1

Structural basis for the function of the β-barrel assembly-enhancing protease BepA

2
3
4

5 Mohammad Shahrizal[1 †], Yasushi Daimon[2 †], Yoshiki Tanaka[1 †], Yugo Hayashi[1],
6 Shintaro Nakayama[1], Shigehiro Iwaki[1], Shin-ichiro Narita[3], Hironari Kamikubo[1], Yoshinori Akiyama[2 *],
7 and Tomoya Tsukazaki[1 *]

8

9 [1]Nara Institute of Science and Technology, Ikoma, Nara 630-0192, Japan
10 [2]Institute for Frontier Life and Medical Sciences, Kyoto University, Kyoto 606-8507, Japan
11 [3]Faculty of Nutritional Sciences, University of Morioka, Iwate 020-0694, Japan

12

13 [†]These authors contributed equally.
14 *Correspondence should be addressed to Y.A. (yakiyama@infront.kyoto-u.ac.jp) and T.T.
15 (ttsukazaki@mac.com)

16

17 **Abstract**

18 The β-barrel–assembly machinery (BAM) complex mediates the assembly of β-barrel membrane proteins in
19 the outer membrane. BepA, formerly known as YfgC, interacts with the BAM complex and functions as a
20 protease/chaperone for the enhancement of the assembly and/or degradation of β-barrel membrane proteins.
21 To elucidate the molecular mechanism underlying the dual functions of BepA, its full-length
22 three-dimensional structure is needed. Here, we report the crystal structure of full-length BepA at 2.6-Å
23 resolution. BepA possesses an N-terminal protease domain and a C-terminal tetratricopeptide repeat (TPR)
24 domain, which interact with each other. Domain cross-linking by structure-guided introduction of disulfide
25 bonds did not affect the activities of BepA *in vivo*, suggesting that the function of this protein does not
26 involve domain rearrangement. The full-length BepA structure is compatible with the previously proposed
27 docking model of BAM complex and TPR domain of BepA.
28

29 **Introduction**

30 Gram-negative bacteria, including *Escherichia coli*, possess two layers of biological membranes,
31 namely the outer membrane and the inner membrane, between which exists the periplasmic space containing
32 a peptidoglycan layer [1]. The two membranes act as a barrier to protect cells from changes in the external
33 environment and hazardous substances. In addition, these membranes are involved in several stress responses,
34 including the activation of the σ^E factor, which occurs on the cytoplasmic side, by monitoring the changes in
35 the states of the outer membrane and the periplasm [2-5]. The outer membrane, which possesses an
36 asymmetric architecture, is mainly composed of phospholipids on the inner leaflet and lipopolysaccharides on
37 the outer leaflet, together with peripherally associated lipoproteins and embedded β-barrel proteins. All
38 β-barrel outer membrane proteins are synthesized in the cytoplasm, translocated across the inner membrane
39 via the Sec complex, and inserted into the outer membrane via the BAM (β-barrel assembly machinery)
40 complex [6-8], where they are converted into folded mature forms.

41 Precursors of outer membrane proteins with an N-terminal signal peptide that are targeted to the inner
42 membrane are initially synthesized by ribosomes and transferred to the Sec machinery, which consists of
43 SecA ATPase, SecYEG complex, and SecDF [9, 10]. Following their transfer, the machinery drives the
44 translocation of unfolded precursor proteins across the inner membrane using the ATP hydrolysis energy and
45 proton motive force. The signal sequence is cleaved by the signal peptidase during this process and the
46 processed proteins are released into the periplasm. At the final stage, the proteins are inserted into the outer
47 membrane via the BAM complex and assembled while undergoing quality control and the assistance of
48 targeting to the membrane by periplasmic chaperones/proteases such as Skp, DegP, SurA, and FkpA,
49 resulting in their conversion into functional β-barrel proteins [11-13]. The central, essential component of the
50 BAM complex is an evolutionarily conserved β-barrel protein BamA [14], which belongs to the Omp85
51 protein family, which includes Sam50 in mitochondria and Toc75 in chloroplast. *In E. coli*, BamA forms a
52 complex with four lipoproteins, namely BamB, BamC, BamD, and BamE [6-8].

53 The σ^E-dependent stress response, which is an essential regulation system in *E. coli,* is activated by the
54 accumulation of misfolded outer membrane proteins. One of 114 genes identified as members of *E. coli* σ^E

55 regulon encodes a periplasmic protein BepA (formerly YfgC) [5]. A BepA-deficient *E. coli* strain was shown
56 to be more sensitive to specific antibiotics than wild-type cells, presumably due to the disrupted barrier
57 functions of the outer membrane. Studies of the biogenesis of a β-barrel outer membrane protein LptD, which
58 is involved in the transport and assembly of lipopolysaccharides, suggest that BepA is related to the
59 biogenesis and quality control of LptD [15, 16]. LptD, which has two intramolecular disulfide bonds in its
60 mature form, undergoes disulfide bond rearrangement during its biogenesis. BepA promotes this
61 rearrangement and degrades incorrectly folded LptD. Impairment in biosynthesis of BepA-dependent LptD
62 would cause reduction of the outer membrane barrier functions, leading to increased drug sensitivity of the
63 cells. In addition, it was proposed that BepA interacts with the BAM complex to act in LptD folding and
64 degrades a stalled LptD on the BAM complex. BepA was also proposed to cleave BamA that is misassembled
65 during its biogenesis under a chaperone (SurA)-depleted condition. [15-17]. These results suggest that BepA
66 functions as a chaperone for outer membrane protein biogenesis and as a protease for the removal of
67 misfolded outer membrane proteins.

68 The mature $BepA_{28-487}$ (MW ~51,000) protein, which is generated by cleavage of the signal peptide from
69 the precursor form, is composed of a N-terminal protease domain and a tetratricopeptide repeat (TPR)
70 domain. According to the MEROPS protein database [18], the protease domain is classified under the
71 metalloprotease M48C subfamily, which shares a HEXXH motif. The active center is proposed to be formed
72 by a Zn atom and residues H136, E137, and H140 of the HEXXH motif and E201 of the following
73 C-terminal side (Fig. S1). Mutations H136R or E137Q in the HEXXH motif affect LptD biogenesis, implying
74 that protease activity is crucial for BepA function [15]. Deletion mutants of the TPR domain abolish the
75 activities of BepA, indicating that the TPR domain is also essential [16]. The crystal structure of the TPR
76 domain revealed that 10 successive anti-parallel alpha-helices formed a large palm structure with a negatively
77 charged pocket and a small palm structure with a cavity [16]. Generally, TPR domains provide a pocket for
78 interactions with partners [19, 20]. However, the interactions between the pocket of the BepA TPR domain
79 and a substrate have not been elucidated to date. In contrast, site-specific photocrosslinking experiments
80 revealed that several points at the edges of the TPR domain interact with Bam proteins and LptD. Based on

81 these results, a model of the BAM complex with the TPR domain of BepA was proposed (Fig. S2) [16].
82 However, in order to unravel the underlying molecular mechanism in detail, the full-length structure of BepA
83 is needed. In this study, we determined the full-length structure of BepA and proposed a functional
84 mechanism for its activity in consideration of the stability of this protein in solution.
85

86 **Results and discussion**

87 **Overall structure of full-length BepA**

88 Initial phases were calculated by molecular replacement with the crystal structures of the BepA TPR
89 domain (PDB ID 5XI8) [16] and a Zn-dependent peptidase Q74D82 (PDB ID 3C37) from *Geobacter*
90 *sulfurreducens*. The crystal structure of $\text{BepA}_{45\text{-}482}$ at 2.6-Å resolution was modeled with $R_{\text{work}} = 21.5$ % and
91 $R_{\text{free}} = 25.9$ % (Fig. 1, Table 1). The asymmetric unit of the *P*1 space group contains six BepA molecules
92 (designated Mol A-F), the superimpositions of which showed that the overall structures exhibit no significant
93 structural change, with a root mean square deviation (RMSD) of 0.47-0.82 Å for the Cα atoms (Fig. S3).
94 Here, we discuss the structure of BepA using Mol B, which has the clearest electron density map among Mol
95 A-F. The crystal structure of BepA consists of two domains: the N-terminal residues 45-288 form a M48C
96 protease domain with eleven α-helices (α1-α11) and a β-sheet with three β-strands (β1-β3) between α2 and
97 α3 (Fig. 1A,B, Fig. S1), similar to the zinc peptidase Q74D82 (PDB ID 3C37); the successive C-terminal
98 residues 305-482 form the TPR domain with ten anti-parallel α-helices (H1-H10) in accordance with the
99 previously reported TPR domain structure [16] (Fig. S4). The α12 tightly links the α11 of the protease
100 domain and the H1 of the TPR domain. The residues 153-159 between α5-α6 and residues 178-191 between
101 α6-α7 were not modeled because of the poor electron density (Fig. 1B dotted line). The α-helices H1, H3,
102 and H5 in the N-terminal region of the TPR domain are strongly associated with α-helices α2, α11, and α12
103 in the protease domain via hydrophobic/hydrophilic interactions, resulting in a 1,126 Å^2 interaction area
104 between the 45-304 and 305-390 residues. The C-terminal regions of the TPR domain show several
105 interactions with the protease domain, including the formation of hydrogen bonds between the O of N68 and
106 Nδ2 of N411, between Nη1 of R61 and the O of L439, and between Oε1 of Q51 and Nη1 of R480, creating a

107 505 Å^2 surface area between the 45-304 and 391-482 residues (Fig. 1B).

108

109 **Active site of BepA**

110 In the present crystal structure of BepA, the active site motif of the protease domain, H^{136}EISH,
111 has one zinc atom that is coordinated by two histidine residues (H136 and H140) of the α4 helix, one
112 glutamate residue (E201) residue of the α7 helix, and one histidine residue (H246) of the α9-α10 loop (Fig.
113 1B). The sequence alignment of full-length BepA with a Zn-dependent peptidase Q74D82 (PDB ID 3C37)
114 indicates that the sequence identity is about 28% (Fig. S1), while the superimposition of BepA and
115 Zn-dependent peptidase Q74D82 structures show 1.77 Å of RMSD for the Cα atoms (Fig. S4). Therefore, the
116 overall architectures of the protease domain of BepA and Q74D82 are similar. The residues coordinating the
117 zinc atom of the active site in Q74D82, whose positioning is also essentially similar to that of BepA, are two
118 histidine residues of H^{106}EINH, E162, and H208 (Fig. S4, Fig. S1). The molecular mechanisms of proteolysis
119 by BepA and Q74D82 are expected to be similar. Although the loop between the α6 and α7 helices of BepA
120 is partially disordered (Fig. 1B red), the C-terminal region of the loop seems to cover the active site. In the
121 crystal structure of BepA, the activity of this protein would be low as the Zn atom of the active site is buried
122 and the surrounding residues sterically interfere with the interaction between the active site and substrates. It
123 is proposed that when BepA exhibits protease activity, the loop is dislocated in response to interactions with
124 substrates or an unidentified factor, which enables unfolded regions of substrates to approach the protease
125 active site. The dynamics of the loop covering the active site may adequately regulate the protease activity.

126

127 **TPR domain and central cavity of BepA**

128 Although the overall architecture of the TPR domain in the full-length structure of BepA is similar
129 to that of the previously reported isolated TPR domain [16], reflecting an RMSD (1.67 Å) of the
130 superimposition in Fig. S4, the TPR domain in the full-length structure is slightly extended compared with
131 that of the isolated TPR domain, presumably due to the interaction with the protease domain (Fig. S4B). In
132 the full-length structure, the negatively charged cavity of the TPR domain faces the protease domain (Fig.

133 1C), resulting in a characteristic cavity that is larger than that of the isolated TPR structure [16]. The
134 hydrophilic cavity, which is about 20.2 Å × 12.5 Å wide and 14.6 Å deep at its maximum, includes D69 in
135 α1-α2 loop of the protease domain, E321 in H1 helix, D347 and D351 in H3 helix, and D408 and D409 in
136 H6-H7 loop (Fig. 1C). In addition, a negatively-charged ditch protruding from the cavity leads to the area
137 near the active site. Because TPR domains generally provide binding surfaces [19, 20], the negativity charged
138 regions of BepA may provide interaction sites for positively charged parts of substrates.

139

140 **The crystal structure of BepA reflects its active form in solution**

141 We performed small-angle X-ray scattering (SAXS) analysis of BepA to verify whether its
142 conformation in an aqueous solution is the same as that of its crystal structure. The experimental scattering
143 curve and the three-dimensional shape model of BepA based on SAXS data were adequately fitted with the
144 BepA crystal structure (Fig. 2A, Fig. S5). However, we cannot exclude the possibility that BepA function
145 involves alteration of the interaction between the protease and TPR domains; the dissociation of its domains
146 may be necessary for its function. To confirm the conformation of BepA *in vivo*, we performed disulfide bond
147 crosslinking analysis using BepA mutants. We selected five pairs of N-terminal amino and C-terminal
148 residues that are proximal within 13.3 Å (Fig. 2B) and introduced single and double cysteines into the
149 indicated positions in Fig. 2B. The formation of intramolecular disulfide bonds between the cysteine residues
150 restricts the flexibility of the N-terminal and C-terminal domains: BepA mutants possessing disulfide bonds
151 can stably exist *in vivo* with a conformation similar to the crystal structure. First, we detected the
152 accumulation and migration of BepA mutants over-expressed in the Δ*bepA* strain (Fig. 2C). The
153 double-cysteine mutants Q51C-F481C, Y55C-F481C, R58C-G441C, and R61C-A440C show highly migrated
154 bands ($BepA^{OX}$) on non-reducing SDS-PAGE. These bands disappeared under reducing conditions and the
155 single-cysteine mutants did not exhibit such highly migrated bands, indicating that the double-cysteine
156 mutants formed intramolecular disulfide bonds *in vivo*. The higher mobility of the BepA mutants possessing
157 an intramolecular disulfide bond would be ascribed to the more compact conformations of these proteins than
158 that of wild type BepA. In addition, any upshifted bands due to disulfide bond formation were not observed

159 (Fig. S6), suggesting that no intermolecular crosslink product between the BepA mutants was formed. In this
160 experiment it is reasonable that two pairs located more than 9 Å-distant in the crystal structure did not form a
161 disulfide bond. Further, we examined the ability of the BepA mutants to enhance LptD-maturation. Although
162 the absence of BepA (vector) caused the accumulation of $LptD^{C}$, a folding intermediate, whose
163 intramolecular disulfide bonds (C31-C173 and C724-C725) are different from those of mature $LptD^{NC}$
164 (C31-C724 and C173-C725) [15, 16, 21], all of the BepA mutants retained the LptD-maturation enhancing
165 activity (Fig. 2C). Next, we examined the properties of erythromycin (EM) sensitivity of Δ*bepA* cells
166 expressing the BepA mutants. The cells without BepA showed increased sensitivity to EM (Fig. 2C vector),
167 whereas all of the cells expressing BepA mutants exhibited EM resistance, similar to that of wild-type BepA
168 (Fig. 2C). Although most of the Q51C-F481C and R58C-G441C mutant BepA molecules spontaneously
169 formed an intramolecular disulfide bond, some amount of the mutant proteins remains reduced. Therefore, it
170 cannot be ruled out that these reduced molecules are responsible for the above observed activities. To make
171 sure that the BepA with a disulfide bond indeed possesses the activities comparable to normal (reduced)
172 BepA, we lowered the accumulation of BepA proteins by adjusting the glucose concentration of the medium
173 to repress expression from the *lac* promoter and examined their activities (Fig 2D). In the presence of 0.1%
174 glucose, the levels of the plasmid-expressed BepA were as low as that of the chromosomally-encoded BepA
175 (chr), but the mutant proteins were able to decrease $LptD^{C}$ accumulation and increase EM resistance of the
176 cells (Fig. 2D, 0.1% glucose). Under this condition, a small amount of the reduced form of the mutant BepA
177 proteins still accumulated. However, their levels were significantly lower than that of WT BepA in the
178 presence of 0.4% glucose (Fig. 2D WT, 0.4% glucose vs mutants, 0.1% glucose) that were not sufficient to
179 exhibit the normal BepA activities. This result suggested that the disulfide-bonded BepA mutants retain their
180 chaperone-like activities. Finally, we confirmed the protease activity of BepA (Fig. 2E), showing that all
181 BepA mutants generated the degradation products of BamA that is probably misassembled in Δ*surA* cells. In
182 addition, a comparison of the R58C/G441D mutant with chromosome-encoded BepA (chr), which showed
183 little BamA degradation, indicates that at least the disulfide-bonded R58C/G441C mutant is active in
184 substrate degradation (Fig. 2E). Taken together, BepA mutants were found to retain the chaperone-like and

185 protease activities regardless of the presence or absence of disulfide bonds between the N-terminal and
186 C-terminal regions; that is, the crystal structure of BepA reflects its active form in solution. It is suggested
187 that BepA can function without the dissociation between the independent protease and TPR domains *in vivo*,
188 and the crystal structure represents both its functional and resting states.
189
190 **Putative functional model of BepA**
191 In this paper, the full-length structure of BepA was determined at 2.6-Å resolution by X-ray
192 crystallography and modeled by SAXS in aqueous solution. Our functional analyses revealed that BepA
193 showed activity *in vivo* without large conformational transitions involving the dissociation of the protease and
194 TPR domains. Daimon et al. previously proposed a three-dimensional model of BepA and the BAM complex
195 based on cross-linking results and the crystal structure of the TPR domain [16] (Fig. S1). Indeed, this
196 previously described model lacks the protease domain. Here, we superimposed the structure of full-length
197 BepA onto the previous model [16] based on positions of each TPR domain. The orientation of BepA was
198 slightly modified to avoid clashes between the BAM complex and BepA (Fig. 3A). The positions of the TPR
199 domains in the previous and new models are similar. Therefore, the interaction sites between the BAM
200 complex and BepA, which were identified by photocrosslinking experiments, are reasonably positioned in the
201 new model as well. The protease domain protrudes to the periplasm side from the outer membrane. Here, we
202 discuss the molecular mechanism of BepA in combination with the BAM complex based on the new model.
203 It has been proposed that substrate proteins interact with BamD, the POTRA domain of BamA, and BepA [16,
204 22-24]. This model shows a space between them, which may temporarily capture a proportion of the unfolded
205 substrates. The flexible POTRA domain [25] may undergo large structural changes during integration of the
206 outer membrane proteins into the outer membrane; accordingly, the site may be larger, and interact with
207 larger proteins than the space in the model. The negatively charged cavity of BepA (Fig. 1C) may interact
208 with the positively charged parts of substrates; however, this remains to be confirmed. The previously
209 reported interaction sites of BepA with LptD are located at the edge of the cavity [16], which is consistent
210 with the idea that the substrate proteins interact with the cavity. The negatively charged cavity and protruded

211 ditch (Fig. 1C) might interact with the histidine, arginine, and lysine residues of the substrate. The unfolded
212 substrate protein may integrate into the membrane from the lateral gate of BamA in a stepwise fashion after
213 interacting with BepA and the POTRA domain of BamA in the budding model, or the integration of the
214 temporarily folded substrate protein into the membrane may be mediated by BepA and the BAM complex in
215 the assisted model [11, 12]. The chaperone activity of BepA may be important during these reactions. If the
216 transportation of a substrate stops midway, the substrate may be stalled for a long time in the area indicated
217 by the asterisks shown in Fig. 3A. The short helices α5, α6, and α9, which includes H246, and loops near the
218 active site of BepA may be flexible (Fig. 1A); it is conceivable that they are sometimes dislocated
219 stochastically or in response to some unidentified signals. The prolonged stalling of the substrate at this
220 position may result in a higher probability of its interaction with the active site resulting from the dislocation
221 of the loops and short helices. Therefore, the long-stalled substrate protein at the same position may be
222 cleaved by the protease domain of BepA. The loop structure on the active site of BepA may play an important
223 role in prolonging the duration of transport of the substrate to the active site; this stalling may result in the
224 substrate being cleaved (Fig. 3B). This notion is consistent with the idea that BepA digests stalled proteins
225 [17]. To elucidate the molecular mechanism underlying BepA activity, it is necessary to perform detailed
226 analysis of the interaction at the amino-acid residue level of BepA and the substrate protein, and to obtain
227 structural information for the full complex of BepA and the BAM complex.
228

229 **ACKNOWLEDGMENTS**

230 We thank K. Abe and S. Suzuki for secretarial assistance, K. Kobayashi and K. Yoshikaie for technical
231 support, T. Lithgow, T. Shiota and S. Matsuyama for antisera, and the beamline scientists at BL32XU of
232 SPring-8 (Hyogo, Japan) for helping with data collection. The synchrotron radiation experiments were
233 performed at BL32XU of SPring-8 with the approval of the Japan Synchrotron Radiation Research Institute
234 (JASRI) (Proposal Nos. 2017A2557, 2017B2557). The SAXS experiments were performed under the
235 approval of the Photon Factory Program Advisory Committee (Proposal No. 2016G077). This work was
236 supported by the JSPS/MEXT KAKENHI (Grant Nos. JP26119007, JP18H02405, JP17H05669, JP15H04350,
237 JP18H02404, JP18KK0197), Mitsubishi Foundation, Noguchi Institute, Naito Foundation, Mochida
238 Memorial Foundation for Medical and Pharmaceutical Research, Joint Usage/Research Center program of
239 Institute for Frontier Life and Medical Sciences Kyoto University.

240

241 **AUTHOR CONTRIBUTIONS**

242 Conceptualization, Y.D., S.Narita, Y.A., T.T; Methodology and Investigation M.M., Y.D., Y.T., Y.H.,
243 S.Nakayama, S.I., H.K.; Writing, M.M., Y.A, T.T.; Supervision Y.A., T.T.

244

245 **ACCESSION NUMBER**

246 Coordinates and structure factors have been deposited in the Protein Data Bank under accession number
247 6AIT.

248

249 **Table 1 | Data Collection and Refinement**

250

251 **Figure 1 | Crystal structure of BepA**

252 A, Schematic representation and amino acid sequence of BepA. α helices (α1-12 and H1-10) and β strands
253 (β1-3) are shown. The protease active site motif, HEXXH, and Zn-coordinated residues are highlighted by
254 red boxes. Residues contributing to the negative cavity are labeled by red ellipses. B, Overall structure and
255 close-up views. The protease domain is indicated in various colors, from blue to green. The connecting helix
256 and TPR domains are colored in gray and orange, respectively. The side chains interacting with other
257 domains are represented as a stick model. The zinc atom in the active center is coordinated by side chains of
258 H136, H140, H246, and E201 (upper left). The C-terminal region of the TPR domain interacts with α1 of the
259 protease domain via hydrogen bonds (lower left, dashed yellow line). C, Surface representation. The surfaces
260 are colored as in B (left). The other surfaces are colored to indicate electrostatic potential ranging from blue
261 (+10 kT/e) to red (-10 kT/e). The orientations are the same as B (left and middle). The acidic residues in the
262 negatively charged cavity are represented as a stick model.

263

264 **Figure 2 | Crystal structure of BepA reflects its active form in solution**

265 A. SAXS analysis of BepA in solution. The crystal structure was superimposed on the shape reconstruction of
266 BepA based on SAXS data. B, Close-up view of the N-terminal and C-terminal regions. The positions that
267 were substituted by Cys were mapped on the crystal structure. C and D, Intramolecular disulfide-bonded
268 BepA mutants retain the chaperone-like activity of BepA. Accumulation of LptD and BepA (middle and
269 bottom) and erythromycin (EM) sensitivity (upper) of the Δ*bepA* cells expressing the individual BepA double
270 cysteine mutants were determined. Wild–type cells carrying pUC18 (chr) or Δ*bepA* cells carrying pUC18
271 (vector), pUC-bepA (WT), or a derivative of pUC-bepA having the single or double cysteine mutations at the
272 indicated positions in B were grown at 30°C in L medium. The medium was supplemented with 0.1% or
273 0.4% glucose in D. Total cellular proteins were analyzed by SDS-PAGE under a reducing (+ME) or a
274 non-reducing (−ME) condition and immunoblotting. The minimum inhibitory concentration (MIC) of EM for

275 each of the cells expressing BepA or its derivative were examined as described in Methods. E, Degradation of
276 BamA by double cysteine BepA mutants. The Δ*surA* cells carrying pUC18 (chr) or the Δ*surA* Δ*bepA* cells
277 carrying pUC18 (vector), pUC-bepA (WT), or a derivative of pUC-bepA harboring the double cysteine
278 mutation were grown at 30°C in M9 medium supplemented with 1 mM IPTG. Total cellular proteins were
279 analyzed as above. $LptD^{NC}$, $LptD^{C}$, and $LptD^{RED}$ indicate LptD with the non-consecutive disulfide bonds (the
280 mature form), LptD with the consecutive disulfide bonds (a folding intermediate), and reduced LptD,
281 respectively. $BepA^{RED}$ and $BepA^{OX}$ indicate reduced BepA and oxidized BepA, respectively. The asterisks
282 indicate non-specific bands serving as a loading control.
283
284 **Figure 3** | Working model of BepA
285 A, Docking model of the BAM complex and BepA. LptD-interaction sites were mapped onto the BepA
286 structure [16]. The asterisks represent a putative substrate-interaction area described in the text. B,
287 Substrate-cleavage model of BepA. The TPR domain might provide the negatively charged interaction site
288 for the substrate. The active site of the protease domain is covered by the flexible loop (red). When substrates
289 interact with BepA for a prolonged duration, the frequency of the contact with the protease active site would
290 increase with time, resulting in the eventual cleavage of the substrate by BepA, which contributes to the
291 quality control of biogenesis for outer membrane proteins.

292 **Methods**

293 **Expression and purification of BepA**

294 The pYD296 plasmid, which was modified from pET-16b-TEV [16], encodes MG-H_{10}-SSGENLYFQG-*E.*
295 *coli* $BepA_{45\text{-}482}$. *E. coli* strain KRX strain cells (Promega) were transformed with pYD296 and stored at -80°C
296 as a glycerol stock until use. The pre-cultured *E. coli* cells in 25 mL of LB Broth, Lennox (Nacalai)
297 supplemented with 50 µg/mL ampicillin and 0.4% glucose at 37°C for 12 h were inoculated into 2.5 L of LB
298 Broth, Lennox supplemented with ampicillin, and cultured at 37°C until OD_{600} ~ 0.6. Then, protein
299 expression was induced by 0.2% rhamnose. The culture was transferred into a 17°C incubator and shaken for
300 12 h. The cells were harvested by centrifugation at 6,000 × *g* for 10 min (Hitachi CR22N) and suspended in
301 10 mM Tris-HCl (pH 7.0). The washed cells were retrieved by centrifugation at 6,000 × *g* for 10 min as a
302 pellet and resuspended in sonication buffer (20 mM Tris-HCl (pH 7.0), 500 mM NaCl, 20 mM imidazole-HCl
303 (pH 7.0), 1 mM 2-mercaptoethanol (ME) and 0.1 mM phenylmethanesulfonyl fluoride (PMSF)), together
304 with 0.5 mM EDTA (pH 8.05). Cells were disrupted by sonication using Q500 (QSONICA) (20% power,
305 pulse rate: 1s on and 1s off, rod: CL 334 for 30 min and centrifuged (20,400 × *g*, 30 min) at 4°C. The
306 supernatant was mixed with 3 mL of Ni Sepharose Excel (GE Healthcare) equilibrated with 20 mM Tris-HCl
307 (pH 7.0), 500 mM NaCl, 20 mM imidazole-HCl (pH 7.0), 1 mM 2-ME and 0.1 mM PMSF and rotated for 1 h
308 at 4°C. The resin was washed with 50 mL of wash buffer (20 mM Tris-HCl (pH 7.0), 500 mM NaCl, 20 mM
309 imidazole-HCl (pH 7.0), 1 mM 2-ME and 0.1 mM PMSF) and eluted with 40 mL of 20 mM Tris-HCl (pH
310 7.0), 500 mM NaCl, 300 mM imidazole-HCl (pH 7.0), 1 mM 2-ME and 0.1 mM PMSF with different batches.
311 TEV protease was mixed with the eluted proteins, with a protein weight ratio of 1:1 (1 mg of TEV protease to
312 1 mg of protein) and dialyzed against 20 mM Tris-HCl (pH 7.0), 500 mM NaCl, 1 mM 2-ME, and 0.1 mM
313 PMSF at 4°C for 12 h using a dialysis membrane (Spectra/Por 7 MWCO 10 kD (SPECTRUM)). The cleaved
314 protein was then incubated with 3 mL of Ni Sepharose Excel (GE Healthcare) by rotation before being
315 washed from the polyhistidine-tag with 20 mM Tris-HCl (pH 7.0), 500 mM NaCl, 20 mM imidazole-HCl (pH
316 7.0), 1 mM 2-ME and 0.1 mM PMSF. The cleaved protein was then concentrated using an Amicon Ultra 50K
317 filter (Millipore) and applied to a Superdex 200 Increase 10/300 column (GE Healthcare) equilibrated with 20

318 mM Tris-HCl (pH 7.0), 500 mM NaCl, 1 mM 2-ME, and 0.1 mM PMSF. For crystallization, the purified
319 sample was concentrated to 17 mg/mL.
320

321 **Crystallization of BepA**

322 A volume of 0.1 μL of 17 mg/mL purified BepA was incubated at 4°C by the sitting drop vapor diffusion
323 method against a reservoir solution of 2 M ammonium sulfate in 96-well Violamo crystallization plates (plate
324 model number: VCP-1). The crystallization plate was set up using Crystal Gryphon (Art Robbins
325 Instruments). Microcrystals appeared after a 24-h incubation at 4°C; then, the plate was incubated at 37°C
326 overnight to dissolve the microcrystals, followed by incubation at 4°C for two days. Single rod-shaped
327 crystals grew to a size of approximately 90 μm × 15 μm in five days. The crystals were collected using
328 crystal harvesting mounts and loops (MiTeGen) and directly cryo-cooled in liquid nitrogen before performing
329 X-ray diffraction experiments.
330

331 **Data collection and determination of structure**

332 The X-ray diffraction dataset of the BepA crystal was collected by the Helical Data Collection Method on
333 beamline BL32XU at SPring-8 using a microbeam [26]. The collected diffraction images were processed with
334 XDS [27]. Initial phase was calculated by molecular replacement with the previously determined BepA
335 tetratricopeptide repeat (TPR) domain (PDB ID 5XI8)[16] and a Zn-dependent peptidase Q74D82 (PDB ID
336 3C37) as templates using Phaser [28]. We found independent six BepA molecule in the asymmetric unit by
337 connecting the peptidase domain and the TPR domain based on the SAXS model. The structural model of
338 BepA was stepwise-refined using COOT [29] and PHENIX [30] to R_{work}/R_{free} = 0.206/0.263 with space group
339 *P*1 at 2.6 Å resolution. A Ramachandran plot was constructed using Molprobity [31], and molecular graphics
340 were generated using CueMol2 (http://www.cuemol.org/).
341

342 **SAXS analysis**

343 The SAXS measurements were carried out at the beamline BL-10C, Photon Factory (Tsukuba, Japan) [32].

344 Scattering images were collected from buffer (containing 20 mM Tris-HCl, 500 mM NaCl, 1 mM DTT, 0.1
345 mM PMSF, pH 7) and BepA solutions at 5 different concentrations (3.3, 6.5, 9.8, 12.7, and 15.9 mg/mL) at
346 20°C. Ovalbumin (45 kDa, Sigma-Aldrich) was also measured to estimate the apparent molecular weight of
347 BepA. The 60 scattering images were collected using a PILATUS3 2M detector (Dectris), where the exposure
348 time for each image is 15 sec. The 2D images were converted into 1D profiles using SAngler software [33].
349 The scattering profile averaging the first 10 frames among the 60 frames, where no radiation damage was
350 observed at every concentration, were subject to the following analysis. The Guinier plots of BepA (Fig.
351 S5A) did not show a steeper curve at the low Q-region, suggesting that no obvious aggregation occurred. The
352 Rg^2 and the forward scattering intensity normalized by the weight per volume (mg/ml), I(0)/conc. were
353 extrapolated to obtain values at zero protein concentration (Fig. S5B). The scattering profile at zero protein
354 concentration was also calculated from extrapolation of the 5 scattering profiles at the different
355 concentrations using the SVD implemented to Igor software (WaveMetrics). The scattering profiles at the
356 higher Q-region (> 0.12 $Å^{-1}$) did not exhibit a clear dependence on concentration. The profile at the highest
357 concentration at the high Q-region was merged into the profile at zero protein concentration. The maximum
358 dimension was calculated using GNOM software [34]. The parameters from the above analyses were
359 summarized in the Table S2. *Ab initio* shape modelling was performed using DAMMINIF to obtain 20
360 models without structural restrictions such as point symmetry and particle anisometry [35]. Using the average
361 model of the 20 models as a starting model, we finally refined the shape model using DAMMIN [36]. The
362 refinement procedures were independently performed three times to confirm reproducibility. Fig. 2A shows
363 the representative model, into which the crystal structure of BepA is superposed using Situs software [37].
364

365 **Functional analyses of BepA**

366 The preparation of media was as described previously [16]. *E. coli* K12 strains and plasmids used in this
367 study are listed in Supporting Information Table S2 [15, 38, 39]. Derivatives of pUC-bepA encoding a mutant
368 form of BepA were constructed by site-directed mutagenesis. For determination of minimum inhibitory
369 concentration (MIC) of erythromycin, overnight cultures were diluted 10^3-fold with L-medium, and 5 μL of

370 these cultures were inoculated on L medium-based agar plates supplemented without or with 0.1% or 0.4%
371 glucose and containing 6.25, 12.5, 25, 50, and 100 μg/mL erythromycin. The plates were incubated for 18-20
372 h at 30°C.
373

374 **Table 1 | Data Collection and Refinement**

	BepA
Wavelength (Ǎ)	1.00
Resolution range	48.18 - 2.598 (2.691 - 2.598)
Space group	*P*1
Unit cell a, b, c (Ǎ) α, β, γ (˚)	85.844 , 104.674 , 104.971 113.606 105.843 104.026
Total reflections	327,914 (31,495)
Unique reflections	89,522 (8,738)
Multiplicity	3.7 (3.6)
Completeness (%)	98.24 (94.90)
Mean *I*/sigma (*I*)	8.88 (0.86)
Wilson B-factor	62.50
R-merge	0.1039 (1.218)
R-meas	0.1219 (1.434)
R-pim	0.06328 (0.7477)
$CC_{1/2}$	0.997 (0.402)
CC*	0.999 (0.757)
Reflections used in refinement	89,487 (8,729)
Reflections used for R-free	2,016 (195)
R-work	0.2064 (0.3456)
R-free	0.2634 (0.3687)
CC (work)	0.964 (0.646)
CC (free)	0.944 (0.511)
Number of non-hydrogen atoms	19,732
macromolecules	19,568
ligands	54
solvent	110
Protein residues	2,485
RMS (bonds)	0.004
RMS (angles)	0.62
Ramachandran favored (%)	97.87
Ramachandran allowed (%)	2.09
Ramachandran outliers (%)	0.04
Rotamer outliers (%)	0.15
Clashscore	9.18
Average B-factor	81.27
ligands	86.44
solvent	56.41
Number of TLS groups	35

375 Statistics for the highest-resolution shell are shown in parentheses.
376

377

378

379 [1] H. Nikaido, Molecular basis of bacterial outer membrane permeability revisited, Microbiol. Mol. Biol. Rev. 67
380 (2003) 593-656.
381 [2] G. Rowley, M. Spector, J. Kormanec, M. Roberts, Pushing the envelope: extracytoplasmic stress responses in
382 bacterial pathogens, Nat. Rev. Microbiol. 4 (2006) 383-94.
383 [3] N.P. Walsh, B.M. Alba, B. Bose, C.A. Gross, R.T. Sauer, OMP peptide signals initiate the envelope-stress response
384 by activating DegS protease via relief of inhibition mediated by its PDZ domain, Cell 113 (2003) 61-71.
385 [4] S. Lima, M.S. Guo, R. Chaba, C.A. Gross, R.T. Sauer, Dual molecular signals mediate the bacterial response to
386 outer-membrane stress, Science 340 (2013) 837-41.
387 [5] S. Bury-Mone, Y. Nomane, N. Reymond, R. Barbet, E. Jacquet, S. Imbeaud, A. Jacq, P. Bouloc, Global analysis of
388 extracytoplasmic stress signaling in Escherichia coli, PLoS Genet. 5 (2009) e1000651.
389 [6] M.G. Iadanza, A.J. Higgins, B. Schiffrin, A.N. Calabrese, D.J. Brockwell, A.E. Ashcroft, S.E. Radford, N.A. Ranson,
390 Lateral opening in the intact beta-barrel assembly machinery captured by cryo-EM, Nat. Commun. 7 (2016) 12865.
391 [7] Y. Gu, H. Li, H. Dong, Y. Zeng, Z. Zhang, N.G. Paterson, P.J. Stansfeld, Z. Wang, Y. Zhang, W. Wang, C. Dong,
392 Structural basis of outer membrane protein insertion by the BAM complex, Nature 531 (2016) 64-9.
393 [8] L. Han, J. Zheng, Y. Wang, X. Yang, Y. Liu, C. Sun, B. Cao, H. Zhou, D. Ni, J. Lou, Y. Zhao, Y. Huang, Structure of
394 the BAM complex and its implications for biogenesis of outer-membrane proteins, Nat. Struct. Mol. Biol. 23 (2016)
395 192-6.
396 [9] A. Tsirigotaki, J. De Geyter, N. Sostaric, A. Economou, S. Karamanou, Protein export through the bacterial Sec
397 pathway, Nat. Rev. Microbiol. 15 (2017) 21-36.
398 [10] T. Tsukazaki, Structure-based working model of SecDF, a proton-driven bacterial protein translocation factor,
399 FEMS Microbiol. Lett. 365 (2018).
400 [11] A. Konovalova, D.E. Kahne, T.J. Silhavy, Outer Membrane Biogenesis, Annu. Rev. Microbiol. 71 (2017) 539-556.
401 [12] N. Noinaj, J.C. Gumbart, S.K. Buchanan, The beta-barrel assembly machinery in motion, Nat. Rev. Microbiol. 15
402 (2017) 197-204.
403 [13] J. De Geyter, A. Tsirigotaki, G. Orfanoudaki, V. Zorzini, A. Economou, S. Karamanou, Protein folding in the cell
404 envelope of Escherichia coli, Nat. Microbiol. 1 (2016) 16107.
405 [14] N. Noinaj, A.J. Kuszak, J.C. Gumbart, P. Lukacik, H. Chang, N.C. Easley, T. Lithgow, S.K. Buchanan, Structural
406 insight into the biogenesis of beta-barrel membrane proteins, Nature 501 (2013) 385-90.
407 [15] S. Narita, C. Masui, T. Suzuki, N. Dohmae, Y. Akiyama, Protease homolog BepA (YfgC) promotes assembly and
408 degradation of beta-barrel membrane proteins in Escherichia coli, Proc. Natl. Acad. Sci. USA 110 (2013) E3612-21.
409 [16] Y. Daimon, C. Iwama-Masui, Y. Tanaka, T. Shiota, T. Suzuki, R. Miyazaki, H. Sakurada, T. Lithgow, N. Dohmae, H.
410 Mori, T. Tsukazaki, S.I. Narita, Y. Akiyama, The TPR domain of BepA is required for productive interaction with
411 substrate proteins and the beta-barrel assembly machinery complex, Mol. Microbiol. 106 (2017) 760-776.
412 [17] G.R. Soltes, N.R. Martin, E. Park, H.A. Sutterlin, T.J. Silhavy, Distinctive Roles for Periplasmic Proteases in the
413 Maintenance of Essential Outer Membrane Protein Assembly, J. Bacteriol. 199 (2017).
414 [18] N.D. Rawlings, A.J. Barrett, P.D. Thomas, X. Huang, A. Bateman, R.D. Finn, The MEROPS database of proteolytic
415 enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database, Nucleic
416 Acids Res. 46 (2018) D624-D632.
417 [19] R.K. Allan, T. Ratajczak, Versatile TPR domains accommodate different modes of target protein recognition and
418 function, Cell Stress Chaperones 16 (2011) 353-67.
419 [20] N. Zeytuni, R. Zarivach, Structural and functional discussion of the tetra-trico-peptide repeat, a protein interaction
420 module, Structure 20 (2012) 397-405.
421 [21] S.S. Chng, M. Xue, R.A. Garner, H. Kadokura, D. Boyd, J. Beckwith, D. Kahne, Disulfide rearrangement triggered
422 by translocon assembly controls lipopolysaccharide export, Science 337 (2012) 1665-8.
423 [22] S. Kim, J.C. Malinverni, P. Sliz, T.J. Silhavy, S.C. Harrison, D. Kahne, Structure and function of an essential
424 component of the outer membrane protein assembly machine, Science 317 (2007) 961-4.
425 [23] P.Z. Gatzeva-Topalova, T.A. Walton, M.C. Sousa, Crystal structure of YaeT: conformational flexibility and substrate
426 recognition, Structure 16 (2008) 1873-81.
427 [24] C.L. Hagan, J.S. Wzorek, D. Kahne, Inhibition of the beta-barrel assembly machine by a peptide that binds BamD,
428 Proc. Natl. Acad. Sci. USA 112 (2015) 2011-6.
429 [25] P.J. Fleming, D.S. Patel, E.L. Wu, Y. Qi, M.S. Yeom, M.C. Sousa, K.G. Fleming, W. Im, BamA POTRA Domain
430 Interacts with a Native Lipid Membrane Surface, Biophys. J. 110 (2016) 2698-2709.
431 [26] K. Hirata, Y. Kawano, G. Ueno, K. Hashimoto, H. Murakami, K. Hasegawa, T. Hikima, T. Kumasaka, M.

432 Yamamoto, Achievement of protein micro-crystallography at SPring-8 beamline BL32XU, J Phys Conf Ser 425 (2013).
433 [27] W. Kabsch, Xds, Acta Crystallogr. D Biol. Crystallogr. 66 (2010) 125-32.
434 [28] A.J. McCoy, R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni, R.J. Read, Phaser crystallographic
435 software, J. Appl. Crystallogr. 40 (2007) 658-674.
436 [29] P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics, Acta Crystallogr. D Biol. Crystallogr. 60
437 (2004) 2126-32.
438 [30] P.D. Adams, R.W. Grosse-Kunstleve, L.W. Hung, T.R. Ioerger, A.J. McCoy, N.W. Moriarty, R.J. Read, J.C.
439 Sacchettini, N.K. Sauter, T.C. Terwilliger, PHENIX: building new software for automated crystallographic structure
440 determination, Acta Crystallogr. D Biol. Crystallogr. 58 (2002) 1948-54.
441 [31] V.B. Chen, W.B. Arendall, 3rd, J.J. Headd, D.A. Keedy, R.M. Immormino, G.J. Kapral, L.W. Murray, J.S.
442 Richardson, D.C. Richardson, MolProbity: all-atom structure validation for macromolecular crystallography, Acta
443 Crystallogr. D Biol. Crystallogr. 66 (2010) 12-21.
444 [32] N. Igarashi, Y. Watanabe, Y. Shinohara, Y. Inoko, G. Matsuba, H. Okuda, T. Mori, K. Ito, Upgrade of the small
445 angle X-ray scattering beamlines at the Photon Factory, J. Phys.: Conf. Ser. 272 (2011) 012026.
446 [33] N. Shimizu, K. Yatabe, Y. Nagatani, S. Saijo, T. Kosuge, N. Igarashi, Software development for analysis of
447 small-angle X-ray scattering data, AIP Conference Proceedings 1741 (2016) 050017.
448 [34] D.I. Svergun, Determination of the regularization parameter in indirect-transform methods using perceptual criteria,
449 J. Appl. Cryst. 25 (1992) 495-503.
450 [35] D. Franke, D.I. Svergun, DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering, J.
451 Appl. Crystallogr. 42 (2009) 342-346.
452 [36] D.I. Svergun, Restoring low resolution structure of biological macromolecules from solution scattering using
453 simulated annealing, Biophys. J. 76 (1999) 2879-86.
454 [37] W. Wriggers, Conventions and workflows for using Situs, Acta Crystallogr. D Biol. Crystallogr. 68 (2012) 344-51.
455 [38] A. Kihara, Y. Akiyama, K. Ito, A protease complex in the Escherichia coli plasma membrane: HflKC (HflA) forms a
456 complex with FtsH (HflB), regulating its proteolytic activity against SecY, EMBO J. 15 (1996) 6122-31.
457 [39] C. Yanisch-Perron, J. Vieira, J. Messing, Improved M13 phage cloning vectors and host strains: nucleotide
458 sequences of the M13mp18 and pUC19 vectors, Gene 33 (1985) 103-19.
459