**Dongchun Ni,^{a,b} Kun Yang^{a*} and
Yihua Huang^{b*}**^aSchool of Public Health, Tianjin Medical University, Tianjin 300070, People's Republic of China, and ^bNational Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, People's Republic of ChinaCorrespondence e-mail:
yangkun@tjmu.edu.cn,
yihuahuang@sun5.ibp.ac.cn

Received 10 December 2013

Accepted 10 February 2014

Refolding, crystallization and preliminary X-ray crystallographic studies of the β -barrel domain of BamA, a membrane protein essential for outer membrane protein biogenesis

In Gram-negative bacteria, the assembly of outer membrane proteins (OMPs) requires a five-protein β -barrel assembly machinery (BAM) complex, of which BamA is an essential and evolutionarily conserved integral outer membrane protein. Here, the refolding, crystallization and preliminary X-ray crystallographic characterization of the β -barrel domain of BamA from *Escherichia coli* (*EcBamA*) are reported. Native and selenomethionine-substituted *EcBamA* proteins were crystallized at 16°C and X-ray diffraction data were collected to 2.6 and 3.7 Å resolution, respectively. The native crystals belonged to space group $P2_12_12$, with unit-cell parameters $a = 118.492$, $b = 159.883$, $c = 56.000$ Å and two molecules in one asymmetric unit; selenomethionine-substituted protein crystals belonged to space group $P4_322$, with unit-cell parameters $a = b = 163.162$, $c = 46.388$ Å and one molecule in one asymmetric unit. Initial phases for *EcBamA* β -barrel domain were obtained from a SeMet SAD data set. These preliminary X-ray crystallographic studies paved the way for further structural determination of the β -barrel domain of *EcBamA*.

1. Introduction

In Gram-negative bacteria, biogenesis of outer membrane proteins (OMPs) is a complicated but highly organized process: the nascently synthesized OMP polypeptides in cytoplasm have to be translocated across the cytoplasmic membrane by the SecYEG channel (Pugsley, 1993), ferried over the inter-membrane space with the assistance of various periplasmic chaperones and finally inserted into the outer membrane by the so-called BAM (β -barrel assembly machinery) complex. The BAM complex consists of an integral outer membrane protein BamA and four accessory lipoproteins BamB, BamC, BamD and BamE (Knowles *et al.*, 2009; Wu *et al.*, 2005; Hagan *et al.*, 2011; Ricci & Silhavy, 2012; Rigel & Silhavy, 2012). Among them, BamA and BamD are two essential components of the BAM complex, whereas BamB, BamC and BamE seem to play regulatory roles in the process of outer membrane protein biogenesis.

BamA is a member of the Omp85 protein family of OMPs, and its homologues are also found in mitochondria and chloroplasts in eukaryotes. It has been postulated that β -barrel outer membrane proteins in both prokaryotes and eukaryotes share a similar mechanism in membrane insertion, which is mediated by BamA and its homologues. BamA from *Escherichia coli* (*EcBamA*) consists of five POTRA domains at the N-terminus and a C-terminal trans-membrane β -barrel domain (Fig. 1). The N-terminal POTRA domains of BamA function in OMP substrate recognition and serve as a scaffold for the assembly of the BAM complex (Hagan *et al.*, 2010), but the exact function of the β -barrel domain of BamA is unclear. Currently, structures are available for BamB (Noinaj *et al.*, 2011; Heuck *et al.*, 2011; Kim *et al.*, 2011), BamC (Albrecht & Zeth, 2011), BamD (Kim *et al.*, 2011; Albrecht & Zeth, 2011; Sandoval *et al.*, 2011; Dong *et al.*, 2012), BamE (Albrecht & Zeth, 2011; Knowles *et al.*, 2011) and for the POTRA domains of BamA (Gatzeva-Topalova *et al.*, 2008, 2010; Kim *et al.*, 2007; Knowles *et al.*, 2008; Zhang *et al.*, 2011). Recently, Noinaj and coworkers reported the crystal structures of two BamA homologues from *Neisseria gonorrhoeae* (*NgBamA*) and *Haemophilus ducreyi* (*HdBamA*) (Noinaj *et al.*, 2013) and suggested that β -strands may be integrated into the outer membrane via lateral opening of the barrel at the interface of the first and the

Figure 1

Domain organization of full-length *EcBamA*. *EcBamA* contains five N-terminal POTRA domains (denoted P1, P2, P3, P4 and P5; grey) and a C-terminal β -barrel (residues 426–810; green).

last β -strands. A sequence alignment of *NgBamA*, *HdBamA* and *EcBamA* is shown in Fig. 2. As most of the functional studies for understanding OMP biogenesis have been performed using *E. coli* as a model strain, a high-resolution structure of BamA from *E. coli* (*EcBamA*) is important for understanding outer membrane protein biogenesis. To this end, we refolded, crystallized and performed preliminary characterization of the crystals of the β -barrel domain of BamA from *E. coli* (including residues 426–810), and these efforts paved the way for further structural determination and structure-based functional analysis of *EcBamA*.

2. Materials and methods

2.1. Expression, purification and refolding of the β -barrel domain of *EcBamA*(426–810)

We initially attempted to express and purify various N-terminally truncated *EcBamA* proteins, which covered the whole C-terminal β -barrel domain plus different numbers of copies of POTRA domains, in their native state from bacterial outer membranes for crystallization trials. With the exception of constructs that expressed the β -barrel domain alone and the fifth POTRA domain plus the β -barrel domain, the *EcBamA* constructs were highly expressed and readily purified to homogeneity for crystallization screening. However, these proteins failed to produce crystals with sufficient quality for X-ray crystallographic characterization, even though various detergents were screened and extensive optimization was pursued. To overcome the difficulty of obtaining sufficient amounts of proteins containing the β -barrel domain of *EcBamA*, we subcloned

EcBamA (including residues 426–810) from *E. coli* (ATCC MG1655) into the pET-24b vector (Novagen) without signal peptide and affinity tags, and attempted to refold the β -barrel domain from inclusion bodies for crystallization studies. Expression was performed in *E. coli* strain BL21(DE3) cells at 37°C in LB medium supplemented with 50 $\mu\text{g ml}^{-1}$ kanamycin. Expression was induced by the addition of 0.8 mM IPTG for 2 h at 37°C when the cell density reached an OD_{600} of 0.8. Inclusion bodies were isolated and solubilized in a buffer consisting of 20 mM Tris–HCl pH 8.0, 8 M urea to a final protein concentration of about 30 mg ml^{-1} . Protein refolding was performed by dropwise addition of urea-solubilized protein solution into a buffer consisting of 20 mM Tris–HCl pH 8.0, 0.5% LDAO at 4°C overnight with vigorous stirring to a final protein concentration of about 0.6 mg ml^{-1} .

The refolded *EcBamA*(426–810) was concentrated using a 50 kDa cutoff spin concentrator and applied onto a Mono Q column (GE Healthcare) for detergent exchange. Protein was eluted from the Mono Q column with a buffer consisting of 20 mM Tris pH 8.0, 0.6% C_8E_4 , 1 M NaCl. The peak fractions were pooled, concentrated and then applied onto a Superdex-200 10/30 size-exclusion column (GE Healthcare) and eluted with 20 mM Tris–HCl pH 8.0, 150 mM NaCl, 0.6% C_8E_4 (Fig. 3). Monodisperse peak fractions were pooled and concentrated to $\sim 10 \text{ mg ml}^{-1}$ for crystallization trials. The expression and purification of SeMet *EcBamA* β -barrel protein were performed in a similar manner.

2.2. Crystallization and data collection

Both native and SeMet *EcBamA* β -barrel domain crystals were obtained at 16°C using the hanging-drop vapour-diffusion method by mixing protein and precipitants in a 1:1 ratio. Crystals were grown at a protein concentration of $\sim 10 \text{ mg ml}^{-1}$ for the native protein and $\sim 8 \text{ mg ml}^{-1}$ for the SeMet-derivative protein in a buffer consisting of 24% PEG 4000, 0.1 M MES, 6% Tacsimate pH 6.0. Native crystals appeared in about one week after setting up the crystallization trays and grew to a final size of about $60 \times 60 \times 45 \mu\text{m}$ within two weeks (Fig. 4).

Figure 2

Multiple sequence alignment using *T-Coffee* for *EcBamA*(426–810), *HdBamA*(424–793) and *NgBamA*(427–792). The sequence identity of *EcBamA*(426–810) to *HdBamA*(424–793) and *NgBamA*(427–792) is 16.1 and 15.4%, respectively. Compared with *HdBamA*(424–793) and *NgBamA*(427–792), *EcBamA*(426–810) has an unusual long L6 loop insertion (70 residues) as marked in the alignment. The secondary structures at the top of the aligned sequences are based on the structural information of *HdBamA* (PDB entry 4k3c) and *NgBamA* (PDB entry 4k3b) (Noinaj *et al.*, 2013).

Figure 3 Gel-filtration profile of *EcBamA*(426–810) refolded from inclusion bodies. *EcBamA*(426–810) has an apparent molecular weight of about 100 kDa on a Superdex 200 10/30 column in elution buffer consisting of 20 mM Tris pH 8.0, 150 mM NaCl, 0.6% C₈E₄. Molecular-weight standards are labelled at the top in kDa. The inset shows the different mobility of *EcBamA*(426–810) on 15% SDS-PAGE for boiled (95°C for 5 min) and non-boiled (25°C) samples. The left lane contains molecular-weight marker (labelled in kDa).

Figure 4 Native *EcBamA*(426–810) protein was crystallized in a condition consisting of 24% PEG 4000, 0.1 M MES, 6% Tacsimate pH 6.0. Crystals normally grew to final dimensions of 60 × 60 × 45 μm in two weeks.

Figure 5 Electron-density map ($2F_o - F_c$ at a contour level of 2.0σ) calculated from a SeMet SAD data set showing the shape of two β-barrels, suggesting that a correct solution was found.

Diffraction data were collected at -173°C on beamline BL17U at the Shanghai Synchrotron Radiation Facility (SSRF; Shanghai, People's Republic of China). All diffraction data were processed with *HKL-2000* (Otwinowski & Minor, 1997). Data-collection statistics are summarized in Table 1. Five out of nine expected selenium sites were found with *SHELXD* (Sheldrick, 2008) and the calculated electron-density map showed a clear β-barrel shape, suggesting that the found selenium sites are correct (Fig. 5).

3. Results and discussion

The calculated monomeric molecular weight of *EcBamA*(426–810) is 43 006.02 Da. In contrast to soluble proteins, membrane proteins have an apparently larger molecular weight on size-exclusion columns compared with their calculated molecular weights based on their amino-acid residue compositions owing to binding of detergent molecules to the protein surface. On a size-exclusion column, the *EcBamA* β-barrel domain elutes as a monomer with an apparent molecular weight of about 100 kDa assuming that the detergent molecules contribute half of the molecular weight of the C₈E₄-bound *EcBamA* β-barrel protein. Boiled (95°C for 5 min) and non-boiled (25°C) samples of the refolded *EcBamA*(426–810) protein exhibited different mobilities on SDS-PAGE, suggesting that *EcBamA*(426–810) refolded well (Fig. 2).

Diffraction data sets were collected from native and SeMet-derivative crystals to resolution limits of 2.6 and 3.7 Å, respectively, on CCD detectors at the SSRF. The space group was determined to be $P2_12_12$ for native crystals, with unit-cell parameters $a = 118.492$, $b = 159.883$, $c = 56.000$ Å. It is estimated that there are two molecules in the crystallographic asymmetric unit, and the Matthews coefficient

Figure 6 Diffraction patterns for native *EcBamA*(426–810) crystals (left) and crystals obtained from HgCl₂-treated native *EcBamA*(426–810) protein (right) showing the difference in spot shapes. The native crystals diffracted to about 2.6 Å resolution, while the crystals obtained from HgCl₂-treated *EcBamA*(426–810) protein only diffracted to about 3.4 Å resolution in most cases, but with a significantly improved diffraction pattern in terms of spot shape.

Table 1

Data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

	Native	SeMet (+ HgCl ₂)
Beamline	BL17U, SSRF	BL17U, SSRF
Wavelength (Å)	1.0062	0.9792
Space group	<i>P</i> 2 ₁ 2 ₁ 2	<i>P</i> 4 ₃ 22
Unit-cell parameters (Å, °)	<i>a</i> = 118.492, <i>b</i> = 159.883, <i>c</i> = 56.0, $\alpha = \beta = \gamma = 90$	<i>a</i> = 163.162, <i>b</i> = 163.162, <i>c</i> = 46.388, $\alpha = \beta = \gamma = 90$
Resolution (Å)	50–2.6 (2.69–2.60)	50–3.7 (3.95–3.70)
<i>R</i> _{merge} (%)	15 (38)	24 (70)
Completeness (%)	97.4 (93.5)	100.0 (100.0)
<i>I</i> / <i>σ</i> (<i>I</i>)	13.97 (2.3)	18.76 (3.4)
Unique reflections	31883	7616
Multiplicity	4.3	20

Figure 7

Diffraction pattern of SeMet *EcBamA*(426–810) crystals (the SeMet-derivative protein was pretreated with 1 mM HgCl₂ overnight). Crystals have a maximum resolution of about 3.7 Å in most cases.

(*V*_M) was calculated to be 3.084 Å³ Da^{−1}, which corresponds to a solvent content of 60.12% (Matthews, 1968). Although the native crystals of *EcBamA*(426–810) diffracted to high resolution in most cases (about 2.6 Å), the diffraction spots were very streaky and smeary (Fig. 6). Efforts to improve the quality of the diffraction pattern were not very successful, and this made it difficult to obtain accurate diffraction data for further phase calculation. To circumvent this difficulty, we pretreated native proteins with 1 mM HgCl₂ overnight and found that the HgCl₂-treated protein crystals gave much better diffraction patterns in terms of spot shape, although the

resolution dropped significantly to only about 3.4 Å in most cases (Fig. 6). Unexpectedly, no mercury sites were found after processing the data collected from many HgCl₂-treated *EcBamA*(426–810) crystals at the peak wavelength of mercury (λ = 1.0063 Å). The HgCl₂-treated protein crystals belonged to a different space group *P*4₃22, with unit-cell parameters *a* = *b* = 163.162, *c* = 46.388 Å. There is one molecule in the asymmetric unit. Based on these observations, we overexpressed SeMet *EcBamA* β -barrel protein and pretreated the SeMet-derivative protein with 1 mM HgCl₂ overnight at 4 °C before crystallization. This process enabled us to collect several data sets at the Se peak wavelength (λ = 0.9792 Å) with sufficiently accurate data for structural determination (Fig. 7). Diffraction data statistics for native and SeMet *EcBamA*(426–810) crystals are given in Table 1. Currently, efforts are under way to build the model of *EcBamA*(426–810).

We are grateful to Xiaomin Hou, Xiou Cao, Li Deng and Maoshui Wang for technical assistance, and to staff scientists from the Beijing Synchrotron Radiation Facility (BSRF) and the Shanghai Synchrotron Radiation Facility (SSRF) for help with data collection. The work was supported by grants from the Ministry of Science and Technology (2012CB917302 and 2013CB910603 to YH) and the National Natural Science Foundation of China (31170698 to YH).

References

- Albrecht, R. & Zeth, K. (2011). *J. Biol. Chem.* **286**, 27792–27803.
- Dong, C., Hou, H.-F., Yang, X., Shen, Y.-Q. & Dong, Y.-H. (2012). *Acta Cryst. D* **68**, 95–101.
- 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.
- Hagan, C. L., Kim, S. & Kahne, D. (2010). *Science*, **328**, 890–892.
- Hagan, C. L., Silhavy, T. J. & Kahne, D. (2011). *Annu. Rev. Biochem.* **80**, 189–210.
- Heuck, A., Schleiffer, A. & Clausen, T. (2011). *J. Mol. Biol.* **406**, 659–666.
- Kim, K. H., Aulakh, S. & Paetzel, M. (2011). *J. Biol. Chem.* **286**, 39116–39121.
- Kim, S., Malinverni, J. C., Sliz, P., Silhavy, T. J., Harrison, S. C. & Kahne, D. (2007). *Science*, **317**, 961–964.
- Knowles, T. J. *et al.* (2011). *EMBO Rep.* **12**, 123–128.
- Knowles, T. J., Jeeves, M., Bobat, S., Dancea, F., McClelland, D., Palmer, T., Overduin, M. & Henderson, I. R. (2008). *Mol. Microbiol.* **68**, 1216–1227.
- Knowles, T. J., Scott-Tucker, A., Overduin, M. & Henderson, I. R. (2009). *Nature Rev. Microbiol.* **7**, 206–214.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- 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.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Pugsley, A. P. (1993). *Microbiol. Rev.* **57**, 50–108.
- Ricci, D. P. & Silhavy, T. J. (2012). *Biochim. Biophys. Acta*, **1818**, 1067–1084.
- Rigel, N. W. & Silhavy, T. J. (2012). *Curr. Opin. Microbiol.* **15**, 189–193.
- Sandoval, C. M., Baker, S. L., Jansen, K., Metzner, S. I. & Sousa, M. C. (2011). *J. Mol. Biol.* **409**, 348–357.
- Sheldrick, G. M. (2008). *Acta Cryst. A* **64**, 112–122.
- Wu, T., Malinverni, J., Ruiz, N., Kim, S., Silhavy, T. J. & Kahne, D. (2005). *Cell*, **121**, 235–245.
- Zhang, H., Gao, Z.-Q., Hou, H.-F., Xu, J.-H., Li, L.-F., Su, X.-D. & Dong, Y.-H. (2011). *Acta Cryst. F* **67**, 734–738.