Article

Structure of a nascent membrane protein as it folds on the BAM complex

https://doi.org/10.1038/s41586-020-2370-1

Received: 18 November 2019

Accepted: 25 March 2020

Published online: 11 June 2020

Check for updates

David Tomasek[1,2], Shaun Rawson[3], James Lee[1,2,5], Joseph S. Wzorek[2,6], Stephen C. Harrison[3,4]✉, Zongli Li[3,4]✉ & Daniel Kahne[1,2,3]✉

Mitochondria, chloroplasts and Gram-negative bacteria are encased in a double layer of membranes. The outer membrane contains proteins with a β-barrel structure[1,2]. β-Barrels are sheets of β-strands wrapped into a cylinder, in which the first strand is hydrogen-bonded to the final strand. Conserved multi-subunit molecular machines fold and insert these proteins into the outer membrane[3–5]. One subunit of the machines is itself a β-barrel protein that has a central role in folding other β-barrels. In Gram-negative bacteria, the β-barrel assembly machine (BAM) consists of the β-barrel protein BamA, and four lipoproteins[5–8]. To understand how the BAM complex accelerates folding without using exogenous energy (for example, ATP)[9], we trapped folding intermediates on this machine. Here we report the structure of the BAM complex of *Escherichia coli* folding BamA itself. The BamA catalyst forms an asymmetric hybrid β-barrel with the BamA substrate. The N-terminal edge of the BamA catalyst has an antiparallel hydrogen-bonded interface with the C-terminal edge of the BamA substrate, consistent with previous crosslinking studies[10–12]; the other edges of the BamA catalyst and substrate are close to each other, but curl inward and do not pair. Six hydrogen bonds in a membrane environment make the interface between the two proteins very stable. This stability allows folding, but creates a high kinetic barrier to substrate release after folding has finished. Features at each end of the substrate overcome this barrier and promote release by stepwise exchange of hydrogen bonds. This mechanism of substrate-assisted product release explains how the BAM complex can stably associate with the substrate during folding and then turn over rapidly when folding is complete.

Previous structures of the BAM complex from bacteria have provided clues to the mechanism of substrate folding[13–17]. A notable finding in some of these structures is that the BamA β-barrel component of the machine (Fig. 1a) is present in an open state with unpaired N- and C-terminal β-strands[13–15,17]. This open seam has previously been proposed to pair with substrate β-strands by hydrogen bonding to facilitate the formation of β-sheets, leading to a general model in which folded portions of the substrate undergo 'budding' through the open seam into the membrane[13,18–20]. Crosslinking experiments have provided support for interaction of β-barrel substrates with the edges of the open seam[10–12], but no model for folding has addressed how the BAM complex catalyses the rapid and repeated assembly of substrates into β-barrels in the absence of an exogenous source of energy[9,21,22]. Paired β-strands in a membrane are so stable that the kinetic barrier to disrupt them is enormous. To explain how the folding process is catalysed requires structures of folding intermediates trapped on the BAM complex.

Capturing substrates on the BAM complex

We developed an approach to trap partially folded substrates on the BAM complex of *E. coli*. To do so, we selected BamA as a substrate of the BAM complex, so that BamA within the machine (hereafter $BamA^{M}$) folds other copies of BamA. We generated a series of substrates (hereafter $BamA^{S}$), each of which lacked one of the eight extracellular loops (L1–L8) (Fig. 1b, Extended Data Fig. 1a). We reasoned that loop deletions could prevent the correct formation of β-hairpins, resulting in slowed folding. To test this possibility, we assessed susceptibility of $BamA^{S}$ variants to the periplasmic protease DegP to determine whether folding stalled in the periplasm before membrane insertion[23]. Removal of loops within the C-terminal half of $BamA^{S}$ resulted in susceptibility to DegP, but removal of loops within the N-terminal half did not (Fig. 1c, d). Likewise, urea extraction experiments—which indicate whether a substrate is properly integrated in the membrane—showed that variants of $BamA^{S}$ that lack loops in the C-terminal half of the protein could be extracted by the denaturant, whereas $BamA^{S}$

[1]Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA, USA. [2]Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA, USA. [3]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA, USA. [4]Howard Hughes Medical Institute, Boston, MA, USA. [5]Present address: Laboratory of Membrane Biophysics and Biology, The Rockefeller University, New York, NY, USA. [6]Present address: Novartis Institutes for BioMedical Research, Cambridge, MA, USA. ✉e-mail: harrison@crystal.harvard.edu; zongli_li@hms.harvard.edu; kahne@chemistry.harvard.edu

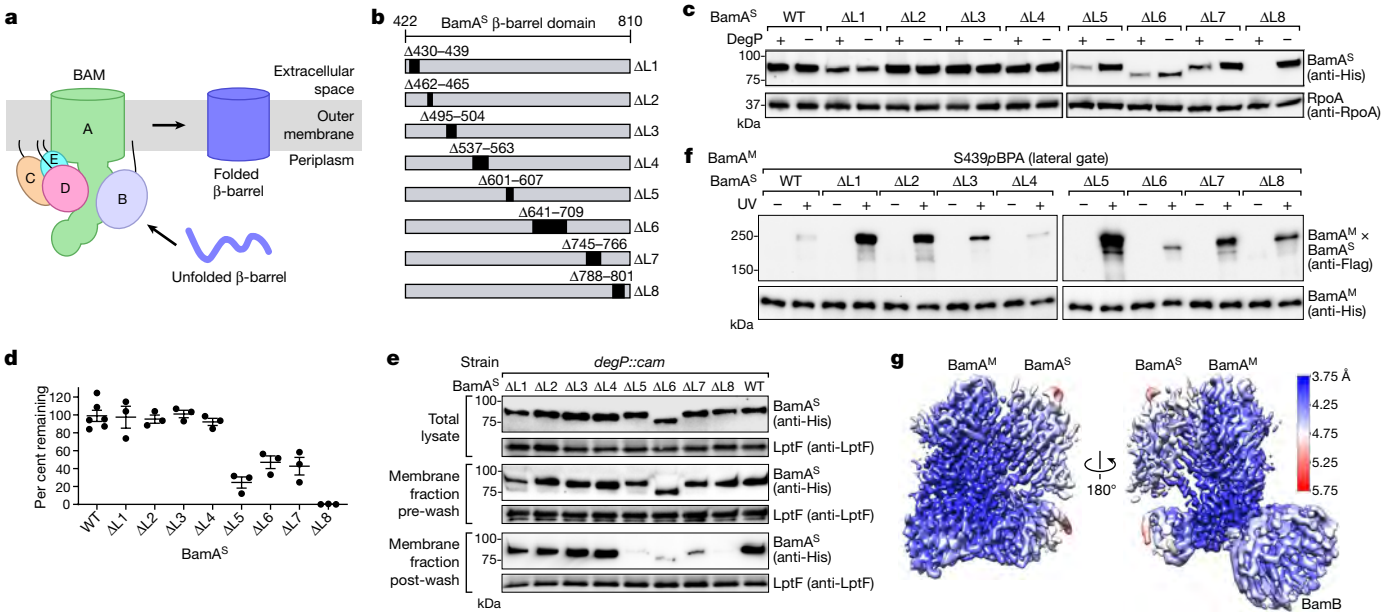


Fig. 1 | Deletion of extracellular loops causes BamA^S to stall at different stages of folding on the BAM complex. a, Overview of β-barrel assembly by the BAM complex in Gram-negative bacteria. Letters A to E designate BamA to BamE, respectively. **b**, Schematic of engineered BamA^S(ΔL1) to BamA^S(ΔL8). The residues deleted in each substrate are indicated. For simplicity, only the β-barrel domain of the substrate is shown. **c**, Expression levels of 6×His-tagged BamA^S in strains with or without *degP*. Anti-RpoA immunoblots are provided as loading controls. WT, wild type. **d**, Quantification of immunoblotting data shown in **c**, with expression levels for each substrate calculated as the percentage that remains when *degP* is expressed. The plotted data represent mean ± s.e.m. from quantification data in **c** and additional independent replicates ($n = 6$ for wild type, $n = 3$ for other substrates). **e**, Urea extraction of 6×His-tagged substrates (BamA^S(ΔL1) to BamA^S(ΔL8)). Samples from total cell lysates (top), membrane fractions before urea incubation (middle) and membrane fractions that remained after urea incubation (bottom) were analysed. Anti-LptF immunoblots are provided as loading controls. Data shown are representative of results from two biological replicates. **f**, In vivo photo-crosslinking of 6×His-tagged full-length BamA^M(S439*p*BPA) to 3×Flag-tagged BamA^S. The strain that lacks *degP* is used to ensure equal expression levels for all mutants, as shown in **c**. POTRA domains 3, 4 and 5 (residues 172–421) are removed in the substrates to prevent them from forming functional BAM complexes if they complete folding. Data shown are representative of results from three biological replicates. **g**, Cryo-EM map of substrate-bound BAM complex, coloured by local resolution.

variants that lack N-terminal loops could not (Fig. 1e). Both of these experiments are consistent with the idea that the C-terminal half (but not the N-terminal half) of the BamA^S β-barrel is needed for integration into the membrane, and is therefore assembled early in the folding process.

To assess whether substrates accumulate on the BAM complex during folding, we incorporated the photo-crosslinkable amino acid *para*-benzoyl phenylalanine (*p*BPA) near the N-terminal β-strand of the seam in BamA^M (at residue S439; hereafter BamA^M(S439*p*BPA)). It has previously been found that upon irradiation with ultraviolet light *p*BPA introduced at this position forms crosslinks to the C-terminal region of a different substrate, LptD, during its folding[12]. Importantly, substrates that fold slowly form stronger crosslinks to BamA^M than do wild-type substrates, because they have a longer residence time. Here we tested crosslinking from BamA^M(S439*p*BPA) to BamA^S variants with a loop deletion; in these substrates, we also removed POTRA domains 3, 4 and 5 because—as has previously been shown[24]—removal of these domains prevents BamA^S from forming BamA–BamB–BamC–BamD–BamE complexes if it does complete folding. We found that BamA^S in which a loop was deleted crosslinked strongly to the N terminus of BamA^M (Fig. 1f). Together, these results show that the removal of extracellular loops from BamA^S leads to the accumulation of the BamA^S variants on the machine outside or within the membrane, depending on the stage in the folding process at which they stall. Our results are consistent with previous reports that have shown that deletions of the extracellular loops of BamA lead to impaired cell growth and lower protein levels[25,26].

Structure of a substrate-engaged BAM complex

We chose BamA^S with loop 1 removed (BamA^S(ΔL1)) for structural studies because it accumulates at a late stage of folding, when it has already integrated into the membrane but remains associated with the BAM complex. To stabilize the interactions between BamA^M and the folding BamA^S(ΔL1), we engineered a disulfide bond between the C-terminal region of BamA^M and N-terminal region of BamA^S (Extended Data Figs. 1b–f, 2a). BamA^S(ΔL1) appears to be capable of adopting a folded state under reducing conditions, although folding occurs slowly (Extended Data Fig. 2b). We purified the BamA–BamB–BamC–BamD–BamE–BamA^S complex (Extended Data Fig. 2c, d) and determined its structure by cryo-electron microscopy (cryo-EM) to an overall resolution of 4.1 Å (Fig. 1g, Extended Data Fig. 3, Supplementary Table 1). The cryo-EM map contains density for all five components of the BAM complex and for BamA^S(ΔL1) (Fig. 2a, b, Extended Data Fig. 4). All of the β-strands of the BamA^M β-barrel were resolved, as were all β-strands of the BamA^S(ΔL1) β-barrel except β1 and β2; extracellular loop 1 would normally connect these two β-strands. The POTRA domains of BamA^S(ΔL1) were also not visible in the structure.

Using previously published X-ray crystallographic structures[15,27], we built atomic models into our cryo-EM map (Fig. 2c). The overall architecture of the BAM complex in its substrate-bound state is similar to that seen in substrate-free states[14–17]; the organization of the periplasmic components most closely resembled that of substrate-free structures that contain a laterally open form of BamA (Extended Data Fig. 5). Structures of the individual components are likewise similar to those already published[14–17] (Extended Data Fig. 6). However, there are notable conformational differences between the N-terminal segment of the BamA^M β-barrel in our substrate-engaged complex and the corresponding segment in any of the substrate-free, laterally open structures[14,15,17] (Fig. 2d). Specifically, the first and second β-strands of BamA^M in the substrate-engaged β-barrel bend outward, separating them more from the C-terminal β-strand than they are in the substrate-free β-barrel.

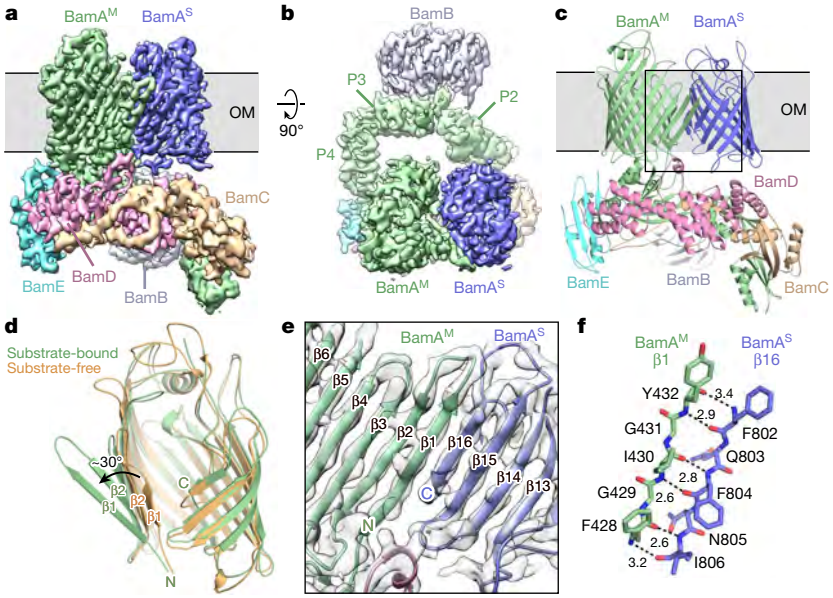


Fig. 2 | The N-terminal β-strand of BamA^M interacts with the C-terminal β-strand of BamA^S by β-sheet augmentation. a, Side view of the cryo-EM reconstruction of the BAM complex bound to BamA^S(ΔL1). Green, BamA^M; silver, BamB; wheat, BamC; pink, BamD; cyan, BamE; blue, BamA^S. This colour scheme is maintained in all figures that show the cryo-EM structure. Approximate boundaries of the outer membrane (OM) are shown. BamC and POTRA domains 1 and 2 of BamA^M are shown at a lower contour level than the rest of the structure. **b**, Top-down view turned 90° from the view in **a**. P2, P3 and P4 refer to POTRA domains 2, 3 and 4 of BamA^M, which are visible in this view. **c**, Side view of the atomic model presented in the same orientation as in **a**. **d**, Overlay of BamA^M from substrate-bound complex (green) and from a substrate-free, laterally open complex (orange) (Protein Data Bank (PDB) code 5EKQ). C, C terminus; N, N terminus. **e**, Enlarged view of boxed region from **c**. The cryo-EM map is shown in grey. **f**, Interactions between β1 of BamA^M and β16 of BamA^S are shown in stick representation. Hydrogen bonds between the β-strands are shown as dashed black lines, and the bond lengths (in Å) are labelled.

Smaller deviations in β3, β4 and β5 allow the larger changes in β1 and β2, but all remaining β-strands superimposed well on previous structures of BamA[14–17] (Fig. 2d, Extended Data Fig. 6).

Similar to the BamA^M β-barrel, the membrane-embedded BamA^S(ΔL1) β-barrel has a lateral opening. The location of BamA^S in the membrane is consistent with our biochemical experiments (Fig. 1). Density was absent for β1 and β2 of BamA^S(ΔL1), presumably because deletion of extracellular loop 1 prevented the correct formation of the corresponding β-hairpin (Extended Data Fig. 7). Additionally, the POTRA domains of the substrate could not be resolved. However, the rest of the β-barrel closely resembled the folded form of BamA, which suggests that the substrate in the structure is an on-pathway intermediate (Extended Data Fig. 7). The N-terminal β-strand of BamA^M and the C-terminal β-strand of BamA^S are paired in an interaction that involves six hydrogen bonds (Fig. 2e, f), creating a BamA^M–BamA^S hybrid barrel with a continuous lumen.

Asymmetric interface of BamA^M–BamA^S

Whereas the C-terminal β-strand of BamA^S is captured by the N-terminal β-strand of BamA^M, the β-strands at the other ends of BamA^S and BamA^M are not hydrogen-bonded (Fig. 3a–c). Instead, these ends curve inwards, so that the β-barrels make contact along their exterior surfaces that would normally contact membrane lipids. Because we used a disulfide bond to tether the C-terminal region of BamA^M to the N-terminal region of BamA^S (to stabilize the complex during purification), it was possible that the asymmetric interface we observed in our structure reflected a nonnative conformation of the complex. To address this, we determined another structure including BamA^S(ΔL1) but used a cysteine tether introduced between N-terminal region of BamA^M and the C-terminal region of BamA^S (Extended Data Fig. 8). This 6.5 Å structure is notably similar to the 4.1 Å structure (Extended Data Fig. 8i), which indicates that using cysteine tethers does not alter the overall conformation of the complex.

The asymmetry of the hybrid β-barrel implies that side chains within BamA^S(ΔL1) that are oriented outward should be exposed to lipopolysaccharide (LPS) molecules in the membrane—except near the N terminus of BamA^S(ΔL1), which should be inaccessible to LPS. We tested this implication by incorporating *p*BPA at four sites that are oriented towards the exterior of BamA^S (Fig. 3c, Extended Data Fig. 9a, b), including at a site on the surface that curves inward (T467) in the hybrid β-barrel. After in vivo photo-crosslinking, we purified the BamA–BamB–BamC–BamD–BamE–BamA^S complex and analysed crosslinks from BamA^S to both LPS and BamA^M (Fig. 3d). The results showed that the T467*p*BPA substitution in BamA^S(ΔL1) yielded a crosslink to BamA^M, whereas the other substitutions (Y531*p*BPA, M741*p*BPA and F804*p*BPA) yielded crosslinks to LPS. These results agree with the finding from our structure that most of the exterior surface of BamA^S(ΔL1) faces the membrane, whereas the N-terminal β-strands interact with BamA^M.

β1 and β2 of BamA^S(ΔL1) are not visible in our structure. To determine their location, we substituted F428 and F440 in BamA^S(ΔL1) with *p*BPA. We observed crosslinks from the substrate to BamA^M (Extended Data Fig. 9c, d). On the basis of where β3 is located in the structure and because the periplasmic loop from β3 to β2 is not long enough to reach the accessory subunits of the BAM complex, we infer that β1 and β2 probably interact with the β-barrel or POTRA domain 5 of BamA^M.

Although there is an unpaired edge at one end each of BamA^M and of BamA^S, inward curvature ensures that these edges face into the lumen of the hybrid barrel, where they are solvated by water. The seal created by the hydrophobic-interaction interface between BamA^M and BamA^S prevents the entry of lipid molecules into the lumen (Fig. 3b, c). Early intermediates may not be able to curve around to form this interface; we did not detect crosslinks to LPS using *p*BPA substitutions within BamA^S(ΔL5) or BamA^S(ΔL8) (Fig. 3d). This finding, combined with the results in Fig. 1, supports the notion that the polypeptide chain takes different routes during early and late folding rather than budding continuously from the BamA^M β-barrel into the membrane[11,20,23,28,29]. Sequential membrane insertion of hairpins would imply that the

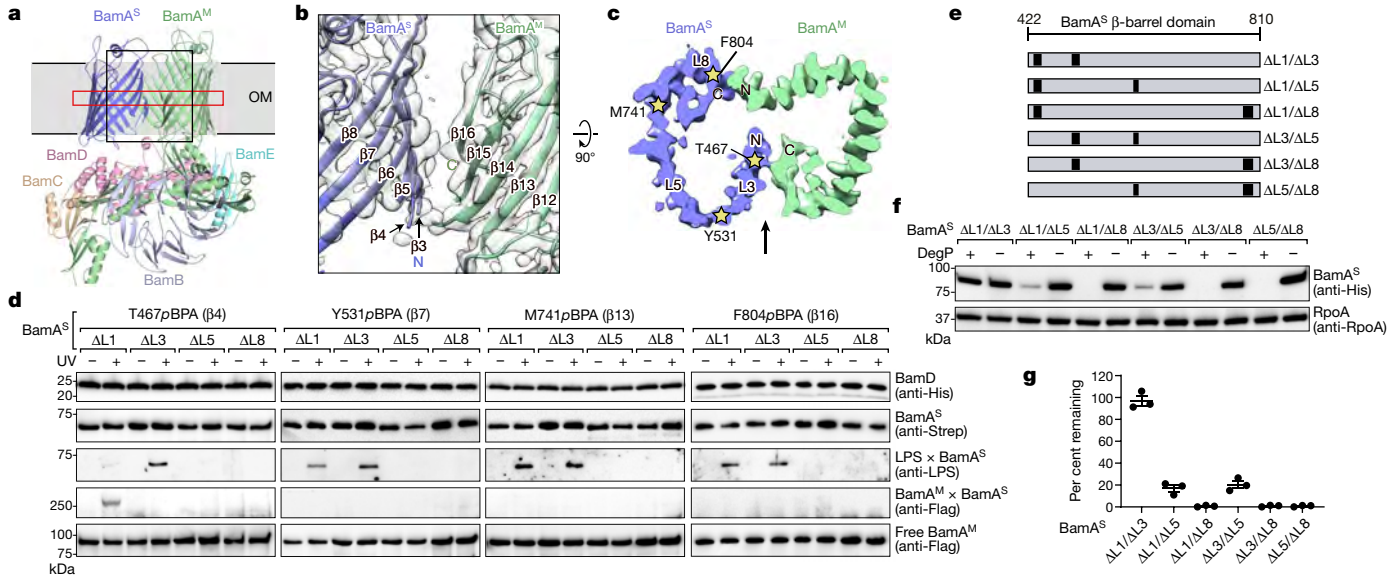


Fig. 3 | The hybrid barrel is asymmetric, as the C-terminal β-strand of BamA^M and the N-terminal β-strand of BamA^S are unpaired in the membrane. a, Side view turned 180° degrees from view in Fig. 2c to show interactions between the C terminus of BamA^M and the N terminus of BamA^S. **b**, Enlarged view of the region outlined in a black box in a. The cryo-EM map is shown in grey. This view shows the 'seal' formed by the interaction of the ends of BamA^M and BamA^S. **c**, Top-down view of a slice through the cryo-EM map, as indicated by the red box in **a**. The N and C termini of BamA^M and BamA^S, and approximate positions of L3, L5 and L8 within BamA^S, are indicated. The approximate positions of *p*BPA substitutions are indicated with stars. The arrow denotes the view of the seal from **b**. **d**, In vivo photo-crosslinking of Strep-tagged substrates (BamA^S(ΔL1) to BamA^S(ΔL8)) that contain *p*BPA at positions T467, Y531, M741 or F804 and a deletion of POTRA domains 3–5. Immunoblotting was performed using anti-His, anti-Strep, anti-LPS and anti-Flag antibodies to detect BamD (loading control), the substrates, substrate-bound LPS and BamA^M, respectively. For the anti-Flag immunoblot, a longer exposure to detect crosslinks (top) and a shorter exposure (bottom) are shown. Data shown are representative of results from two biological replicates. **e**, Schematic of BamA^S in which two extracellular loops are removed. The residues removed are as given in Fig. 1b. **f**, Expression levels of BamA^S containing two loop deletions in strains with or without *degP*. An anti-RpoA immunoblot is provided as a loading control. **g**, Quantification of immunoblotting data shown in **f**, performed as in Fig. 1d. The plotted data are mean ± s.e.m. from quantification of data in **f** and additional independent replicates ($n = 3$ for each substrate).

nonpolar side of any hairpin would always face the lipid (Fig. 3c, d). Specifying the mechanism of early folding will require visualizing an early intermediate.

The interaction between the N-terminal β-strand of the BamA^M β-barrel with the C-terminal edge of the BamA^S β-barrel has direct implications for the folding mechanism. By binding the exposed N-terminal edge of BamA^M as an extended β-strand, the C terminus of the substrate forms a new edge to template binding of the subsequent β-strand, in a process known as β-strand augmentation[30–32]. Therefore, folding will be directional and proceed from the C terminus to the N terminus of the substrate. On the basis of crosslinking experiments, a sequential C-to-N-terminal folding model has previously been proposed for the mitochondrial homologue Sam50[10]. We probed directional folding by generating six variants of BamA^S in which two extracellular loops were removed (Fig. 3e). In each construct, removal of the loop closer to the C terminus determined the overall susceptibility to protease digestion, which reports on membrane integration (Fig. 3f, g). These differences in proteolytic susceptibility enable us to conclude that the region including loop 8 folds before that including loop 5, and so on. This finding establishes C-to-N-terminal directional folding and is consistent with sequential β-strand augmentation.

Mechanism of substrate release

The six hydrogen bonds at the interface between the N-terminal β-strand of the BamA^M β-barrel and the C-terminal β-strand of the BamA^S β-barrel (Fig. 2e, f) prevent premature product release, but the strength of that interaction raises the question of how release occurs at all. In a membrane, disrupting a hydrogen bond requires approximately 4 kcal mol^{-1} (refs. [33,34]). Thus, the kinetic barrier to disrupting all six hydrogen bonds at once would be insurmountably high, even though the process is thermodynamically favourable.

Two unexpected features of the structure suggest a mechanism for substrate β-barrel closure and release. First, the edges of the two β-barrels at the folding end of BamA^S do not engage in hydrogen bonding but instead turn inward towards the water-filled lumen of the hybrid β-barrel. Thus, no polar bonds to the machine need to be broken at that edge. Second, the register of the hydrogen bonds between the N-terminal β-strand of the BamA^M β-barrel and the C-terminal β-strand of the BamA^S β-barrel is not the same as it is between the N- and C-terminal β-strands of a folded BamA β-barrel[13,35]. The hydrogen-bonded residues on the BamA^M side are largely the same as those at the seam of a closed β-barrel, but these residues interact with different residues in the C-terminal β-strand of the trapped BamA^S. In our structure, BamA^M splays open at the extracellular end of its N-terminal β-strand (Fig. 2d), which facilitates pairing with residues in the C-terminal β-strand of BamA^S that precede the residues that are paired with the N-terminal β-strand in the fully folded structure (Figs. 2f, 4a). This pairing creates an overhang at the end of the BamA^S C-terminal β-strand (Fig. 4b, c) by steric clash with the loop between β14 and β15 of BamA^S. These observations suggest a stepwise model for how release occurs (Fig. 4d). When the N-terminal β-strand of BamA^S adds to the β-barrel of BamA^S, it approaches the C-terminal overhang (Fig. 4d). The N and C termini of BamA^S begin to form hydrogen bonds, sequentially disrupting bonds between BamA^S and BamA^M. Unlike concerted disruption, sequential displacement is energetically feasible. As hydrogen bonds form to close the β-barrel of BamA^S, the C-terminal β-strand of BamA^S peels away stepwise from its interaction with the N-terminal β-strand of the BamA^M β-barrel. The accompanying relaxation of strain in the splayed N-terminal hairpin might accelerate the strand exchange.

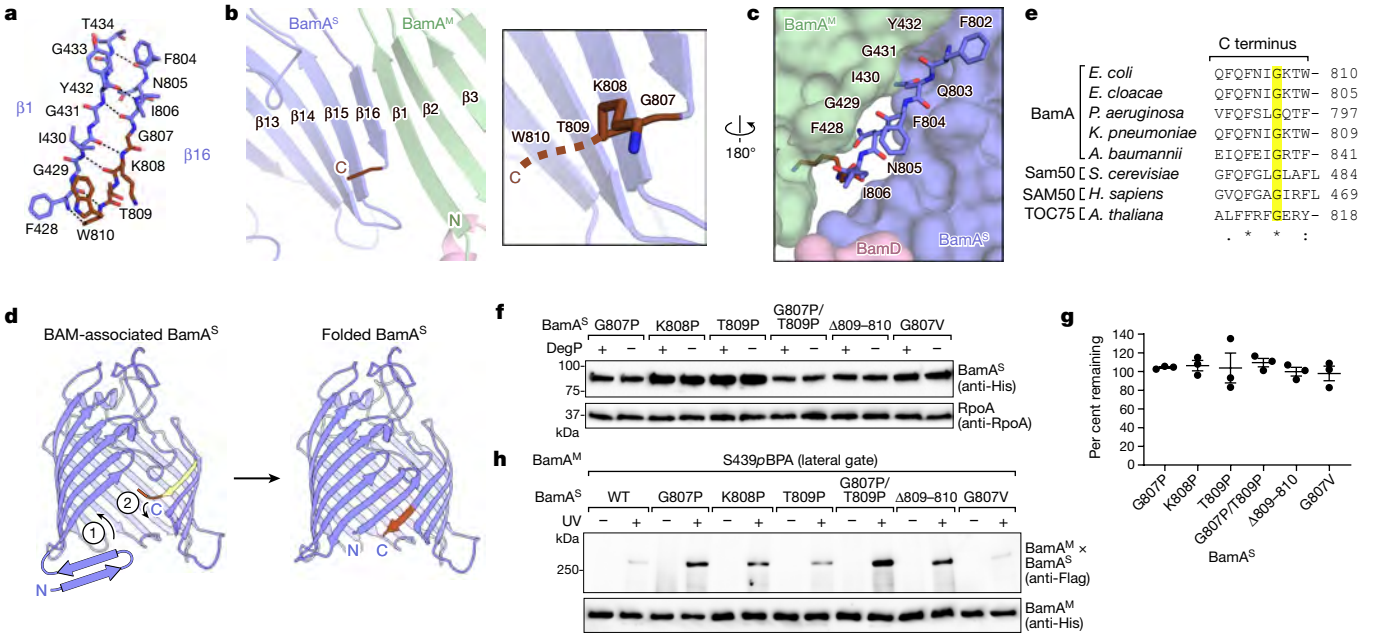


Fig. 4 | Release from BamA^M requires structural features of the N and C termini of BamA^S that allow stepwise hydrogen-bond exchange. **a**, Interactions between β1 and β16 in the folded form of BamA (PDB code 4N75). Residues G807 to W810, which form an overhang in BAM-complex-bound BamA^S (shown in **b**–**d**), are coloured in brown. **b**, Interaction between BamA^M and BamA^S (left). The view is from the interior of the two β-barrels. Residues within the C-terminal overhang of BamA^S are in brown. The final two residues of BamA^S could not be resolved, and their expected location is indicated with a dashed brown line (right). **c**, View of interactions between BamA^M and BamA^S, turned 180° from the view in **b** and similar to the view in Fig. 2e. With the exception of the C-terminal β-strand (β16) and overhang of BamA^S, proteins are shown in semi-transparent surface representation. Residues in β1 of BamA^M and β16 of BamA^S are labelled. **d**, Model of substrate release involving pairing between the C-terminal overhang and N-terminal β-strand of BamA^S. The C-terminal overhang of BamA^S is in brown, and residues that interact with BamA^M in the BAM-complex-associated BamA^S are in yellow. Step 1, addition of N-terminal β-strands; step 2, formation of hydrogen bonds between β1 and the C-terminal overhang displaces the C terminus from BamA^M. **e**, Alignment of the C-terminal regions of BamA from different species of Gram-negative bacteria (*E. coli*, *Enterobacter cloacae*, *Pseudomonas aeruginosa*, *Klebsiella pneumoniae* and *Acinetobacter baumannii*) and other members of the Omp85 superfamily (from *Saccharomyces cerevisiae*, *Homo sapiens* and *Arabidopsis thaliana*). The invariant glycine is highlighted in yellow. **f**, Expression levels of 6×His-tagged BamA^S in strains with or without *degP*. Anti-RpoA immunoblots are provided as loading controls. **g**, Quantification of immunoblotting data shown in **f**, performed as in Fig. 1d. The plotted data represent mean ± s.e.m. from quantification of data in **f** and additional independent replicates ($n = 3$ for each substrate). **h**, In vivo photo-crosslinking of 6×His-tagged full-length BamA^M(S439*p*BPA) to 3×Flag-tagged BamA^S with POTRA domains 3–5 removed. Data shown are representative of results from three biological replicates.

We tested the implication of the strand-exchange mechanism as follows. First, we ruled out a special role for an invariant glycine present in BamA (G807 in *E. coli*) and homologues in mitochondria and chloroplasts, which is just at the boundary of the C-terminal overhang in our structure (Fig. 4e). Changing G807 to valine has previously been reported to cause outer membrane defects[36]. We made the G807V substitution to determine whether it prevents release of an otherwise wild-type BamA^S, but found no defects in BamA^S assembly (Fig. 4f–h) as judged by DegP sensitivity and by crosslinking. These results indicate that the G807V substitution alters BamA function rather than assembly.

Second, we made substrate variants with proline substitutions at positions 807, 808 or 809 in the overhanging C-terminal segment. Proline distorts the conformation and hydrogen bonding in a β-sheet. DegP did not degrade the proline-substituted BamA^S (Fig. 4f, g), which implies that these variants undergo membrane integration. In vivo photo-crosslinking to BamA^M (Fig. 4h) showed that these BamA^S variants accumulate on BamA^M in a membrane-integrated state. Moreover, deleting the C-terminal residues (809 and 810) also caused late-stage assembly defects (Fig. 4f–h). These results suggest that forming hydrogen bonds between the C-terminal overhang and N-terminal β-strand of BamA^S facilitates the release of BamA^S from BamA^M.

Discussion

The structures reported here support the central idea of models in which substrates pass through the side of the BamA β-barrel into the membrane[13,18–20] (Extended Data Fig. 10). It is also consistent with crosslinking studies that have shown that the C-terminal β-strand of the folding substrate interacts strongly with the N-terminal edge of the seam of the BamA^M β-barrel[10–12]. Moreover, the two unexpected features of our structure—the unpaired edges at one junction and the hydrogen-bond register and overhang at the other—resolve problems posed by models that postulate hydrogen-bonded arrays at both edges of the machine. Sequential replacement of each substrate–machine hydrogen bond by a substrate–substrate hydrogen bond yields a succession of rapid steps rather than a single slow one. We believe that the β-barrel closure and release mechanism described here is general: other substrates of the BAM complex could have C-terminal overhangs (Supplementary Table 2) that—when bound to the machine—would protrude inside the hybrid β-barrel to engage their own N termini. With this information, it may be possible to design antibiotics that bind to features of BamA and inhibit β-barrel folding and release[37–40].

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-2370-1.

1. Walther, D. M., Rapaport, D. & Tommassen, J. Biogenesis of β-barrel membrane proteins in bacteria and eukaryotes: evolutionary conservation and divergence. *Cell. Mol. Life Sci.* **66**, 2789–2804 (2009).
2. Hagan, C. L., Silhavy, T. J. & Kahne, D. β-Barrel membrane protein assembly by the Bam complex. *Annu. Rev. Biochem.* **80**, 189–210 (2011).
3. Paschen, S. A. et al. Evolutionary conservation of biogenesis of β-barrel membrane proteins. *Nature* **426**, 862–866 (2003).
4. Wiedemann, N. et al. Machinery for protein sorting and assembly in the mitochondrial outer membrane. *Nature* **424**, 565–571 (2003).
5. Wu, T. et al. Identification of a multicomponent complex required for outer membrane biogenesis in *Escherichia coli*. *Cell* **121**, 235–245 (2005).
6. Voulhoux, R., Bos, M. P., Geurtsen, J., Mols, M. & Tommassen, J. Role of a highly conserved bacterial protein in outer membrane protein assembly. *Science* **299**, 262–265 (2003).
7. Malinverni, J. C. et al. YfiO stabilizes the YaeT complex and is essential for outer membrane protein assembly in *Escherichia coli*. *Mol. Microbiol.* **61**, 151–164 (2006).
8. Sklar, J. G. et al. 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 (2007).
9. Hagan, C. L., Kim, S. & Kahne, D. Reconstitution of outer membrane protein assembly from purified components. *Science* **328**, 890–892 (2010).
10. Höhr, A. I. C. et al. Membrane protein insertion through a mitochondrial β-barrel gate. *Science* **359**, eaah6834 (2018).
11. Doyle, M. T. & Bernstein, H. D. Bacterial outer membrane proteins assemble via asymmetric interactions with the BamA β-barrel. *Nat. Commun.* **10**, 3358 (2019).
12. Lee, J. et al. Formation of a β-barrel membrane protein is catalyzed by the interior surface of the assembly machine protein BamA. *eLife* **8**, e49787 (2019).
13. Noinaj, N. et al. Structural insight into the biogenesis of β-barrel membrane proteins. *Nature* **501**, 385–390 (2013).
14. Bakelar, J., Buchanan, S. K. & Noinaj, N. The structure of the β-barrel assembly machinery complex. *Science* **351**, 180–186 (2016).
15. Gu, Y. et al. Structural basis of outer membrane protein insertion by the BAM complex. *Nature* **531**, 64–69 (2016).
16. Han, L. et al. Structure of the BAM complex and its implications for biogenesis of outer-membrane proteins. *Nat. Struct. Mol. Biol.* **23**, 192–196 (2016).
17. Iadanza, M. G. et al. Lateral opening in the intact β-barrel assembly machinery captured by cryo-EM. *Nat. Commun.* **7**, 12865 (2016).
18. Kim, K. H., Aulakh, S. & Paetzel, M. The bacterial outer membrane β-barrel assembly machinery. *Protein Sci.* **21**, 751–768 (2012).
19. Noinaj, N., Kuszak, A. J., Balusek, C., Gumbart, J. C. & Buchanan, S. K. Lateral opening and exit pore formation are required for BamA function. *Structure* **22**, 1055–1062 (2014).
20. Schiffrin, B., Brockwell, D. J. & Radford, S. E. Outer membrane protein folding from an energy landscape perspective. *BMC Biol.* **15**, 123 (2017).
21. Hagan, C. L. & Kahne, D. The reconstituted *Escherichia coli* Bam complex catalyzes multiple rounds of β-barrel assembly. *Biochemistry* **50**, 7444–7446 (2011).
22. Plummer, A. M. & Fleming, K. G. BamA alone accelerates outer membrane protein folding in vitro through a catalytic mechanism. *Biochemistry* **54**, 6009–6011 (2015).
23. Wzorek, J. S., Lee, J., Tomasek, D., Hagan, C. L. & Kahne, D. E. Membrane integration of an essential β-barrel protein prerequires burial of an extracellular loop. *Proc. Natl Acad. Sci. USA* **114**, 2598–2603 (2017).
24. Hagan, C. L., Westwood, D. B. & Kahne, D. Bam lipoproteins assemble BamA in vitro. *Biochemistry* **52**, 6108–6113 (2013).
25. Gu, Y., Zeng, Y., Wang, Z. & Dong, C. BamA β16C strand and periplasmic turns are critical for outer membrane protein insertion and assembly. *Biochem. J.* **474**, 3951–3961 (2017).
26. Browning, D. F. et al. Mutational and topological analysis of the *Escherichia coli* BamA protein. *PLoS ONE* **8**, e84512 (2013).
27. Albrecht, R. & Zeth, K. Structural basis of outer membrane protein biogenesis in bacteria. *J. Biol. Chem.* **286**, 27792–27803 (2011).
28. Lee, J. et al. Characterization of a stalled complex on the β-barrel assembly machine. *Proc. Natl Acad. Sci. USA* **113**, 8717–8722 (2016).
29. Sikdar, R., Peterson, J. H., Anderson, D. E. & Bernstein, H. D. Folding of a bacterial integral outer membrane protein is initiated in the periplasm. *Nat. Commun.* **8**, 1309 (2017).
30. Harrison, S. C. Peptide-surface association: the case of PDZ and PTB domains. *Cell* **86**, 341–343 (1996).
31. Remaut, H. & Waksman, G. Protein–protein interaction through β-strand addition. *Trends Biochem. Sci.* **31**, 436–444 (2006).
32. Kim, S. et al. Structure and function of an essential component of the outer membrane protein assembly machine. *Science* **317**, 961–964 (2007).
33. White, S. H. & Wimley, W. C. Membrane protein folding and stability: physical principles. *Annu. Rev. Biophys. Biomol. Struct.* **28**, 319–365 (1999).
34. Roseman, M. A. Hydrophobicity of the peptide C| |O···H| |N hydrogen-bonded group. *J. Mol. Biol.* **201**, 621–623 (1988).
35. Ni, D. et al. Structural and functional analysis of the β-barrel domain of BamA from *Escherichia coli*. *FASEB J.* **28**, 2677–2685 (2014).
36. Lundquist, K., Bakelar, J., Noinaj, N. & Gumbart, J. C. C-terminal kink formation is required for lateral gating in BamA. *Proc. Natl Acad. Sci. USA* **115**, E7942–E7949 (2018).
37. Storek, K. M. et al. Monoclonal antibody targeting the β-barrel assembly machine of *Escherichia coli* is bactericidal. *Proc. Natl Acad. Sci. USA* **115**, 3692–3697 (2018).
38. Hart, E. M. et al. A small-molecule inhibitor of BamA impervious to efflux and the outer membrane permeability barrier. *Proc. Natl Acad. Sci. USA* **116**, 21748–21757 (2019).
39. Luther, A. et al. Chimeric peptidomimetic antibiotics against Gram-negative bacteria. *Nature* **576**, 452–458 (2019).
40. Imai, Y. et al. A new antibiotic selectively kills Gram-negative pathogens. *Nature* **576**, 459–464 (2019).

Methods

No statistical methods were used to predetermine sample size. The experiments were not randomized and investigators were not blinded to allocation during experiments and outcome assessment.

SDS–PAGE and immunoblotting

Homemade Tris-HCl 4–20% polyacrylamide gradient gels or Mini-PROTEAN TGX 7.5% precast gels (BioRad) were used with Tris-glycine running buffer. The 2× SDS sample buffer refers to a mixture containing 125 mM Tris (pH 6.8), 4% (w/v) SDS, 30% (v/v) glycerol, 0.005% bromophenol blue, and 5% (v/v) β-mercaptoethanol. SDS–PAGE analyses were performed at 200 V for 45 to 60 min. To analyse purified protein complexes for cryo-EM, SDS–PAGE was performed and followed by staining with Coomassie blue (Alfa Aesar). Coomassie-stained SDS–PAGE gels were imaged using the Gel feature of an Azure Biosystems C400 imager. For western blotting, proteins were transferred onto Immun-Blot PVDF membranes (Bio-Rad) and then incubated with appropriate antibodies. All HRP conjugates were visualized with the Amersham ECL Prime western blotting detection reagent (GE Healthcare). Western blots were imaged using the lowest sensitivity setting of the Chemi feature of an Azure Biosystems C400 imager. Uncropped immunoblots are available in Supplementary Fig. 1.

Plasmids, strains and oligonucleotides

Plasmids, strains and oligonucleotides used in this study are reported in Supplementary Tables 3, 4 and 5, respectively.

Analysis of cellular BamA levels

Derivatives of a plasmid containing 6×His–BamA cloned into the pZS21 vector (pSK476) were generated to contain single loop deletions (pDT175-pDT182), double loop deletions (pDT260-pDT265), or C-terminal overhang mutations (pDT521, pDT536-pDT538, pDT550). MC4100 cells[41] and MC4100 *degP::cam* cells[23] were transformed with these plasmids. The resulting strains were grown in LB supplemented with 50 μg/ml kanamycin (for MC4100) or 50 μg/ml kanamycin and 30 μg/ml chloramphenicol (for MC4100 *degP::cam*) (37 °C, 220 rpm). Once an optical density at 600 nm (OD_{600}) of about 1.0 was reached, the cells from a 1-ml sample were collected by centrifugation (5,000*g*, 10 min, 4 °C). The resulting cell pellets were resuspended in 80 μl of a 1:1 mixture of 2× SDS sample buffer and buffer containing 20 mM Tris-HCl pH 8.0 and 150 mM NaCl. After boiling for 10 min, the samples were applied to SDS–PAGE and analysed via western blotting. The 6×His–BamA was detected by using a penta-His (HRP) antibody (Qiagen). RpoA was detected using a mouse anti-RpoA (*E. coli* RNA polymerase-α) primary antibody (BioLegend, clone 4RA2) followed by a sheep anti-mouse (HRP) secondary antibody (GE Healthcare).

Membrane extraction with urea

Membrane extraction was performed as previously described[42]. MC4100 *degP::cam* strains were generated containing the pZS21 vector that encodes 6×His–BamA or derivatives containing single- or double-loop deletions. These strains were grown overnight in LB supplemented with 50 μg/ml kanamycin and 30 μg/ml chloramphenicol (37 °C, 220 rpm). These cultures were used to inoculate 100 ml of LB with the same additives via 1:100 dilutions. The resulting cultures were grown (37 °C, 220 rpm) until an OD_{600} of about 1.0 was reached. The cells were collected via centrifugation (5,000*g*, 10 min, 4 °C) and resuspended in 4 ml of resuspension solution (50 mM Tris-HCl pH 8.0, 10 μg/ml DNase). Cells were lysed via sonication (on and off at 10-s intervals for a total of 3 min on and 3 min off). The resulting lysate was centrifuged to pellet cell debris (5,000*g*, 10 min, 4 °C). The supernatant was used to prepare normalized cell lysate suspensions to an absorbance at 280 nm value of about 24 (assayed by using NanoDrop). An aliquot of each normalized lysate sample was removed, mixed 1:1 with 2× SDS sample buffer (containing β-mercaptoethanol), and boiled for 10 min. For each sample, 800 μl of the remaining lysate was transferred to 70.1 Ti ultracentrifuge tubes (Beckman Coulter) along with an additional 5 ml of 50 mM Tris-HCl pH 8.0. Membrane fractions were then isolated via ultracentrifugation (100,000*g*, 25 min, 4 °C) using an Optima XE-90 ultracentrifuge (Beckman Coulter). The supernatant was discarded, and the membrane pellet was resuspended in 800 μl of 6 M urea. The resuspensions were incubated at 37 °C for 1 h. An aliquot of each sample, representative of total membrane protein content, was removed, mixed 1:1 with 2× SDS sample buffer (containing β-mercaptoethanol), and boiled for 10 min. The post-wash membrane pellet was isolated from the remainder of each sample via ultracentrifugation (100,000*g*, 45 min, 20 °C) after addition of 5 ml of 50 mM Tris-HCl pH 8.0. The pellets were resuspended in 500 μl of water. An aliquot of each sample was removed, mixed 1:1 with 2× SDS sample buffer (containing β-mercaptoethanol), and boiled for 10 min. The samples were applied to SDS–PAGE and analysed via western blotting. The 6×His–BamA was detected by using a penta-His (HRP) antibody (Qiagen). LptF was detected using LptF antiserum followed by a donkey anti-rabbit (HRP) secondary antibody (GE Healthcare). The source of rabbit anti-LptF antiserum has previously been reported[43].

In vivo photo-crosslinking of BamA to substrates

Photo-crosslinking experiments are based on previously described techniques[44,45], with modifications. The plasmid pSup-BpaRS-6TRN encodes an orthogonal tRNA and aminoacyl-tRNA synthetase to incorporate the unnatural amino acid *p*BPA at amber (TAG) stop codons[46]. This plasmid contains a chloramphenicol resistance cassette, so is not compatible with the MC4100 *degP::cam* strain. To circumvent this, we generated a variant of the pSup-BpaRS-6TRN plasmid that instead contains a spectinomycin resistance cassette. In brief, pSup-BpaRS-6TRN minus the chloramphenicol resistance cassette was amplified, and the spectinomycin resistance cassette was amplified from the pCDFDuet vector (EMD Millipore). The pSup-BpaRS-6TRN backbone and the spectinomycin resistance cassette were joined via Gibson assembly[47] to generate pSup-BpaRS-6TRN($spec^R$) (pDT504).

The forms of 6×His–$BamA^M$ containing different *p*BPA substitutions were cloned into the pZS21 vector to generate pJL77. The 3×Flag-tagged substrates were cloned into the pTrc99a vector to generate pDT209 (wild-type $BamA^S$); pDT201-pDT208 (single-loop deletions); pDT268, pDT270 and pDT272 (double-loop deletions); and pDT526, pDT558-pDT560 and pDT562 (C-terminal overhang mutants). The 3×Flag-tagged substrates also contain deletion of POTRA domains 3, 4 and 5 (Δ172–421) to avoid the possibility of these substrates forming BAM complexes if they completed folding. The deletion does not prevent folding of an otherwise wild-type BamA[24], and allows us to assess $BamA^S$ mutants in terms of their ability to fold rather than function within a BAM complex. In these constructs, POTRA domains 1 and 2 are retained because deletion of all five POTRA domains prevents an otherwise wild-type BamA from folding properly.

MC4100 *degP::cam* strains were generated, each containing pSup-BpaRS-6TRN($spec^R$), pJL77 and one of the substrate-encoding plasmids. These strains were grown overnight in LB supplemented with 50 μg/ml carbenicillin, 50 μg/ml kanamycin, 50 μg/ml spectinomycin and 0.2% (w/v) glucose (37 °C, 220 rpm). These overnight cultures were diluted 1:100 into 100 ml of fresh LB containing the same additives minus glucose but supplemented with 0.8 mM *p*BPA (Bachem), and were grown (37 °C, 220 rpm) to an OD_{600} of about 0.35. After normalization by optical density, each culture was split in half, with one half used directly for irradiation with UV light at 365 nm for 10 min (on ice). Photo-crosslinking was performed by using a UVP Blak-Ray B-100AP high-intensity UV lamp with a 100-W spot bulb. All cells were then pelleted by centrifugation (5,000*g*, 10 min, 4 °C).

For protein purification, pellets were resuspended in 5 ml TBS (20 mM Tris-HCl pH 8.0, 300 mM NaCl) supplemented with 1% (w/v)

Anzergent 3-14 (Anatrace), 20 mM imidazole (pH 8.0), 100 μg/ml lysozyme (Sigma-Aldrich), 1 mM PMSF (Sigma-Aldrich), and 50 μg/ml DNase I (Sigma-Aldrich). Cells were lysed via sonication (on and off at 10-s intervals for a total of 1 min and 30 s on and 1 min and 30 s off). The resulting lysate was centrifuged (10,000*g*, 10 min, 4 °C). An aliquot of the supernatant was taken for analysis of 3×Flag–BamA^{S} expression levels via SDS–PAGE and immunoblotting using a monoclonal anti-Flag M2-peroxidase (HRP) mouse antibody (Sigma-Aldrich). The remainder of the supernatant was incubated with Ni-nitrilotriacetic acid (NTA) resin on a rocking platform (1 h, 4 °C). After removal of unbound proteins, the resin was washed twice with 50 column volumes TBS containing 0.02% Anzergent 3-14 and 20 mM imidazole (pH 8.0). Samples were eluted with 5 column volumes TBS containing 0.02% Anzergent 3-14 and 200 mM imidazole (pH 8.0). Eluates were supplemented with 10% TCA by volume and incubated on ice for 20 min. Precipitated proteins were pelleted via centrifugation (21,130*g*, 10 min, 4 °C). All samples were resuspended in a 1:1 mixture of 2× SDS loading dye (containing β-mercaptoethanol) and 1 M Tris pH 8.0. After boiling for 20 min, each sample was analysed by SDS–PAGE and western blotting. The 6×His–BamA^{M} was detected using a penta-His (HRP) antibody (Qiagen). The 3×Flag–BamA^{S} that was pulled down with 6×His–BamA^{M} was detected by using a monoclonal anti-Flag M2-peroxidase (HRP) antibody (Sigma-Aldrich).

In vivo photo-crosslinking of BamA to LPS

The forms of 6×His–BamA containing different *p*BPA substitutions were cloned into the pZS21 vector to generate pDT411, pDT416 and pDT421-pDT430. MC4100 strains were generated, each containingSup-BpaRS-6TRN (containing a chloramphenicol resistance cassette) and one of the plasmids encoding *p*BPA-substituted BamA. These strains were grown overnight in LB supplemented with 50 μg/ml kanamycin and 30 μg/ml chloramphenicol. These overnight cultures were diluted 1:100 into 100 ml of fresh LB containing the same additives supplemented with 0.8 mM *p*BPA (Bachem), and were grown (37 °C, 220 rpm) to an OD_{600} of about 0.35. The protocols for UV irradiation and protein purification are the same as described in 'In vivo photo-crosslinking of BamA to substrates'.

In vivo photo-crosslinking of substrates to LPS or BamA^{M}, followed by DSP crosslinking to the BAM complex

To observe crosslinking from substrates to LPS or to BamA^{M}, a separate crosslinking protocol was devised to ensure that only substrates stalled on the BAM complex were assessed. In brief, we crosslinked the exterior surface of substrates to their surroundings (using incorporated *p*BPA) and then crosslinked stalled substrates to the BAM complex (using an amine-to-amine crosslinker). This was followed by purification of BAM–substrate complexes and release of the substrate with reducing agent (which breaks the amine-to-amine crosslinks), allowing assessment of UV-dependent crosslinks by SDS–PAGE and western blotting.

The pJH114 plasmid[48] was modified to encode 3×Flag–BamA^{M} and BamD–8×His (while eliminating the 8×His tag on BamE), generating pDT340. The 2×Strep-tagged BamA^{S} substrates containing deletion of POTRA domains 3–5, deletion of an extracellular loop (L1, L3, L5 or L8), and a *p*BPA substitution (at T467, Y531, M741 or F804) were cloned into the pCDF vector to generate pDT436-pDT439 (BamA^{S}(T467*p*BPA)), pDT451-pDT454 (BamA^{S}(Y531*p*BPA)), pDT471-pDT474 (BamA^{S}(M741*p*BPA)), and pDT476-pDT479 (BamA^{S}(F804*p*BPA)). BL21(DE3) strains were generated that contained pSup-BpaRS-6TRN, pDT340, and one of the substrate-encoding plasmids.

These strains were grown overnight in LB supplemented with 50 μg/ml carbenicillin, 50 μg/ml spectinomycin, 30 μg/ml chloramphenicol and 0.2% (w/v) glucose (28 °C, 220 rpm). These overnight cultures were diluted 1:100 into 100 ml of fresh LB containing the same additives, and were grown (37 °C, 220 rpm) to an OD_{600} of about 0.35. At this point, isopropyl β-D-1-thiogalactopyranoside (IPTG) (VWR) and 0.8 mM *p*BPA (Bachem) were added to final concentrations of 0.1 mM and 0.8 mM, respectively. The strains were grown for an additional 90 min. After normalization by optical density, each culture was split in half, with one half used directly for irradiation with UV light at 365 nm for 10 min (on ice). All cells were then pelleted by centrifugation (5,000*g*, 10 min, 4 °C). Pellets were resuspended in 20 ml PBS (20 mM NaH_2PO_4 pH 7.2, 150 mM NaCl) and the amine-to-amine crosslinker dithiobis(succinimidyl propionate) (DSP) (Thermo Fisher Scientific) was added to a final concentration of 0.5 mM. After incubation on a rocking platform (30 to 60 min, room temperature), the crosslinking reaction was quenched via addition of Tris-HCl to a final concentration of 20 mM. Cells were centrifuged (5,000*g*, 10 min, 4 °C) and the pellets were frozen at −80 °C before subsequent purification.

Protein purification was performed similarly as described in 'In vivo photo-crosslinking of BamA to substrates' for testing site-specific in vivo crosslinking without DSP crosslinking, with a few differences. The Ni-NTA wash step was performed using 50 column volumes TBS containing 0.1% Triton X-100 (Sigma-Aldrich), 0.1% SDS (J. T. Baker), and 40 mM imidazole (pH 8.0). The Ni-NTA elution step was performed using 5 column volumes TBS containing 200 mM imidazole (pH 8.0) and no detergent. Instead of using TCA precipitation, each sample was concentrated using an Amicon Ultra 0.5-ml 10-kDa molecular-weight cutoff centrifugal concentrator (EMD Millipore). All samples were supplemented with an equal volume of 2× SDS loading dye (containing β-mercaptoethanol). After boiling for 10 min, each sample was analysed by SDS–PAGE and western blotting. The 8×His–BamD was detected by using a penta-His (HRP) antibody (Qiagen). The 2×Strep–BamA^{S} substrates were detected using a monoclonal Strep-Tag II (HRP) antibody (EMD Millipore). The 3×Flag–BamA^{M} was detected using a monoclonal anti-Flag M2-peroxidase (HRP) antibody (Sigma-Aldrich). LPS was detected using a mouse monoclonal anti-LPS core primary antibody (Hycult Biotech, clone WN1 222-5) followed by a sheep anti-mouse (HRP) conjugate secondary antibody (GE Healthcare).

Assessing cysteine-to-cysteine crosslinking

The 6×His–BamA^{M} containing a cysteine substitution was cloned into the pZS21 vector to generate pDT511-pDT513 (C-proximal cysteine mutant) and pDT400 (N-proximal cysteine mutant). The 3×Flag-tagged substrates (BamA^{S} containing a deletion of POTRA domains 3–5, deletion of loop 1 and a cysteine substitution) were cloned into the pTrc99a vector to generate pDT514-pDT516 (N-proximal cysteine mutant) and pDT481 and pDT486-pDT490 (C-proximal cysteine mutant). Substrates containing the above, but with POTRA domains 3–5 intact, were cloned to generate pDT566-pDT568. MC4100 cells were transformed with one BamA^{M}-encoding plasmid and one substrate-encoding plasmid. The resulting strains were grown overnight in LB supplemented with 50 μg/ml carbenicillin, 50 μg/ml kanamycin and 0.2% (w/v) glucose (37 °C, 220 rpm). These overnight cultures were diluted 1:100 into 100 ml of fresh LB containing the same additives minus glucose, and were grown (37 °C, 220 rpm) to an OD_{600} of about 0.5. Cells were then collected by centrifugation (4,200*g*, 10 min, 4 °C). Cell pellets were resuspended in PBS (20 mM NaH_2PO_4 pH 7.2, 150 mM NaCl). TCEP-HCl (VWR) was then added at a final concentration of 2 mM, and cells were incubated on a rocking platform (20 min, room temperature). Cells were then centrifuged (5,000*g*, 10 min, 4 °C) and again resuspended in PBS. To test for cysteine crosslinking, 1,2-bis(maleimido)ethane (BMOE) (Thermo Fisher) was added at a final concentration of 0.5 mM. After incubation on a rocking platform (40 min, room temperature), the crosslinking reaction was quenched via addition of L-cysteine hydrochloride monohydrate (Alfa Aesar) to a final concentration of 10 mM. Cells were centrifuged (5,000*g*, 10 min, 4 °C) and the pellets were frozen at −80 °C before subsequent purification. To test disulfide bond formation in the absence of crosslinker, cells were frozen after the initial centrifugation step. In each sample, 6×His–BamA^{M} was purified and

crosslinked adducts were detected by western blotting as described in 'In vivo photo-crosslinking of BamA to substrates' for site-specific in vivo photo-crosslinking experiments. β-Mercaptoethanol was not added to samples used to assess disulfide formation in the absence of crosslinker.

Assessing heat modifiability of substrates

The 2×Strep–BamA^{S} containing deletion of POTRA domains 3–5 and the C690S, C700S and T467C mutations was cloned into the pTrc99a vector to generate pDT534. Another plasmid containing the same mutations, but with deletion of loop 1 (Δ430–439) was cloned to generate pDT535. The equivalent constructs with the E800C mutation were cloned to generate pDT509 (loop 1 intact) and pDT510 (loop 1 deleted). MC4100 cells were transformed with pDT534, pDT535, pDT509 or pDT510. The resulting strains were grown overnight in LB supplemented with 50 μg/ml carbenicillin and 0.2% (w/v) glucose (37 °C, 220 rpm). These overnight cultures were diluted 1:100 into 5 ml of fresh LB containing the same additives minus glucose, and were grown (37 °C, 220 rpm) to an OD_{600} of about 0.4. Cells were then collected by centrifugation (4,200*g*, 10 min, 4 °C) and resuspended in 2× SDS loading dye (containing β-mercaptoethanol). Each sample was divided into two, with one half boiled for 5 min and the other half remaining unboiled. All samples were analysed by SDS–PAGE and western blotting. The 2×Strep–BamA^{S} was detected using a monoclonal Strep-Tag II (HRP) antibody (EMD Millipore).

Expression and crosslinking of substrate-bound BAM complex for cryo-EM

The plasmid pJH114 was modified such that BamA^{M} contained the C690S and C700S substitutions, in addition to F804C, generating pDT517. The 2×Strep–BamA^{S} containing deletion of POTRA domains 3-5, deletion of loop 1 and the C690S, C700S and T467C substitutions was cloned into the pBAD33 vector to generate pDT518. BL21(DE3) cells containing pDT517 and pDT518 were grown overnight (37 °C, 220 rpm) in LB supplemented with 50 μg/ml carbenicillin, 30 μg/ml chloramphenicol and 0.2% glucose. This overnight culture was diluted 1:100 into 3× 1.5 l of LB supplemented with 50 μg/ml carbenicillin and 30 μg/ml chloramphenicol. The resulting cultures were grown (37 °C, 220 rpm) until an OD_{600} of about 0.7 was reached. At this point, the temperature was turned down to 30 °C, and cells were allowed to continue shaking. After 20 min, protein expression was induced with IPTG (VWR) and L-(+)-arabinose (Alfa Aesar) at final concentrations of 0.2 mM and 0.1% (w/v), respectively. After three hours of additional shaking, cells were collected via centrifugation (4,200*g*, 10 min, 4 °C). Cell pellets were resuspended in PBS (20 mM NaH_2PO_4 pH 7.2, 150 mM NaCl) and incubated with 0.3 mM copper(II) sulfate (Acros Organics) and 0.3 mM 1,10-phenanthroline (Sigma-Aldrich) for 30 min. Cells were centrifuged (5,000*g*, 10 min, 4 °C) and the pellets were frozen at −80 °C before subsequent purification.

For the low-resolution cryo-EM structure, the purification was performed similarly, but using plasmids pDT397 and pDT500 containing the S439C and E800C mutations, respectively. Cells were treated with BMOE as described in 'Assessing cysteine-to-cysteine crosslinking' rather than with copper(II) sulfate and 1,10-phenanthroline.

Purification of substrate-bound BAM complex for cryo-EM

Cell pellets expressing BAM–substrate complex as described in 'Expression and crosslinking of substrate-bound BAM complex for cryo-EM' were thawed and resuspended in buffer containing 20 mM Tris-HCl pH 8.0, 150 mM NaCl, 100 μg/ml lysozyme (Sigma-Aldrich), 1 mM PMSF (Sigma-Aldrich), 50 μg/ml DNase I (Sigma-Aldrich) and 2.5 mM $MgCl_2$ (Sigma-Aldrich). Cells were lysed using an Emulsiflex C3 (Avestin) at a pressure of 10,000 to 15,000 psi. After lysis, cell debris was removed via centrifugation (5,000*g*, 10 min, 4 °C). Membrane fractions were isolated via ultracentrifugation using a 45 Ti rotor (Beckman Coulter) (37,000 rpm, 45 min, 4 °C) and an Optima XE-90 ultracentrifuge (Beckman Coulter). The membrane pellet was resuspended in buffer containing 20 mM Tris-HCl pH 8.0, 150 mM NaCl, 100 μg/ml lysozyme and 1 mM PMSF. Membrane fractions were solubilized via incubation with 0.75% *n*-dodecyl-β-D-maltopyranoside (DDM) (Anatrace) and 0.5% glyco-diosgenin (GDN) (Anatrace) on a rocking platform (2 h, 4 °C). Unsolubilized material was then isolated via ultracentrifugation in a 70 Ti rotor (Beckman Coulter) (37,000 rpm, 30 min, 4 °C). The supernatant, consisting of solubilized membrane proteins, was removed and supplemented with imidazole (pH 8.0) to a final concentration of 5 mM.

The supernatant was incubated with Ni-nitrilotriacetic acid (NTA) resin (Qiagen) that had been pre-washed with 10 colume volumes buffer W1 (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 10 mM imidazole pH 8.0, 0.02% GDN). After batch binding on a rocking platform (1 h, 4 °C), the resin was washed with 10 column volumes buffer W1. Elution was performed via addition of 5 column volumes buffer E1 (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 200 mM imidazole pH 8.0, 0.02% GDN).

The Ni-NTA eluate was immediately incubated with Strep-Tactin XT Superflow resin (IBA Lifesciences) that had been pre-washed with 10 column volumes buffer W2 (100 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.02% GDN). After batch binding on a rocking platform (1 h, 4 °C), the resin was washed with 15 column volumes buffer W2. Elution was performed via addition of 9 column volumes buffer E2 (100 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM EDTA, 50 mM D-biotin, 0.02% GDN).

The Strep resin eluate was concentrated using an Amicon Ultra 4-ml 100-kDa molecular-weight cutoff centrifugal concentrator (EMD Millipore). The sample was then applied to an ÄKTA Pure (GE Healthcare Life Sciences) for purification via size-exclusion chromatography using a Superdex 200 Increase 10/300 GL column. The protein was eluted in buffer containing 20 mM Tris-HCl pH 8.0, 150 mM NaCl and 0.02% GDN. After elution, protein corresponding to the centre peaks of the chromatogram was concentrated to 5 mg/ml using an Amicon Ultra 0.5-ml 100-kDa molecular-weight cutoff centrifugal concentrator (EMD Millipore). Regarding the purifications performed to obtain both the 4.1 Å and 6.5 Å structures, a final yield of approximately 0.1 mg of complex per litre of bacterial culture could be obtained.

Electron microscopy data collection

Purified substrate-bound BAM complex as described in 'Purification of substrate-bound BAM complex for cryo-EM' was applied to glow-discharged Quantifoil R 2/1 holey carbon 400-mesh copper grids (Quantifoil). Grids were blotted for 4–5 s at 100% humidity with the blot force set to 16, and flash-frozen by liquid-nitrogen-cooled liquid ethane using a Thermo Fisher Scientific Vitrobot Mark IV (Thermo Fisher Scientific). The grid was then loaded onto a Titan Krios G3i electron cryo-microscope (Thermo Fisher) operated at 300 kV accelerating voltage. Image stacks (movies) were recorded on a Gatan Bioquantum K3 Imaging Filter (Gatan) using the super-resolution counting mode and the calibrated magnification of 58717× using SerialEM[49]. The slit of the energy filter was set to 25 eV, with a defocus range between 1.1 and 2.8 μm. The electron dose rate was 17 e^- per physical pixel per second, and the subframe time was set to 0.06 s. A total exposure time of 3 s resulted in 50 subframes per image stack. The total electron dose was 70 e^- per Å^2 (about 1.4 e^- per Å^2 per subframe). The multishot scheme in SerialEM was used for data collection, with settings of four holes per stage move, and five shots per hole, which greatly sped up the data collection. The data collection for both structures (4.1 Å and 6.5 Å) was performed in the same manner.

Image processing and 3D reconstruction

The movie frames were motion-corrected and dose-weighted by MotionCor2[50] and contrast transfer function (CTF) parameters were estimated by CTFFIND4[51]. Particle picking was carried out using crYOLO[52] giving 2,054,956 initial particles. Following successive rounds

of 2D and 3D classification within Relion[53], 516,419 particles were then polished through the Bayesian polishing process[54]. These particles were then subjected to further 2D and 3D classification within CryoSPARC[55], after which 223,353 particles were selected (see classification flowchart in Extended Data Fig. 3) and subjected to nonuniform refinement in CryoSPARC, which led to the final reconstruction at 4.1 Å resolution. Masking strategies did not yield an improved structure. Maps used for figures were filtered according to local resolution with *B*-factor sharpening within CryoSPARC.

For the low-resolution (6.5 Å) structure, the movie frames were motion-corrected and dose-weighted by MotionCor2[50] and CTF parameters were estimated by CTFFIND4[51]. Particle picking was carried out using crYOLO[52] giving 690,143 initial particles. Following successive rounds of 2D and 3D classification within Relion[53], 233,064 particles were then polished through the Bayesian polishing process[54]. Refinement in Relion led to the final reconstruction at 6.5 Å resolution. Masking strategies did not yield an improved structure. Maps used for figures were filtered according to local resolution with *B*-factor sharpening within Relion. Structural biology applications used in this project (other than CryoSPARC) were compiled and configured by SBGrid[56].

Model building, refinement and validation

The atomic model was generated using available structures of the BAM-complex components. For $BamA^{M}$–BamB–BamD–BamE, an initial model was generated by rigid-body-fitting these components into the 4.1 Å electron microscopy map using UCSF Chimera[57]. The $BamA^{M}$ POTRA domains and BamE were obtained from PDB code 5D0O. The $BamA^{M}$ β-barrel and BamD were obtained from PDB code 5D0Q. BamB was obtained from PDB code 2YH3. For the substrate, a homology model generated with I-TASSER was used[58–60].

The initial modelling was followed by manual adjustments using Coot[61]. All selenomethiones were replaced with methionines. The improved model was then refined in real space against the cryo-EM map using real space refinement in PHENIX[62] with secondary structure restraints. Iterative rounds of manual and automated refinement in Coot and PHENIX, respectively, generated a model that included all components except for BamC and the POTRA domains 1 and 2 of $BamA^{M}$. Owing to the lower resolution of these components, the atomic models were generated by docking previously determined structures into our map (from 5D0Q for BamC and 5D0O for the POTRA domains) without further refinement. For the entirety of POTRA domains 1 and 2 of $BamA^{M}$ (residues 1–171) and the C-terminal domain of BamC (residues 89–209), only the main chain atoms are included in the model. For the substrate, the two POTRA domains could not be resolved, so they are not included in the model. The final model was visually inspected for general fit to the map, and geometry was further evaluated using MolProbity[63]. Cryo-EM data collection, refinement and validation statistics are summarized in Supplementary Table 1. Figures depicting the structure were prepared in Chimera or PyMOL (Schrödinger, https://www.pymol.org).

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

The electron microscopy maps have been deposited into the Electron Microscopy Data Bank (EMDB) under accession codes EMD-20969 (4.1 Å) and EMD-21313 (6.5 Å). The atomic model, built from the 4.1 Å map, has been deposited into the Protein Data Bank (PDB) under accession code 6V05. All data are available in the Article and its Supplementary Information.

41. Casadaban, M. J. Transposition and fusion of the *lac* genes to selected promoters in Escherichia coli using bacteriophage lambda and Mu. *J. Mol. Biol.* **104**, 541–555 (1976).
42. Leo, J. C., Oberhettinger, P. & Linke, D. Assessing the outer membrane insertion and folding of multimeric transmembrane β-barrel proteins. *Methods Mol. Biol.* **1329**, 157–167 (2015).
43. Villa, R. et al. The *Escherichia coli* Lpt transenvelope protein complex for lipopolysaccharide export is assembled via conserved structurally homologous domains. *J. Bacteriol.* **195**, 1100–1108 (2013).
44. Freinkman, E., Chng, S.-S. & Kahne, D. The complex that inserts lipopolysaccharide into the bacterial outer membrane forms a two-protein plug-and-barrel. *Proc. Natl Acad. Sci. USA* **108**, 2486–2491 (2011).
45. Hagan, C. L., Wzorek, J. S. & Kahne, D. Inhibition of the β-barrel assembly machine by a peptide that binds BamD. *Proc. Natl Acad. Sci. USA* **112**, 2011–2016 (2015).
46. Ryu, Y. & Schultz, P. G. Efficient incorporation of unnatural amino acids into proteins in *Escherichia coli*. *Nat. Methods* **3**, 263–265 (2006).
47. Gibson, D. G. et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
48. Roman-Hernandez, G., Peterson, J. H. & Bernstein, H. D. Reconstitution of bacterial autotransporter assembly using purified components. *eLife* **3**, e04234 (2014).
49. Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* **152**, 36–51 (2005).
50. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
51. Rohou, A. & Grigorieff, N. CTFFIND4: fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
52. Wagner, T. et al. SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM. *Commun. Biol.* **2**, 218 (2019).
53. Scheres, S. H. W. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012).
54. Zivanov, J., Nakane, T. & Scheres, S. H. W. A Bayesian approach to beam-induced motion correction in cryo-EM single-particle analysis. *IUCrJ* **6**, 5–17 (2019).
55. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
56. Morin, A. et al. Collaboration gets the most out of software. *eLife* **2**, e01456 (2013).
57. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
58. Yang, J. et al. The I-TASSER suite: protein structure and function prediction. *Nat. Methods* **12**, 7–8 (2015).
59. Zhang, Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* **9**, 40 (2008).
60. Roy, A., Kucukural, A. & Zhang, Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protocols* **5**, 725–738 (2010).
61. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
62. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
63. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66**, 12–21 (2010).
64. Lee, J. et al. Substrate binding to BamD triggers a conformational change in BamA to control membrane insertion. *Proc. Natl Acad. Sci. USA* **115**, 2359–2364 (2018).

Acknowledgements We thank S. Sterling, R. Walsh and M. Chambers for assistance with electron microscopy; R. Losick, S. Walker, A. Kruse, M. May, K. Pahil, R. Taylor and M. Mandler for providing feedback on the manuscript; and H. Bernstein (NIH) for the kind gift of the pJH114 plasmid. Cryo-EM data were collected at the Harvard Cryo-EM Center for Structural Biology at Harvard Medical School. This work was supported by NIH grants F31GM116210 (to J.L.), F32GM108258 (to J.S.W.), and R01GM066174 and R01AI081059 (to D.K.). S.C.H. is an Investigator in the Howard Hughes Medical Institute.

Author contributions D.T. designed and all performed biochemical experiments with assistance from J.L. and J.S.W. at early stages of the project. D.T. performed protein purification and screened samples using negative-stain electron microscopy. Z.L. collected cryo-EM data. Z.L. and S.R. processed cryo-EM data. D.T. and S.R. built the atomic models. D.K. and S.C.H. supervised the project. D.T. and D.K. wrote the manuscript and all authors contributed to editing.

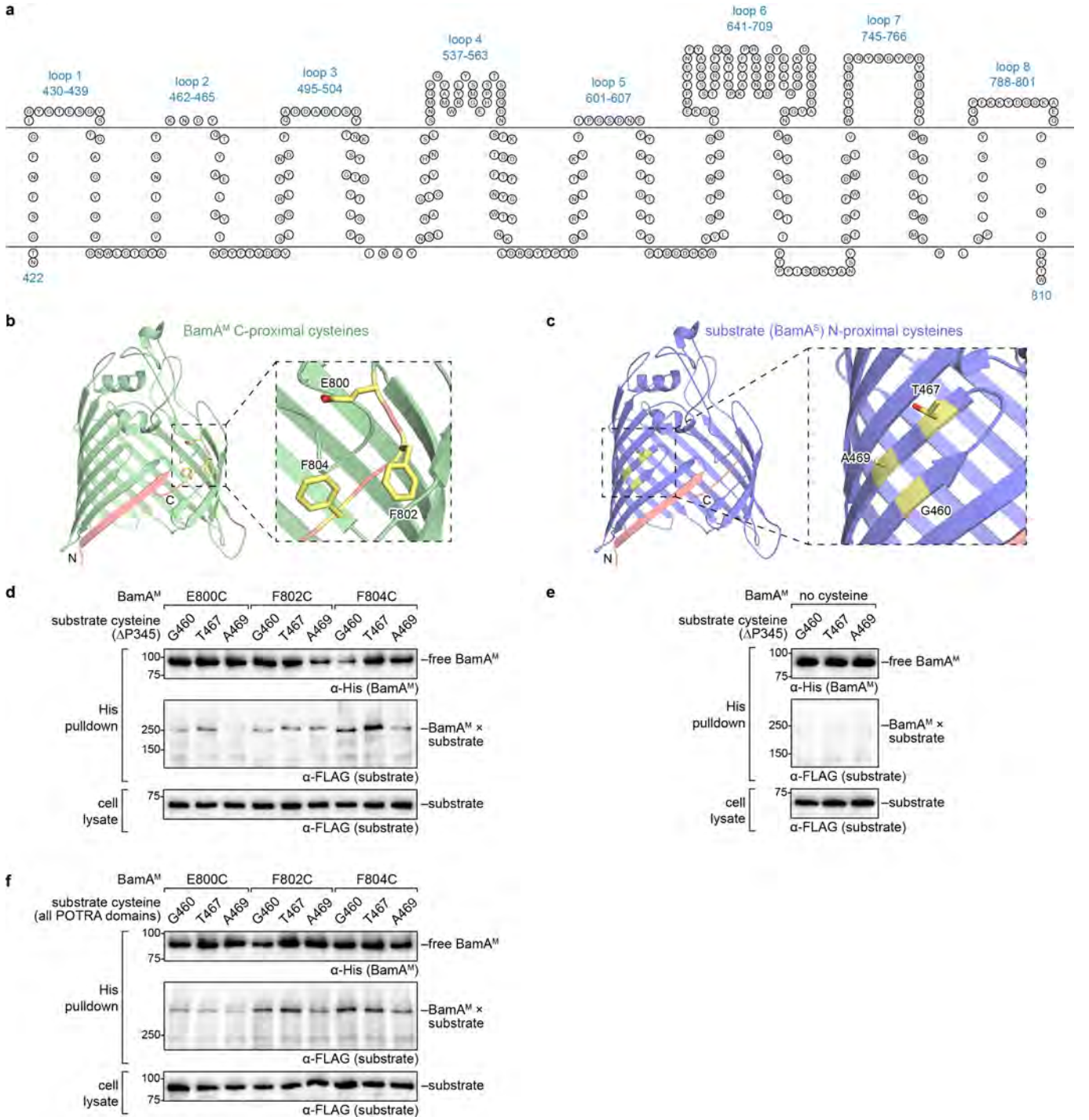
Competing interests The authors declare no competing interests.

Additional information

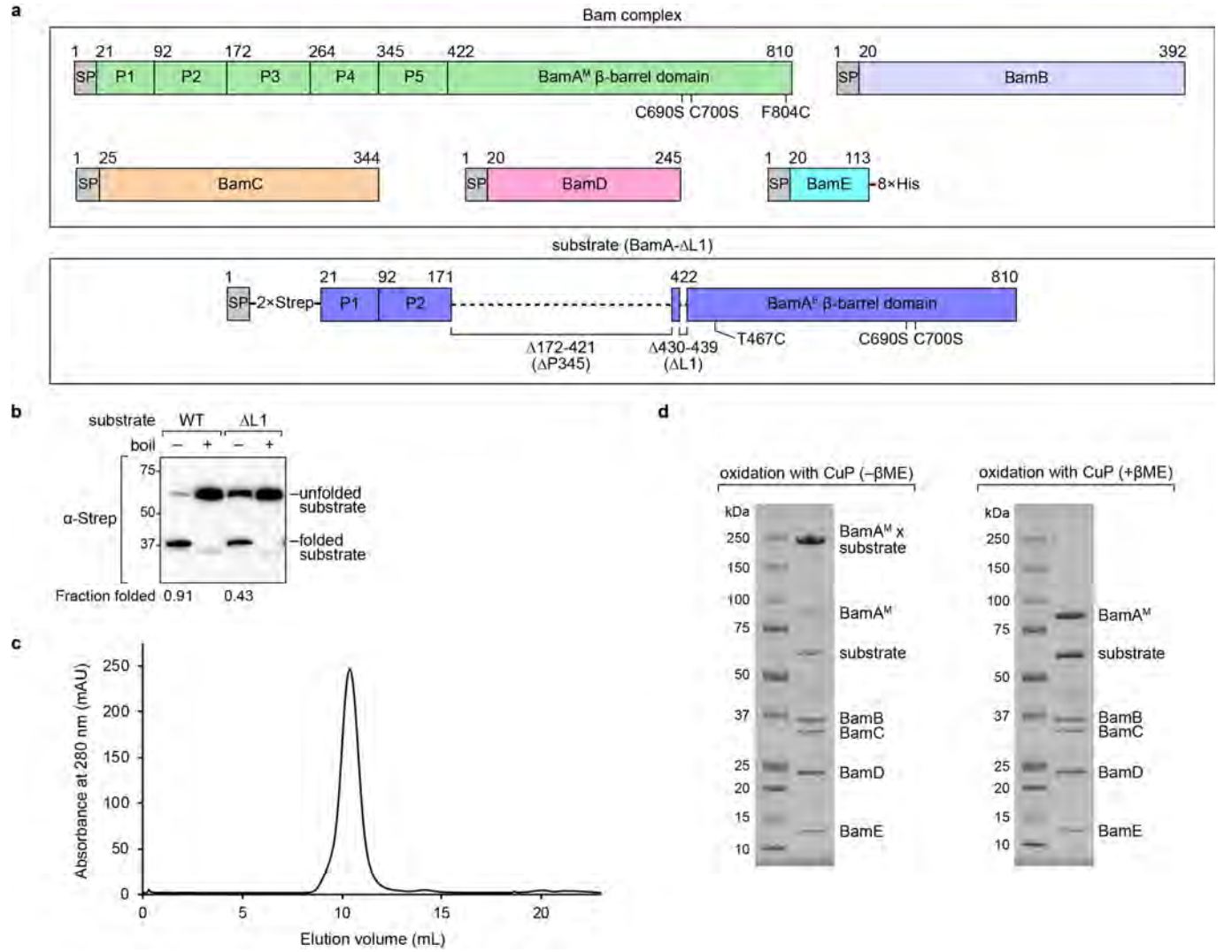
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-020-2370-1.

Correspondence and requests for materials should be addressed to S.C.H., Z.L. or D.K.

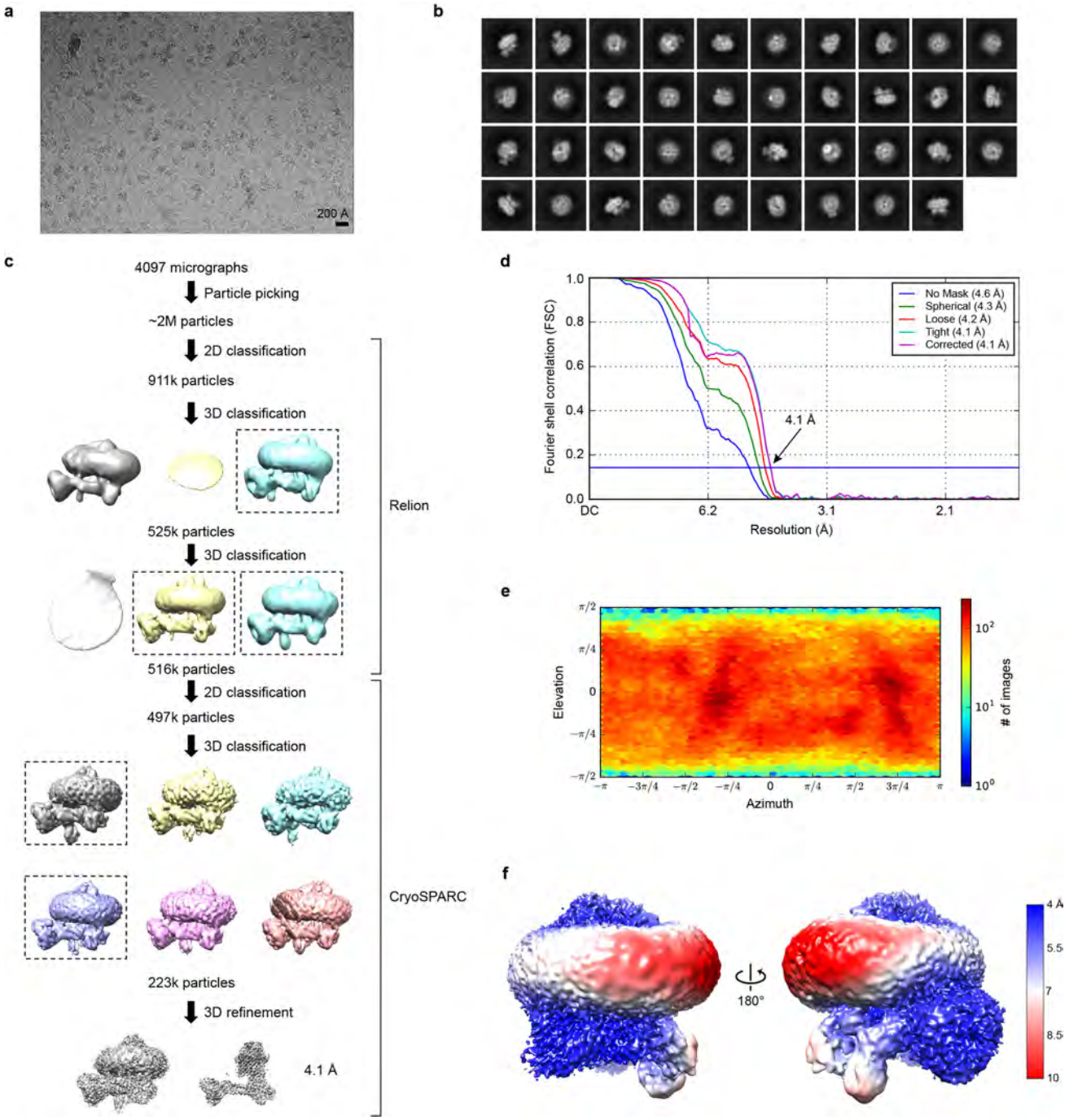
Peer review information *Nature* thanks Yihua Huang, Nikolaus Pfanner, Bert van den Berg and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.



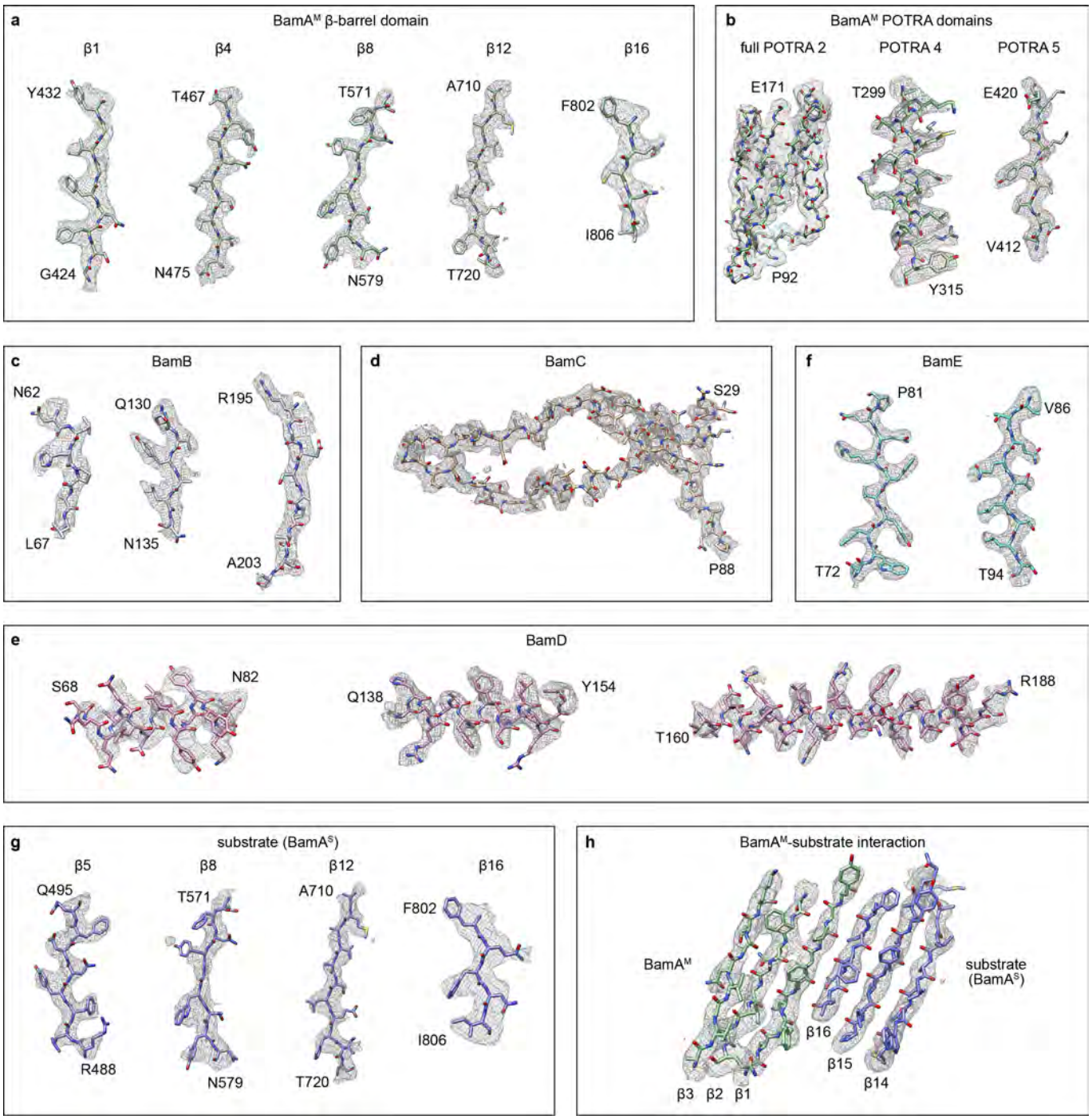
Extended Data Fig. 1 | Capturing BamA^S(ΔL1) on the BAM complex using disulfide-bond formation. a, Topological map of the β-barrel domain of BamA with extracellular loops labelled. **b**, Residues near the C terminus of the β-barrel domain of BamA^M that were substituted with cysteine (yellow sticks) shown on the structure of the β-barrel domain of BamA (green) (PDB code 5D0O). The N- and C-terminal β-strands are labelled and shown in salmon. **c**, As in **b**, but showing residues near the N terminus of the β-barrel domain of BamA^S(ΔL1) (blue) that were substituted with cysteine. The positions of these cysteine substitutions were selected on the basis of reports that the C-terminal region of BamA^M is in proximity to the N-terminal region of the substrate during folding[10–12]. **d**, Disulfide-bond formation between 6×His-tagged BamA^M and 3×Flag-tagged BamA^S(ΔP345/ΔL1) containing the cysteine substitutions shown in **b** and **c**. The presence of a BAM-complex–substrate disulfide bond is detected as a high-molecular-weight adduct on the anti-Flag (middle) immunoblot. The adducts generated via disulfide-bond formation were not present in an amount high enough for detection with the anti-His antibody but could be detected with the anti-Flag antibody. An anti-Flag immunoblot of total cell lysates (bottom) shows that the BamA^S variants containing different cysteine substitutions are expressed at similar levels. In this experiment, no oxidizing agent (for example, copper(II) sulfate and 1,10-phenanthroline) was added. **e**, As in **d**, but without any cysteine substitutions introduced into BamA^M. Disulfide-bond formation is not observed between BamA^M and BamA^S when no cysteine is introduced into BamA^M. **f**, As in **d**, but with BamA^S containing all five POTRA domains. Disulfide-bond formation between BamA^M and BamA^S containing all POTRA domains indicates that the stalling of BamA^S in **d** is not due to the deletion of POTRA domains 3, 4 and 5. Data shown in **d**–**f** are representative of results from two biological replicates.



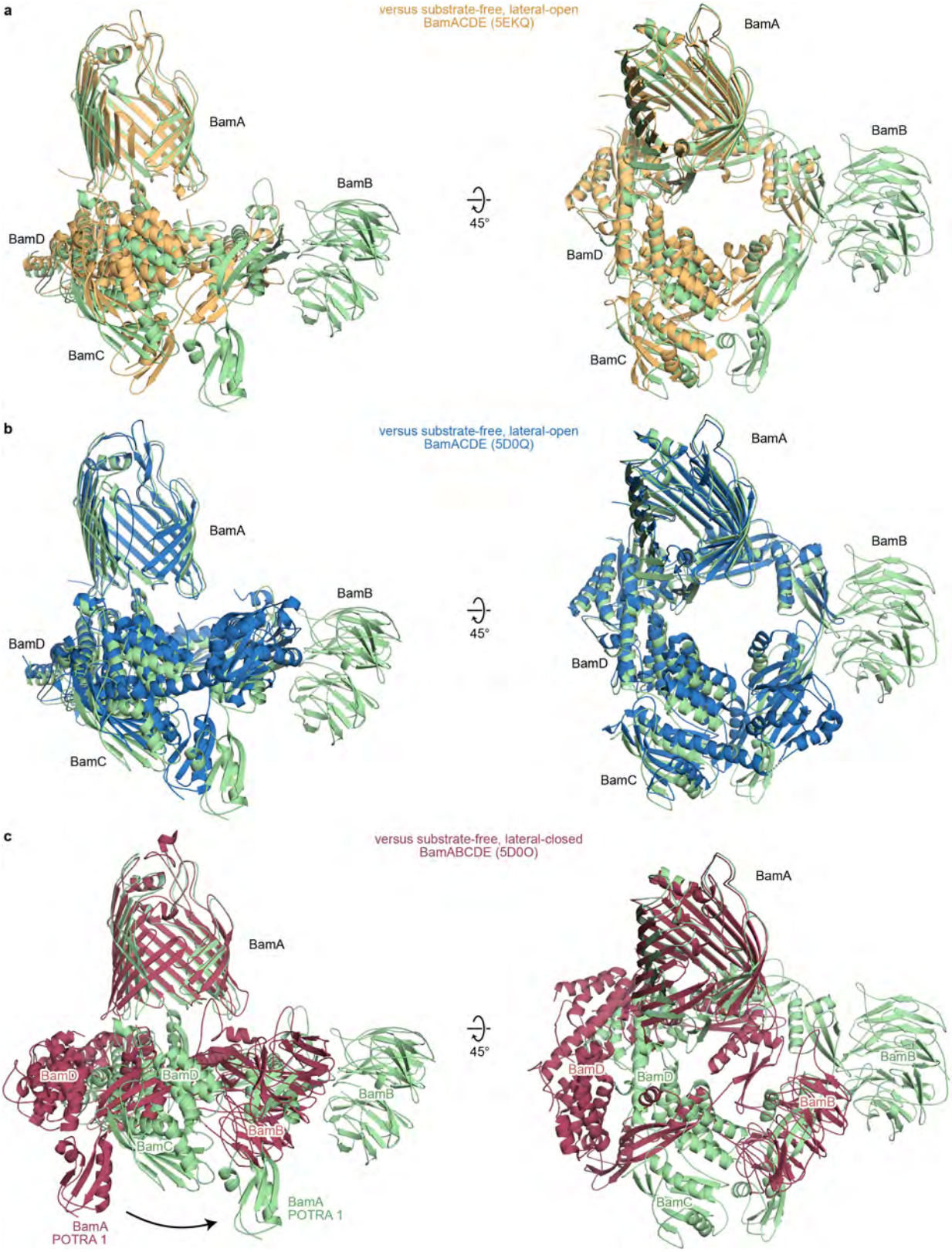
Extended Data Fig. 2 | Expression and purification of substrate-bound BAM complex for cryo-EM. a, Constructs used for protein expression. Mutations and affinity tags introduced into each protein are indicated. SP, signal peptide; P1 to P5, POTRA domains of BamA. The BamA^S variants also contain deletion of POTRA domains 3, 4 and 5 (Δ172–421) to avoid the possibility of them forming BAM complexes if they complete folding. The deletion does not prevent folding of an otherwise wild-type BamA, as shown in **b**. **b**, Assessment of heat modifiability of BamA^S(ΔL1) used for cryo-EM (ΔL1; right half of blot) that contains the T467C mutation, and BamA^S that contains loop 1 and the T467C mutation (WT; left half of blot). The samples in the unboiled lanes can be used to calculate the fraction of each substrate that is folded in vivo (indicated below the blot). BamA^S(ΔL1) accumulates on the BAM complex because of the deletion of loop 1. **c**, Representative size-exclusion chromatogram of the substrate-bound BAM complex in which disulfide-bond formation between the N terminus of BamA^S(ΔL1) and the C terminus of BamA^M was induced with the oxidizing agent copper sulfate and 1,10-phenanthroline (CuP). **d**, SDS–PAGE gels showing peak fractions from size-exclusion chromatography of the complex in **c**. The left gel shows the complex without addition of β-mercaptoethanol (βME), and the right gel shows the complex after addition of βME to break the disulfide bond. Data shown in **b**–**d** are representative of results from two biological replicates.



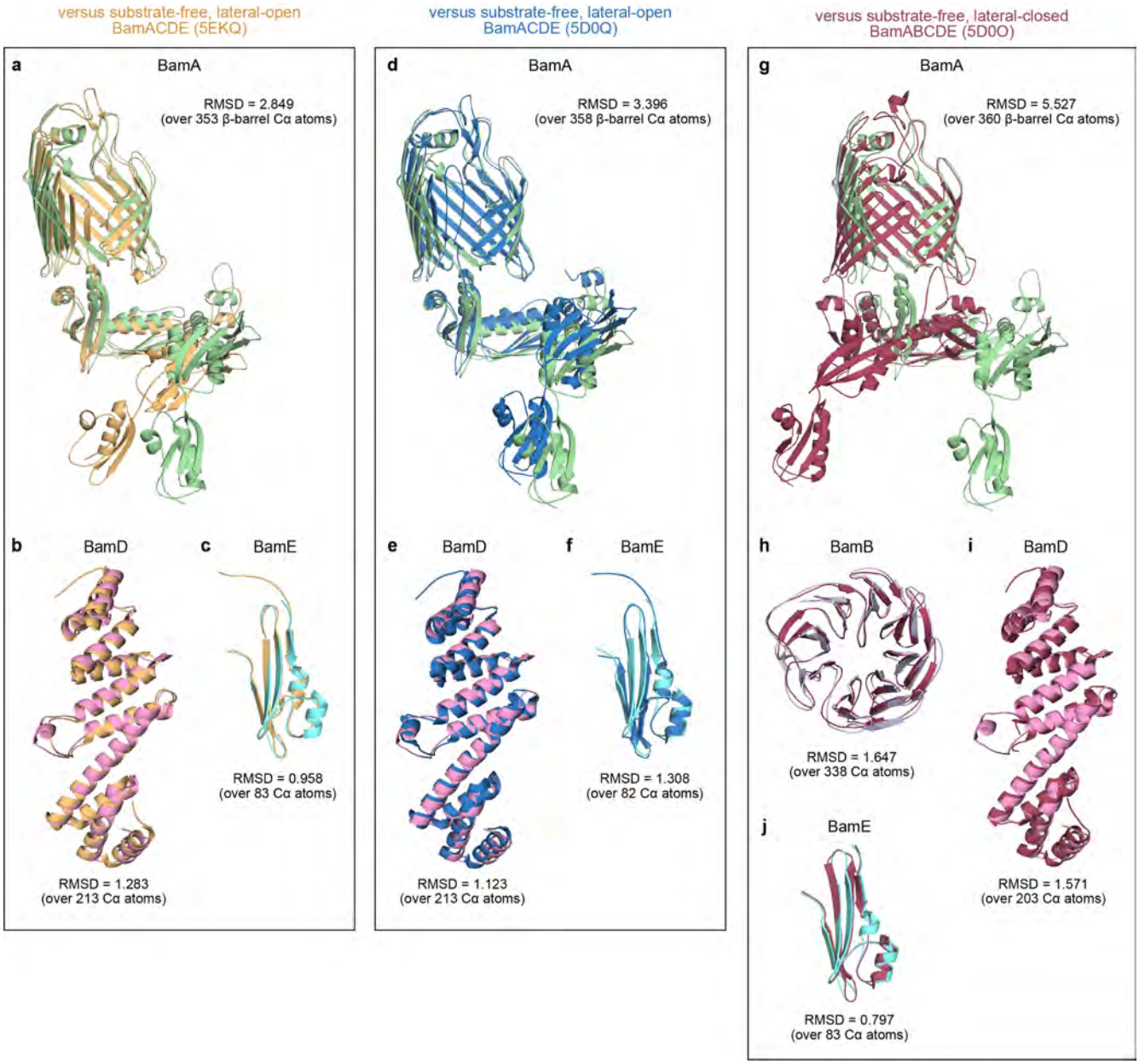
Extended Data Fig. 3 | Cryo-EM data processing and analysis for the substrate-bound BAM complex. a, Representative cryo-EM micrograph of the substrate-bound BAM complex embedded in vitreous ice. **b**, Selected two-dimensional class averages of cryo-EM particle images. **c**, Scheme of three-dimensional classification and refinement of cryo-EM particle images. **d**, Gold-standard Fourier shell correlation (FSC) curves calculated with different masks in cryoSPARC. The resolutions were determined at FSC = 0.143 (horizontal blue line). The final corrected mask gave an overall resolution of 4.1 Å. **e**, Distribution of orientations over azimuth and elevation angles for particles included in the calculation of the final map. **f**, Cryo-EM map coloured by local resolution as shown in Fig. 1g, but at a lower contour level.



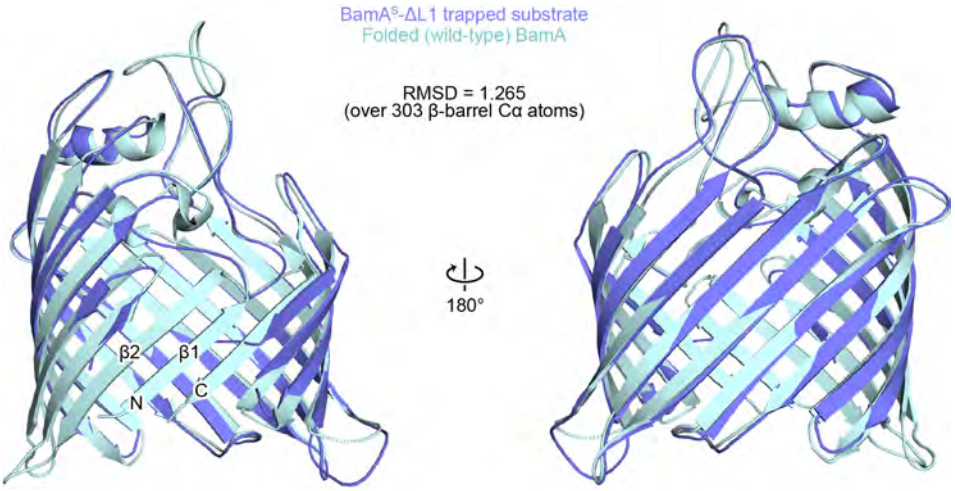
Extended Data Fig. 4 | Fit of the atomic model into the cryo-EM map. **a**–**h**, The atomic model (in stick representation) shown with the corresponding portion of the cryo-EM map (shown in grey mesh) for selected regions in the BamA^{M} β-barrel domain (**a**); the BamA^{M} POTRA domains (**b**); BamB (**c**); BamC (**d**); BamD (**e**); BamE (**f**); BamA^{S} (**g**); and the BamA^{M}–BamA^{S} interaction (**h**). Side-chain densities are visible, and individual β-strands can be resolved. Images were prepared in UCSF Chimera using a 2 Å carve radius.



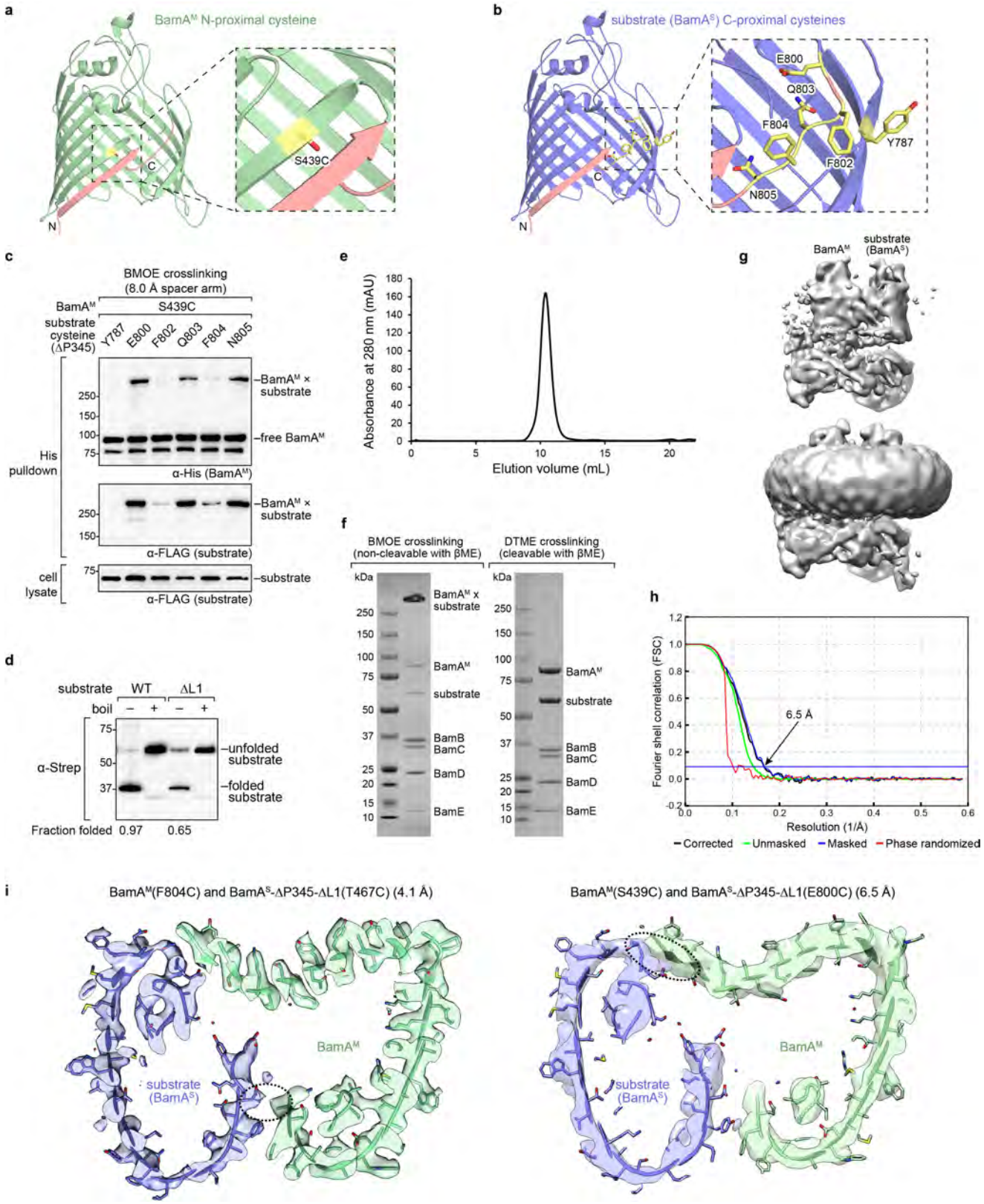
Extended Data Fig. 5 | Comparison of the architecture of the substrate-bound BAM complex to that of substrate-free BAM complexes. **a**–**c**, Alignments of the substrate-bound BAM complex (all components shown in green) with published structures[14,15] of the BAM complex (shown in different colours). Alignments are performed using the β-barrel domains of BamA.



Extended Data Fig. 6 | Conformational differences between BAM-complex components from substrate-bound and substrate-free complexes. **a**–**c**, Alignments of atomic models of BamA^{M} (**a**), BamD (**b**) and BamE (**c**) from our substrate-bound complex (coloured as in Fig. 2) with the corresponding components from a substrate-free, laterally open BAM complex (light orange) (PDB code 5EKQ). **d**–**f**, As in **a**–**c**, but using a different substrate-free, laterally open BAM complex (blue) (PDB code 5D0Q). **g**–**j**, Alignments of atomic models of BamA^{M} (**g**), BamB (**h**), BamD (**i**) and BamE (**j**) from our substrate-bound complex with the corresponding components from a substrate-free, laterally closed BAM complex (dark red) (PDB code 5D0O). Alignments of BamA^{M} are performed using the β-barrel domains. Alignments of BamB are not shown for complexes in which BamB is absent. Alignments of BamC are not shown because BamC in our structure was obtained by docking BamC as a rigid body from PDB code 5D0Q. The root mean square deviation (RMSD) value for each alignment, obtained with PyMOL, is indicated.

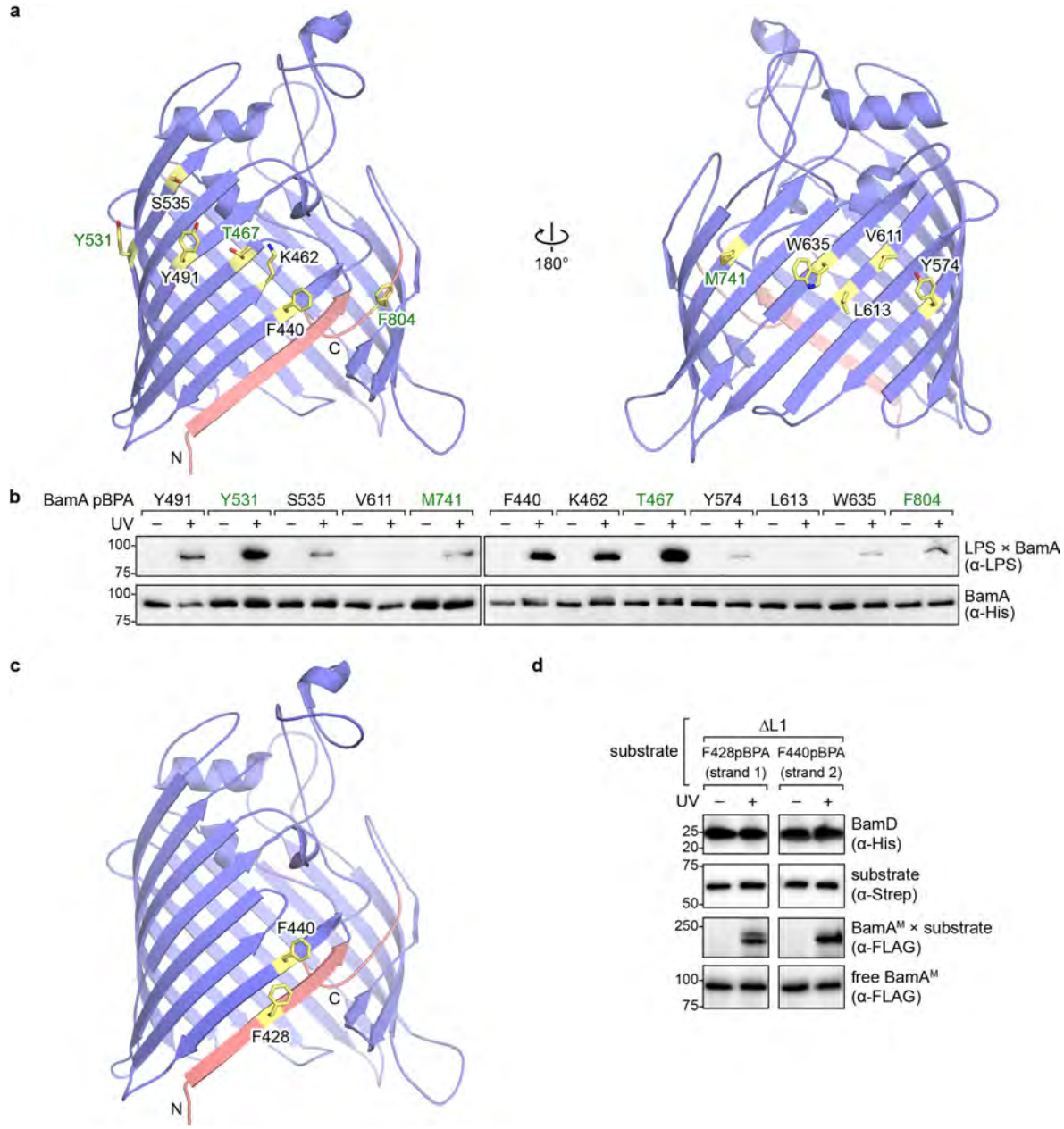


Extended Data Fig. 7 | BamA^S(ΔL1) accumulates on the BAM complex in a largely folded state. Alignment of the atomic model of the BamA^S(ΔL1) substrate (blue) with that of folded wild-type BamA (cyan) (PDB code 4N75). The first two β-strands (β1 and β2, labelled on the structure of wild-type BamA) of BamA^S(ΔL1) in our structure could not be resolved. The similarity between the trapped BamA^S(ΔL1) and folded wild-type BamA suggests that the former represents an on-pathway folding intermediate.

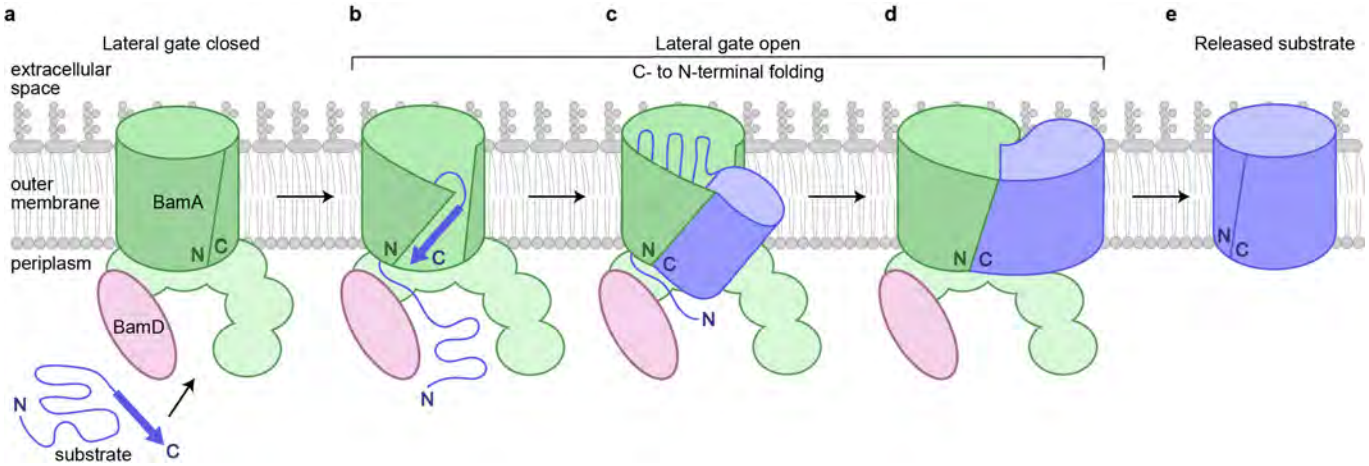


Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | Low-resolution cryo-EM structure of a substrate-bound BAM complex with an alternative cysteine crosslink. a, Residue near the N terminus of the β-barrel domain of $BamA^{M}$ that was substituted with cysteine (yellow stick) shown on the structure of the β-barrel domain of BamA (green) (PDB code 5D0O). The N- and C-terminal β-strands are labelled and shown in salmon. **b**, As in **a**, but showing residues near the C terminus of the β-barrel domain of $BamA^{S}$(ΔL1) (blue) that were substituted with cysteine. **c**, Cysteine crosslinking between 6×His-tagged $BamA^{M}$ and 3×Flag-tagged $BamA^{S}$(ΔP345/ΔL1) containing the cysteine substitutions shown in **a** and **b**. The crosslinker 1,2-bis(maleimido)ethane (BMOE), which has an 8.0 Å spacer arm, was used. The presence of a BAM-complex–substrate crosslink is detected as a high-molecular-weight adduct on anti-His (top) and anti-Flag (middle) immunoblots. An anti-Flag immunoblot of total cell lysates (bottom) shows that $BamA^{S}$ containing different cysteine substitutions is expressed at similar levels. On the basis of these results, E800C was selected for use within $BamA^{S}$(ΔL1) for cryo-EM and S439C was introduced into $BamA^{M}$. **d**, Assessment of heat modifiability of $BamA^{S}$(ΔL1) used for cryo-EM (ΔL1; right half of blot) that contains the E800C substitution, and $BamA^{S}$ that contains loop 1 and the E800C substitution (WT; left half of blot). The samples in the unboiled lanes can be used to calculate the fraction of each substrate that is folded in vivo (indicated below the blot). $BamA^{S}$ containing the E800C mutation retains the ability to fold in vivo. **e**, Representative size-exclusion chromatogram of the substrate-bound BAM complex in which cysteine-crosslink formation between the C terminus of $BamA^{S}$(ΔL1) and the N terminus of $BamA^{M}$ was induced with BMOE. **f**, SDS–PAGE gels showing peak fractions from size-exclusion chromatography. The left gel shows the complex from **e** with addition of βME, which cannot break the crosslink formed by BMOE. The right gel shows the complex crosslinked with dithiobismaleimidoethane (DTME), which is similar to BMOE but has a 13.3 Å spacer arm and is cleavable in the presence of βME to visualize the individual components of the complex. **g**, Cryo-EM map obtained using the $BamA^{M}$(S439C) and $BamA^{S}$(ΔP345/ΔL1/E800C) cysteine pair crosslinked with BMOE. The two images show the same view at different contour levels. $BamA^{M}$ and $BamA^{S}$ are labelled. **h**, Gold-standard Fourier shell correlation (FSC) curves for the low-resolution cryo-EM structure calculated in Relion. The final corrected mask gave an overall resolution of 6.5 Å at FSC = 0.143 (horizontal blue line). **i**, Comparison of the 4.1 Å structure and the 6.5 Å structure that contains the alternative cysteine pair in $BamA^{M}$ and $BamA^{S}$. Left, a top-down slice through the 4.1 Å cryo-EM map is shown (similar to the slice in Fig. 3c) and is overlaid with the atomic model generated from that map. Right, a slice, similar to that on the left, through the 6.5 Å cryo-EM map is shown and is overlaid with the atomic model generated from the 4.1 Å map after fitting of the atomic model into the 6.5 Å map as a rigid body. The dashed oval represents the approximate location of the cysteine tether introduced into each complex. The use of different cysteine pairs yielded similar cryo-EM structures, showing that the presence of the disulfide bond in the main (4.1 Å) structure did not introduce a nonnative conformation into the complex. Data shown in **c**–**f** are representative of results from two biological replicates.



Extended Data Fig. 9 | In vivo photo-crosslinking with wild-type BamA and BamA^{S}(ΔL1). a, Residues in BamA substituted with the photo-crosslinkable amino acid *p*BPA (yellow sticks) shown on the structure of the β-barrel domain of BamA (blue) (PDB code 5D0O). These *p*BPA substitutions were used to observe crosslinks to LPS. The N and C termini of BamA are shown in salmon. All highlighted residues have side chains oriented outward, towards the membrane environment. **b**, In vivo photo-crosslinking of BamA to LPS. In **a** and **b**, the *p*BPA substitutions that were subsequently used to test crosslinking of stalled substrates to LPS (in Fig. 3d) are indicated in green. **c**, Residues in β1 and β2 of BamA^{S}(ΔL1) substituted with *p*BPA shown on the structure of the β-barrel domain of BamA. Colours are as described in **a**. These positions were selected because β1 and β2 of the substrate are not visible in the cryo-EM structure (Extended Data Fig. 7). **d**, In vivo photo-crosslinking of Strep-tagged BamA^{S}(ΔL1) containing *p*BPA at positions F428 or F440 and deletion of POTRA domains 3–5. Immunoblotting was performed using anti-His, anti-Strep and anti-Flag antibodies to detect BamD (loading control), BamA^{S}(ΔL1) and BamA^{M}, respectively. For anti-Flag immunoblot, a longer exposure (top) to detect crosslinks, and a shorter exposure (bottom) are shown. Although β1 and β2 of BamA^{S}(ΔL1) are not visible in the cryo-EM structure, the in vivo photo-crosslinking results show that residues within these β-strands are in proximity to BamA^{M}. The experiment in **b** was performed once, and the *p*BPA substitutions indicated in green were tested again in Fig. 3d. Data shown in **d** are representative of results from two biological replicates.



Extended Data Fig. 10 | Proposed model of β-barrel assembly by the BAM complex. a, The substrate is recruited to the BAM complex. $BamA^{M}$ is in a closed state before the substrate-induced opening of its lateral gate[64]. For simplicity, only $BamA^{M}$, the substrate and BamD are shown. **b**–**d**, The C terminus of the substrate interacts with the exposed N-terminal edge of $BamA^{M}$, and β-strands or β-hairpins of the substrate are added sequentially from the C terminus to the N terminus. Early folding may occur within the interior of the $BamA^{M}$ β-barrel[12], and folded portions of the substrate may then be released outward. Full membrane integration could occur after a substantial amount of folding. The steps shown in **b** and **c** represent intermediate stages to that shown in **d**, which corresponds to the cryo-EM structure of the substrate-bound BAM complex. **e**, The substrate is released into the membrane environment once its N- and C-terminal ends are joined.