Crystal structure of β-Barrel Assembly-Enhancing Protease (BepA) for biogenesis of outer membrane proteins

Bin Mohammad Umar Mohammad Shahrizal

Nara Institute of Science And Technology

Graduate School of Biological Sciences

Structural Life Science Laboratory

Professor Tomoya Tsukazaki

Submitted on 2021/03/11

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

BepA	β-Barrel Assembly-Enhancing Protease
BAM complex	β-barrel–assembly machinery complex
LPS	Lipopolysaccharide
Lpt	LPS-assembly protein
OM	Outer membrane
OMP	Outer membrane protein
βME	2-mercaptoethanol
EDTA	Ethylenediaminetetraacetic acid
PMSF	Phenylmethanesylfonyl fluoride
SAXS	Small-angle X-ray scattering
POTRA	Polypeptide transport associated domains
EM	Erythromycin
TPR	Tetratricopeptide repeat
Zn	Zinc
MIC	Minimum inhibitory concentration

ABSTRACT

Gram-negative bacteria can cause many types of infections such as foodborne diseases, cholera, pneumonia and gonorrhea. They have developed resistance towards antibiotics, for examples vancomycin and erythromycin. The inner and outer membranes in Gram-negative bacteria function as barriers to protect the bacteria from hazardous substances, such as chemicals and antibiotics. Lipopolysaccharide (LPS) is responsible for these functions, especially in the outer membrane. LPS is the major constituent of the outer membrane in bacteria which contributes to the structural integrity of bacteria and also protects the membrane from chemical attack. LPS is translocated by outer membrane proteins (OMPs). This means that if there is a lack of OMP, there could be a lack LPS in the outer membrane. For OMPs to be functional, proper folding and translocation of OMPs to the outer membrane are needed. Thus, the process of OMPs' translocation from the cytoplasm to the outer membrane is very crucial for bacteria, which is regulated by periplasmic chaperones. One of the chaperones is known as β-barrel assembly enhancing protease (BepA). When OMPs are misfolded and cannot carry out their functions, a stress response in the bacteria will activate BepA which degrades the misfolded OMPs, thus keeping the integrity of the outer membrane . The other function of BepA is to assemble OMPs which will then assemble LPS. Despite having important roles, the functions and molecular mechanisms of BepA are relatively unclear. Therefore, the full-length protein structure of BepA solved by X-ray crystallography might offer important insights to BepA's molecular mechanisms contributing to its functions.

In order to study the full-length structure, pYD296 plasmid encoding full-length *E.coli* BepA was transformed into *E. coli* KRX cells. The cells were cultured and protein expression was induced at 17°C with the addition of 0.2% rhamnose. The cells were harvested and the protease

activity of the protease domain was suppressed by the addition of EDTA. After cells disruption, Ni-NTA chromatography was performed to purify BepA protein. His-tag was removed by TEV protease while being dialyzed. In order to increase the purity of the protein, gel-filtration chromatography was performed. Purified BepA protein was pooled and concentrated to 17 mg/mL and crystallized by vapor-diffusion method. The obtained crystals were brought to SPring-8 in Hyogo prefecture to collect X-ray diffraction data. From the diffraction data, full-length crystal structure of BepA protein was constructed. Then, structural analysis was performed.

The full-length structure of BepA protein at 2.6-Å resolution reveals that there are 11 alpha helices (α1-α11) and 3 strands of β-sheets in the N-terminal protease domain, while there are 10 anti-parallel alpha helices (H1-H10) in the C-terminal TPR domain, with α12 helix connects the two domains together. The structure also reveals the presence of a negatively charged pocket and an active site coordinated by a zinc atom within the conserved HEXXH motif. Residues between α5-α6 and α6-α7 could not be modeled due to poor electron density, suggesting the region might be flexible. α9 helix is also assumed to cover the active site due to its short plug-like length and its convenient position near the zinc atom. The pocket and the active site of BepA probably serve as a substrate-binding site. Since the pocket is negatively charged, substrates of BepA might be positively charged. Kratky plot generated from previous small-angle X-ray scattering (SAXS) experiment showed that histidine tag might be stabilizing the structure of BepA by inducing oligomer formation of BepA. Other than that, important residues coordinating the pocket and active site were also directed for mutational analyses for functional studies. Given the structure and collaborative functional analyses data of BepA, a working model of BepA's functional mechanism was proposed. This knowledge could be very crucial in drugs development against antibiotics-resistant bacteria.

CHAPTER 1 INTRODUCTION

1.1 Cell membranes.

Cell membranes play important roles in Gram-negative bacteria[1]. A cell envelope is made up of an inner and an outer membrane (OM). About 150 Å of space between the inner and outer membranes is called as periplasmic space and holds a layer of peptidoglycan[2]. These structures are protecting Gram-negative bacteria from stresses and toxins from outer environment[3]. Both lipid bilayers of the inner membrane are made up of phospholipids, while the outer membrane is a unique structure consisting of lipopolysaccharides (LPS) which is stabilized by the presence of divalent ions like calcium and magnesium[4]. The OM of Gram-negative bacteria serves as a barrier that prevents many antibiotics from reaching their intracellular targets[5]. How this barrier protects the cells from antibiotics is yet to be elucidated.

1.2 Lipopolysaccharides.

LPS is a molecule which is composed of three components: core Lipid A and a variable O-antigen carbohydrate chain[6]. Lipid A consists of the hydrophobic acyl chains that tether the molecule to the OM, whereas the core contains up to 200-sugar-length oligosaccharide[6]. The O-antigen are hydrophilic and protrude from the cell surface[6]. LPS of bacteria are the major components of outer surface membrane in most of the Gram-negative bacteria[7]. LPS plays important roles not only in bacteria but also in eukaryotes from insects to humans by stimulating innate and natural immunity in eukaryotes[7]. Lipid A is an endotoxin that can give rise to fever and serious illness in humans and animals[8].

There are several antibiotics used for treatments of diseases caused by Gram-negative bacteria. Most notably are hydrophobic antibiotics that are able to enter into the cytoplasm by penetrating through the outer membrane bilayer[9]. These are aminoglycosides, for example gentamycin and kanamycin, macrolides such as erythromycin, novobiocin, rifamycins, fusidic acid and cationic peptides[9]. Tetracylcine and quinolones penetrate the membrane by a lipid-mediated and porin-mediated pathways, respectively[9]. LPS provides resistance to antibiotics in Gram-negative bacteria[10]. The core region of LPS protects the cells from compounds, especially the hydrophobic antibiotics giving Gram-negative bacteria the resistance to antibiotics[10].

1.3 Translocation of lipopolysaccharides by outer membrane proteins.

To retain the structural integrity of Gram-negative bacteria, both membrane proteins and LPS must be frequently translocated from the cytoplasm and inserted into the outer membrane[11]. During the synthesis, Lipid A and O-antigen are synthesized independently in the cytoplasm[12]. Then, they are transported to the outer layer of the inner membrane and assembled together, forming a mature LPS[12]. Subsequently, LPS is released into the periplasmic space and transported to the outer membrane[12]. At the outer membrane, the LPS transport system LptABCDEFG will insert LPS into the outer membrane[13]. LPS is inserted into the outer membrane by the membrane protein LptDE complex which made up of an integral membrane protein LptD, and a globular lipoprotein known as LptE[14].

1.4 The BAM complex.

The BAM complex is a membrane-located molecular machinery which coordinates the integration of outer membrane proteins (OMPs) into the outer membrane[15]. The most important component

of the BAM complex is a β-barrel protein known as BamA[16]. BamA serves as a scaffold for the other components of the BAM complex, that is BamB, BamC, BamD and BamE[16]. BamA has five soluble polypeptide transport associated domains (POTRA) serving as the core of the BAM complex machinery[16]. In each POTRA domain there is a β1–α1–α2–β2–β3 architecture[17]. The POTRA domains were proposed to interact with OMPs by β-strand augmentation mechanism[17].

1.5 Biogenesis of outer membrane proteins by the BAM complex.

In the biogenesis of OMPs, they must first travel through the inner membrane and then the periplasmic space[18]. Since the periplasmic space lacks ATP, premature OMPs are folded to their final and mature structure at the OM[18]. As the precursors of OMPs emerge through the Sec translocon located at the inner membrane, they are recognized and transferred to the BAM complex by periplasmic chaperones (**Fig. 1, Fig. 2**). The OMPs in unfolded conformations are translocated through the inner membrane and then integrated into the OM with the aid from the BAM complex [18]. Previous studies proposed that the opening of lateral gate of BamA is caused by the stepwise rotation of BAM complex's ring-like region at the periplasmic side [19]. This opening drives the gradual integration of OMPs into the outer membrane[19]. After correct folding and insertion assisted by the BAM complex and chaperones, these OMPs continue to involve in a number of pathogenic and drug resistance functions[19].

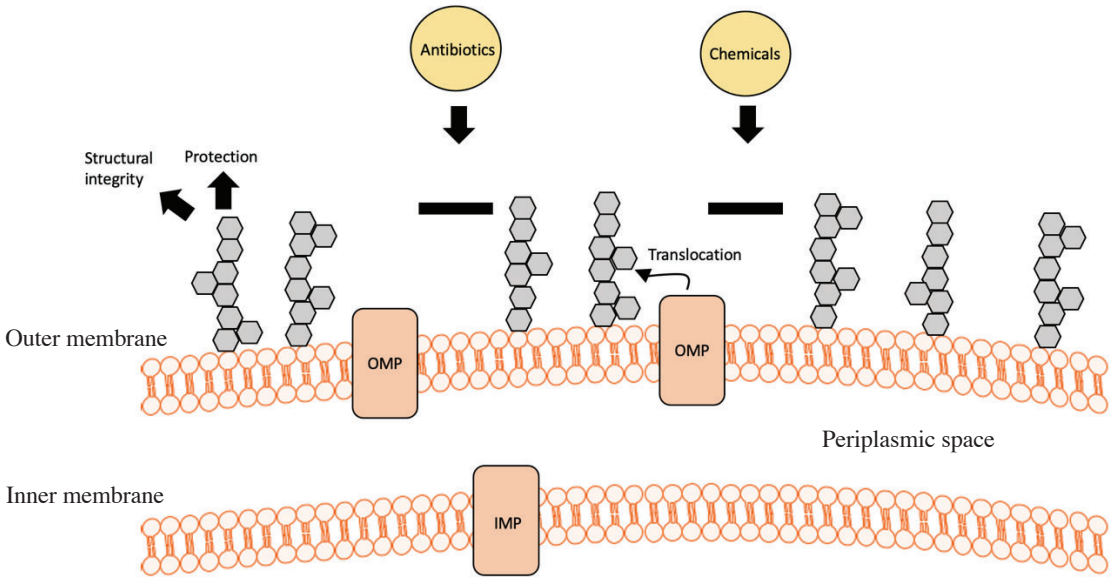

Figure 1. Translocation of an OMP.

BepA translocates OMP which then translocates LPS in bacteria. LPS then protects the membrane from antibiotics and chemicals attack which contributes to antibiotics resistance in gram-negative bacteria.

1.6 Roles of chaperones in translocation of outer membrane proteins.

As previously mentioned, the translocation of OMPs precursors to the BAM complex are assisted by periplasmic chaperones. The maintenance of the quality of OMPs' biogenesis does not only depend on chaperones, but also periplasmic proteases[20]. Chaperones assist the proteins to be correctly folded, while proteases degrade misfolded OMPs[20]. Accumulated misfolded OMPs in the periplasm can disrupt the integrity of the membranes[20]. Thus, both periplasmic chaperons and proteases are important.

In *Escherichia coli* (*E. coli*), the three most characterized periplasmic chaperones are Survival protein A (SurA), Seventeen kilodalton protein (Skp) and DegP. Skp and SurA were shown to have interactions with OMPs such as outer membrane protein A (OmpA) [21]. In addition to these

three major *E. coli* chaperones listed, there are other periplasmic chaperones that assist the translocation of OMPs to the OM. β-barrel assembly enhancing protease (BepA), formerly known as YfgC, has been shown to be involved in the OMP pathway[22] and has been recently heavily studied but remain poorly characterized.

1.7 Beta-barrel assembly enhancing protease (BepA) maintaining membrane integrity.

BepA is an *E. coli* periplasmic protein which was speculated to have important roles in the translocation of OMPs to the BAM complex and also in the integration of OMPs into the OM, thus maintaining the integrity of the OM (**Fig. 1**, **Fig.2**). BepA was shown to be upregulated in cells lacking SurA [22], suggesting that it may have a biological significance in maintaining the cellular membrane integrity during the absence of conventional chaperones.

BepA was also shown to be important for cells, especially for gram-negative bacteria. Cells lacking BepA protein (Δ*bepA*) exhibited lower erythromycin resistance compared to the wild-type protein [23]. Erythromycin, rifampicin, vancomycin, and novobiocin, are examples of antibiotics with a high molecular mass that are not easy to pass through an intact outer membrane [22]. However, when BepA could not perform its functions, LPS could not be inserted into the outer membrane due to the misfolding of Lpt proteins that assemble LPS. This causes the outer membrane to become more permeable, thus allowing the entries of the high molecular mass antibiotics that are fatal to Gram-negative bacteria [22]. Another study also showed Δ*bepA* cells had increased sensitivity to vancomycin [24]. Other than that, Δ*bepA* showed lower degradation of overexpressed membrane proteins that failed to form complex with their partners, most notably LptD, meanwhile cells with wild-type BepA showed higher LptD degradation when LptE was absent [25]. Δ*bepA* cells also lack in the ability to correctly fold LptD proteins [23]. These phenotypic attributes showed that

BepA plays a role in maintaining the integrity of membrane by mediating the correct folding of membrane proteins and their interaction with other membrane protein. The correct folding and integration to the membrane will thus keep the coherence of the cell membrane that provides protection for the cells.

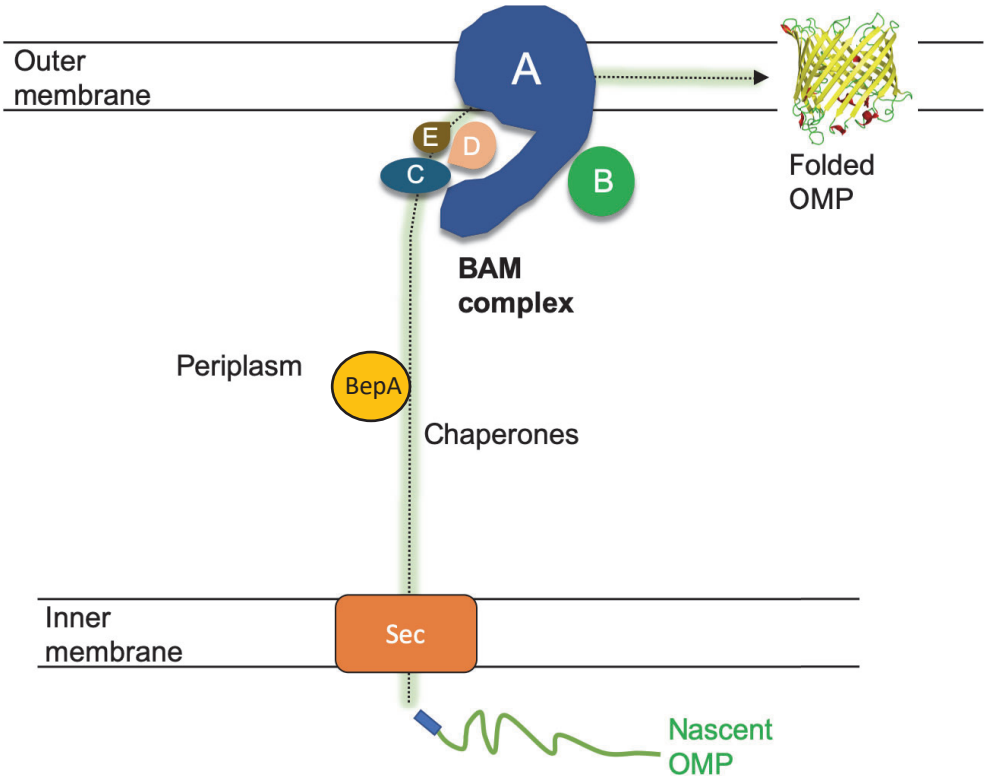

Figure 2. Proposed roles of BepA in the translocation of OMP into the OM.

Precursor OMP is translocated into periplasm region through the Sec translocon, and then to the outer membrane assisted by periplasmic chaperones and the BAM complex.

1.7.1 Activation of BepA by the σ^E pathway.

In bacteria, there are various signaling pathways that sense and respond to different stimulants from outside environment around them [26]. The bacterial envelope serves as the first defense line against most external stress factors [26]. The components in bacterial membranes sense irregularities and transduce signals to induce transcription [26]. This then leads to the response against these irregularities [26]. In *E. coli*, five well-known response pathways (Bae, Cpx, Psp, Rcs, and σ^E stress pathways) are activated in response to these stress factors[26]. 114 genes are known to be involved in the σ^E pathway including four homologs of periplasmic protease: DegP, PtrA, YhiJ, and BepA [22]. This indicates that BepA is one of the proteins expressed when *E. coli* cells experience envelope stress. Out of these four proteins, BepA is the only one speculated to have dual domains [22]; one of the domains as a chaperone or substrate attachment site, and the other domain that performs proteolytic activity. This unique characteristic might offer a mechanism of correlation between the two domains which lacks further information.

1.7.2 BepA as a chaperone.

Previous studies have shown that BepA may act as a chaperone for the assembly of OMPs. When LptD is correctly folded, two disulfide bonds are formed at non-consecutive cysteine residue pairs: C31-C724 and C173-C725. These disulfide bonds connect the periplasmic and the β-barrel domains of LptD, termed as $LptD^{NC}$ [22]. In the absence of BepA, however, LptD does not fold properly and disulfide bonds are formed between consecutive cysteine pairs (C31–C173 and C724–C725) termed as $LptD^{C}$ [22]. This hinders LptD from accumulating at the membrane which is undesirable for the cell. A study also showed that SDS-PAGE's band intensity of properly folded LptD increased with increasing concentration of BepA, and no maturation of LptD could be found

in cells lacking BepA [22]. Other than LptD, BepA was also reported to interact with BamA of the BAM complex. A study showed that BamA also interacts with the TPR domain of BepA. A study showed that residue E224 of the POTRA3 domain of BamA interact with the TPR domain of BepA [23]. BepA has also been reported to be involved in the proper recruitment of LoiP (YggG) to OM [27]. These results suggest BepA's chaperone-like activity promoting the correct folding of OMP (**Fig. 3**).

1.7.3 BepA as a protease.

In addition to the chaperon activity of BepA mentioned above, BepA also has a protease activity that cleaves misfolded OMPs to maintain the integrity of the OM [22]. BepA has a role in the maintenance of LptD by proteolysis [22]. In the case that LptD fails to mature due to the absence of LptE, $LptD^C$ accumulates in the periplasm [22]. BepA then degrade these accumulated $LptD^C$ [22]. The overexpression of wild-type BepA was also shown to improve the degradation of $LptD^C$ [22]. In addition, $LptD^C$ accumulates in the strain expressing mutant form of BepA lacking proteolytic activity [22]. It was also reported that proper assembly of BamA is perturbed without SurA which stimulates BepA[22]. Collectively, these observations suggest that BepA degrades non-functioning misfolded OMPs *in vivo* (**Fig. 3**) [22]. These results also suggest that the protease activity of BepA is crucial in maintaining the integrity of the OM of bacteria.

BamA contains polypeptide transport-associated (POTRA) domains that are periplasmically exposed [22]. The lipoprotein subunits of the BAM complex and substrate OMPs bind to the POTRA domains of BamA [22]. A previous study showed that BepA cleaves BamA in cells lurking SurA, suggesting that BamA is a proteolytic substrate of BepA [22]. In BepA cells with mutated active site, which will be elaborated in another section, no degradation of BamA could be

detected by SDS-PAGE [22]. The mechanism of degradation of BepA substrate, including the substrate specificity requires more investigation.

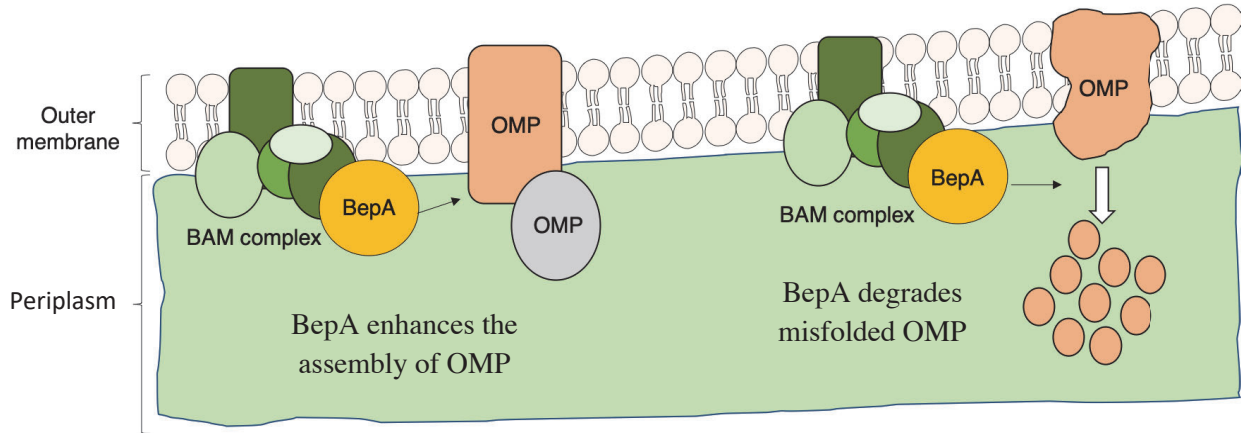

Figure 3. BepA acts as a chaperone and a protease.

Left: Proposed function of BepA in enhancing the assembly of OMPs by interacting with the BAM complex. Right: Proposed function of BepA in degrading misfolded OMPs.

1.7.4 Domain architecture of BepA.

As a sequence, BepA consists of an N-terminal protease domain and a C-terminal tetratricopeptide repeat (TPR) domain (**Fig. 4**). The amino acid sequence of BepA suggests that BepA is a peptidase of the M48 family possessing a zinc metalloproteinase active site motif, HEXXH [22]. In the HEXXH motif at the active site, there are usually histidine residues with the presence of a third amino acid coordinating a zinc ion [28]. The TPR motif of BepA is a structural motif that is often involved in protein–protein interactions [25]. A TPR domain generally comprises of three to sixteen repeats of about 34 amino acids which forms tandemly aligned anti-parallel α-helices [25].

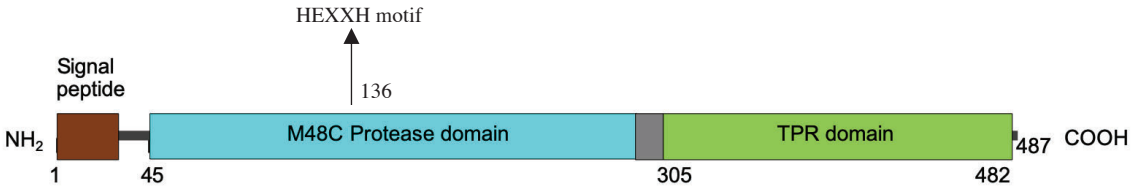

Figure 4. Domain arrangement of BepA.

BepA consists of protease domain at its N-terminal side and a tetratricopeptide repeat (TPR) domain at its C-terminal side with an approximate size of 51 kDa.

1.7.5 Comparison of sequences of BepA and a Zn-dependent peptidase Q74D82.

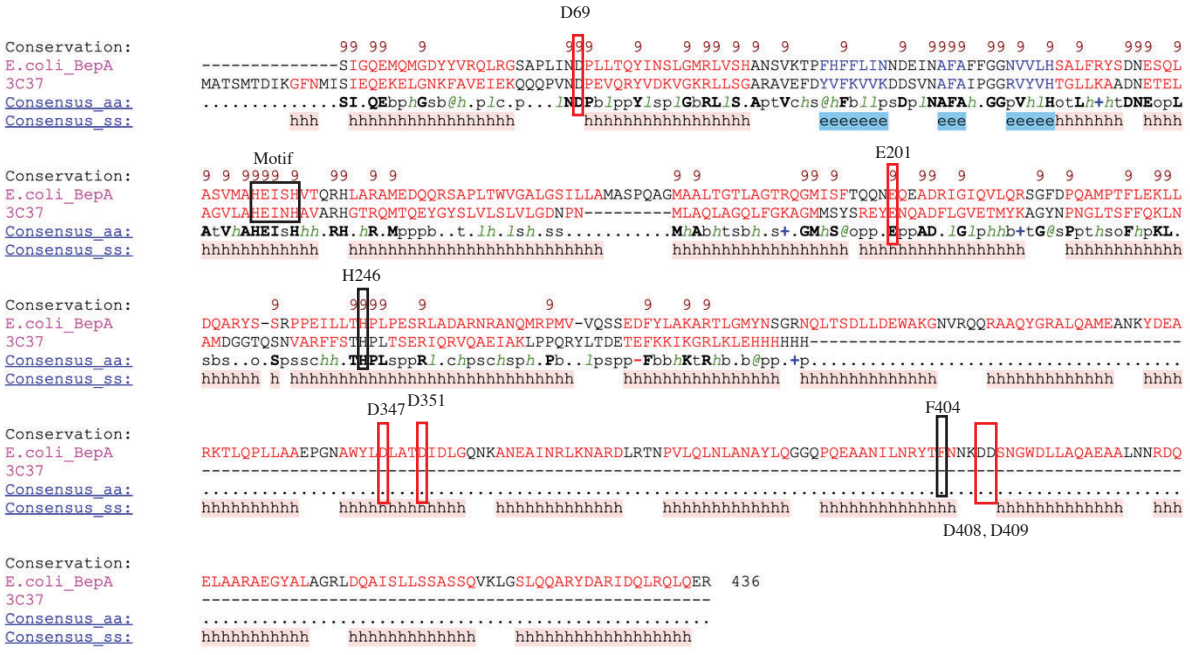

Figure 5. Sequence comparison of BepA and a Zn-dependent peptidase Q74D82.

Sequences of BepA (UniProt ID: P66948) and a M48 peptidase (UniProt ID: Q74D82) are aligned using ESPript 3.0. Helices are indicated as “h” highlighted in red, while β-sheets are labelled as ‘e’ highlighted in blue. Conserved residues are shown in bold and labelled as “9”. Important residues are labelled; the conserved HEXXH motif is indicated by black box; H246, which also coordinates the zinc atom, is also labelled; F404, which was showed to mediate the interaction of BepA with LptD, is also labelled. Residues contributing to negatively charged cavity are labelled in red boxes.

The sequences of the protease domain of BepA and Zn-dependent peptidase Q74D82 are relatively similar, with a 28% sequence identity. The HEXXH motif is also conserved since both of the proteases are of the M48 family metalloprotease.

Some residues of the conserved region were proven to be crucial for BepA's functions, especially the residues in the HEXXH motif. A study showed that when the conserved E137 residue was mutated, BepA lost its proteolytic function [22–25]. Other than that, cells with BepA's E137 mutation also had increased susceptibility to antibiotics, such as vancomycin [24]. BepA's H 136 and H246 might be important too, as its mutation to asparagine increased cells susceptibility to antibiotics as well [22,24]. Mutation of F404 residue in the TPR domain of BepA also led to cells' decreased resistance to antibiotics [25]. Other than that, F404 mutations also decreased the ability of BepA to cross-link with partner proteins, suggesting that F404 residue might be important for protein-protein interaction [25].

1.7.6 The tetratricopeptide repeat (TPR) domain.

As for the TPR domain of BepA, X-ray crystal structure was recently reported at a resolution of 1.7 Å (PDB ID: 5XI8) [25]. In the structure, BepA TPR domain is composed of ten anti-parallel α-helices (H1 to H10) (**Fig. 6**)[25]. The structure also showed that BepA TPR domain was composed of two subdomains: an N-terminal subdomain and a C-terminal subdomain. The N-terminal subdomain has a concave surface which is negatively charged, termed as a "pocket" which is common in other TPR domains [25]. The TPR domain of BepA might serve as interacting site for positively charged substrates through this pocket. The C-terminal region has a small "cavity" that is located at the opposite side of the concave surface of the N-terminal region [25].

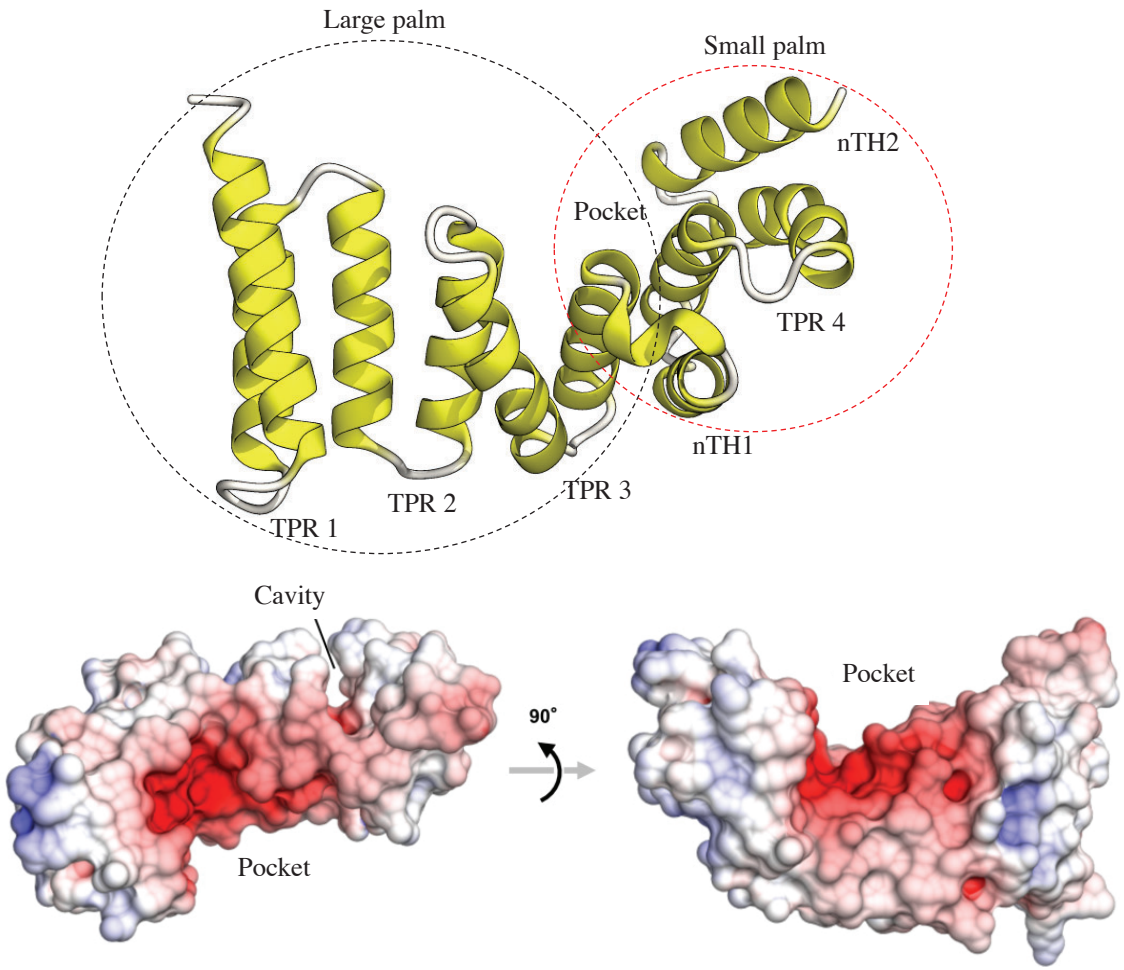

Figure 6. The TPR domain of BepA.

BepA's TPR domain consists of an N-terminal subdomain which forms a negatively charged pocket and a C-terminal region which forms a small cavity. Electrostatic potential ranging from blue (+10 kT/e) to red (-10 kT/e).

By employing an *in vivo* photo-cross-linking approach, Daimon *et al*. also confirmed that BepA's TPR domain is the interacting site for the BAM complex [25], as previously assumed since TPR domains are usually exploited as interacting site with other proteins [29].

BepA was proposed to interact with the BAM complex while carrying out its functions. It was speculated that this interaction is achieved when the TPR domain of BepA is integrated into

the central region of the ring-like structure facing the periplasmic space of the BAM complex (**Fig. 7**) [25]. In the model structure of BepA TPR domain-BAM complex, the TPR domain of BepA interacts with the BAM complex by anchoring BamA's POTRA 1 to POTRA 5 domains, BamC and BamD (**Fig. 7**).

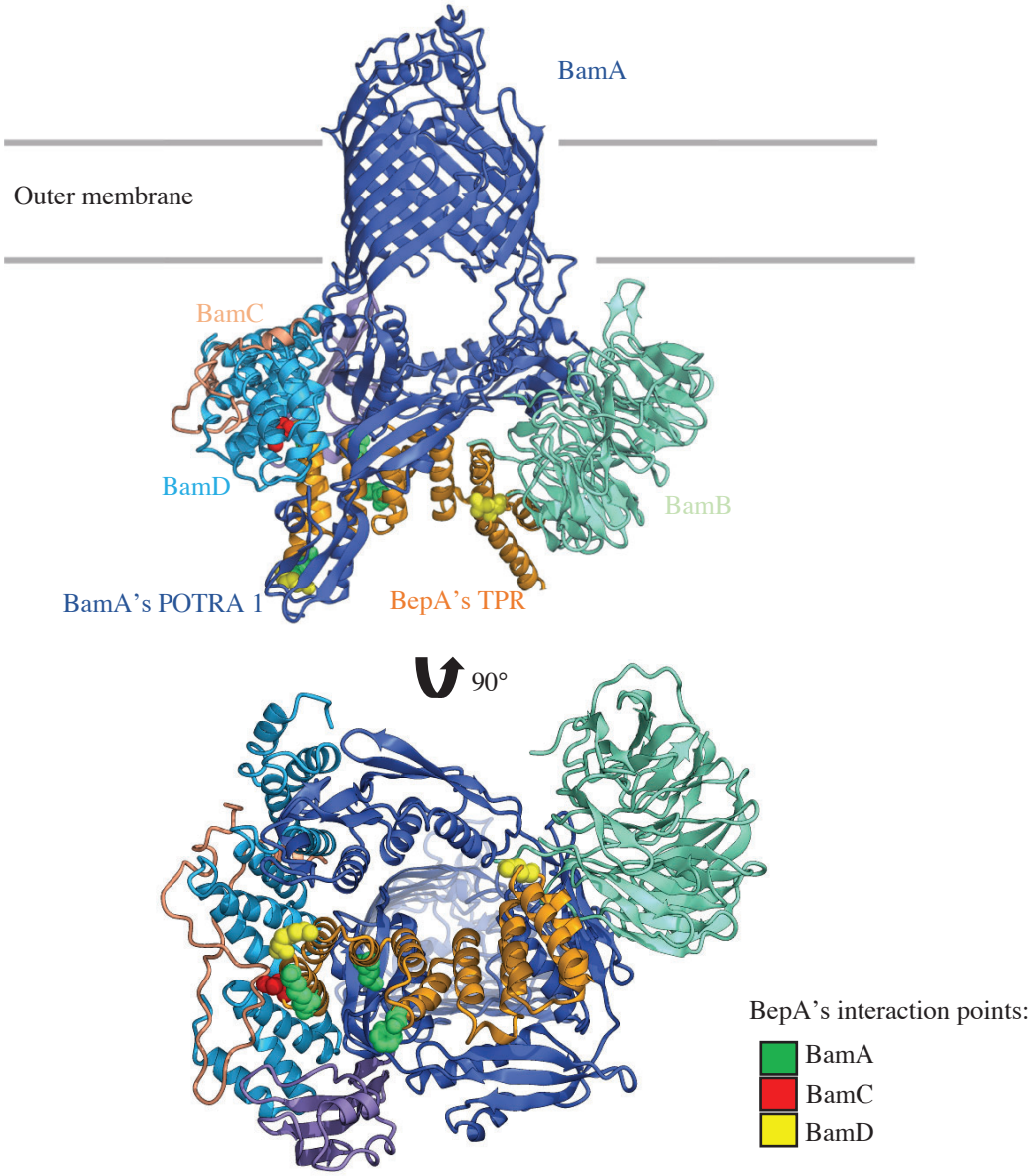

Figure 7. Interaction of TPR domain and the BAM complex.

The TPR domain of BepA interacts with the BAM complex by anchoring BamA's POTRA 1 to POTRA 5 domains, BamB and BamC.

1.8 Unclear molecular mechanisms of BepA.

Even though the structure of TPR domain of BepA was solved, the full-length structure of BepA is still required since studying only the TPR domain is not sufficient to propose the functions of BepA in detail. While there are four types of M48 metalloproteases in *E. coli*, BepA is unique in the sense that it has both protease and TPR domains, suggesting an interesting dual-function mechanism that is still unexplored. It is uncertain whether the two domains of BepA act dependently or independently of each other. The possibility of the two domains to interact with each other to maintain the structure of BepA is also debatable. The role of the H^{136} EXXH at the protease domain of BepA in cleaving substrates remained to be answered since there is also a binding pocket and cavity at the TPR domain which were also proposed to be involved in substrate attachment. To answer these questions, I aimed to obtain crystal structure of full-length BepA structure. By understanding the nature of important residues in the full-length BepA structure, mutational analyses could also be performed to study BepA's activities, thus giving a possibility for us to fully understand its functions. Other than that, by exploring the properties of BepA's active site or cavities, the properties of BepA's substrates could also be proposed. This would open opportunities to study BepA's interaction with a substrate.

CHAPTER 2 EXPERIMENTAL PROCEDURES

2.1 Expression and purification of BepA.

E.coli BepA sequence encoding 45-482 amino acids region without the signal peptide was cloned into pET-16b-TEV plasmid using BamHI site[23]. The resulting plasmid (pYD296 plasmid) encodes MG-H_{10}-SSGENLYFQG-$BepA_{45-482}$. The plasmid was then transformed into *E. coli* KRX strain cells (Promega) and stored at -80°C as a glycerol stock. In order to prepare the pre-culture, the glycerol stock of KRX/pYD296 was inoculated into 25 mL LB Broth, Lennox (Nacalai) to prepare the pre-culture. Then, 50 μg/mL ampicillin and 0.4% glucose were added to the medium and the cells were cultured overnight at 37°C. Then, the pre-culture was added together with 50 μg/mL ampicillin to large-scale culture of 2.5 L and further cultured at 37°C until the OD_{600} reached ~0.6. After that, 0.2% rhamnose was added to induce protein expression. The culture was transferred to 17°C incubator and shaken for 12 hours.

In order to harvest the cells, centrifugation at 6 000g was performed for 10 min at 4 degree (Hitachi CR22N, R9A2 rotor). The supernatant was discarded and the pellet was suspended in 10 mM Tris-HCl (pH 7.0). The solution was re-pelleted by centrifugation at 6,000g for 10 min and the pellet was re-suspended in Wash Buffer [20 mM Tris-HCl (pH 7.0), 500 mM NaCl, 20 mM imidazole-HCl (pH 7.0), 1 mM 2-mercaptoethanol (ME), and 0.1 mM phenylmethanesylfonyl fluoride (PMSF)]. In order to suppress the protease activity of one of the domains, 0.5 mM EDTA (pH 8.05) was added. The cells were disrupted by sonication for 30 min on ice using Q500 (QSONICA) by the following conditions; 20% power, pulse rate: 1 s pulse on and 1 s pulse off, sonicator rod: CL 334). Disrupted cells were separated by centrifugation at 20,400 g for 30 min at 4°C (Hitachi CR22N, R20A2 rotor).

For purification of BepA, the supernatant was mixed with 3 mL of Ni Sepharose Excel (GE Healthcare) pre-equilibrated with Wash Buffer. The mixture was rotated for 1 h in the cold room. After rotation, the resin was transferred to the column. The resin was then washed with 50 mL of Wash Buffer to remove unbound proteins and His-tagged proteins were eluted using 5 mL of Elute Buffer [20 mM Tris-HCl (pH 7.0), 500 mM NaCl, 300 mM imidazole-HCl (pH 7.0), 1 mM 2-ME, and 0.1 mM PMSF] for by 8-step gradient elution. SDS-PAGE was performed to confirm the presence of His-BepA protein in the eluted fractions.

In order to remove the polyhistidine-tag, the eluted fractions were pooled together into a 50 mL-Falcon tube and TEV protease was added so that the protein weight ratio would be 1:20 (1 mg of TEV protease to 20 mg of protein). After several rotations, the protein solution was transferred into a dialysis membrane [Spectra/Por 7 MWCO 10 kD (SPECTRUM)] and dialysed against Dialysis Buffer [20 mM Tris-HCl (pH 7.0), 500 mM NaCl, 1 mM 2-ME, and 0.1 mM PMSF] for overnight at 4 degree. The cleaved BepA protein was checked by SDS-PAGE. The dialysed sample was then subjected to a second Ni-NTA chromatography with the similar conditions to remove uncut His-BepA and TEV protease with His-tag. Cut BepA was then used for further purifications.

Cleaved BepA protein was concentrated to ~500 μL using the Amicon Ultra 10 K filter (Millipore) to be further purified by gel-filtration chromatography. Samples were loaded onto Superdex 200 Increase 10/300 column (GE Healthcare) pre-equilibrated with Gel-filtration Buffer [20 mM Tris-HCl (pH 7.0), 500 mM NaCl, 1 mM 2-ME, and 0.1 mM PMSF] and separated using AKTA explorer 10S (GE Healthcare). SDS-PAGE was performed according to the gel-filtration profile to confirm the purity of BepA protein.

2.2 Crystallization of BepA.

Purified BepA protein was pooled and concentrated to 17 mg/mL by the Amicon Ultra 10 K filter (Millipore) for crystallization experiments. 0.1 μL of protein drops only, and 0.1 μL of protein drops with 0.1 μL crystallization buffer drops were automatically prepared using the Crystal Gryphon (Art Robbins Instruments). Crystallization of BepA protein was performed by vapor diffusion method in 96-well Violamo crystallization plates (VCP-1) at 4°C and 20°C against solutions with different concentrations of 2-Methyl-2,4-pentanediol (MPD), ammonium sulfate $(NH_4)_2SO_4$ and sodium chloride (NaCl). Rod-shaped crystals were obtained in 2 M $(NH_4)_2SO_4$ condition after an overnight incubation at 4°C followed by an incubation at 37°C for overnight and incubation at 4°C for another five days. The crystals grew to a size of approximately 90 µM x 15 µM (**Fig. 8**). Then, the crystals were collected using harvesting mounts and loops (MiTeGen), and flash-frozen in liquid nitrogen in order to be transported to SPring-8 for X-ray diffraction data collection.

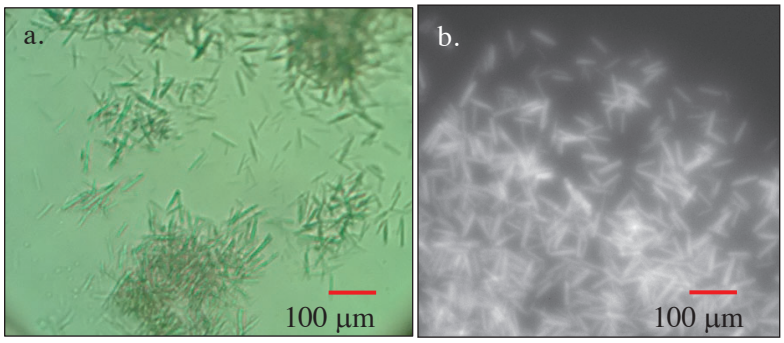

Figure 8. BepA crystals under bright-field microscope (a) and under UV exposure (b).

2.3 X-ray data collection, phasing and structure refinement.

2.3.1 Diffraction data collection.

X-ray diffraction data of BepA crystals was collected on beamline BL32XU at SPring-8 by the Helical Data Collection Method using a microbeam. The collected data was processed by X-ray Detector Software (XDS).

2.3.2 Molecular replacement.

Molecular replacement was performed by Phaser [30] to calculate the initial phases. For molecular replacement, previously determined BepA TPR domain (PDB: 5XI8) [25] and zinc-dependent peptidase Q74D82 (PDB: 3C37) were used as templates.

2.3.3 Structure refinement.

BepA initial structure was refined in a stepwise fashion using COOT [31] and PHENIX [32] to R_{work}/R_{free}= 0.206/0.263 at 2.6 Å resolution. Detailed refinement statistics are listed in **Table 1**. The space group was *P*1 and the molecular graphics were designed using CueMol2 (https://www.cuemol.org/).

Table 1. Refinement statistics of BepA structure.

Aspects	Values
Wavelength	1
Resolution range	48.18 - 2.598 (2.691 - 2.598)
Space group	*P*1
Unit cell	85.844 104.674 104.971 113.606 105.843 104.026
Total reflections	327914 (31495)
Unique reflections	89522 (8738)
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)
CC1/2	0.997 (0.402)
CC*	0.999 (0.757)
Reflections used in refinement	89487 (8729)
Reflections used for R-free	2016 (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	19732
macromolecules	19568
ligands	54
solvent	110
Protein residues	2485
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
macromolecules	81.39
ligands	86.44
solvent	56.41
Number of TLS groups	35

CHAPTER 3 RESULTS

The structure of BepA was modelled at 2.6 Å resolution with R_{work}/R_{free} = 0.21/0.26 as described in the Chapter 2. There are six BepA molecules (Mol 1 to 6) in the asymmetric unit (**Table 2, Fig. 9**), with Mol 2 has the clearest electron density. The results of superimpositions of the other five molecules to Mol 2 are shown in **Table 2**, giving an average of RMSD = 0.48 Å, indicating no significant structural change.

Table 2. Superimpositions of the six BepA molecules in the asymmetric unit.

Superimposition	RMSD
Mol 2 to Mol 1	0.57
Mol 2 to Mol 3	0.54
Mol 2 to Mol 4	0.60
Mol 2 to Mol 5	0.47
Mol 2 to Mol 6	0.69
Average	**0.48**

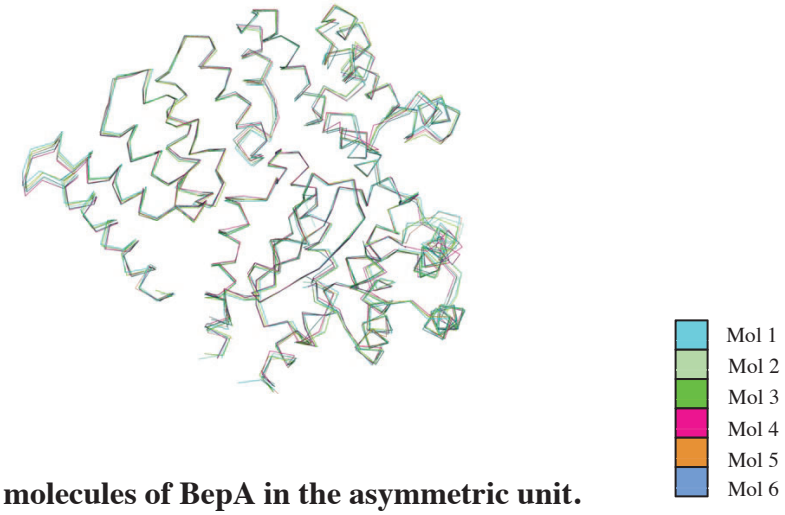

Figure 9. Comparison of six molecules of BepA in the asymmetric unit.

The superimpositions of Cα atoms of each BepA molecule in the asymmetric unit showed no significant structural change.

3.3 Structure of BepA.

3.3.1 Overall structure of BepA.

In the crystal structure, there is a N-terminal protease domain from residues 45-288 and a C-terminal TPR domain from residues 305-482 (**Fig. 10**). In the protease domain, there are 11 alpha-helices (α1-α11) and a β-sheet with three β-strands (β1-β3) between α2 and α3 (**Fig. 10**). In the TPR domain, there are 10 anti-parallel α-helices (H1-H10) (**Fig. 10**) which is consistent with the previously reported TPR domain structure of BepA (PDB ID: 5XI8) [25]. α12 connects the two domains together (**Fig. 10**). The regions between α5-α6 (153-159) and α6-α7 (178-191) could not be modeled due to poor electron density in those sections (**Fig. 10**, red dashed line).

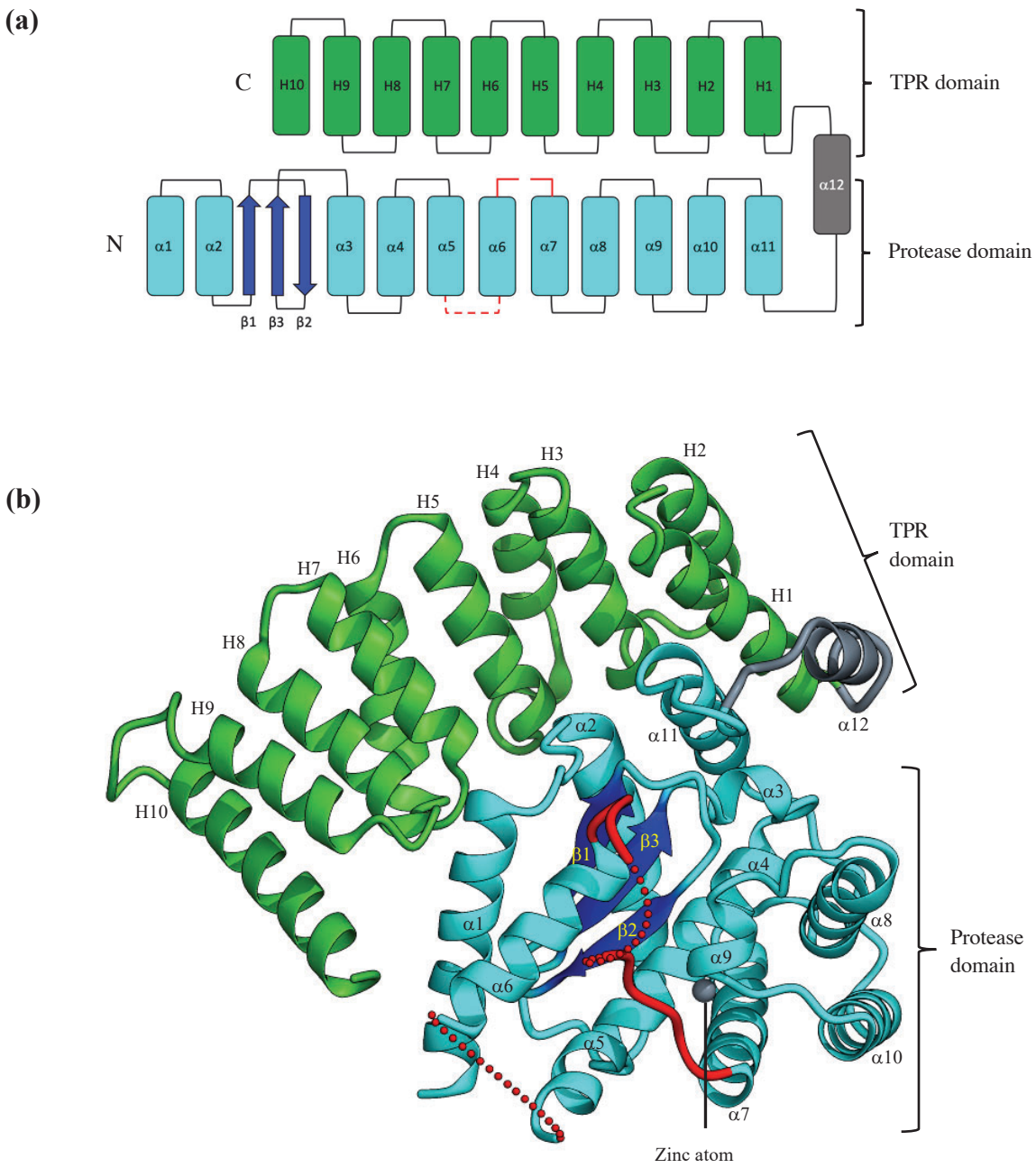

Figure 10. Crystal structure of BepA.

(a) Schematic representation of the crystal structure of BepA. There are 11 N-terminal alpha helices (α1-α11) and 3 strands of β-sheets in the protease domain, while there are 10 C-terminal anti-parallel alpha helices (H1-H10) in the TPR domain. α12 helix connects the two domains together. Red dashed-lines represent unmodeled regions of the structure.

(b) The N-terminal protease domain is colored in aqua while the C-terminal TPR domain is colored in green. β-sheets are colored in blue. The α12 helix helix colored in grey acts as a connecting helix. Residues between α5-α6 and α6-α7 are colored in red dashed-lines.

3.3.2 The protease domain.

The protease domain of BepA in the crystal structure is relatively similar to the structure of zinc-dependent peptidase Q74D82 (PDB ID：3C37), as expected from its sequence with a sequence identity of 28% (**Fig. 5**, **Fig. 10**).

3.3.2.1 The active site motif of the protease domain.

In the protease domain of the crystal structure of BepA, the active site motif H^{136}EISH holds a zinc atom, which is coordinated by three histidine residues and one glutamate residue. Two histidine residues (H136 and H140) are located on α4 helix and a glutamate residue (E201) reside at α7 helix and the third histidine residue (H246) is placed at α9-α10 loop (**Fig. 11**). In the current full-length BepA structure, the active site of the protease domain was covered by the loop between α6-α7 (**Fig. 11**), with the zinc atom embedded deeply between α5, α6, α7 and α9. Although the loops connecting to α6 were not visualized in the current BepA crystal structure (**Fig. 11**), this entire region may act as a cap for the active site.

In the active site, the distances between the zinc atom and coordinating residues are as follows: H136, 2.11 Å; H140, 2.29 Å; H246, 2.13 Å; and E201, 2.02 Å (**Fig. 11**). For comparison, in the previously solved zinc-dependent protease Q74D82 (PDB ID：3C37), the distances between the zinc atom and coordinating residues are as follows; H106 with 2.07 Å and H110 with 2.06 Å, followed by E162 with 1.95 Å and H208 with 2.04 Å (**Fig. 12**). In both crystal structures of the present study and Q74D82 protease, both zinc atoms have the shortest distance with coordinating glutamate residues which are almost similar (2.02 Å in the present structure and 1.95 Å in Q74D82 structure). In order to further analyze the differences between the two structures, the most common method to quantitatively measure similarity between two or more protein structures is to use the

root mean square distance deviation (RMSD) of superposed atoms [33]. This method evaluates structure superpositions and the lower the RMSD, the better the similarity between the two structures [33]. When the protease domain and Q74D82 structures were superimposed, the RMSD was 1.77 Å for the Cα atoms, indicating that the overall architecture of the two structures is not significantly different (**Fig. 12**). The position of the active site motif of both structures is also similar (**Fig. 12**).

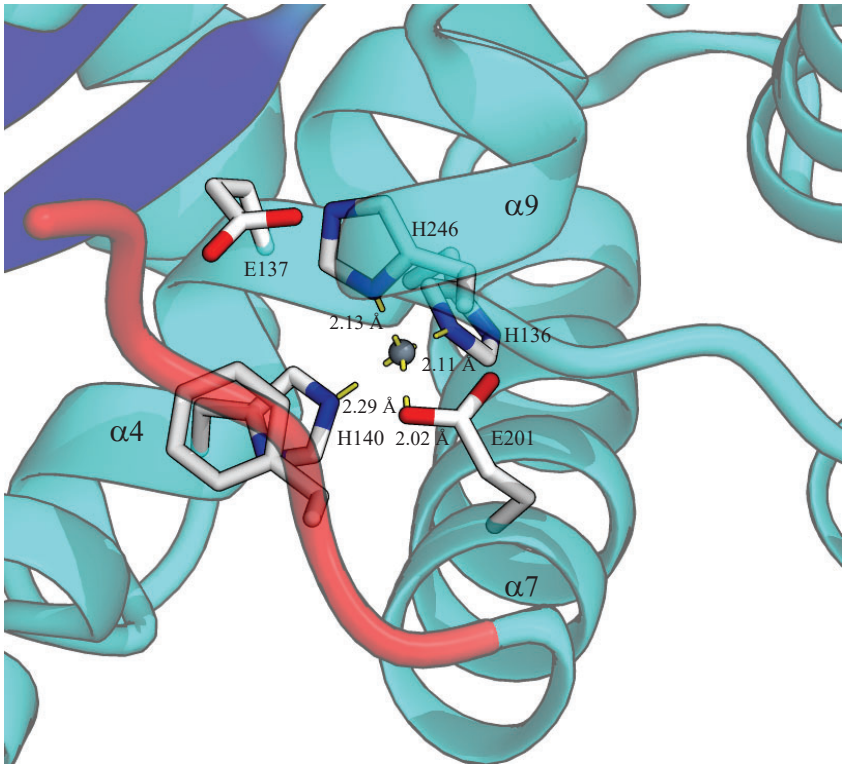

Figure 11. The active site motif H[136]EISH.

In the current BepA structure, the zinc atom is coordinated by three histidine residues and one glutamate residue; two histidine residues (H136 and H140) reside at α4 helix, one glutamate residue (E201) at α7 helix and last histidine residue (H246) at α9-α10 loop.

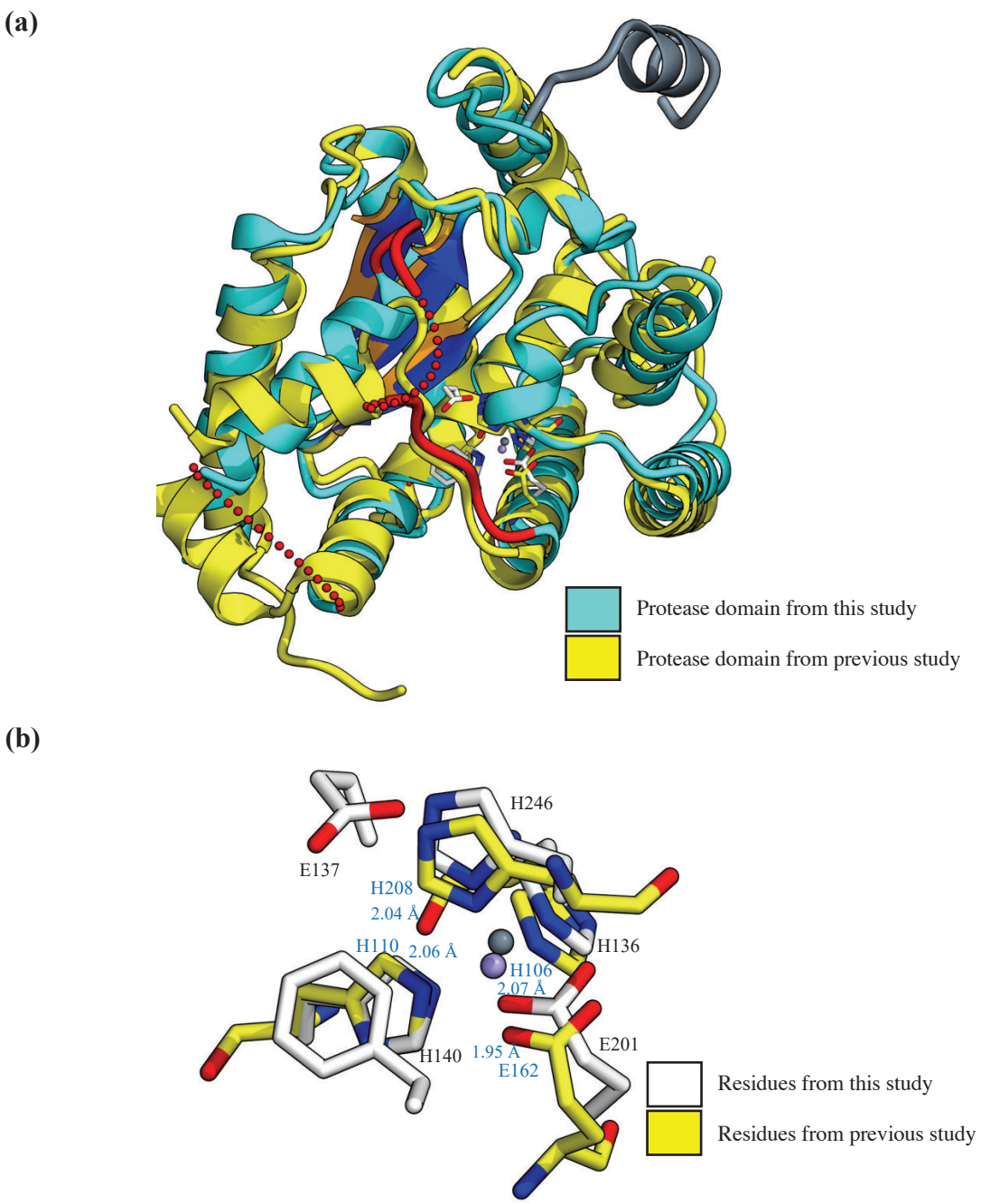

Figure 12. Superimposition of BepA's protease domain and Q74D82 protease.

(a) Overall superimposition of BepA's protease domain and Q74D82 protease. There is no significant structural difference between BepA's protease domain (aqua) and Q74D82 protease (yellow). α12 in the BepA structure is absent in Q74D82 protease structure.

(b) Superimposition of the active site motif HEXXH of BepA's protease domain and Q74D82 peptidase. The coordinating residues from BepA's protease domain are colored in white and labelled in black, and yellow for those from Q74D82 protease and labelled in blue.

3.3.3 The TPR domain.

There are 10 α-helices in the BepA crystal structure, which is consistent with the previously reported structure of isolated TPR domain of BepA [25]. When TPR domain region from the full-length BepA structure was superimposed with the previous TPR domain structure, there is no significant global change between the two structures, with an RMSD of 1.67 Å (**Fig. 13**). When the length of TPR domain is measured, there is a slight extension observed in the TPR domain of full-length BepA structure (43.50 Å) compared to that of the previous TPR domain structure (39.71 Å) (**Fig. 13**). This extension is probably due to the association of the TPR domain with the protease domain [25].

When surface presentation according to electrical charges was calculated, there was a negatively charged pocket in the TPR domain that is facing toward the protease domain with the size of 20.2 Å × 12.5 Å wide and a depth of 14.6 Å (**Fig. 14**) consistent with previous TPR domain alone structure **(Fig. 6)**. The negative charge in the pocket is contributed by D69 of the protease domain, E321 from H1 helix, D347 and D351 from H3 helix, and D408 and D409 from H6–H7 loop (**Fig. 14**). From the surface presentation, a ditch that protrudes from the pocket to the active site was also present (**Fig. 14**). This ditch is also thought to work as the binding site for the substrates.

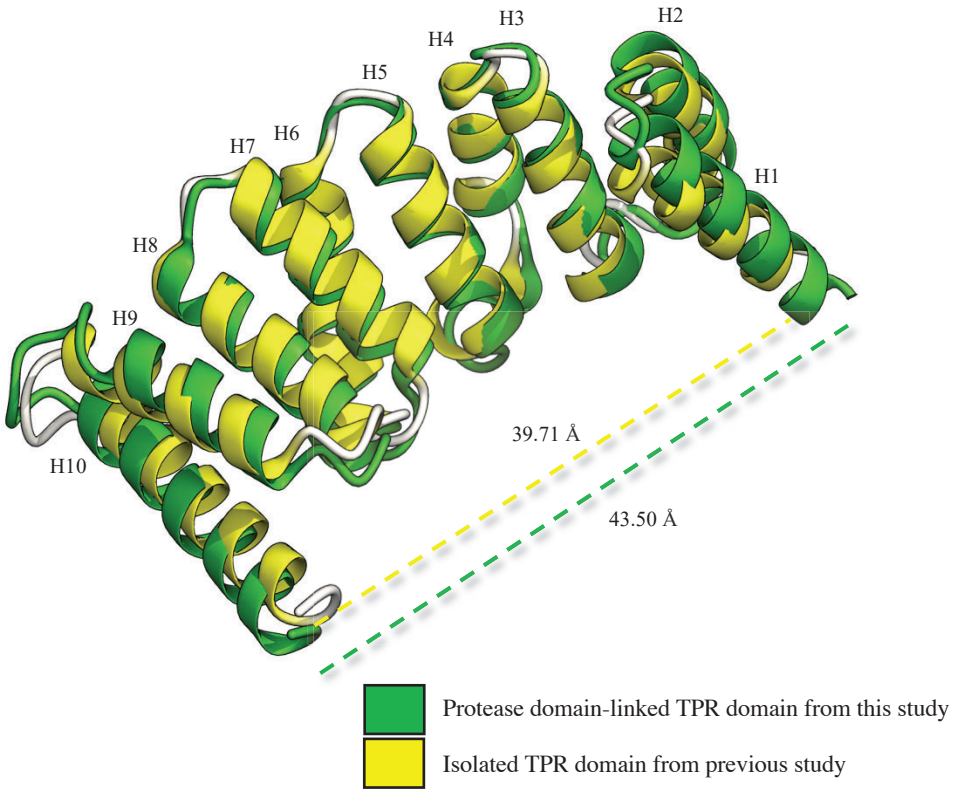

Figure 13. Comparison of the TPR domains from two crystal structures.

Superimposition of the TPR domain region from the current BepA structure (green) with the previously solved structure of TPR domain (yellow). In the current crystal structure, the TPR domain is slightly extended compared to the previously solved TPR domain.

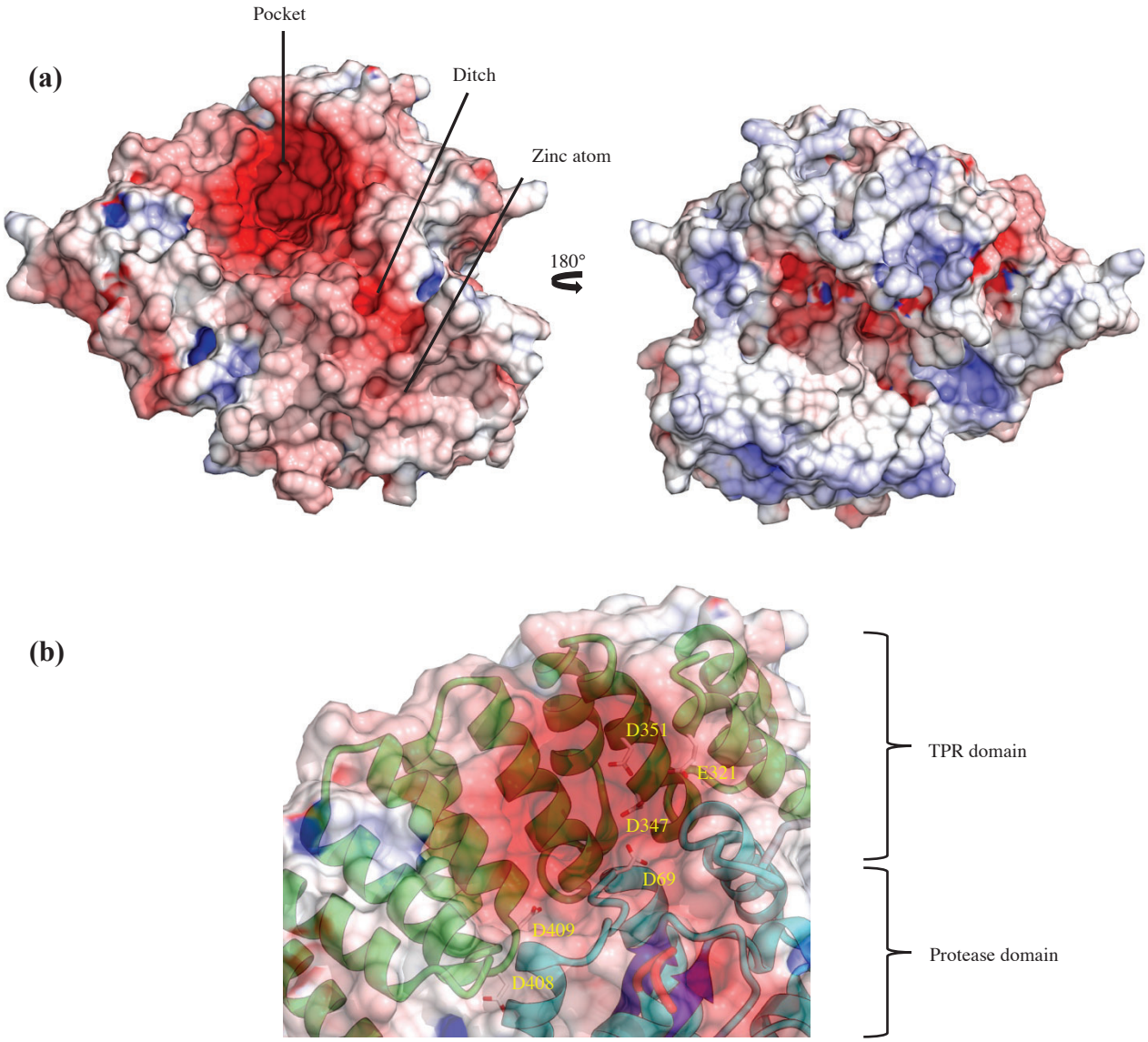

Figure 14. Surface charge distribution of BepA.

(a) There is a negatively charged pocket in the TPR domain and a protruding ditch from the pocket to the active site in the protease domain. Surfaces are colored based on electrostatic potential ranging from blue (+10 *kT/e*) to red (−10 *kT/e*).

(b) D69, E321, D347, D351, D408 and D409 are located at the negatively charge pocket.

3.3.4 Interaction between the domains.

In the crystal structure of BepA, the protease and TPR domains were interacting with each other. α-helices H1, H3 and H5 of the TPR domain at N-terminal side are tightly associated with α-helices α2, α11, and α12 of the protease domain by both hydrophobic and hydrophilic interactions (**Fig. 15**). In addition to this, there are hydrogen bonds between the C-terminal region of the TPR domain with protease domain: O of N68 and Nδ2 of N411, Nη1 of R61 and the O of L439, and Oε1 of Q51 and Nη1 of R480 (**Fig. 15**).

3.3.5 Interaction of BepA with substrate.

Previous small-angle X-ray scattering (SAXS) experiment performed by our laboratory implied that a histidine tag (His-tag) induces an oligomeric form of BepA. The generated Kratky plot, in which linear plots of functions of the scattered intensity I(Q) plotted against functions of the scattering variable Q (**Figure A1**). The SAXS analysis showed that BepA with His-tag forms larger oligomeric state compared to BepA without His-tag. I speculated that His-tag may mimic a substrate to BepA probably due to its positively charge surface. The generated Kratky plot also showed the presence of additional peaks, or shoulders, at low q, indicating that BepA forms ununified conformations.

3.3.6 Model of interaction of BepA and the BAM complex.

BepA's TPR domain was proposed to interact with BAM complex while performing its functions [25] (**Fig. 7**). Here, model of interaction between full-length BepA and BAM complex was generated by fitting the full-length BepA structure into this model (**Fig. 16**). With the current configuration of TPR domain and protease domain in the full-length BepA structure, protease domain can be located facing the periplasmic region without clashing the BAM complex (**Fig. 16**). This further supports the previous model of interaction of BepA's TPR domain and BAM complex.

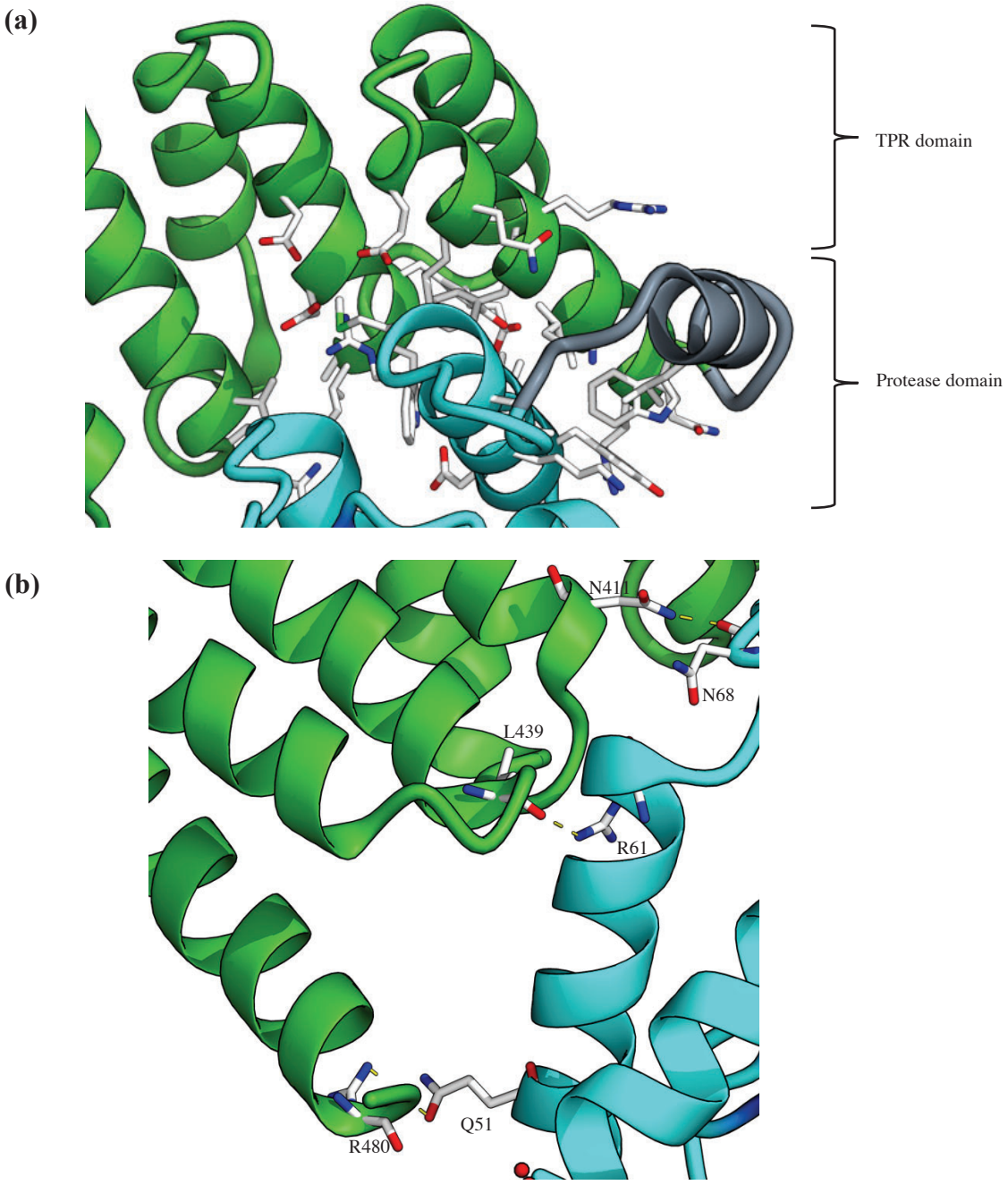

Figure 15. Interaction between the domains of BepA.

(a) Hydrophobic/hydrophilic interactions between the protease domain of BepA and the N-terminal region of the TPR domain of BepA.

(b) Hydrogen bonds formed between the C-terminal region of the TPR domain and the protease domain.

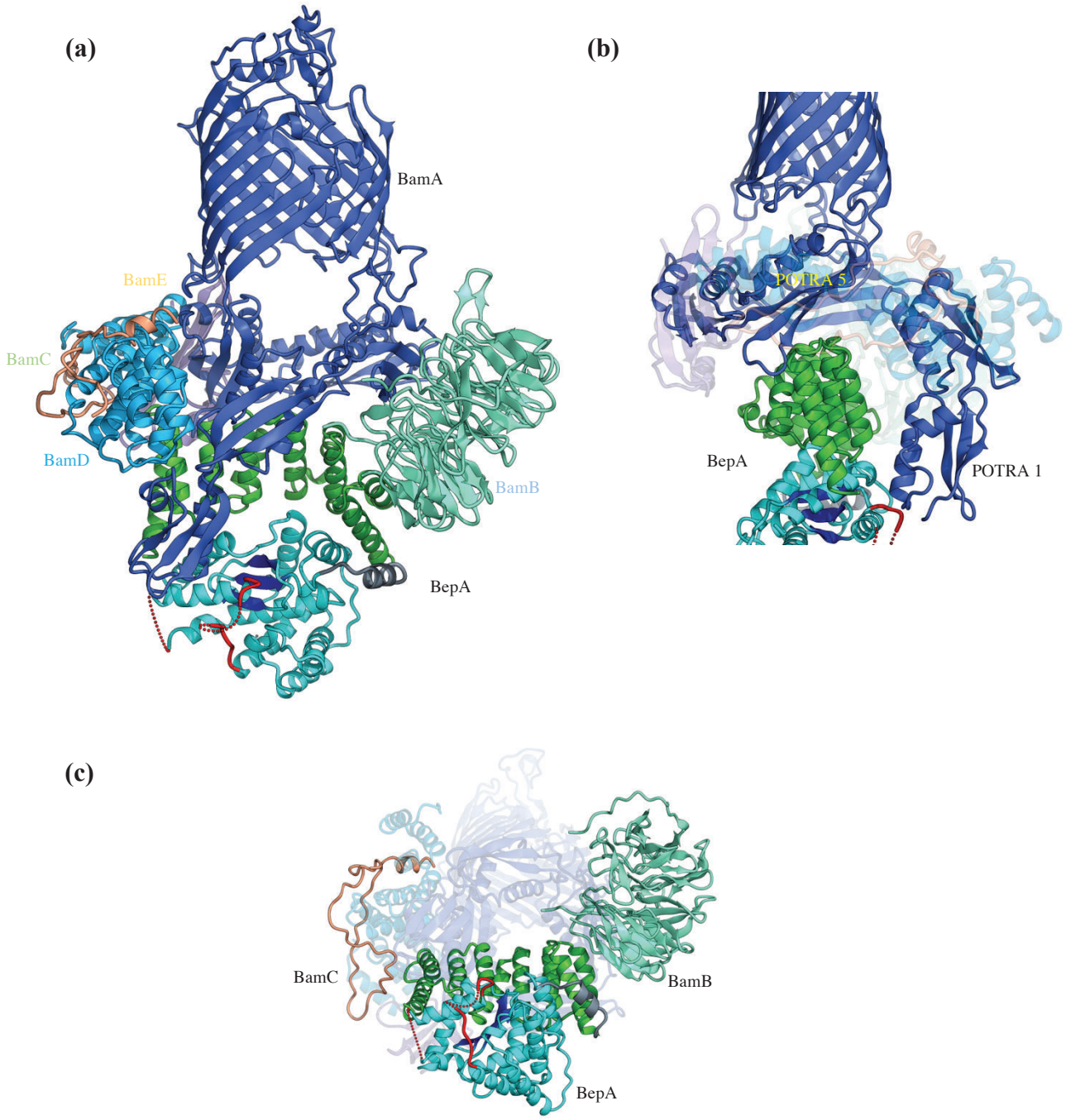

Figure 16. Model of interacting scheme of BepA with the BAM complex.

The TPR domain is thought to be involved in the overall interaction and protease domain was located at the opposite side from BAM complex. (a) Overall superimposition of full-length BepA with BAM complex showing the protease domain faces the periplasmic side. (b) The TPR domain was proposed to interact with the POTRA domains. (c) The TPR domain was also proposed to interact with BamC.

CHAPTER 4 DISCUSSION

The crystal structure of full-length BepA was determined at 2.6 Å resolution. The full-length crystal structure shows the presence of two domains, the N-terminal protease domain and the C-terminal TPR domain.

In order to test whether the overall structure of BepA in solution has the same conformation as that of the current crystal structure, additional experiments were performed with collaborators. First, the conformation of BepA in solution was tested by SAXS analysis[23] (**Fig. A1**). The three-dimensional shape model of BepA together with the experimental scattering curve and based on SAXS data showed that BepA in solution had the same conformation as in the crystal structure. The two domains were probably locked together by the aforementioned interactions. Next, to confirm this assumption, cysteine residues were introduced to the two domains so that they would form disulfide bond. Derivatives of BepA's mutants were constructed by site-directed mutagenesis. Disulfide bonds can be formed under oxidizing conditions, between the thiol groups of cysteine residues by the process of oxidative folding[34]. Residues that have a close proximity between each other (<15 Å) at the protease and TPR domains of BepA were selected and mutated to cysteine residues. The selections include Q51, Y55, V57, R58, and R61 from the protease domain, and A440, G441, R442, F481 from the TPR domain[23] (**Fig. 17**, **Fig. A2**). This would lock the conformation of the BepA structure as the conformation of the crystal structure.

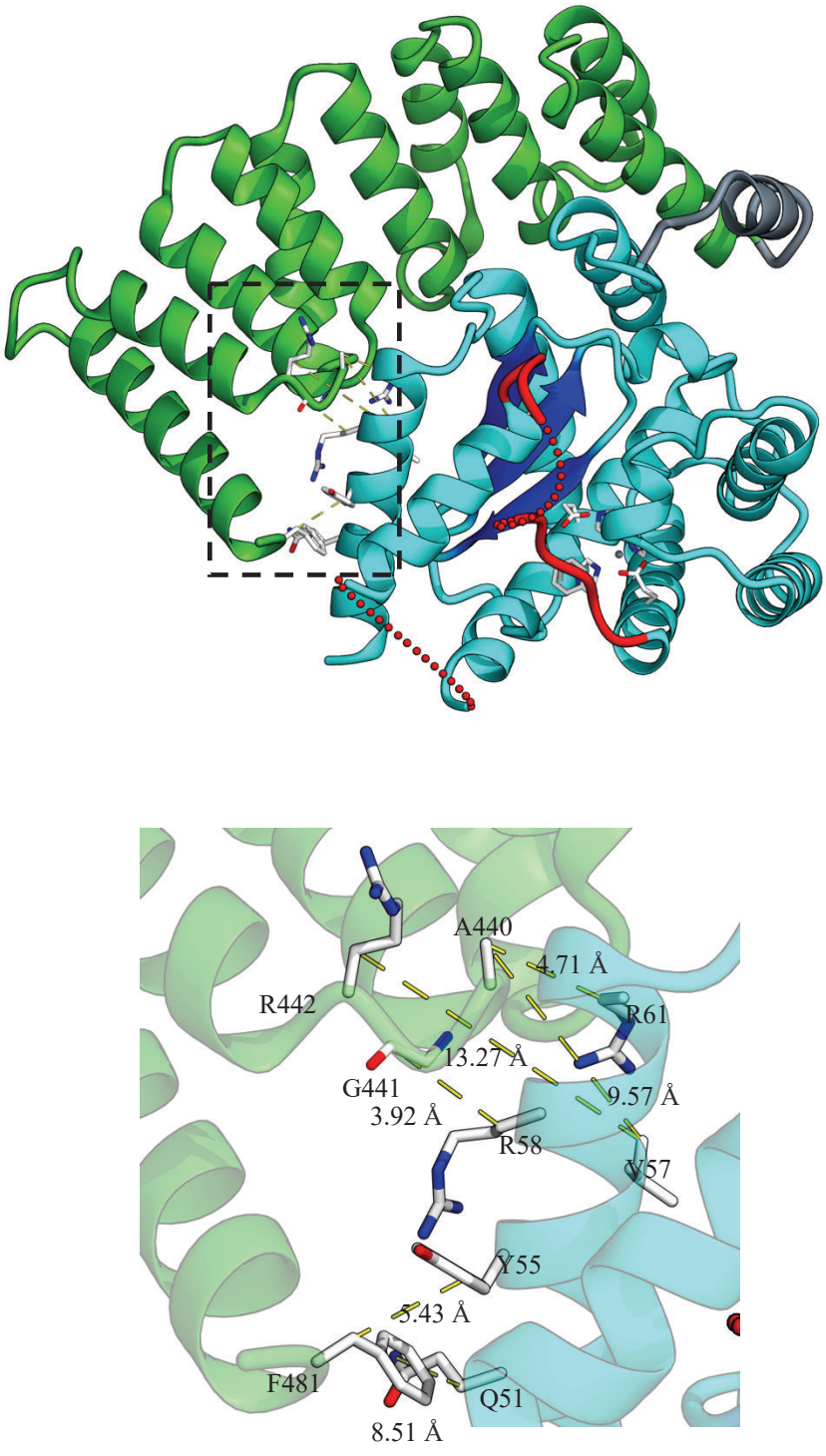

Figure 17. Disulfide bonds introduction between the protease and TPR domains.

Residues mutated to cysteine residues are mapped onto the crystal structure. The distance between each mutated residue pair is also shown.

In order to confirm the presence of disulfide bonds in BepA mutants, the migration of BepA bands was inspected in SDS-PAGE [23] (**Fig. A2**). Proteins with disulfide bonds are more compact and have lesser resistance during their migration on SDS-PAGE, thus they are able to migrate farther than non-bonded proteins [34]. Based on this principle, the presence of disulfide bonds in -BepA mutants were tested. Mutants with two cysteine residues, Q51C–F481C, Y55C–F481C, R58C–G441C, and R61C–A440C showed band shift which were farther than single cysteine mutants under oxidizing condition on SDS-PAGE gel[23] (**Fig. A2**). This confirmed the occurrence of disulfide bonds in the BepA mutants. The performance of the BepA mutants was then tested by their ability to fold LptD properly[23] (**Fig. A2**). BepA was shown to enhance the folding of LptD by shifting the disulfide bonds arrangement of LptD assembly intermediate ($LptD^{C}$) from C31–C173 and C724–C725 to a mature LptD ($LptD^{NC}$) with the disulfide bond arrangements of C31–C724 and C173–C725 [22]. Based on this principle, the capability of LptD folding by BepA mutants was tested[23] (**Fig. A2**). As the negative control, as expected, cells without BepA exhibited accumulation of $LptD^{C}$ [23] (**Fig. A2**). For BepA mutants in locked form the folding capability of $LptD^{C}$ to $LptD^{NC}$ was retained [23] (**Fig. A2**). This shows that BepA does not require a large conformational change in order to correctly fold LptD.

BepA has functions in excluding antibiotics with high molecular mass, such as erythromycin (EM), vancomycin , novobiocin and rifampicin [22]. Based on this characteristic, the performance of the BepA mutants was also tested by their sensitivity to erythromycin by calculating the minimum inhibitory concentration (MIC) of EM (**Fig. A2**). Cells lacking BepA were more sensitive to EM with the MIC of around 6.25 μg/mL while WT cells can tolerate up to around 50 μg/mL EM[23] (**Fig. A2**).The cells with BepA mutants also showed similar EM resistance as that of wild-type BepA with the MIC of around 50 μg/mL[23] (**Fig. A2**). This suggests the locking

of the two domains does not affect the activity of BepA proteins and mutated BepA are still able to perform their functions. This shows that BepA mutants were still able to retain their chaperone activities in a locked conformation of the TPR and the protease domains.

Then, BepA mutants were tested with their capability to degrade BamA in Δ*surA* cells[23] (**Fig. A3**). BamA was shown to be degraded by BepA in Δ*surA* cells[22]. Cells without BepA did not exhibit BamA's degradation products[23] (**Fig. A3**). In cells with BepA mutants, however, they were still able to degrade BamA which was comparable to the wild-type cells and single cysteine mutants[23] (**Fig. A3**). This suggests that protease activity of crosslinked BepA mutants was also still retained despite the limitation of the dynamics of the domains.

Recently, another group reported the crystal structure of full-length BepA [24] (**Fig. 18**). In the crystal structure by Bryant *et al*., the overall conformation of BepA's crystal structure was similar with the full-length BepA structure obtained in this work (**Fig. 18**). In the latest BepA structure, which I will name as Bryant's BepA structure from now on, the HEXXH motif was also positioned in the H4 helix (α4 in my structure) **(Fig. 18)**. In Bryant's BepA structure, the zinc atom is also regulated by H136, H140, E201 and H246 **(Fig. 18)**. In Bryant's BepA structure, there is an active site plug between helices H6 and H8, which is labelled as H7 plug (α9 in my structure). H246 is positioned in the active site plug, which is similar to my structure. Since the active site plug in both of my and Bryant's BepA structures seems to cover the zinc atom, these two structures of BepA might represent BepA in its inactive form.

In Bryant's BepA structure, α5 and α6 helices could not be modeled due to poor electron density in those regions. However, α5 and α6 helices were able to be modeled in my structure, indicating these helices might be flexible **(Fig. 18)**. This speculation is further supported by the

observation of the poor density between α5 and α6 helices in my structure **(Fig. 18)**. Bryant *et al*. also proposed the position of a putative lid that offers access to the active site **(Fig. 18).**

In the crystal structures of full-length BepA, the two domains interact with each other (**Fig. 18**). The TPR domain's N-terminal side interacts with the protease domain by hydrophobic/hydrophilic interactions, while the TPR domain's C-terminal side interacts with the protease domain by hydrogen bonds (**Fig. 17**). It is also probably due to these domain interactions that keep the two domains together to retain the functions of BepA *in vivo*. Since the overall conformation of the crystal structures of the full-length BepA are relatively the same [24], these interactions are thought to be very stable. By taking all the observations obtained in this study, the SAXS analysis, the functional analyses and the structure reported by Bryant *et al*., it is highly likely that the conformations of BepA in the crystal structure reflect the functional conformation *in vivo* and it does not involve conformational changes when performing the typical functions of BepA.

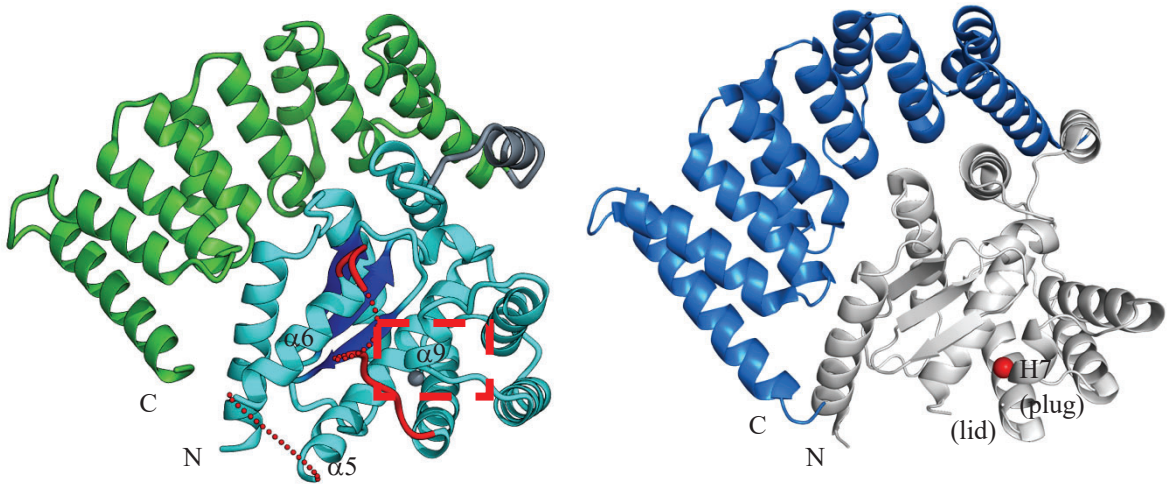

Figure 18. Comparison of the full-length BepA crystal structures.

The conformations of the two crystal structures are relatively similar, except for the missing α5 and α6 helices in the most recent structure (right). The putative lid and the H7 plug, which is the α9 plug in my structure (left), is also labelled.

In the protease domain, there is an extra α-helix (α12) compared to the previously determined Q74D82 peptidase (PDB ID 3C37). α12 acts as a linker connecting the protease domain to the TPR domain. Other than this, the overall structure of the protease domain of current crystal structure and that of Q74D82 peptidase is relatively the same (**Fig. 12**). The active site motif $H^{136}EXXH$ is also similarly positioned (**Fig. 12**). Thus, the proteolysis mechanisms of the protease domain of BepA might be similar with those of Q74D82 peptidase. In the protease domain, there are residues between α5-α6 (153-159) and α6-α7 (178-191) which could not be modelled due to poor electron density (**Fig. 10**). In the crystal structure of Q74D82 peptidase, α6-α7 loop was able to be modelled, suggesting that the loop might be flexible in BepA which explains the

poor density at the regions. However, α5-α6 loop which is absent in the current crystal structure was also absent in Q74D82 peptidase structure.

As mentioned above, recently reported full-length BepA structure by Bryant *et al*. lacked the entire α5 and α6 helices region. Another molecular surface representation according to electrical charges was calculated. The surface electrical charges showed some differences (**Fig. 19**). When the structure lacks α5 and α6 helices, the position of the original negatively charged ditch was different compared to the BepA's structure having α5 and α6 helices (**Fig. 19**). This different 'route' was more extended beyond the active site compared to that of the BepA's structure having α5 and α6 helices (**Fig. 19**). The extended ditch is also negatively charged (**Fig. 19**). In my BepA structure, α5 and α6 helices were able to be modelled while the surface electrical charges did not yield negatively charged surface beyond the active site. This indicates that substrates probably could not bind to the active site when α5 and α6 helices are present. However, when α5 and α6 helices are dislocated, the ditch opens to a new route, more extended compared to the original ditch and is also negatively charged. This raises the possibility of substrates attachment also to this extended ditch when α5 and α6 helices are absent.

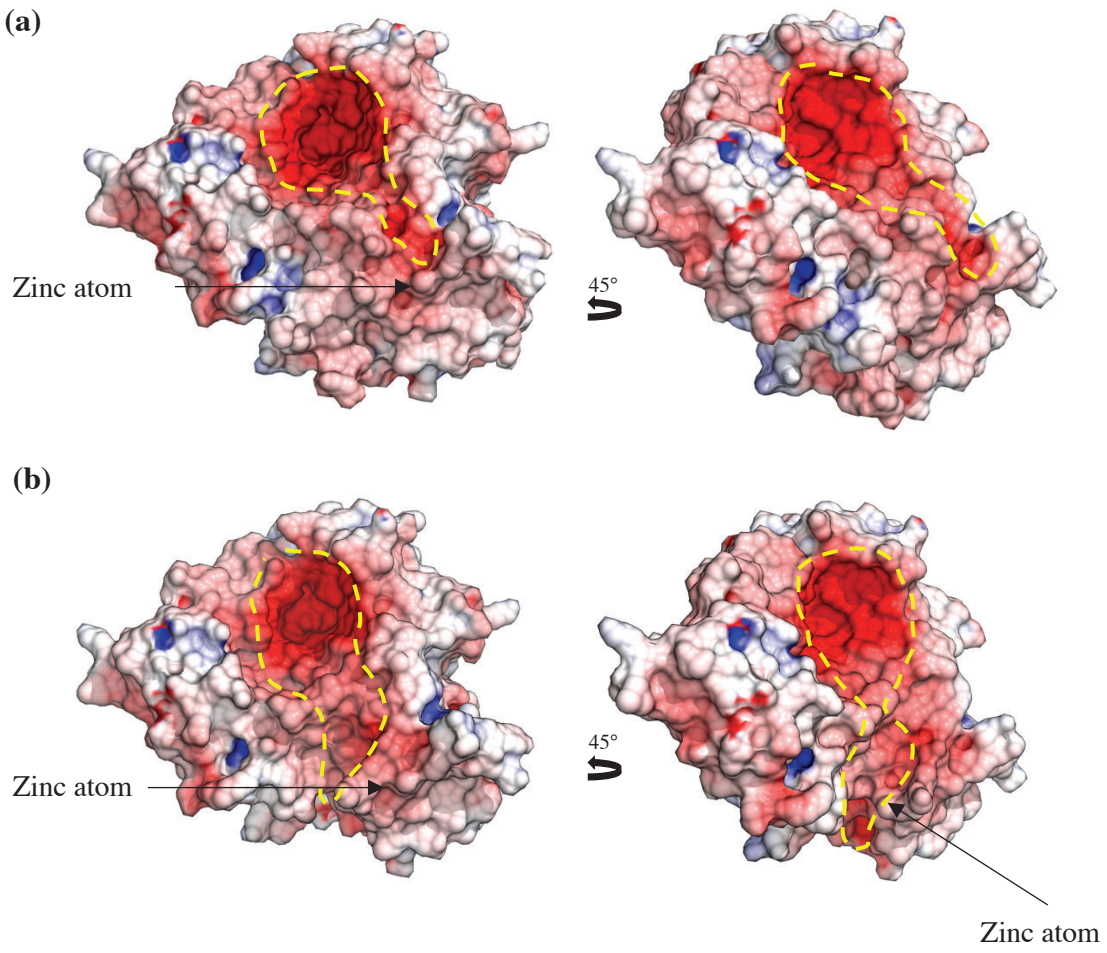

Figure 19. Comparison of surface electrical charges.

The negatively charged ditch in the BepA's structure without α5 and α6 helices **(b)** is different from that of the BepA's structure with α5 and α6 helices **(a)**. When α5 and α6 helices are absent, the negatively charged ditch extended beyond the location of the active site. This raises the possibility of substrates attachment to this extended region, too. Surfaces are colored based on electrostatic potential ranging from blue (+10 *kT/e*) to red (−10 *kT/e*).

Probably α5 and α6 helices are flexible, which allows some space for substrate to attach to this extended region. The size of the substrate would matter too. Whether or not these helices, especially α6 helix, which seems 'detached' from the rest of the protease domain due to the diordered loops (**Fig. 18**), is flexible and takes part in covering the active site, remains a matter to be resolved.

The TPR domain of the current crystal structure is relatively similar to the previously determined TPR domain of BepA (PDB ID: 5XI8) (**Fig. 13**) [25]. The TPR domain in the current study was found to be slightly extended compared to the previous structure (**Fig. 13**). This was presumably due to the association of the protease domain with the TPR domain, since the previous structure was of an isolated TPR domain (**Fig. 13**) [25]. The presence of a negatively charged pocket at the TPR domain suggests that the pocket catches positively charged substrates. In general, TPR domains are exploited in protein-protein interactions [29] and the pocket might be the interaction site. Other than that, there is a negatively charged ditch that protrudes from the pocket to the active site at the protease domain. This suggests that the two domains are dependent on each other: any interaction of ligands or substrates at the negatively charged pocket at the TPR domain might affect the conformation of the protease domain since a part of the substrate would also bind to the ditch at the protease domain.

In the crystal structure by Bryant *et al*., the active site at the protease domain was also covered by a helix, similar to the current study (**Fig. 18**) [24]. In the current structure and Bryant's BepA structure, same helix, α9 in the naming of current study, that appears to cover the active site (**Fig. 18**). H246 residue is also coordinating the zinc atom in the current crystal structure. α9 helix might be covering the active site during BepA's inactive state as seen in the both full-length BepA structures. α9 helix would then dislocate in response to specific signals which will also dislocate

α9-α10 loop which contains the H246 residue. This residue might have an important role during the dislocation of α9 from the active site. It was shown that when BepA's structure was superimposed with nuclear membrane zinc metalloprotease ZMPSTE24 structure from human with a trapped substrate (PDB ID: 2YPT), residue H246 on the active-site plug of BepA clashed with the substrate in the ZMPSTE24 structure [35]. This suggests that the dislocation of α9 helix would dislocate the positioning of residue H246 in the α9-α10 loop, giving space for the attachment of a substrate at the active site. Other than that, H246 residue was also shown to be conserved by sequence alignment **(Fig. 5)**. Hence, residue H246 might be crucial. In a very recent study conducted by Daimon *et al*., when H246 was mutated to alanine (H246A), an elevated normal pathway of proteolytic activity of LptD was observed [34]. The expression of BepA(Δα9) also yielded the same observation [34]. This was probably caused by the inability or the absence of the α9 plug to dislocate to the active site to inhibit the proteolytic activity of the protease domain of BepA. From this observation, the α9 was proposed to be as an ON/OFF switch which regulates the protease activity of BepA [36]. Similar observation was also seen in mutational analyses performed by Bryant *et al*., in which H246N BepA mutation showed increased sensitivity to vancomycin, even though the effect was not as severe as in E137Q BepA mutation [24]. The study also showed that H246N mutation caused increased deregulated proteolytic activity of BepA protein [24]. These results from Daimon *et al*. and Bryant *et al*., together with the observation of the positioning of the H246 residue in both of the crystal structures, suggest that H246 may play an important role in the regulation of the proteolytic activity of BepA.

In the TPR domain, residues R466 and R470 were shown to be conserved [24]. These residues might be important in substrate recognition due to their positions in the cavity and high level of

conservation. Even though these residues have high conservation, their importance will require further studies.

When the substrate detaches, α9-α10 loop would then dislocate which will move α9 helix to cover the active site back again, forming an auto-inhibition mechanism. Therefore, the residues involved in the catalytic activity and the residues of the α9 helix might play crucial roles in keeping the functionality of the protease domain. Mature LptD ($LptD^{NC}$) was not susceptible to the mutation of H246 and the expression of BepA(Δα9) [36]. This suggests that the chaperone activity of BepA is not affected by the Δα9 or H246A mutant.

The flexibility of α9 plug is also probably important. If the proteolytic activity is not controlled by the auto-inhibitory mechanism proposed earlier, that is by dislocating α9 plug back to the active site after the protease domain has performed its function, the integrity of BepA is at stake. In order to test this hypothesis, Bryant *et al*. locked the active site plug by introducing a disulfide bond between a residue from the plug; E241 and E103 residue from β1-β3 loop. This made the plug immobile, always covering the active site. Then the performance of these mutants were tested. It was shown that double mutations of E103C and E241C had an impact on the sensitivity of cells with the mutants against vancomycin [24]. Cells with the mutants had lost their resistance against antibiotics, suggesting that the flexibility of the active site plug is important.

A parallel observation was also observed in another study, in which the immobilization of BepA's α9/H246 loop by disulfide cross-linking was tested to investigate whether it would affect the activity of the protease domain [34]. The residues which are in close proximity around the loop were mutated to cystein residues (E241 of the α9 helix and E103 of the β1-β2 loop) [34]. The immobilization of the α9 plug was shown to strongly inhibit the protease activity of BepA to degrade LptD [34]. This was mostly caused by the covering of the active site of the protease domain

by the α9 plug that inhibited the proteolytic activity of BepA. Therefore, these findings suggest that the α9 plug has a very crucial role in the autoinhibition of the protease activity of BepA that is flexible and reversible.

From the current structure, the disordered α6-α7 loop also seems to cover the active site (**Fig. 10**). α6 helix was not modelled in the structure by Bryant *et al*., due to poor electron density. Therefore, the disordered α6-α7 loop could not be seen in the other crystal structure. The disorder was probably due to the flexibility of the α6-α7 loop. In my crystal structure of BepA, α6 helix and α6-α7 loop is embedding the active site, suggesting that the crystallized BepA was probably in its fully inactive form. When α9 helix dislocates, α6-α7 loop might also dislocate to expose the active site. This might correspond to fully-active form of BepA.

When the residues contributing to the formation of the negatively charged pocket at the TPR domain were mutated, BepA's resistance to vancomycin decreased [24]. This probably affected the formation of the pocket, thus affecting substrates attachment to BepA. Whether or not BepA with mutations of the residues at the pocket would still perform its chaperone functions remains to be a question.

In this study, when BepA mutants were performing the degradation of BamA, probably it only involved the dislocation of the loop or the active site plug which exposed the active site of the protease domain. Significant flexibility of the domains is probably not necessary, since BepA mutants were still able to retain chaperone and protease activities. Taken together, the functional analyses suggested that crosslinked BepA mutants were still able to perform the two basic functions as a chaperone and a protease[23] (**Fig. A2**, **Fig. A3**).

BepA was proposed to interact with the BAM complex through the TPR domain [25]. In the previous model, the interaction sites between BepA's TPR domain and the BAM complex, which

were determined by photocrosslinking experiments, were labelled (**Fig. 7**) [25]. BepA's TPR domain was shown to interact with BamA, BamC and BamD of the BAM complex [25]. It was assumed that the small palm of the TPR domain interacts with the BAM complex by inserting into the interior of the ring-like structure (**Fig. 7**) [25]. After the TPR domain was manually docked into the crystal structure of the BAM complex, the interaction points between the TPR domain and the BAM complex matched with the crosslinking results of the TPR domain with the individual Bam components [25]. However, residue N323 of the TPR domain was positioned far from BamD even though cross-linking analysis showed an interaction between the two [25]. This suggests there might be conformational changes when BepA's TPR domain interacts and inserts into the BAM complex.

In addition, BepA's TPR domain was also shown to be the interaction site for membrane proteins such as LptD [25]. Previously, it was shown that F404 residue in the TPR domain of BepA can be cross-linked either to LptD or BamA, and mutations of the residue affect the functional performance of BepA [25]. This suggested that the interaction with LptD and BamA mediated by F404 residue of the TPR domain was important for the BepA functions [25]. As mentioned earlier, N323 residue of the TPR domain of BepA was shown to crosslink with BamD [25]. When N323 residue was mapped to the current crystal structure and surface representation according to electrical charges was calculated, N323 residue of the TPR domain was positioned near to the negatively charged pocket (**Fig. 20**). Thus, there is a possibility that BamD interacts with the TPR domain of BepA by binding to the pocket.

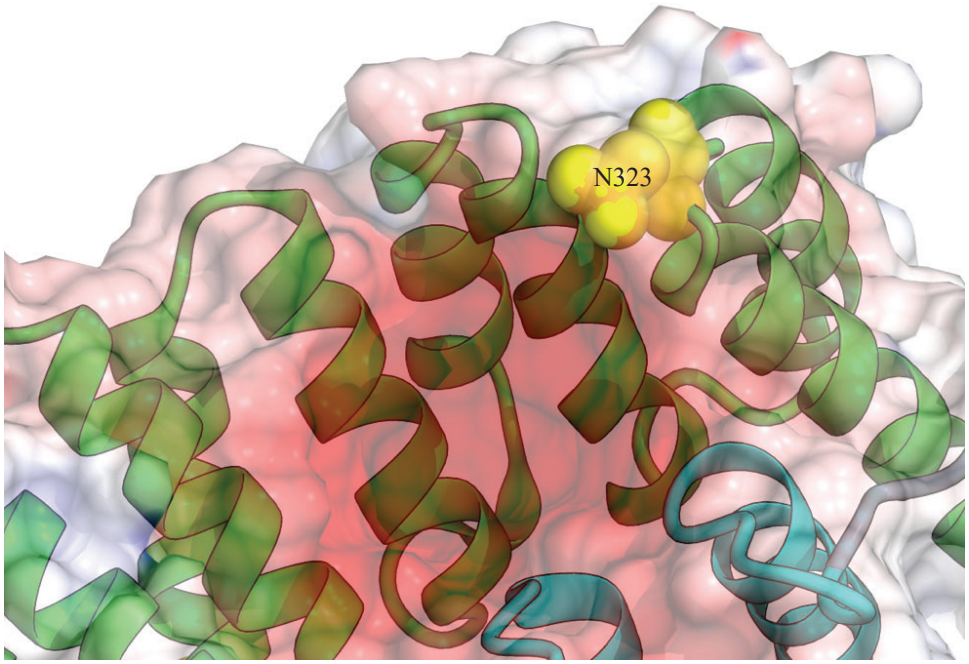

Figure 20. N323 residue positioned near the hydrophilic pocket of TPR domain.

N323 residue (yellow) was previously proposed to be able to crosslink with BamD. The residue was mapped to the current crystal structure of BepAwith the electrical surface representation calculated. Surfaces are colored based on electrostatic potential. Blue (+10 kT/e) to red (−10 kT/e).

Based on the previous TPR domain-BAM complex model, a new model of BepA-BAM complex structure was constructed (**Fig. 21**). I manually docked the crystal structure of BepA from this study to that of the previous model, and the proposed interaction sites between the TPR domain and the BAM complex are positioned similarly in the new model as well. The protease domain faces the periplasmic side, without interacting with the BAM complex. Since the protease domain of BepA seems to be left out from the interactions between the BAM complex and BepA's TPR domain, protease domain is thought to be able to perform independent role to attach to different substrates from the TPR domain, probably in proteins degradation. In the new model, the position

of N323 of the TPR domain of BepA is also distant from BamD, similar to the observation seen in the previous model (**Fig. 21**) [25]. Even though N323 of the TPR domain is distant from BamD, photocrosslinking experiments suggested that N323 residue could indeed crosslink with BamD [25]. This suggests the flexibility of BepA or the BAM complex during the docking process which require further investigations. With this new model of BepA-BAM complex, substrates bound to the negatively charged pocket and ditch can be handed over to BAM complex (Fig. 20), presenting a working model of BepA-BAM complex at the OM.

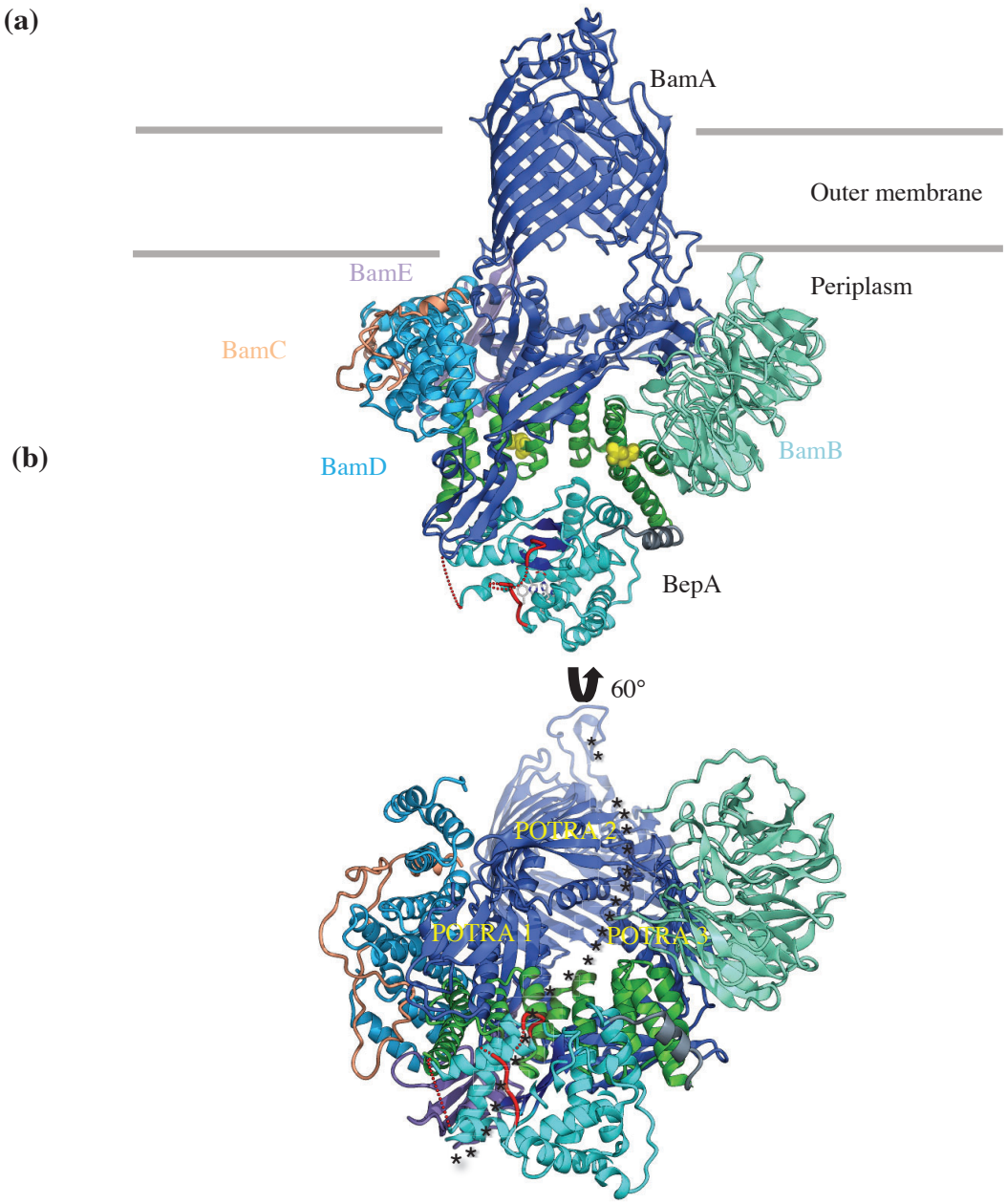

Figure 21. Working model of BepA-BAM complex.

(a) BepA's TPR domain interacts with the BAM complex. The protease domain is facing to the periplasmic side away from the outer membrane. N323 (yellow, right) and F404 (yellow, left) residues which were shown to crosslink with BamD and BamA, LptD, respectively, were mapped on the crystal structure. N323 residue is distant from BamD despite photocrosslinking experiments linking them two, suggesting flexibility of either BepA or the BAM complex during docking. **(b)** Proposed substrate interaction site is shown, attaching at the ditch and pocket at the TPR domain and to the POTRA domain of BamA.

It is proposed that during substrate binding, the substrate first gets captured at negatively charged pocket at the TPR domain, which acts as a protein interactor. When the substrate is stalled at the cavity, it would be possible that the stalling of the substrate would lead to conformational changes that would dislocate α6–α7 loop and the α9 helix from the active site due to some unknown signal. The substrate is thought to be then passed on to the active site through the negatively charged ditch that connects the pocket to the active site. The α6–α7 loop and the α9 helix dislocations will then expose the HEXXH motif at the protease domain. BepA then performs its function as a protease to cleave the substrate, thus releasing the substrate from the pocket and the ditch. This will stimulate the dislocation the α6–α7 loop and α9 helix are thought to re-cover the active site, initiating auto-inhibition of the protease domain. Based on this hypothetical mechanism, a working model of BepA is proposed (**Fig. 22**).

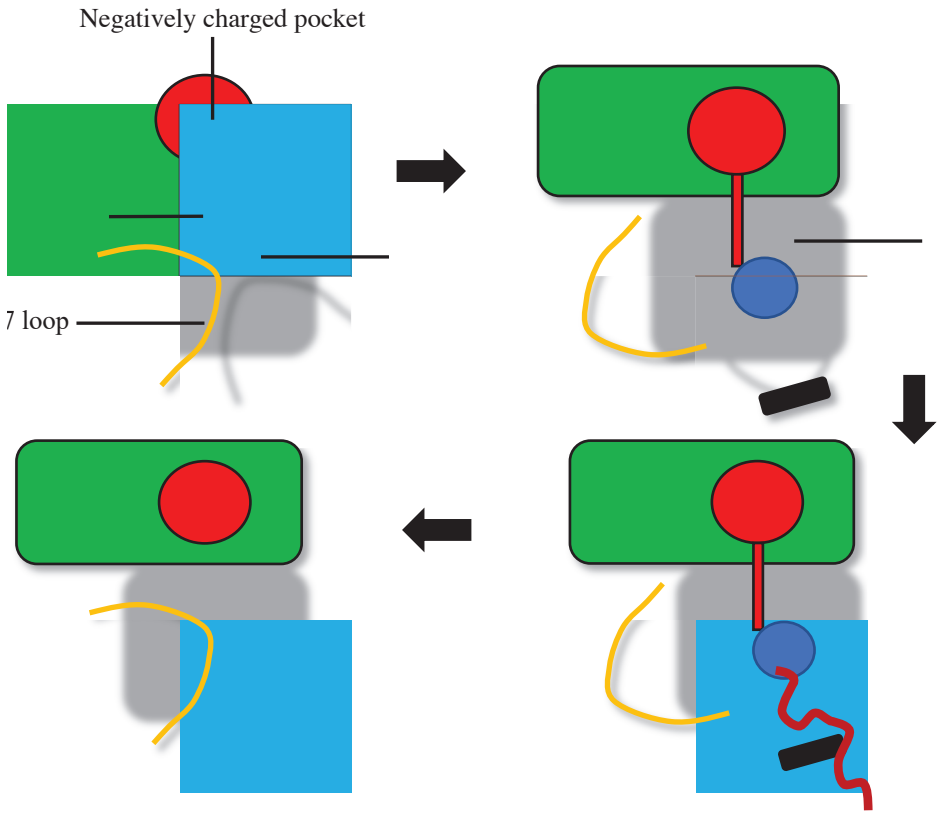

Figure 22. Working model of BepA.

(a) BepA in its inactive form. The α6–α7 loop and α9 helix cover the active site of the protease domain. (b) When a substrate binds to the pocket and the ditch, this would trigger the loop and α9 helix to dislocate, thus exposing the active site. (c) The exposed active site enables the protease domain to cleave the substrate. (d) When the cleaved substrate is released from the pocket and the ditch, the α6–α7 loop and α9 helix might re-cover the active site.

In conclusion, this study gives insights about the domain arrangement, interaction between the two domains of BepA and interaction mechanism of BepA with substrates. By solving the crystal structure of full-length BepA, the presence of a negatively charged pocket and a ditch gives some hints about the properties of BepA, particularly regarding the interaction of BepA with positively charged substrates. BepA is a unique chaperone that depends on the regulation between the two domains for its functions. The TPR domain might offer as a site for substrate attachment

to be inserted into the outer membrane. When there are misfolded proteins that could not be inserted into the outer membrane, the misfolded protein would stall at the binding site and this would trigger the protease domain to trigger its function. This regulation offers a balance that maintain the integrity of the outer membrane by degradation of misfolded proteins that offer no functions other than being a liability to the membrane. This, indirectly, will protect the cells from chemical attack due to the integrity of the outer membrane.

However, to fully understand the mechanisms of BepA, this would need more experiments such as isothermal titration calorimetry (ITC) experiments or co-crystallization of BepA with positively charged substrates to gain insight into substrate binding scheme. Co-crystallization with positively-charged substrate might also reveal the function of the negatively charged pocket of TPR domain located at the opposite side of pocket. Residues F404 and N323 was shown to be important for BepA's interaction with substrates, especially N323 residue since it is positioned conveniently near the negatively charged pocket. Future studies focusing on further mutational analyses on these residues might offer some insights of BepA's interactions with substrates. Since the residues R466 and R470 in the TPR domain were shown to be conserved, mutational analysis of these residues might offer some insights regarding BepA's interactions with substrates. Other than that, to fully understand BepA's interaction with substrates, BepA structure with deleted α9 helix could be studied, and co-crystallization might be performed with this derivative of BepA structure. In addition, if we can get the crystal structure of full-length BepA construct lacking α9 loop, it would present the activation mechanisms of BepA. Regardless, the full-length crystal structure of BepA provides useful information to elucidate the mechanism of membrane protein translocations in Gram-negative bacteria.

ACKNOWLEDGEMENT

This project is a product of contributions and support from important people. First of all, I would like to express sincere gratitude to my supervisor, Professor Dr. Tomoya Tsukazaki for his patience, guidance and his motivations to inspire me to complete this project. Under his guidance, I have published my first ever research paper in an impact journal. He also has advised a lot regarding my future plans. I also would like to thank Assistant Professor Dr. Yoshiki Tanaka for his advices and suggestions for my experiments, which have been very useful.

I want to also thank my advisors, Professor Dr. Shiro Suetsugu and Professor Hiroshi Takagi for their fruitful comments throughout my study. Their comments have helped me in how to improve my experiments, especially with the trouble-shootings.

I would like to thank Dr. Yasushi Daimon and Prof. Yoshinori Akiyama from Institute for Frontier Life and Medical Sciences, Kyoto University, Prof. Hironari Kamikubo and Assistant Prof. Yugo Hayashi from NAIST, and Prof. Shin-ichiro Narita from the Faculty of Nutritional Sciences, University of Morioka for their collaborations in completing this research. I want to thank Shintaro Nakayama for his guidance for the BepA's project.

I want to acknowledge Shigehiro Iwaki for his role as a tutor when I first arrived at NAIST. Since then, he has become a good friend who talks to me about life in Japan. He has also guided me to do experiments I was not familiar with. Sometimes, he was a partner for me to practice speaking in Japanese.

I would like to acknowledge MEXT scholarship and NAIST for the financial supports and daily life supports, respectively. Last but not least, I would like to thank my wife for her support while I was completing my study.

BIBLIOGRAPHY

1. 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).
2. Miller, S. I. & Salama, N. R. The gram-negative bacterial periplasm: Size matters. *PLoS Biol.* **16,** e2004935 (2018).
3. Bos, M. P., Robert, V. & Tommassen, J. Biogenesis of the gram-negative bacterial outer membrane. *Annu. Rev. Microbiol.* **61,** 191–214 (2007).
4. Yamashita, S. & Buchanan, S. K. Solute and ion transport: outer membrane pores and receptors. *Ecosal Plus* **4,** (2010).
5. May, K. L. & Grabowicz, M. The bacterial outer membrane is an evolving antibiotic barrier. *Proc. Natl. Acad. Sci. USA* **115,** 8852–8854 (2018).
6. Le Brun, A. P. *et al.* Structural characterization of a model gram-negative bacterial surface using lipopolysaccharides from rough strains of Escherichia coli. *Biomacromolecules* **14,** 2014–2022 (2013).
7. Ma, H. *et al.* Modeling diversity in structures of bacterial outer membrane lipids. *J. Chem. Theory Comput.* **13,** 811–824 (2017).
8. Trent, M. S., Stead, C. M., Tran, A. X. & Hankins, J. V. Diversity of endotoxin and its impact on pathogenesis. *J Endotoxin Res* **12,** 205–223 (2006).
9. Savage, P. B. Multidrug-resistant bacteria: overcoming antibiotic permeability barriers of gram-negative bacteria. *Ann. Med.* **33,** 167–171 (2001).
10. Yamasaki, S., Nagasawa, S., Fukushima, A., Hayashi-Nishino, M. & Nishino, K. Cooperation of the multidrug efflux pump and lipopolysaccharides in the intrinsic antibiotic resistance of Salmonella enterica serovar Typhimurium. *J. Antimicrob. Chemother.* **68,** 1066–1070 (2013).
11. Dong, H. *et al.* Structural basis for outer membrane lipopolysaccharide insertion. *Nature* **511,** 52–56 (2014).
12. Li, X., Gu, Y., Dong, H., Wang, W. & Dong, C. Trapped lipopolysaccharide and LptD intermediates reveal lipopolysaccharide translocation steps across the Escherichia coli outer membrane. *Sci. Rep.* **5,** 11883 (2015).
13. May, J. M., Sherman, D. J., Simpson, B. W., Ruiz, N. & Kahne, D. Lipopolysaccharide transport to the cell surface: periplasmic transport and assembly into the outer membrane. *Philos. Trans. R. Soc. Lond. B, Biol. Sci.* **370,** (2015).
14. Qiao, S., Luo, Q., Zhao, Y., Zhang, X. C. & Huang, Y. Structural basis for lipopolysaccharide insertion in the bacterial outer membrane. *Nature* **511,** 108–111 (2014).
15. Ricci, D. P. & Silhavy, T. J. The Bam machine: a molecular cooper. *Biochim. Biophys. Acta* **1818,** 1067–1084 (2012).
16. Albrecht, R. & Zeth, K. Crystallization and preliminary X-ray data collection of the Escherichia coli lipoproteins BamC, BamD and BamE. *Acta Crystallogr. Sect. F, Struct. Biol. Cryst. Commun.* **66,** 1586–1590 (2010).
17. Guérin, J. *et al.* Dynamic interplay of membrane-proximal POTRA domain and conserved loop L6 in Omp85 transporter FhaC. *Mol. Microbiol.* **98,** 490–501 (2015).
18. Hussain, S. & Bernstein, H. D. The Bam complex catalyzes efficient insertion of bacterial outer membrane proteins into membrane vesicles of variable lipid composition. *J. Biol. Chem.* **293,** 2959–2973 (2018).

19. Gu, Y. *et al.* Structural basis of outer membrane protein insertion by the BAM complex. *Nature* **531,** 64–69 (2016).
20. Rollauer, S. E., Sooreshjani, M. A., Noinaj, N. & Buchanan, S. K. Outer membrane protein biogenesis in Gram-negative bacteria. *Philos. Trans. R. Soc. Lond. B, Biol. Sci.* **370,** (2015).
21. Sklar, J. G., Wu, T., Kahne, D. & Silhavy, T. J. Defining the roles of the periplasmic chaperones SurA, Skp, and DegP in Escherichia coli. *Genes Dev.* **21,** 2473–2484 (2007).
22. Narita, S., Masui, C., Suzuki, T., Dohmae, N. & Akiyama, Y. Protease homolog BepA (YfgC) promotes assembly and degradation of β-barrel membrane proteins in Escherichia coli. *Proc. Natl. Acad. Sci. USA* **110,** E3612-21 (2013).
23. Shahrizal, M. *et al.* Structural Basis for the Function of the β-Barrel Assembly-Enhancing Protease BepA. *J. Mol. Biol.* **431,** 625–635 (2019).
24. Bryant, J. A. *et al.* Structure-function characterization of the conserved regulatory mechanism of the Escherichia coli M48-metalloprotease BepA. *J. Bacteriol.* (2020). doi:10.1128/JB.00434-20
25. Daimon, Y. *et al.* The TPR domain of BepA is required for productive interaction with substrate proteins and the β-barrel assembly machinery complex. *Mol. Microbiol.* **106,** 760–776 (2017).
26. Bury-Moné, S. *et al.* Global analysis of extracytoplasmic stress signaling in Escherichia coli. *PLoS Genet.* **5,** e1000651 (2009).
27. Lütticke, C. *et al.* E. coli LoiP (YggG), a metalloprotease hydrolyzing Phe-Phe bonds. *Mol. Biosyst.* **8,** 1775–1782 (2012).
28. Vallee, B. L. & Auld, D. S. Cocatalytic zinc motifs in enzyme catalysis. *Proc. Natl. Acad. Sci. USA* **90,** 2715–2718 (1993).
29. Cortajarena, A. L. & Regan, L. Ligand binding by TPR domains. *Protein Sci.* **15,** 1193–1198 (2006).
30. McCoy, A. J. *et al.* Phaser crystallographic software. *J Appl Crystallogr* **40,** 658–674 (2007).
31. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. Sect. D, Biol. Crystallogr.* **60,** 2126–2132 (2004).
32. Adams, P. D. *et al.* PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. Sect. D, Biol. Crystallogr.* **58,** 1948–1954 (2002).
33. Armougom, F., Moretti, S., Keduas, V. & Notredame, C. The iRMSD: a local measure of sequence alignment accuracy using structural information. *Bioinformatics* **22,** e35-9 (2006).
34. Robinson, P. J., Pringle, M. A., Woolhead, C. A. & Bulleid, N. J. Folding of a single domain protein entering the endoplasmic reticulum precedes disulfide formation. *J. Biol. Chem.* **292,** 6978–6986 (2017).
35. Bryant, J. A. *et al.* Structure-function analysis of the *Escherichia coli* β–barrel assembly enhancing protease BepA suggests a role for a self-inhibitory state. *BioRxiv* (2019). doi:10.1101/689117
36. Daimon, Y. *et al.* Reversible autoinhibitory regulation of Escherichia coli metallopeptidase BepA for selective β-barrel protein degradation. *Proc. Natl. Acad. Sci. USA* **117,** 27989–27996 (2020).