipoprotein component of the BAM complex. The crystal structure shows that BamB has an eight-bladed β-propeller structure, with four β-strands making up each blade. Mapping onto the structure the residues previously shown to be important for BamA interaction reveals that these residues, despite being far apart in the amino acid sequence, are localized to form a continuous solvent-exposed surface on one side of the β-propeller. Found on the same side of the β-propeller is a cluster of residues conserved among BamB homologs. Interestingly, our structural comparison study suggests that other proteins with a BamB-like fold often participate in protein or ligand binding, and that the binding interface on these proteins is located on the surface that is topologically equivalent to where the conserved residues and the residues that are important for BamA interaction are found on BamB. Our structural and bioinformatic analyses, together with previous biochemical data, provide clues to where the BamA and possibly a substrate interaction interface may be located on BamB.

Introduction

The outer membrane is a common feature in Gram-negative bacteria as well as in eukaryotic organelles of endosymbiotic origin, namely, the mitochondria and chloroplast.[1] The outer membrane acts as a physical barrier and a molecular sieve that regulates the traffic of solutes into and out of the cell (in bacteria) or an organelle (in eukaryotes).[1,2] In both bacteria and eukaryotes, the proteins embedded in the outer membrane are essential for maintaining the physical and functional integrity of the outer membrane.[1–3] These proteins all share a β-barrel structure and are collectively known as outer membrane proteins (OMPs).[2,4] In Gram-negative bacteria, OMPs fulfill a diverse range of biological functions such as nutrient uptake, antibiotic resistance, cell adhesion, and maintenance of membrane-selective permeability.[1–3] It is therefore critical for cell viability that OMPs are properly assembled and inserted into the outer membrane.[1,2,4]

OMP biogenesis begins in the cytosol where OMPs are first synthesized. The N-terminal signal sequence then targets the newly synthesized OMPs to the inner membrane, and the protein is translocated across the

*Corresponding author. E-mail address: mpaetzel@sfu.ca.

Abbreviations used: BAM, β-barrel assembly machinery; OMP, outer membrane protein; SAD, single-wavelength anomalous dispersion.

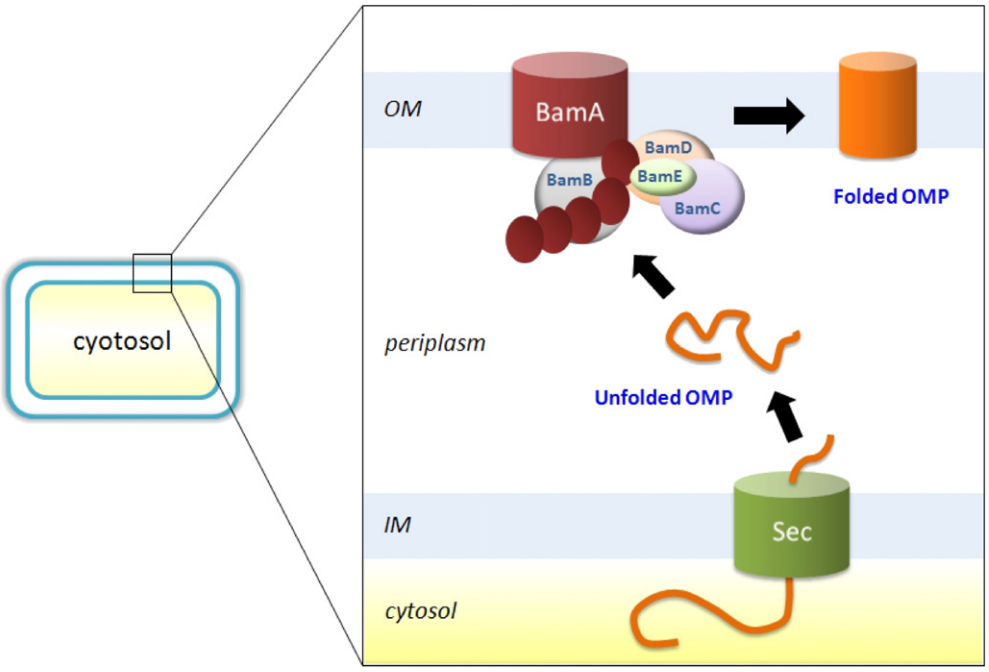

Fig. 1. A schematic diagram of the OMP secretion and assembly pathway in *E. coli*. Following their synthesis in the cytosol, OMPs (orange) are translocated across the inner membrane in an unfolded state via the Sec translocation system (green). The OMPs are then released into the periplasm and subsequently escorted by chaperones to the BAM complex, which is a multicomponent protein complex consisting of BamA, BamB, BamC, BamD, and BamE. By an unknown molecular mechanism, the BAM complex facilitates the assembly and insertion of the OMPs into the outer membrane lipid bilayer.

inner membrane via the Sec translocation system.[2] Following the removal of the N-terminal signal sequence by signal peptidase, the processed protein is released into the periplasm and subsequently carried to the outer membrane in an unfolded state by the periplasmic chaperones SurA, Skp, and DegP.[2,5] At the outer membrane, the folding and membrane insertion of OMPs are catalyzed by the multicomponent BAM (β-barrel assembly machinery) complex (Fig. 1).[6–9] Systems homologous to the BAM complex are also found in the mitochondrial and chloroplast outer membrane and function in the assembly of their β-barrel proteins.[6,10,11]

In *Escherichia coli*, the BAM complex consists of the following five proteins with their commonly known names within brackets: BamA (YaeT/Omp85), BamB (YfgL), BamC (NlpB), BamD (YfiO), and BamE (SmpA).[12–14] BamA is a β-barrel-type integral membrane protein (i.e., it itself is an OMP), and the remaining members of the complex are lipoproteins anchored to the periplasmic face of the outer membrane via a lipid moiety at their N termini.[5,9] While loss of a gene encoding BamA or BamD completely halts OMP biogenesis and leads to cell death,[13,15,16] deletion of the genes coding for BamB, BamC, or BamE results in decreased efficiency of OMP folding and assembly.[12,17,18] The detailed molecular mechanism of how these five different proteins work together as a BAM complex to assist in OMP folding and membrane insertion remains poorly understood.

The BAM complex is currently visualized as a large molecular machine in which BamA is the major structural and functional component, with BamB, BamC, BamD, and BamE serving as its accessory proteins to enhance its efficiency.[5] BamA is a β-barrel protein with a long periplasmic tail consisting of five *po*lypeptide *tr*ansport *a*ssociated (POTRA) motifs.[6,7,9] BamB–BamE proteins are lipoproteins, and therefore they are expected to be found anchored to the periplasmic side of the outer membrane.[6] Thus far, the only structural information available for the BAM complex is the partial POTRA domain structure of *E. coli* BamA[19–22] and the structure of a BamA homolog, *Bordetellapertussis* FhaC`.[23] Kim *et al.*[22] and Knowles *et al.*[19] separately solved the crystal structure of the four N-terminal POTRA domains (POTRA1–POTRA4), while Gatzeva *et al.*[20] solved the solution NMR structure of the first two N-terminal POTRA domains (POTRA1 and POTRA2). Kim *et al.*[22] have also shown by a combination of mutagenesis and pull-down studies that POTRA motifs of BamA serve as docking sites for the lipoproteins BamB–BamE. The POTRA motifs also seem to be important for initial substrate (i.e., unfolded OMP) recognition and chaperone-like

activity.[24–26] In comparison to BamA, the lipoprotein components of the BAM complex are much less well characterized.

Among all the lipoprotein members of the BAM complex, BamB is currently the best-characterized lipoprotein, although its exact role within the BAM complex still remains to be elucidated. Previous studies have shown that the simultaneous absence of BamB and the periplasmic chaperone/protease DegP or the absence of BamB and fkpA (another periplasmic chaperone) results in a conditional synthetic lethal phenotype.[18,28] It has also been shown that SurA (periplasmic chaperone) and BamB deletion mutants are almost indistinguishable from each other in phenotype.[27] Taken together, these experimental results suggest that BamB may work in the SurA pathway to facilitate the delivery of substrates to the essential components of the BAM complex, namely, BamA and BamD.

In order to gain insights into the workings of the BAM complex, we have initiated structural studies on the members of the BAM complex. This paper specifically focuses on BamB, the largest lipoprotein subunit of the BAM complex. Here, we present the first structure of the BAM complex lipoprotein BamB.

Results

Crystal structure of BamB

We have produced a soluble construct of BamB that encompasses the entire wild-type sequence immediately following the cleavable N-terminal signal sequence and the conserved lipidation residue, Cys20 (Ser21-Arg392). The protein was successfully crystallized following purification, and its structure was solved by the single-wavelength anomalous dispersion (SAD) phasing method and then refined to 2.6-Å resolution.

General structural features of BamB

The overall shape of BamB resembles that of a short cylinder with a funnel-shaped channel running down the cylindrical axis (Fig. 2). The protein is approximately 48 Å in diameter and 28 Å in height, and the opening of the channel in the center of the protein is ~18 Å in diameter at one end and ~10 Å at the opposite end. The solvent-exposed cavity created by this channel is approximately 5845 $Å^3$ in volume (Fig. 2d).

The ring-like structure of BamB arises from the eight-bladed β-propeller fold of the protein. Each blade of BamB consists of four antiparallel β-strands, and the blades are arranged around a pseudo-8-fold axis. The last blade (blade 8) is formed by one N-terminal-most β-strand and three C-terminal-most β-strands, and this '1 + 3 velcro closure' arrangement of the terminal β-strands closes the ring structure. By convention,[30] each strand in a blade is labeled A through D starting with strand A as the innermost strand (Fig. 2b and c). The β-strands within each blade are connected by short turns, while the blades are connected to each other via a long loop between the outermost β-strand (strand D) of the previous blade and the innermost β-strand (strand A) of the next blade (Fig. 2b and c). These long connecting loops, referred to as the d-a loops hereinafter, are arranged in a radial manner, with each loop crossing from the outer edge of each blade towards the common center of the propeller structure. The d-a loops together form a continuous molecular surface on one side of the β-propeller.

Analysis of the electrostatic properties of BamB reveals that the protein has a predominantly negatively charged surface (Fig. 2e). This observation is in agreement with the low theoretical p*I* value of 4.7 for the region of the BamB construct seen in the crystal structure. The solvent-filled central channel of the protein is especially negatively charged, with several aspartate and glutamate residues (Glu197, Asp246, Asp248, Asp288, and Asp370) lining the channel surface (Fig. 2f).

Mapping conserved residues onto the structure of BamB

Sequence comparisons of *E. coli* BamB with its functional homologs using ClustalW [31] indicate that a large part of the protein is well conserved throughout different species of Gram-negative bacteria (Fig. 3a). The alignment result shows that there are in total 74 invariant residues, corresponding to 20% of the protein. These invariant residues are spread out throughout the amino acid sequence, but many are found adjacent to each other when the conservation is mapped onto the molecular surface of BamB (Fig. 3b and c).

Most of the solvent-exposed invariant residues (Arg195, Glu240, Asp242, Arg243, Asp246, Arg325, and Glu343) are localized onto a surface of BamB at one side of the β-propeller formed by the d-a loops (Fig. 2c). Located on the same side of the propeller, and both within and in close proximity of this region, are also the residues that were previously determined by mutagenesis to be important for interaction with BamA (Fig. 3; discussed in more detail later).

There are also several residues buried inside the protein that are conserved throughout BamB homologs. Interestingly, many of these buried invariant residues are found repeatedly in each blade, suggesting that the BamB blades are not only structurally but also sequentially homologous to

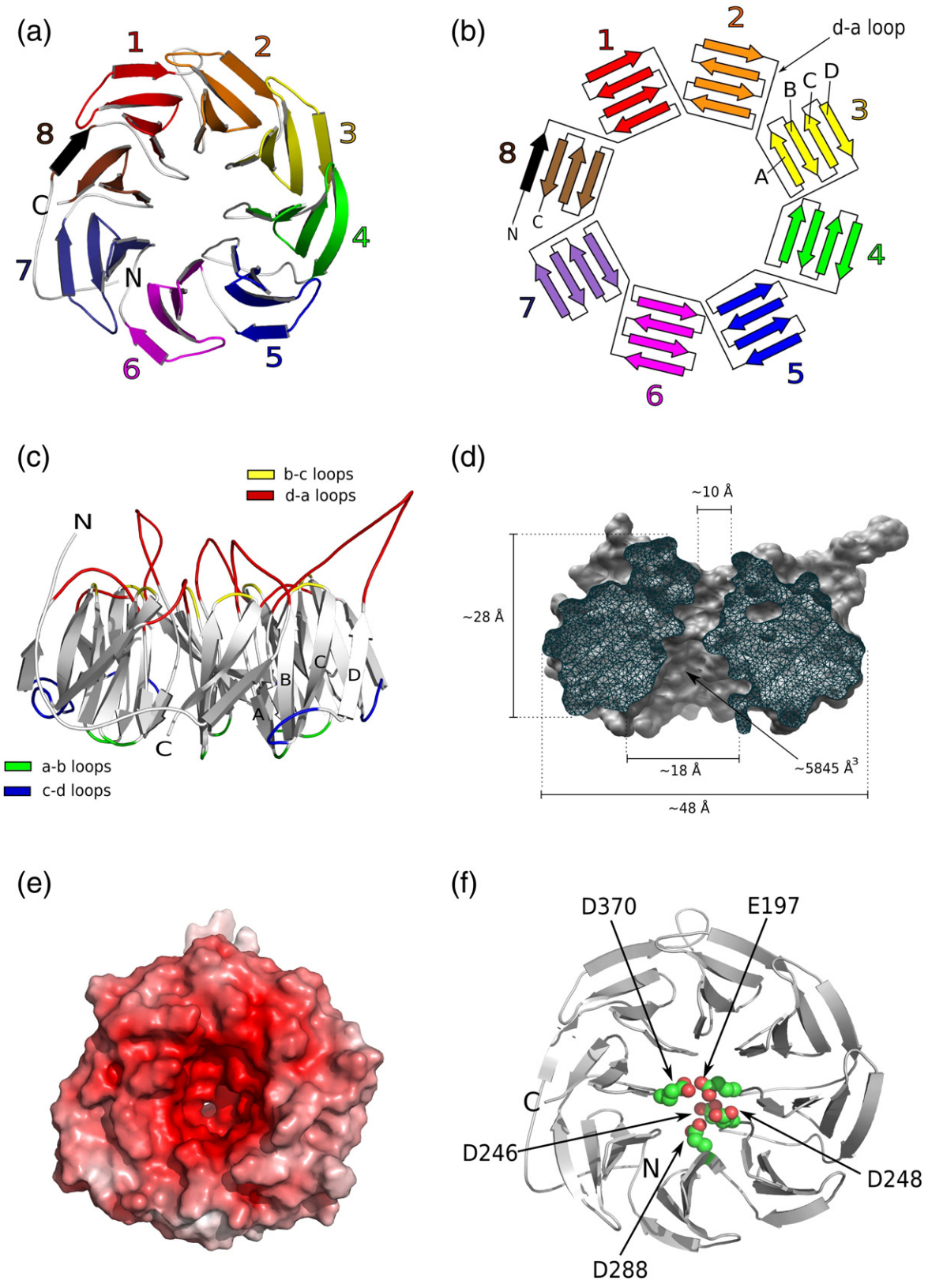

Fig. 2 (*legend on next page*)

each other. Alignment of the sequences from all blades (except blade 8, which is topologically different from the other blades) shows that there are several conserved hydrophobic amino acids as well as an invariant glycine residue (Supplementary Fig. S1). When mapped onto the BamB structure, the hydrophobic residues seem to promote inter-blade contacts by forming interlocking hydrophobic interactions, whereas the conserved glycine residues seem to play a structural role in the connecting loop between strands C and D of each blade.

Mapping residues previously identified as important for BamA interaction onto the BamB structure

As mentioned earlier, some residues found within and near the conserved area on the solvent-exposed surface of BamB have previously been shown to be important for interaction with BamA. Five residues involved in the interaction were identified by Vuong *et al.*[29] using a series of mutagenesis studies, and they are Leu192, Leu194, Arg195, Asp246, and Asp248. When these residues are mapped onto the BamB crystal structure, all five residues are found in a continuous linear patch on the surface of the protein (Fig. 4). While Leu192, Leu194, and Arg195 are located on the d-a loop that connects blade 3 and blade 4, Asp246 and Asp248 are located in the d-a loop connecting blade 4 and blade 5. As these two loops are located adjacent to each other, the five residues are brought into close proximity to form a continuous patch on the surface of BamB.

Structural homologs of BamB

A search for structural homologs using the DALI[32] server identified several proteins that have a significant degree of similarity in protein topology and architecture with BamB (Fig. 5). Proteins with a BamB-like fold include an alcohol dehydrogenase from *Pseudomonas putida* known as ADH IIB[33] (PDB ID: 1KV9) (Fig. 5b), *Homo sapiens* Fbw7[34] (part of an ubiquitin ligase complex; PDB ID: 2QVR) (Fig. 5c), *Paracoccus pantotrophus* cytochrome cd_1 nitrite reductase[35] (PDB ID: 1HJ5) (Fig. 5d), and *Saccharomyces cerevisiae* Sif2p[36] (a transcriptional co-repressor of meiotic genes; PDB ID: 1R5M) (Fig. 5e). All these proteins have an eight-bladed β-propeller structure, and most of them have additional domains positioned on either side of the β-propeller domain. In all four structurally homologous proteins, protein binding sites (the peptide binding site of Fbw7 and the conserved putative protein binding surface of Sif2p) or ligand binding sites (PQQ binding pocket of AD IIB and the heme binding site of nitrite reductase) are found in the β-propeller domain, more specifically on the surface of the domain formed by the d-a loops that are equivalent to the surface of BamB where the conserved residues and the five residues important for BamA interaction are located.

Discussion

In *E. coli*, the BAM complex plays an essential role in OMP folding and assembly in the outer membrane.[9,10] The multicomponent complex is made up of one integral membrane protein, BamA, and four lipoproteins, BamB, BamC, BamD, and BamE.[12–14] Here, we have presented the first crystal structure of BamB, the largest lipoprotein component of the BAM complex.

Analysis of the eight-bladed β-propeller structure of BamB provides us with clues on how BamB may interact with BamA and possibly with protein substrates (i.e., unfolded OMPs). Kim *et al.*[22] have previously shown through a series of deletion and mutagenesis studies that BamB requires four C-terminal POTRA motifs (POTRA2–POTRA5) in order to be co-purified with BamA. In other words, deletion of any of the four POTRA motifs abolishes the observable interaction between the two proteins. Furthermore, they showed, through mutagenesis, that one of the β-strands in POTRA3 (Asp241-Leu247) is important for the BamA–BamB interaction.[22] The authors suggested that the mode of this interaction is probably a β-augmentation (protein–protein interaction through β-strand addition), based on their observation that the C-terminal tail of the crystallized POTRA domain construct was bound to the POTRA3 β-sheet of a neighboring molecule in the crystal lattice.[22] Similarly, Gatzeva-Topalova *et al.*[20] also observed β-augmentation of the POTRA3 β-sheet by the C-terminal tail of a

Fig. 2. General structural features of BamB. (a) A ribbon diagram of the BamB structure, with each blade of the β-propeller structure numbered and shown in different colors. The protein is oriented in such a way that the d-a loop end is facing away from the reader. (b) A topology diagram of BamB. By convention, the β-strands of each blade are labeled A through D starting from the innermost strand (strand A) towards the outermost strand (strand D). As an example, the strands in blade 3 are labeled. (c) A side view of BamB. The different loops found in each blade are colored and labeled. (d) A cross section through the middle of the BamB molecule reveals the dimensions of the funnel-shaped channel of BamB. The overall outer dimensions of the BamB molecule are also given. (e) Electrostatic properties of the BamB molecular surface. The electrostatic potential is mapped onto the solvent-accessible surface of BamB. Red and white represent negative and hydrophobic potentials, respectively. (f) The negatively charged amino acids lining the pore of the protein are shown as spheres on a ribbon model of BamB. The protein is shown in the same orientation as the surface diagram in (c).

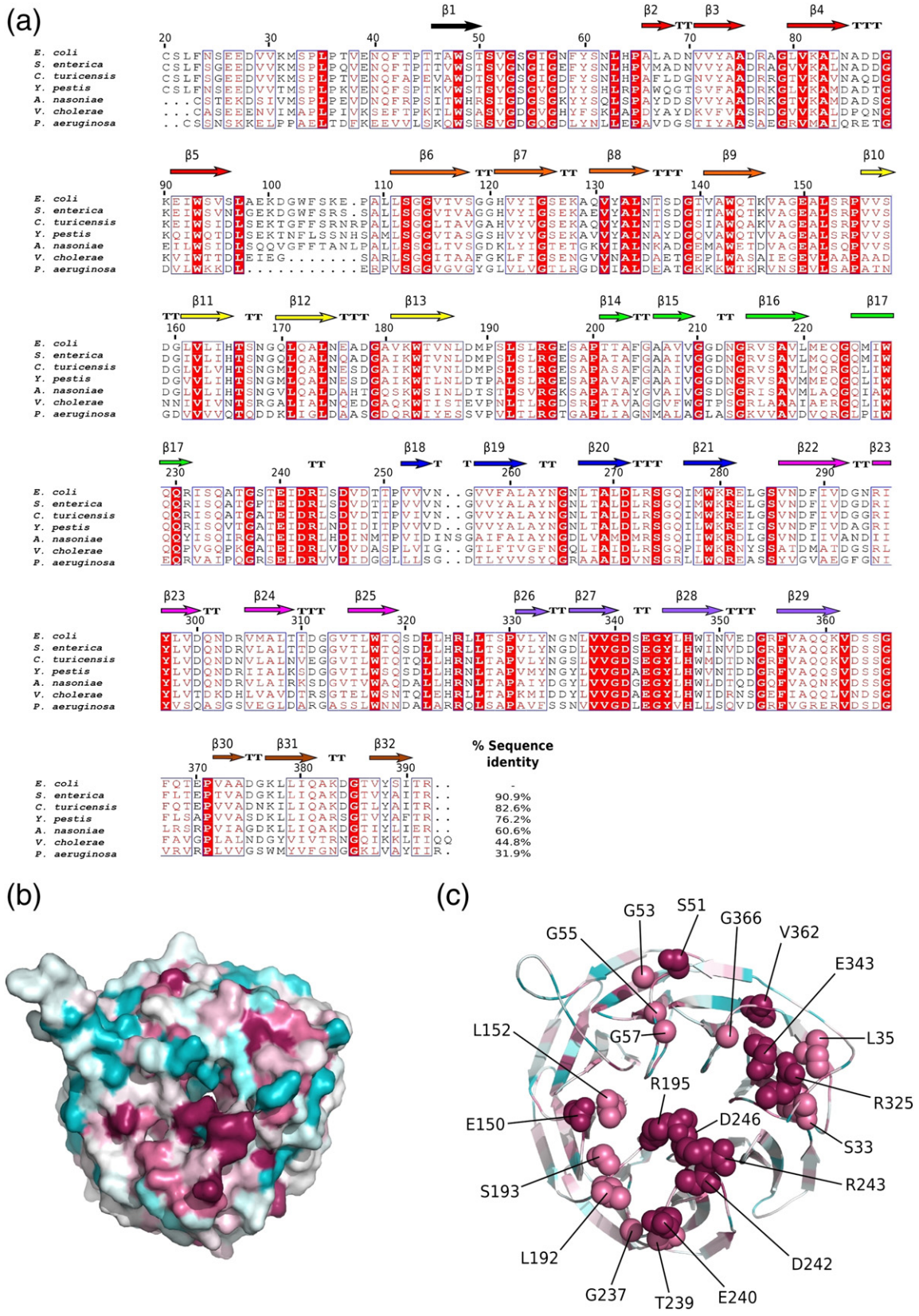

Fig. 3 *(legend on next page)*

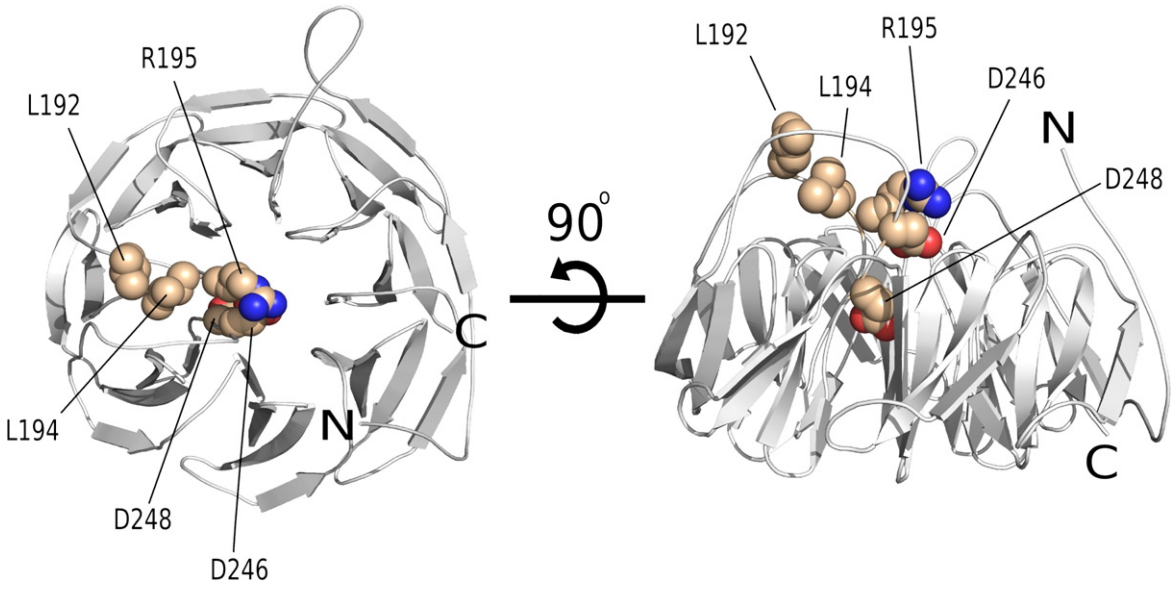

Fig. 4. A potential BamA interaction surface. The five amino acid residues (L192, L194, R195, D246, and D248) of BamB that have previously been identified to be important for interaction with BamA[29] are shown as spheres on a ribbon model of the BamB backbone structure.

neighboring molecule in their crystal structure, despite different crystallization condition and crystal packing. In a separate study, Vuong *et al.*[29] identified five amino acids of BamB that are essential for interaction with BamA. We have shown here that many of these residues are highly conserved among BamB homologs, and have also mapped where these residues are found on our solved crystal structure of BamB. The five residues—Leu192, Leu194, Arg195, Asp246, and Asp248—are found in a continuous patch formed by two d-a loops on the solvent-exposed surface of BamB. Further experiments are necessary to determine whether this region of BamB is where POTRA3 of BamA binds.

In addition to the binding surface identified in this study, there may be additional sites of interaction with BamA on the BamB molecular surface, as all four POTRA motifs (POTRA2–POTRA5) have been shown to interact with BamB.[22] One possible additional binding surface could be the negatively charged pore, which is located adjacent to the identified binding surface. Another possible site of protein–protein interaction is the propeller rim of the BamB structure. If the POTRA motifs, each of which contains a three-stranded β-sheet, interact with BamB indeed via β-augmentation as Kim *et al.*[22] hypothesized, then it seems plausible that the outermost β-strand (i.e., strand D) in each blade of BamB may serve as a binding surface. The strand D from each blade is exposed to the solvent and available for hydrogen-bonding interactions. Similarly, if the function of BamB involves substrate binding and delivery, as suggested in some studies,[18,27] this may be how it interacts with the unfolded OMP substrates. Future biochemical studies and a co-crystal structure of BamB with the POTRA motifs or with an OMP substrate will be needed to confirm this β-augmentation hypothesis.

Even though BamB by itself is not absolutely essential for cell viability, it has been shown that it is essential for proper biogenesis and assembly of the type III secretion system in *Salmonella enterica*

Fig. 3. Conserved amino acids within BamB homologs in Gram-negative bacteria. (a) Sequence alignment of *E. coli* BamB with homologous proteins. The sequences represent those of the mature proteins after their N-terminal signal sequences are removed. The secondary structure of *E. coli* BamB as classified by DSSP[43] is shown above the alignment. Absolutely conserved residues are shown in red boxes, similar residues in red, and stretches of residues that are similar across the group of sequences in blue boxes. The protein sequences were acquired from the Swiss-Prot database: *Escherichia coli* (P77774), *Salmonella enterica* (B5Q1L2), *Arsenophonus nasoniae* (D2TXC8), *Yersinia pestis* (D1TUC6), *Shigella dysenteriae* (B3WYD7), *Vibrio cholerae* (D7HLH6), *Chronobacter turicensis* (C9XY24), and *Pseudomonas aeruginosa* (A6V0W3). (b) A view of BamB conservation mapped onto the BamB surface generated using the above alignment. The protein is shown viewing down the axis with the d-a loops closest to the reader. Individual amino acid residues are colored according to the degree to which they are conserved; absolutely conserved residues are shown in maroon, while highly variable residues are shown in cyan. (c) The side chains of the surface-exposed invariant (maroon) and highly conserved (dark pink) residues are shown as spheres on the ribbon model.

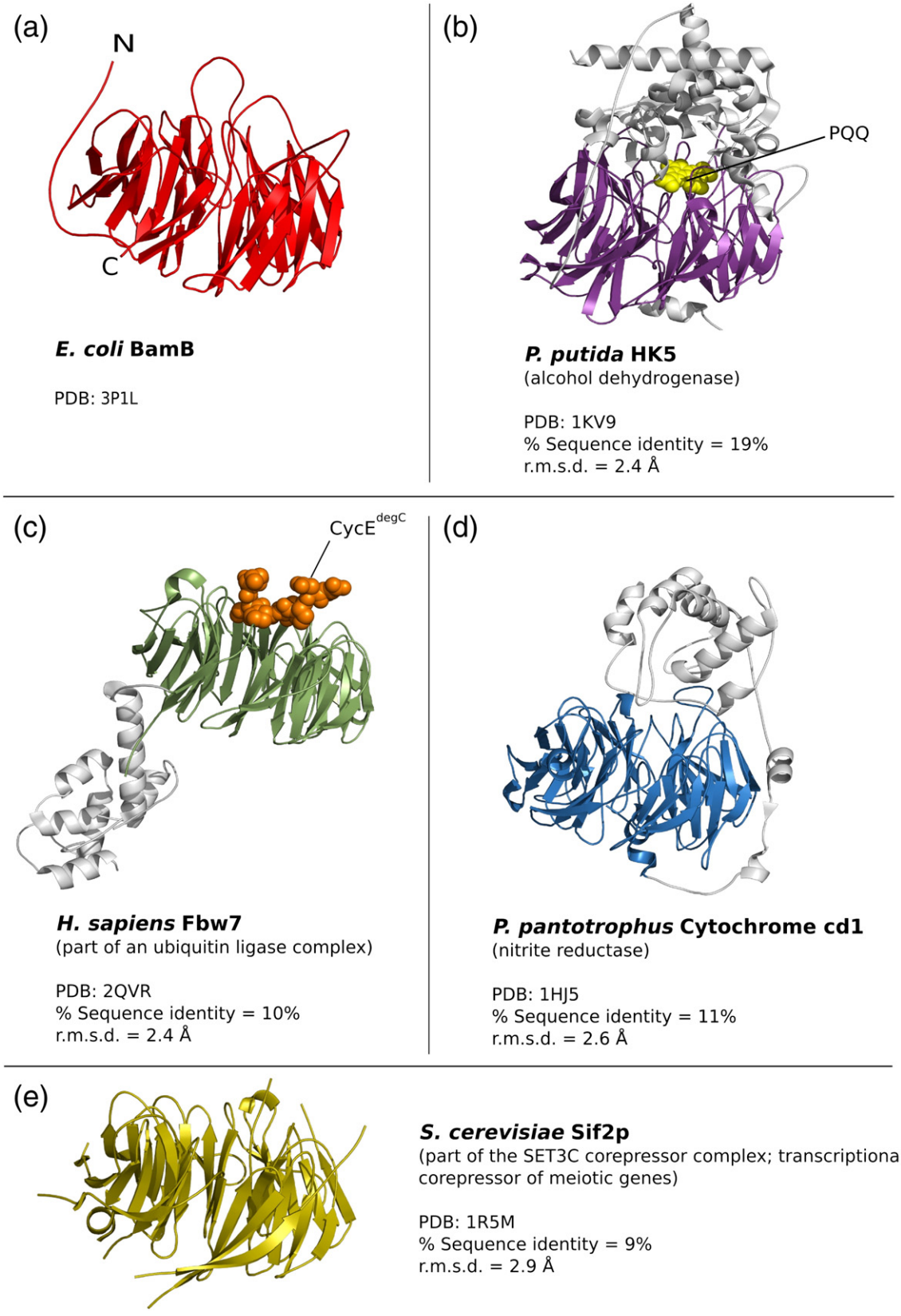

Fig. 5 *(legend on next page)*

Table 1. Crystallographic data collection, phasing, and refinement statistics

	Native	Se-Met incorporated
Crystal parameters		
Space group	$P4_32_12$	$P4_32_12$
a, *b*, *c* (Å)	101.7, 101.7, 108.8	101.6, 101.6, 109.8
% Solvent	63.2	63.5
Data collection statistics		
Wavelength (Å)	0.9794	0.9792
Resolution (Å)	48.0–2.6 (2.7–2.6)	60.1–2.8 (3.0–2.8)
Total reflections	173,441	210,506
Unique reflections	17,912	14,763
R_{merge}[a]	0.075 (0.399)	0.023 (0.127)
Mean $(I)/\sigma I$	20.1 (6.0)	25.8 (6.7)
Completeness	99.1 (99.7)	100.0 (100.0)
Redundancy	9.7 (9.8)	14.3 (14.6)
Phasing statistics		
No. of sites		6 (out of possible 8)
Overall FOM[b]		0.26
Overall FOM (after density modification)		0.42
Refinement statistics		
Protein molecules in asymmetric units	1	
Residues	361	
Water molecules	80	
Total no. of atoms	2711	
R_{cryst}[c]/R_{free}[d]	0.20/0.23	
Average *B*-factor (Å^2)		
All atoms	49.2	
Protein atoms	49.5	
Main-chain atoms	48.4	
Water molecules	39.7	
rmsd on angles (°)	1.99	
rmsd on bonds (Å)	0.02	

The data collection statistics in parentheses are the values for the highest-resolution shell. Se-Met is selenomethionine.

[a] Rmerge = $\sum_{hkl}\sum_i |I_i(hkl) - \bar{I}(\overline{hkl})| / \sum_{hkl}\sum_i(hkl)$, where $I_i(hkl)$ is the observed intensity and $\bar{I}(\overline{hkl})$ is the average intensity obtained from multiple observations of symmetry-related reflections after rejection.

[b] FOM is figure of merit; $\langle |\Sigma P(\alpha) e^{i\alpha}/\Sigma P(\alpha)| \rangle$, where α is the phase angle and $P(\alpha)$ is the phase probability distribution.

[c] $R_{cryst} = \Sigma ||F_o| - |F_c|| / \Sigma |F_o|$, where F_o and F_c are the observed and calculated structure factors, respectively.

[d] R_{free} is calculated using 5% of the reflections randomly excluded from refinement.

serovar Enteritidis.[37] Fardini *et al.*[37] showed that the absence of BamB renders the bacteria avirulent due to its inability to secrete various virulence factors via the type III secretion system.[38] This suggests that BamB may serve as a novel drug target. Further structural and functional studies of BamB and other components of the BAM complex will not only enhance our understanding of how β-barrel proteins assemble and insert into the outer membrane but may also contribute to the development of novel antibiotics.

Materials and Methods

Cloning

A 1116-bp DNA fragment, coding for residues 21–392 of *E. coli* BamB, was amplified from *E. coli* K-12 genomic DNA using the forward primer 5′-ATA TAC ATA TGT CGC TGT TTA ACA GC-3′ and the reverse primer 5′-ATA TAC TCG AGT TAT TAA CGT GTA ATA GA-3′, which contain the restriction sites NdeI and XhoI, respectively. The PCR product was ligated into vector pET28a (Novagen), and the resulting construct encodes BamB (21–392) with a cleavable N-terminal hexa-histidine affinity tag. Subsequent DNA sequencing (Macrogen) confirmed that the BamB insert matched the sequence reported in the Swiss-Prot database (P77774).

Protein expression and purification

The expression plasmid was transformed into *E. coli* BL21(λDE3) and used to inoculate (1:100 back dilution) 3 liters of Luria-Bertani medium containing kanamycin (50 μg/mL). Cultures were grown at 37 °C until the $OD_{600\ nm}$ reached 0.6. The culture was then induced with 1 mM IPTG for 3 h. Cells were harvested by centrifugation and subsequently lysed using an Avestin Emulsiflex-3C cell homogenizer in buffer A [20 mM Tris-HCl (pH 8.0), 100 mM NaCl]. The resulting lysate was clarified by centrifugation (45,000*g*) for 30 min at 4 °C, and the expressed protein was initially purified by Ni^{2+} affinity chromatography. The protein was eluted from the nickel–nitriolotriacetic column (Qiagen) with a step gradient (100–500 mM imidazole in buffer A in 100-mM increments). The fractions containing the protein were pooled and concentrated to approximately 10 mg/mL using an Amicon ultracentrifugal filter device (Millipore) and were then further purified by size-exclusion chromatography (Sephacryl S-100 HiPrep 26/60 column) in buffer A on an AKTA Prime system (GE Health Care). The final protein (Ser21-Arg392 with an N-terminal histidine tag) is 393 residues in length and has a calculated molecular mass of 42,044 Da and a calculated isoelectric point of 5.1. The protein was then concentrated to 30 mg/mL for crystallization. The protein concentration was determined using the Nanodrop ND-100 spectrophotometer and an extinction coefficient of 62,910 M^{-1} cm^{-1}.

Se-Met-incorporated BamB was prepared by growing an overnight culture of BL21(λDE3) transformed with the expression plasmid carrying the BamB gene in M9 minimal medium supplemented with 50 μg/mL kanamycin. Thirty

Fig. 5. *E. coli* BamB (a) has a similar topology and architecture to a variety of other proteins (b–e). Only the regions of the proteins that share structural homology with BamE are shown in color. Nonhomologous regions are shown as grey ribbons. The bound cofactor in (b) and the peptide in (c) are shown as colored spheres and labeled. The percent sequence identity and rmsd for three-dimensional superposition are given for each structure in comparison to BamB.

milliliters of overnight culture was used for each liter of medium to inoculate a total of 3 liters of M9 minimal medium (50 μg/mL kanamycin), which was grown at 37 °C until the $OD_{600\ nm}$ reached 0.6. Each 1-liter culture was then directly supplemented with a mixture of the following amino acids: 100 mg of lysine, phenylalanine, and threonine; 50 mg of isoleucine, leucine, and valine; and 60 mg of selenomethionine. After 15 min, protein expression was induced with 1 mM IPTG (final concentration) for 3 h at 37 °C. The purification procedure for Se-Met-incorporated BamB was the same as that used for the native protein.

Crystallization

The crystals used for SAD data collection were grown by the sitting drop vapor diffusion method. The crystallization drops were prepared by mixing 1 μL of protein (30 mg/mL) suspended in buffer A with 1 μL of reservoir solution and then equilibrating the drop against 1 mL of reservoir solution. The BamB construct yielded crystals in the space group $P4_32_12$ with unit cell dimensions of 101.6, 101.6, and 109.8 Å. The crystals have one molecule in the asymmetric unit with a Matthews coefficient of 3.2 $Å^3$ Da^{-1} (63.5% solvent). The optimal crystallization reservoir condition was 0.1 M citric acid and 2 M sodium chloride. Crystallization was performed at room temperature (~22 °C). The cryo-solution condition contained 0.1 M citric acid, 2 M sodium chloride, and 20% glycerol. Crystals were washed in the cryo-solution before being flash-cooled in liquid nitrogen.

Data collection

Diffraction data were collected on selenomethionine-incorporated crystals at beamline 08ID-1 at the Canadian Macromolecular Crystallography Facility of the Canadian Light Source, using a MarMosaicRayonix MX300 CCD X-ray detector. The crystal-to-detector distance was 320 mm. A total of 120 images were collected with 1° oscillations, and each image was exposed for 1.5 s. The diffraction data were processed with the program iMosfilm.[39] The data collection strategy for the native data was the same as that for the SAD data set. See Table 1 for data collection statistics.

Structure determination and refinement

The BamB structure was solved by SAD using a data set collected at the peak wavelength (0.9792 Å), using the programs Autosol and Autobuild within PHENIX version 1.6.4.[40] Autosol found six of the possible eight selenium sites. The program Autobuild automatically constructed 90% of the polypeptide chain and performed density modification (histogram matching and solvent flattening). The rest of the model was built using the program Coot.[41] The structure was then refined against a higher-resolution native data set (~2.6 Å) by molecular replacement using AutoMR (within PHENIX) followed by restrained refinement using the program Refmac5.[42] The structure includes all but the 10 most amino-terminal residues of the construct. The structure also includes 80 ordered water molecules and one sodium ion. The data collection, phasing, and refinement statistics are summarized in Table 1. The structural coordinate and structure factors for BamB (residues 31–392) have been deposited in the RCSB Protein Data Bank (PDB ID: 3P1L).

Structural analysis

Secondary structural analysis was performed with the programs DSSP[43] and PROMOTIF.[44] The program Coot was used to overlap coordinates for structural comparison. The stereochemistry of the structure was analyzed with the program PROCHECK,[45] and the DALI[32] server was used to find proteins with similar protein folds. The surface electrostatics analysis was performed with the adaptive Boltzmann–Poisson solver plug-in using dielectric constants of 2 and 80 for solute and solvent, respectively. The programs CASTp[46] and SURFACE RACER 1.2[47] were used to detect potential substrate binding sites and measure solvent-exposed molecular surface and cavities, respectively. A probe radius of 1.4 Å was used in the calculation.

Figure preparation

Figures were prepared using PyMOL.[48] The alignment figure was prepared using the programs ClustalW[31] and ESPript.[49]

Accession numbers

PDB accession numbers: *E. coli* BamB (3P1L)

Supplementary materials related to this article can be found online at doi:10.1016/j.jmb.2010.12.020

Acknowledgements

This work was supported in part by the Canadian Institute of Health Research, the National Science and Engineering Research Council of Canada, the Michael Smith Foundation for Health Research, and the Canadian Foundation of Innovation. We thank the staff at the macromolecular beam line 08ID-1, Canadian Light Source, Saskatoon, Canada, for their technical assistance with data collection. The Canadian Light Source is supported by NSERC, NRC, CIHR, and the University of Saskatchewan.

References

1. Costerton, J. W., Ingram, J. M. & Cheng, K. J. (1974). Structure and function of the cell envelope of gram-negative bacteria. *Bacteriol. Rev.* **38**, 87–110.
2. Bos, M. P., Robert, V. & Tommassen, J. (2007). Biogenesis of the gram-negative bacterial outer membrane. *Annu. Rev. Microbiol.* **61**, 191–214.

3. Delcour, A. H. (2009). Outer membrane permeability and antibiotic resistance. *Biochem. Biophys. Acta*, **1794**, 808–816.
4. Walther, D. M., Rapaport, D. & Tommassen, J. (2009). Biogenesis of β-barrel membrane proteins in bacteria and eukaryotes: evolutionary conservation and divergence. *Cell. Mol. Life Sci.* **66**, 2789–2804.
5. Gatsos, X., Perry, A. J., Anwari, K., Dolezal, P., Wolynec, P. P., Likic, V. A. *et al.* (2008). Protein secretion and outer membrane assembly in Alphaproteobacteria. *FEMS Microbiol. Rev.* **32**, 995–1009.
6. Gentle, I. E., Burri, L. & Lithgow, T. (2005). Molecular architecture and function of the Omp85 family of proteins. *Mol. Microbiol.* **58**, 1216–1225.
7. Voulhoux, R. & Tommassen, J. (2004). Omp85, an evolutionarily conserved bacterial protein involved in outer-membrane-protein assembly. *Res. Microbiol.* **155**, 129–135.
8. Wu, T., Malinverni, J., Ruiz, N., Kim, S., Silhavy, T. J. & Kahne, D. (2005). Identification of a multicomponent complex required for outer membrane biogenesis in *Escherichia coli*. *Cell*, **121**, 235–245.
9. Knowles, T. J., Scott-Tucker, A., Overduin, M. & Henderson, I. R. (2009). Membrane protein architects: the role of the BAM complex in outer membrane protein assembly. *Nat. Rev. Microbiol.* **7**, 206–214.
10. Gentle, I., Gabriel, K., Beech, P., Waller, R. & Lithgow, T. (2004). The Omp85 family of proteins is essential for outer membrane biogenesis in mitochondria and bacteria. *J. Cell Biol.* **164**, 19–24.
11. Voulhoux, R., Bos, M. P., Geurtsen, J., Mols, M. & Tommassen, J. (2003). Role of a highly conserved bacterial protein in outer membrane protein assembly. *Science*, **299**, 262–265.
12. Sklar, J. G., Wu, T., Gronenberg, L. S., Malinverni, J. C., Kahne, D. & Silhavy, T. J. (2007). Lipoprotein SmpA is a component of the YaeT complex that assembles outer membrane proteins in *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **104**, 6400–6405.
13. Malinverni, J. C., Werner, J., Kim, S., Sklar, J. G., Kahne, D., Misra, R. & Silhavy, T. J. (2006). YfiO stabilizes the YaeT complex and is essential for outer membrane protein assembly in *Escherichia coli*. *Mol. Microbiol.* **61**, 151–164.
14. Anwari, K., Poggio, S., Perry, A., Gatsos, X., Ramarathinam, S. H., Williamson, N. A. *et al.* (2010). A modular BAM complex in the outer membrane of the α-proteobacterium *Caulobacter crescentus*. *PLoS One*, **5**, e8619.
15. Doerrler, W. T. & Raetz, C. R. (2005). Loss of outer membrane proteins without inhibition of lipid export in an *Escherichia coli* YaeT mutant. *J. Biol. Chem.* **280**, 27679–27687.
16. Fitzpatrick, D. A. & McInerney, J. O. (2005). Evidence of positive Darwinian selection in Omp85, a highly conserved bacterial outer membrane protein essential for cell viability. *J. Mol. Evol.* **60**, 268–273.
17. Ryan, K. R., Taylor, J. A. & Bowers, L. M. (2010). The BAM complex subunit BamE (SmpA) is required for membrane integrity, stalk growth and normal levels of outer membrane β-barrel proteins in Caulobacter crescentus. *Microbiology*, **156**, 742–756.
18. Charlson, E. S., Werner, J. N. & Misra, R. (2006). Differential effects of yfgL mutation on *Escherichia coli* outer membrane proteins and lipopolysaccharide. *J. Bacteriol.* **188**, 7186–7194.
19. Knowles, T. J., Jeeves, M., Bobat, S., Dancea, F., McClelland, D., Palmer, T. *et al.* (2008). Fold and function of polypeptide transport-associated domains responsible for delivering unfolded proteins to membranes. *Mol. Microbiol.* **68**, 1216–1227.
20. Gatzeva-Topalova, P. Z., Walton, T. A. & Sousa, M. C. (2008). Crystal structure of YaeT: conformational flexibility and substrate recognition. *Structure*, **16**, 1873–1881.
21. Misra, R. (2007). First glimpse of the crystal structure of YaeT's POTRA domains. *ACS Chem. Biol.* **2**, 649–651.
22. Kim, S., Malinverni, J. C., Sliz, P., Silhavy, T. J., Harrison, S. C. & Kahne, D. (2007). Structure and function of an essential component of the outer membrane protein assembly machine. *Science*, **317**, 961–964.
23. Clantin, B., Delattre, A. S., Rucktooa, P., Saint, N., Meli, A. C., Locht, C. *et al.* (2007). Structure of the membrane protein FhaC: a member of the Omp85-TpsB transporter superfamily. *Science*, **317**, 957–961; 24.
24. Bos, M. P., Robert, V. & Tommassen, J. (2007). Functioning of outer membrane protein assembly factor Omp85 requires a single POTRA domain. *EMBO Rep.* **8**, 1149–1154.
25. Meli, A. C., Hodak, H., Clantin, B., Locht, C., Molle, G., Jacob-Dubuisson, F. & Saint, N. (2006). Channel properties of TpsB transporter FhaC point to two functional domains with a C-terminal protein-conducting pore. *J. Biol. Chem.* **281**, 158–166.
26. Robert, V., Volokhina, E. B., Senf, F., Bos, M. P., Van Gelder, P. & Tommassen, J. (2006). Assembly factor Omp85 recognizes its outer membrane protein substrates by a species-specific C-terminal motif. *PLoS Biol.* **4**, e377.
27. Ureta, A., Endres, R., Wingree, N. & Silhavy, T. (2007). Kinetic analysis of the assembly of the outer membrane protein LamB in *Escherichia coli* mutants each lacking a secretion or targeting factor in a different cellular compartment. *J. Bacteriol.* **189**, 446–454.
28. Onufryk, C., Crouch, M. L., Fang, F. C. & Gross, C. A. (2005). Characterization of six lipoproteins in the regulon. *J. Bacteriol.* **187**, 4552–4561.
29. Vuong, P., Bennion, D., Mantei, J., Frost, D. & Misra, R. (2008). Analysis of YfgL and YaeT interactions through bioinformatics, mutagenesis, and biochemistry. *J. Bacteriol.* **190**, 1507–1517.
30. Fülöp, V. & Jones, D. (1999). Beta propellers: structural rigidity and functional diversity. *Curr. Opin. Struct. Biol.* **9**, 715–721.
31. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H. *et al.* (2007). Clustal W and Clustal X version 2.0. *Bioinformatics*, **23**, 2947–2948.
32. Holm, L., Kaariainen, S., Rosenstrom, P. & Schenkel, A. (2008). Searching protein structure databases with DaliLite v.3. *Bioinformatics*, **24**, 2780–2781.
33. Chen, Z., Matsushita, K., Yamashita, T., Fujii, T., Toyama, H., Adachi, O. *et al.* (2002). Structure at 1.9 Å resolutin of a quinohemoprotein alcohol dehydrogenase from *Pseudomonas putida* HK5. *Structure*, **10**, 837–849.

34. Hao, B., Oehlmann, S., Sowa, M., Harper, J. & Pavletich, N. (2007). Structure of a Fbw7-Skp1-cyclin E complex: multisite-phosphorylated substrate recognition by SCF ubiquitin ligases. *Mol. Cell*, **26**, 131–143.
35. Sjogren, T. & Hajdu, J. (2001). Structure of the bound dioxygen species in the cytochrome oxidase reaction of cytochrome cd_1 nitrite reductase. *J. Biol. Chem.* **276**, 13072–13076.
36. Cerna, D. & Wilson, D. (2005). The structure of Sif2p, a WD repeat protein functioning in the SET3 corepressor complex. *J. Mol. Biol.* **351**, 923–935.
37. Fardini, Y., Chettab, K., Grepinet, O., Rocherau, S., Troterau, J., Harvey, P. *et al.* (2007). The YfgL lipoprotein is essential for type III secretion system expression and virulence of *Salmonella enterica* serovar Enteritidis. *Infect. Immun.* **75**, 358–370.
38. Amy, M., Velge, P., Senocq, D., Bottreau, E., Mompart, F. & Virlogeux-Payant, I. (2004). Identification of a new *Salmonella enterica* serovar Enteritidis locus involved in cell invasion and in the colonization of chicks. *Res. Microbiol.* **155**, 543–552.
39. Powell, H. (1999). The Rossmann Fourier autoindexing algorithm in MOSFLM. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **55**, 1690–1695.
40. Adams, P., Afonine, P., Bunkoczi, G., Chen, B., Davis, I., Echools, N. *et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **66**, 213–221.
41. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **60**, 2126–2132.
42. Murshudov, G., Vagin, A. & Dodson, E. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr.* **30**, 1022–1025.
43. Kabsch, W. & Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, **22**, 2577–2637.
44. Hutchinson, E. G. & Thornton, J. M. (1996). PROMOTIF—a program to identify and analyze structural motifs in proteins. *Protein Sci.* **5**, 212–220.
45. Laskowski, R., Macarther, M., Moss, D. & Thornton, J. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291.
46. Liang, J., Edelsbrunner, H. & Woodward, C. (1998). Anatomy of protein pocket and cavities: measurement of binding site geometry an implications for ligand design. *Protein Sci.* **7**, 1884–1897.
47. Tsodikov, O., Record, M. & Sergeev, Y. (2002). Novel computer program for fast exact calculation of accessible and molecular surface areas and average surface curvature. *J. Comput. Chem.* **23**, 600–609.
48. DeLano, W. L. (2002). *The PyMOL Molecular Graphics System.*
49. Gouet, P., Courcelle, E., Stuart, D. I. & Metoz, F. (1999). ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics*, **15**, 305–308.