Antibiotic darobactin mimics a β-strand to inhibit outer membrane insertase

Hundeep Kaur[+,1], Roman P. Jakob[+,1], Jan K. Marzinek[2], Robert Green[3], Yu Imai[3], Jani Reddy Bolla[4], Elia Agustoni[1], Carol V. Robinson[4,5], Peter J. Bond[2,6], Kim Lewis*[,3], Timm Maier*[,1], Sebastian Hiller*[,1]

[1]Biozentrum, University of Basel, Switzerland

[2]Bioinformatics Institute (A*STAR), 30 Biopolis Street, #07-01 Matrix, Singapore 138671, Singapore

[3]Antimicrobial Discovery Center, Department of Biology, Northeastern University, Boston MA, USA

[4]Physical and Theoretical Chemistry Laboratory, University of Oxford, South Parks Road, Oxford OX1 3QZ, United Kingdom

[5]The Kavli Institute for Nanoscience Discovery, Oxford, OX1 3QU

[6]Department of Biological Sciences, National University of Singapore, 14 Science Drive 4, Singapore 117543, Singapore.

[+] Equal contribution

* Correspondence: k.lewis@norteastern.edu; timm.maier@unibas.ch; sebastian.hiller@unibas.ch

Keywords: membrane protein structure, antibiotics, outer membrane protein biogenesis, β-strand augmentation, inhibition, cyclic peptides, molecular dynamics simulations, native mass spectrometry, X-ray crystallography, cryo-electron microscopy

Abstract

Antibiotics with novel modes of action targeting Gram-negative bacteria are needed to resolve the antimicrobial resistance crisis[1-3]. These pathogens are protected by an additional outer membrane, rendering proteins on the cell surface attractive drug targets[4,5]. The natural compound darobactin targets the insertase BamA[6], the central unit of the essential BAM complex, which facilitates folding and insertion of outer membrane proteins[7-13]. BamA lacks a typical catalytic center, and it is not obvious how a small molecule such as darobactin might inhibit its function. Here, we resolve the darobactin mode of action at the atomic level by a combination of cryo-electron microscopy, X-ray crystallography, native mass spectrometry, *in vivo* experiments and molecular dynamics simulations. Two unique cyclizations pre-organize the darobactin peptide in a rigid β-strand conformation. This creates a mimic of the recognition signal of native substrates with a superior ability to bind to the lateral gate of BamA. Upon binding, darobactin replaces a lipid molecule from the lateral gate to use the membrane environment as an extended binding pocket. Because the interaction between darobactin and BamA is largely mediated by backbone contacts, it is particularly robust against potential resistance mutations. Our results identify the lateral gate as a functional hotspot in BamA and open the path for rational design of antibiotics targeting this bacterial Achilles heel.

The BAM complex was purified from *E. coli* outer membranes (OMs), reconstituted in n-dodecyl maltoside (DDM) micelles and incubated with darobactin A (darobactin). The cryo-EM reconstruction at 3.0 Å resolution revealed the position of a bound darobactin molecule (Fig. 1a, Extended Data Fig. 1, Supplementary Table 1). BamA features a lateral gate facing the membrane, formed by strands β1 and β16 through a kink in strand β16 at residue Gly807. Previous work showed that substrate-free BamA exists in two interchanging conformations with the gate either being open or being closed by the β16-strand straightening to zip up against

β1[14-16]. Darobactin binds to the open form of the BamA lateral gate and essentially does not perturb the structure of the BAM complex (Extended Data Fig. 2, Supplementary Table 2). To better resolve the interaction, we crystallized the BamA β-barrel domain (BamA-β) with darobactin in detergent micelles, and determined its structure at 2.3 Å resolution (Fig. 1b, Supplementary Table 3). The position and bound conformation of darobactin in the cryo-EM and crystal structures were essentially identical (Extended Data Fig. 2).

Darobactin A is a 965 Da heptapeptide with the linear sequence W^1-N^2-W^3-S^4-K^5-S^6-F^7 that is cyclized twice. Darobactin binds to the BamA lateral gate in antiparallel β-sheet conformation to the β1-strand via a series of backbone hydrogen bonds (Fig. 1c). These start at Ile430 of BamA pairing to W^1 of darobactin and continue all the way to Gly424 paired to F^7. The C-terminal carboxyl group of darobactin contacts the side chain of Asn422, while the N-terminus forms three backbone hydrogen bonds with Ile430 in strand β1, Gly807 in β16 and Leu780 in turn 7. Additionally, the side chain of N^2 of darobactin forms a backbone hydrogen bond with Lys808 and a side chain hydrogen bond with Asn427 in β1 (Figs. 1c, 2a). The observation that darobactin effectively seals the gate-open state of BamA was also confirmed in aqueous solution by a comparison of 2D [^{15}N,^{1}H]-TROSY NMR spectra of BamA (Extended Data Fig. 3)[6,14,15]. Taken together, these data show that darobactin A is a molecular mimic of a β-strand that seals the open lateral gate of BamA upon binding.

The dominance of backbone contacts in the interaction suggests that the binding affinity will be only weakly sensitive to local changes in the BamA sequence. To test this hypothesis, we carried out an alanine scan of the darobactin binding site (Asn422–Phe428 on strand β1) and determined for each mutant the affinity to darobactin by isothermal titration calorimetry (ITC) (Fig. 2b, Extended Data Fig. 4, Supplementary Table 4). The affinity of darobactin to wild type BamA-β was K_D=0.6 μM, in good agreement with previous measurements[6]. Supporting our hypothesis, the alanine mutants T423A, G424A, S425A and F426A, which do

not feature side chain contacts with darobactin, show a K_D essentially identical to wild type. The mutant N422A was also found to have wild-type affinity to darobactin, suggesting that its side chain contact is not relevant for the interaction. Accordingly, none of these five mutations had a clear effect on the functionality of BamA in a complementation assay in living cells (Extended Data Fig. 3d). In contrast, residue Asn427 forms a relevant side chain interaction with darobactin as evidenced by a 10-fold weaker binding to an N427A mutant. This *E. coli* mutant, however, is not viable and therefore does not provide a way for bacteria to readily acquire antibiotic resistance. These data show that darobactin is largely insensitive to amino acid mutations in its binding site, which is unique for an antibiotic. Accordingly, darobactins are variable in their amino acid composition[6] and the amino acid sequence at the darobactin binding site is not conserved among BamA homologues (Fig. 2c). Altogether, these findings suggest a high robustness of the darobactin antibiotic efficacy against resistance mutations.

We further note that the binding pocket of darobactin is of a unique physico-chemical nature, comprising a lipidic, a peptidic and an aqueous component. In the bound state, ~36% of the darobactin surface is in contact with BamA, while ~26% is in contact with the lipid and ~38% with the aqueous solution (Fig. 2d). The structures of darobactin in solution determined by NMR spectroscopy[6] and bound to BamA are perfectly overlapping for the first five residues (Fig. 2e), demonstrating that the cyclizations render it a pre-ordered β-strand mimetic. In agreement with this, the linear peptide WNWSKSF, for which only one of many conformations will fit into this site, does not bind to BamA[6]. Furthermore, the tryptophan residues of darobactin are positioned at the height of the aromatic head groups, in the region where membrane proteins typically have aromatic girdles[17-19]. Overall, darobactin has thus evolved to perfectly read out the conformation of the peptido-lipidic environment at the BamA lateral gate.

To resolve the functional role of lipids interacting with the BamA–darobactin complex, we performed explicitly solvated all-atom molecular dynamics (MD) simulations run in triplicates of 1 μs each of BamA-β in an *E. coli* OM-like asymmetric lipidic environment. In the absence of darobactin, BamA-β adopted an open-gate conformation as the dominant state (Fig. 3a,b, Extended Data Fig. 5). Strand β16 was highly dynamic across all replicas, as revealed by its time-dependent root mean square deviation (RMSD) with respect to the initial structure. Cardiolipin (CL), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine (PE) or 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoglycerol (PG) lipid molecules spontaneously entered the lateral gate region, forming direct contacts between their acyl tails and residues Ile430 and Leu780. The presence of lipid or detergent molecules in the gate has also been observed in crystal structures of BamA-β stabilized by nanobodies[14] and its homolog TamA[20]. Additionally, interaction of lipopolysaccharide (LPS) lipid head groups with the extracellular loop L6 of BamA was observed in the simulations (Fig. 3b). The loop appears to drag one or two neighboring LPS molecules deeper into the bilayer, and this might represent a lower energy pathway for permeation of darobactin along the BamA surface to its target site.

The MD simulations of the BamA–darobactin complex exhibited a very stable conformation, as indicated by the RMSD of strand β16 (Extended Data Fig. 5). A total of 16 residues of BamA formed contacts with darobactin (Arg421, Thr423–Ile430, Val444, Leu780, Gly781 and Gly807–Trp810, Extended Data Fig. 5b) out of which six residues (Phe426–Ile430, Leu780) maintained the interactions throughout the entire simulation time across all replicas. This implies that the lipid interaction mediated by Ile430 and Leu780 is replaced by the tighter interaction with darobactin (Fig. 3c). Darobactin binding significantly decreased the density of lipid phosphates at the dynamic gate region and loop-L6 in BamA-β. Additionally, the K^5 side chain of bound darobactin, which points away from BamA (Fig. 3d), interacted primarily with the negatively charged phosphate moieties of PG and CL lipids.

We then used native mass spectrometry to validate the interplay of lipid and darobactin interactions on the intact BAM complex. In three separate experiments we incubated 3 µM of the BAM complex with 5 µM of one of the lipids PE, PG, or CL and recorded native mass spectra (Extended Data Fig. 6). Lipid adduct peaks were observed for the entire BAM complex and its subcomplexes. Notably, the negatively charged lipids PG and CL apparently have a higher affinity, confirming the results from the MD simulations and previous observations that the BAM complex analyzed from native membranes has CL-bound peaks[21]. We then incubated these BAM:lipid complexes with 1 µM darobactin. Darobactin binds both to lipid-bound and lipid-free species of BAM. This is in perfect agreement with the observation from the MD simulations that a ternary complex of BAM, darobactin and CL forms, and that darobactin binding to the BAM:CL complex is significantly higher than for BAM:PG or BAM:PE (Fig. 3e). To determine the selectivity for specific lipids, we incubated BAM–darobactin with mixtures of PE + CL or PG + CL. In both cases, darobactin significantly enriched on CL-bound species, with an additional increase observed for 2xCL bound species. Taken together, the combination of MD simulations and native mass spectrometry shows that CL binds as a preferred lipid to both the ligand-free and the darobactin-bound BAM complex and enhances the interaction of BamA with darobactin.

We next sought to elucidate the mode of resistance of three darobactin-resistant mutants that were previously identified[6], by *in silico* mutations within MD simulations and by *in vitro* affinity measurements. Each of the mutants contains a double or a triple mutation (Fig. 4a, Extended Data Fig. 7, Supplementary Table 5). In MD simulations of strain 1, the bulky side chain addition G429V disrupted the interactions between the darobactin N^2 side chain and Lys808, followed by loss of contacts between the BamA C-terminal region of strand β-16 and darobactin (Extended Data Fig. 7a). The single mutation G807V, in turn, led to a 100-fold decreased affinity for darobactin. Correspondingly, each of the single point mutations

increased the MIC by 16-fold and their combination by more than 32-fold. In strain 2, the E435K mutation induced an allosteric effect by interacting with the neighboring acidic residues Asp498 and Asp500, destabilizing hairpin β1/β2 and leading to a ~20% decrease in darobactin contacts compared to wild type (Extended Data Fig. 7b). Accordingly, these mutations increased the MIC by more than 32-fold. For strain 3, the single mutation Q445P led to a kink in strand β2, which resulted in a highly dynamic hairpin β1/β2, leading to a loosening of the interaction with darobactin *in silico* and a complete loss of the affinity *in vitro* (Extended Data Fig. 7c). The same mutation increased the *in vivo* MIC 2-fold, and the additional mutation T434A mutation led to another eight-fold increase in MIC. Importantly, while all of these mutations confer resistance to darobactin in rich medium, all of them come at the cost of compromising *E. coli* virulence[6].

Notably, the data obtained so far does not allow to distinguish whether darobactin kills its target bacteria by acting as an inhibitor of BamA, or by creating a toxic entity in complex with BamA. To distinguish between the two possibilities, we performed experiments with diploid strains carrying both wild-type BamA and a resistant mutant protein (Supplementary Table 6). In several different combinations, the resulting MICs corresponded in each case to those of the resistant strain and not to the lower wild-type MIC, clearly ruling out the scenario of a toxic entity, and demonstrating that darobactin acts as an inhibitor of BamA function.

An important feature of darobactin towards its practical application is its inactivity against gut symbionts including the Gram-negative *Bacteroides*[6,22]. We wondered whether their resistance might be explained by the composition of the darobactin interaction site (Extended Data Fig. 8). The amino acid sequence comparison of BamA from different species indicates that the Gly424 of β1-strand in *E. coli* is conserved in all Gram-negatives except for *Bacteroides*. The single point mutant G424D, which locally converts the *E. coli* sequence to *Bacteriodes*, showed a slightly reduced *in vitro* affinity to darobactin compared to wild type

(Fig. 2b). We then tested chimeric strains where the binding-site sequence of strand β1 (GSFNFG in *E. coli*) was replaced with either that of *Acinetobacter baumannii* (GTTTLA) or *Bacteroides fragilis* (DQVEFS). But even with such drastic changes in the sequence of β1-strand, the *in vitro* affinity to darobactin changed only around 10-fold, similar to the single point mutation N427A, and accordingly, the associated change in MIC was increased by only 2- or 4-fold for *E. coli* with *Bacteroide*s or *A. baumannii* sequences, respectively (Supplementary Table 5). These data indicate that the insensitivity of *Bacteroides* to darobactin cannot be explained solely by the primary sequence of strand β1, but that other factors must also contribute, such as the sequence differences in other regions, a different permeability of the outer membrane, the involvement of other molecules, or a reduced essentiality of the targeted BAM in *Bacteroides*. Similar considerations hold for the so far unexplained resistance of the darobactin producer strain.

The physiological role of the BAM complex is to fold β-barrel outer membrane proteins (OMPs) into the OM[23,24]. Most BAM substrates contain a consensus signal sequence at their C-terminus including a conserved terminal aromatic residue[25,26]. The first step of the insertase-activity mediated by BamA likely involves anchoring of this signal sequence to the lateral gate region, as evidenced by chemical cross-linking[8,27], and structurally trapped folding intermediates[28,29]. In a recently reported structure of a late-stage intermediate[29], a substrate BamA binds with its C-terminal strand β16 to strand β1 in such a way that the C-terminal aromatic residue W810 ends up in the same position where the C-terminal Phe of darobactin binds (Fig. 4b, Extended Data Fig. 9a–f). Darobactin resembles typical β-signal sequences both by its alternating hydrophobicity and by the presence of a C-terminal phenylalanine. Albeit not being the most prototypic β-signal, the darobactin sequence is within the peptide space spanned by all known β-signals (Extended Data Figure 9 g–j).

Together, the data suggest that darobactin blocks the first step of the insertion reaction catalyzed by BamA (Fig. 4c). While the β-signal sequences of substrates have evolved to attach only transiently to the gate region of BamA, darobactin's preformed β-strand binds with high affinity, presumably locally outcompeting all cognate substrates. We tested this model in competition binding experiments between darobactin and β-signal peptides from four different OMPs, as well as a β-signal consensus sequence. Each of the five peptides interacted only weakly with BamA-β, with dissociation constants at least three orders of magnitude higher than for darobactin (Extended Data Figure 10, Supplementary Table 7). Even if the total affinity of OMP substrates to BAM *in situ* is enhanced by additional interactions, this huge affinity difference represents strong evidence that darobactin will locally outcompete β-signal peptides also in the native context.

Altogether, our findings highlight the lateral gate as a key target site for antibiotics and darobactin as a promising lead compound. The structure of the BAM-darobactin complex also adds information for understanding other reported BamA inhibitors[6,30-32] and provides a path forward for *de novo* design and optimization of compounds to target an essential surface protein of Gram-negative pathogens.

References

1 Brown, E. D. & Wright, G. D. Antibacterial drug discovery in the resistance era. *Nature* **529**, 336-343, (2016).

2 Lewis, K. The science of antibiotic discovery. *Cell* **181**, 29-45, (2020).

3 Tacconelli, E. *et al.* Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. *Lancet Infect Dis* **18**, 318-327, (2018).

4 Epand, R. M., Walker, C., Epand, R. F. & Magarvey, N. A. Molecular mechanisms of membrane targeting antibiotics. *Biochim Biophys Acta* **1858**, 980-987, (2016).

5 Srinivas, N. *et al.* Peptidomimetic antibiotics target outer-membrane biogenesis in *Pseudomonas aeruginosa*. *Science* **327**, 1010-1013, (2010).

6 Imai, Y. *et al.* A new antibiotic selectively kills Gram-negative pathogens. *Nature* **576**, 459-464, (2019).

7 Bakelar, J., Buchanan, S. K. & Noinaj, N. The structure of the β-barrel assembly machinery complex. *Science* **351**, 180-186, (2016).

8 Doyle, M. T. & Bernstein, H. D. Bacterial outer membrane proteins assemble via asymmetric interactions with the BamA β-barrel. *Nat Commun* **10**, 3358, (2019).

9 Gu, Y. *et al.* Structural basis of outer membrane protein insertion by the BAM complex. *Nature* **531**, 64-69, (2016).

10 Iadanza, M. G. *et al.* Lateral opening in the intact β-barrel assembly machinery captured by cryo-EM. *Nat Commun* **7**, 12865, (2016).

11 Konovalova, A., Kahne, D. E. & Silhavy, T. J. Outer membrane biogenesis. *Annu Rev Microbiol* **71**, 539-556, (2017).

12 Lee, J. *et al.* Formation of a beta-barrel membrane protein is catalyzed by the interior surface of the assembly machine protein BamA. *Elife* **8**, (2019).

13 Noinaj, N. *et al.* Structural insight into the biogenesis of β-barrel membrane proteins. *Nature* **501**, 385-390, (2013).

14 Kaur, H. *et al.* Identification of conformation-selective nanobodies against the membrane protein insertase BamA by an integrated structural biology approach. *J Biomol NMR*, (2019).

15 Hartmann, J. B., Zahn, M., Burmann, I. M., Bibow, S. & Hiller, S. Sequence-specific solution NMR assignments of the β-barrel insertase BamA to monitor its conformational ensemble at the atomic level. *J Am Chem Soc* **140**, 11252-11260, (2018).

16 Ni, D. *et al.* Structural and functional analysis of the β-barrel domain of BamA from Escherichia coli. *FASEB J* **28**, 2677-2685, (2014).

17 Hong, H., Park, S., Jimenez, R. H., Rinehart, D. & Tamm, L. K. Role of aromatic side chains in the folding and thermodynamic stability of integral membrane proteins. *J Am Chem Soc* **129**, 8320-8327, (2007).

18 Lee, A. G. Lipid-protein interactions in biological membranes: a structural perspective. *Biochim Biophys Acta* **1612**, 1-40, (2003).

19 Schulz, G. E. The structure of bacterial outer membrane proteins. *Biochim Biophys Acta* **1565**, 308-317, (2002).

20 Gruss, F. *et al.* The structural basis of autotransporter translocation by TamA. *Nat Struct Mol Biol* **20**, 1318-1320, (2013).

21 Chorev, D. S. *et al.* Protein assemblies ejected directly from native membranes yield complexes for mass spectrometry. *Science* **362**, 829-834, (2018).

22 Wexler, H. M. Bacteroides: the good, the bad, and the nitty-gritty. *Clin Microbiol Rev* **20**, 593-621, (2007).

23 Knowles, T. J., Scott-Tucker, A., Overduin, M. & Henderson, I. R. Membrane protein architects: the role of the BAM complex in outer membrane protein assembly. *Nat Rev Microbiol* **7**, 206-214, (2009).
24 Noinaj, N., Rollauer, S. E. & Buchanan, S. K. The β-barrel membrane protein insertase machinery from Gram-negative bacteria. *Curr Opin Struct Biol* **31**, 35-42, (2015).
25 Robert, V. *et al.* Assembly factor Omp85 recognizes its outer membrane protein substrates by a species-specific C-terminal motif. *PLoS Biol* **4**, e377, (2006).
26 Stubenrauch, C. *et al.* Effective assembly of fimbriae in *Escherichia coli* depends on the translocation assembly module nanomachine. *Nat Microbiol* **1**, 16064, (2016).
27 Hohr, A. I. C. *et al.* Membrane protein insertion through a mitochondrial β-barrel gate. *Science* **359**, (2018).
28 Xiao, L. *et al.* Structures of the β-barrel assembly machine recognizing outer membrane protein substrates. *The FASEB Journal* **35**, e21207, (2021).
29 Tomasek, D. *et al.* Structure of a nascent membrane protein as it folds on the BAM complex. *Nature* **583**, 473-478, (2020).
30 Hart, E. M. *et al.* A small-molecule inhibitor of BamA impervious to efflux and the outer membrane permeability barrier. *Proc Natl Acad Sci USA* **116**, 21748-21757, (2019).
31 Luther, A. *et al.* Chimeric peptidomimetic antibiotics against Gram-negative bacteria. *Nature* **576**, 452-458, (2019).
32 Storek, K. M. *et al.* Monoclonal antibody targeting the β-barrel assembly machine of *Escherichia coli* is bactericidal. *Proc Natl Acad Sci USA* **115**, 3692-3697, (2018).

Figure Legends

Figure 1. Structural basis of darobactin function. (a) Three-dimensional structure of the BAM complex in DDM micelles with bound darobactin, resolved by cryo-EM to a resolution of 3.0 Å. The five proteins BamA–BamE are shown in cartoon mode with colors as annotated. Darobactin (blue) is shown in stick representation. (b) Crystal structure of the BamA β-barrel in detergent micelles with bound darobactin, resolved to 2.3 Å resolution. Colors as in **a**. (c) Detailed view of the BamA–darobactin interaction site. Amino acid residues involved in binding are shown in stick representation and direct contacts are shown as dashed lines. Darobactin residues are labeled $W^1N^2W^3S^4K^5S^6F^7$.

Figure 2. The unique proteo-lipidic binding pocket of darobactin. (a) Map of the hydrogen bonds between darobactin and BamA. Canonical β-strand hydrogen-bonds and side chain interactions are shown by red and blue lines, respectively. (b) Affinity of darobactin for different BamA β-barrel variants in LDAO micelles as determined by ITC. Individual measurements are shown as grey dots and the resulting average as a black bar. n=3 independent experiments per condition, n=2 for wild type. (c) BamA β-barrel in sausage representation, with sequence conservation ranging from low (thin, green) to high (thick, red). The binding site of darobactin is indicated. (d) Three-dimensional visualization of the darobactin binding pocket formed by lipids, protein and water, obtained by extracting the most dominant conformation from MD simulations. (e) Superposition of the structures of free darobactin in solution as determined by NMR spectroscopy (beige) and darobactin in its conformation bound to BamA (blue).

Figure 3. Lipid dynamics of BamA–darobactin in the *E. coli* membrane. (a, b) Representative snapshots from MD simulations of the darobactin-free state, illustrating: (a) a

CL molecule anchored to the gate region by Ile430 and Leu780; (b) the interaction of lipid phosphates (orange spheres) with BamA. Two lipid molecules translocated from their canonical positions are highlighted with arrows. (c, d) Representative snapshots from MD simulations of the BamA-darobactin complex, illustrating: (c) darobactin bound tightly to the gate region interacting with Ile430 and Leu780. (d) K^5 of darobactin interacting with the negatively charged CL headgroup. (e) Mass spectra of BAM complex with CL and darobactin. CL, darobactin and their combination adduct peaks are highlighted in the zoom-in section of the charge state 23+. Bar chart of relative darobactin bound peak intensities indicate that darobactin co-bound with CL is observed to a greater extent than bound alone. Bars represent mean ± s.d., points show data from three independent experiments, $^{**}P$=0.0051. $^{*}P < 0.05$; $^{**}P < 0.01$; two-tailed unpaired Student's t-test.

Figure 4. Mechanism of darobactin action. (a) Location of amino acid mutations from three darobactin-resistant strains 1, 2 and 3 shown on the BamA–darobactin structure in yellow, cyan and green, respectively. (b) Surface view of BamA-β highlighting the binding site of darobactin F^7. (c) In its native function, BamA recognizes the consensus β-signal in OMP nascent chains and folds and inserts them into the membrane (left panel). Darobactin binds to the dynamic lateral gate of BamA, blocking the β-signal binding site and thus BamA function (right panel).

Online Methods

Protein expression and purification

The barrel domain of BamA, here referred to as BamA-β, (residues 421–810, C690S, C700S) with a C-terminal hexa-His-tag was overexpressed and purified as described previously[14,33]. Briefly, the protein was overexpressed in inclusion bodies using BL21(DE3) containing the Lemo System. The cells were resuspended in the lysis buffer (50 mM Tris–HCl, 300 mM NaCl, pH 8.0) and lysed by sonication to obtain BamA-β containing inclusion bodies. The inclusion bodies were solubilized in 50 mM Tris–HCl, 300 mM NaCl, 2 (pH 8.0), buffer with 6 M Guanidine-HCl and unfolded BamA-β was purified using Ni-NTA chromatography. The purified protein was precipitated overnight by dialyzing against ultrapure H_2O. The precipitates were pelleted by centrifugation at 30,000 g and resolubilized in the 6M Guanidinium hydrochloride buffer. This highly concentrated and pure unfolded BamA-β was refolded by adding it dropwise in buffer with 0.5% w/v LDAO micelles (50 mM Tris–HCl, 300 mM NaCl, 500 mM Arginine, 0.5% w/v LDAO, pH 8.0). The refolded protein was dialysed against 50 mM Tris, pH 8.0 overnight. Properly folded protein was further obtained using anion ion exchange chromatography using a linear gradient starting from no NaCl to 0.5 M NaCl. Correctly folded protein peak was collected and concentrated to exchange it into the desired buffer using size exclusion chromatography. For crystallization, the protein was expressed in Luria Broth (LB) and purified using the above-mentioned protocol. The buffer composition for size exclusion chromatography (SEC) prior to crystallization was 20 mM Tris pH 7.5, 150 mM NaCl, 0.05% w/v LDAO and 0.35% w/v C_8E_4. The same protocol was used for all the BamA-β barrel mutants.

For the expression of the BAM complex, one plasmid containing all five *E. coli* BamA–E genes was chemically synthesized (GenScript), in which BamE was C-terminally linked to a His_6-tag and BamB C-terminally linked to a StrepII-tag. The BAM complex was expressed in *E. coli*

BL21(DE3) C43 cells in LB medium in the presence of 100 µg/mL ampicillin. Cells were grown at 37°C until an OD_{600} of 0.6, then induced with 0.1 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) and further cultivated over night at 20°C. Cells were resuspended in ice cold Tris-buffered saline (TBS; 50 mM Tris pH 8.0, 300 mM NaCl) homogenized with a douncer, lysed using a microfluidizer and pelleted by ultracentrifugation with a 45 Ti rotor (220,000 g, 2h, 4°C). The BAM complex was extracted from the pellet with TBS containing 3% (v/v) Elugent (Calbiochem). After ultracentrifugation with a 45 Ti rotor (220,000 g, 30 min, 4°C), the supernatant was loaded on a Ni Sepharose FF column, washed with 20CV of TBS containing 0.05% DDM and eluted with TBS containing 0.05% (w/v) DDM and 500 mM imidazole. The eluate was directly loaded on a Strep-Tactin XT column (IBA) and eluted with TBS containing 0.05% DDM and 10 mM biotin. Finally, the BAM complex protein was purified on a 16/600 Superdex 200 column (GE Healthcare, Supplementary Figure 1) using 20 mM Tris pH 8.0, 150 mM NaCl, 0.05% (w/v) n-dodecyl β-D-maltopyranoside (DDM) and concentrated to 5–10 mg/ml and directly flash frozen in liquid nitrogen.

Crystallization, X-ray data collection and structure determination

The BamA-β-darobactin complex was assembled by mixing BamA-β (10 mg/ml) with a 1.5-fold molar excess of darobactin. All crystallization experiments were carried out in sitting-drop vapor diffusion experiments at room temperature mixing equal volumes of protein and reservoir solution. Rod-like crystals (space group I2, a=69.99, b=82.56, c=94.30, α=90°, β=108.64°, γ=90) were grown in 0.06 M Magnesium chloride/Calcium Chloride, 0.1 M imidazole/2-(N-morpholino)ethanesulfonic acid pH 6.5, 12.5% v/v 2-Methyl-2,4-pentanediol, 12.5% w/v Polyethylene glycol 1,000, and 12.5% w/v Polyethylene glycol 3,350 and directly flash frozen in liquid nitrogen. Rhomboid crystals (space group I2, a=79.61, b=79.81, c=89.34, α=90°, β=106.42°, γ=90°) were grown slowly over several weeks in 0.1 M Magnesium

chloride, 0.03 M Tris pH 8.0, 19% w/v Polyethylene glycol 4,000 and were cryopreserved by the addition of 20% ethylene glycol (v/v) and flash-cooled with liquid nitrogen. All data were collected at the SLS beamline X06SA (Swiss Light Source, Paul Scherrer Institute, Switzerland) at 100 K. Date were integrated, indexed with XDS[34] and scaled using aimless[35]. The structures were solved by molecular replacement using the crystal structure of 6QGW.pdb[14] as search model with the program PHASER[36]. Model building was performed with Coot[37], ligand restraints are derived from PRODRG[38] and the structures were refined with PHENIX[39]. MolProbity[40] was used to evaluate the finale model and PyMOL[41] for protein model visualization. Data and Refinement statistics are summarized in Supplementary Table 3.

Electron microscopy sample preparation and data collection

A 4 µl aliquot of BAM complex in 20 mM Tris pH 8.0, 150 mM NaCl, 0.05% (w/v) n-dodecyl β-D-maltopyranoside (DDM) at a concentration of 5.5 mg/ml with a 2-fold molar excess of darobactin was applied to a C-Flat CF-1.2/1.3-2C (Electron Microscopy Sciences) allowed to incubate at room temperature and 100% humidity for 30 s, then manually blotted with a Whatman filter paper. A second 4 µl aliquot of BAM solution was added, and the grid blotted and vitrified by plunging into liquid ethane using a Vitrobot (FEI, Vitrobot Mark IV). Grid quality was analyzed using a Glacios Cryo-TEM cryo-transmission electron microscope, operated at an acceleration voltage of 200 kV. Final data was collected using a Titan Krios electron microscope (FEI) operated at 300 kV, a GIF Quantum LS imaging filter (Gatan) and a K2 Summit (Gatan) operating in counting mode using SerialEM[42]. Images were acquired at 0.8–2.0 µm defocus and a nominal magnification of 165,000×, corresponding to a pixel size of 0.82 Å (Supplementary Table 1). Movies were collected with a total dose of approximately 50 $e^{-}/Å^{2}$ per 9 s exposure, fractionated over 45 frames, using beam-image shift to record 4 images

per hole except for initial test images.

EM data processing and analysis

Micrographs were corrected for beam-induced drift using patch motion and the contrast transfer function (CTF) parameters for each micrograph were determined using Patch CTF in cryoSPARC[43]. 5527 movies were selected after curation for further processing. Early steps of processing up to heterogeneous classification were carried out in two batches based on data collection time. A simplified and summed representation of data processing is provided in Ext. Data Fig. 1. Particles were picked with the blob picker function in cryoSPARC and subjected to reference-free 2D classification. Ab initio reconstruction calculated on a subset of data (and later refinement) revealed agreement with the previously determined BAM complex structure 5D0O (ref. 9). Heterogenous refinement was done in three classes. Heterogeneity between the classes was mainly observed for the POTRA domains. The class with the most structural features was used for non-uniform refinement[43], particle box sizes were optimized by re-extraction and non-uniform-homogeneous refinement, followed by local refinement using an automatically determined global mask, which resulted in a reconstruction at 3.1 Å resolution. Using global and local CTF refinement followed by two additional rounds of non-uniform and local refinements, a final map at a global resolution of 3.0 Å was obtained. Modelling started by manual fitting of subunits of the BamABCDE crystal structure (5D0O.pdb)[9] into the EM density map in UCSF Chimera[44]. Model building was performed in Coot[37] and real-space refinement was carried out in PHENIX[39,45]. Validation was done using the cryo-EM validation tools in PHENIX[39,45]. The map resolution range was determined from local resolution calculation in Cryosparc, and the model resolution range was determined calculating a map at 6 Å resolution in Chimera from the final model. This map was converted into a mask by

thresholding and applied to the local resolution map, to obtain the histogram of local resolution in the model region.

NMR spectroscopy

Protein samples for NMR spectroscopy were expressed in M9 medium, supplemented with perdeuterated water and 1 g/l ^{15}N-ammonium chloride. Samples in 20 mM NaPi pH 7.5, NaCl 150 mM, 0.1% w/v LDAO were concentrated to a concentration of 300 µM. 2D [^{15}N, ^{1}H]-TROSY experiments were measured on a 700 MHz Bruker spectrometer equipped with a cryogenic probe. 128 transients were accumulated at a sample temperature of 37 °C with 2048 and 256 complex points in the ^{1}H and ^{15}N dimensions, respectively. The data were processed using Topspin and analysed using ccpnmr[46]. Chemical shift perturbation of the amide group was calculated as:

$$\Delta\delta(\mathrm{HN}) = \sqrt{\Delta\delta_H^2 + (0.14 \cdot \Delta\delta_N^2)}$$

$\Delta\delta_H$ is the difference in the ^{1}H chemical shift of darobactin bound BamA-β and apo/nanobody bound BamA-β. $\Delta\delta_H$ is the difference in the ^{15}N chemical shift of darobactin bound BamA-β and apo/nanobody bound BamA-β.

Isothermal titration calorimetry

The isothermal titration calorimetry experiments of darobactin binding to BamA-β variants were carried out at 25°C with the Microcal ITC200 instrument in duplicates. BamA-β (100 µM) in 20 mM NaPi, 150 mM NaCl, 0.1% w/v LDAO, pH 7.5 was placed in the sample cell and 1 mM darobactin in the same buffer was placed in the syringe. Successive 20 injections of darobactin with a spacing of 120 s with a stirring rate of 500 rpm were introduced into the protein solution. Reference experiments in the absence of BamA-β were carried out under identical conditions. The resulting data were analyzed and fit to the independent binding model

using AFFINImeter web, with standard errors determined by Jacknife resampling[47]. The direct binding and competition experiments of β-signal peptides were done with 60 µM BamA-β and 500 µM darobactin. The competition data was fit globally with the constrained parameters K_D and $\Delta H^{binding}$ for darobactin, and for each peptide with the free parameters K_D, $\Delta H^{binding}$, darobactin concentration correction and heat of dilution. The 95% confidence interval was calculated by error-surface projection, using the Fisher F-test in the software Sedphat[48,49].

In vivo complementation assay

The *in vivo* functionality of all BamA β-barrel mutants were tested in JCM166 cells. The cells were transformed to contain full-length BamA in a pET3b plasmid. The full-length BamA construct has an N-terminal His10-tag which was mutated to the respective single point mutants. In JCM166 cells, the genomic gene of BamA is under control of an arabinose promotor. In the presence of 0.1% L-(+)-arabinose, the BamA gene is expressed and the cells show a normal phenotype. The absence of arabinose is lethal and cells only survive by transformation with a BamA plasmid which is able to complement native BamA function. pET3b with $BamA^{+9}$ plasmid and pET3b with wild type BamA were used as negative[15] and positive control, respectively. All transformations were streaked out on agar plates without and with 0.1% L-(+)-arabinose and incubated at 37 °C overnight.

Mapping of conserved regions

To map BamA β-barrel sequence conservation 103 full length BamA sequences (Supplementary Table 8) were aligned with Clustal Omega[50]. The alignment was manually truncated to remove POTRA domains. The final alignment was used as input for AL2CO[51] to map conservation onto the *E. coli* BamA β-barrel structure. For better visualization of the data,

the flexible loop of residues 680–697, which is not resolved in the crystal structure, was arbitrarily modeled.

Construction of *bamA* mutants in *E. coli* MG1655

For *E. coli bamA* recombinant mutant 2, *bamA* fragment containing F394V, E435K, G443D mutations from spontaneous-darobactin-resistant *E. coli* strain 2 was amplified by PCR using the primers bamA-rec-F (5'-ACTATCTGGATCGCGGTTATGC-3') and bamA-rec-R (5′-TTCACAGCAGTCTGGATACGAG-3′), and the transformation was performed as previously described[6]. To construct DNA fragments containing the different *bamA* mutations, overlap extension PCR was performed using the common primer pair bamA-rec-F/bamA-rec-R. The different mutations were created using primer pairs mut-R/mut-F (5'-GGTAGCTTCAACTTTGTTATTGGTTACGGTACTG-3'/5'-CAGTACCGTAACCAATA ACAAAGTTGAAGCTACC-3' for G429V, 5'-GTTCCAGTTTAACATCGTTAAAACCTG GTAAGTG-3'/5'-CACTTACCAGGTTTTAACGATGTTAAACTGGAAC-3' for G807V, 5'-GGTACTACTACATTAGCGATTGGTTACGGTACTGAAAG-3'/5'-CGCTAATGTAG TAGTACCGGTGTTGCGCTCTTTTACC-3' for GSFNFG>GTTTLA, 5'-GACCAGGTA GAGTTCTCTATTGGTTACGGTACTGAAAG -3'/5'-AGAGAACTCTACCTGGTCGGTG TTGCGCTCTTTTACC-3' for GSFNFG > DQV EFS). 5'- GGTATTGGTTACGGTGCTGA AAGTGGCGTGAG-3'/5'-CTCACGCCACTTTCAGCACCGTAACCAATACC-3' for T434A, 5'- CTTCCAGGCTGGTGTGCCGCAGGATAACTGGTTAGG-3'/5- CC TAACCA GTTATCCTGCGGCACACCAGCCTGGAAG-3' for Q445P). Genomic DNA from Q445P was used as template to construct the double mutant T434A, Q445P. The triple mutant T434A, Q445P, A705T was described previously[6]. For each mutation, two DNA fragments were amplified using the *E. coli* MG1655 chromosome as template with primers bamA-rec-F/mut-R and mut-F/bamA-rec-R. The resulting DNA fragments were joined together by PCR using

primers bamA-rec-F/bamA-rec-R, and used to transform electrocompetent cells of *E. coli* MG1655-pKD46 for λ red recombination[52]. The transformation protocol was adapted from the "Quick and Easy *E. coli* Gene Deletion Kit" (GeneBridges, Heidelberg), as described previously[6]. The transformant clones were first selected on 2x MIC darobactin (8 μg ml^{-1}). Several clones were then restreaked for resistance to darobactin (8 μg ml^{-1}), as well as sensitivity to ampicillin (100 μg ml^{-1}, at 30 °C). Each mutation in *bamA* was confirmed via sequencing of the *bamA* locus. The MIC of darobactin against the *E. coli* MG1655 and *bamA* mutants was determined by microbroth dilution, as previously described[6].

Construction of the Bam diploid mutants and growth inhibition assay

The plasmids carrying Bam with wild type or mutated BamA (pBAM-WT, no mutation; pBAM-M2, F394V, E435K and G443D; pBAM-M3, T434A, Q445P and A705T) were used for this study[6]. Plasmids were transformed into *E. coli* MG1655, *bamA* recombinant mutant 2 (F394V, E435K and G443D), mutant 3 (T434A, Q445P and A705T) *bamA6* mutant, *bamA66* mutant or *bamA101* mutant by electroporation[53-55]. Strains were inoculated into Mueller-Hinton II broth (MHIIB) with (for mutants with plasmid which carry Bam) or without (parental strain) 50 μg ml^{-1} of ampicillin at 37 °C with aeration at 220 rpm. After overnight cultivation, bacterial culture was diluted 1:100 in MHIIB with or without 50 μg ml^{-1} of ampicillin and 0.1 mM of IPTG and incubated at 37 °C with aeration at 220 rpm. Exponential culture (OD_{600} of 0.2–0.3) was diluted to an OD_{600} of 0.001 (approximately 5×10^5 c.f.u. ml^{-1}) in MHIIB with or without 50 μg ml^{-1} of ampicillin and 0.1 mM of IPTG. 98 μl aliquots were transferred into round-bottom 96 well plates containing 2 μl of darobactin diluted serially twofold. After overnight cultivation at 37 ˚C, the MIC of darobactin was determined.

Native mass spectrometry on the BAM complex

BAM complex was overexpressed from plasmid pJH114 in BL21(DE3) cells (NEB), grown in LB medium containing carbenicillin (100 µg ml^{-1}). When the culture reached an optical density at 600 nm of ~0.5, expression was induced with 0.5mM isopropyl β-thiogalactopyranoside (IPTG) for 3h at 37 °C. Cells were pelleted by centrifugation at 5,000g and resuspended in a buffer containing 50 mM Tris-HCl (pH 8.0), 300 mM NaCl and EDTA-free protease inhibitor cocktail (Roche). The cells were then disrupted by a microfluidizer (Microfluidics). After centrifugation (20,000 g for 20 min), the supernatant was filtered and loaded onto a 5 ml His Trap-HP column in case of soluble proteins while for membrane proteins, the supernatant was ultracentrifuged (200,000 g), and the membrane fractions were collected. The proteins were solubilized from the membrane fraction with 20 mM Tris (pH 8.0), 150 mM NaCl, 20% glycerol, 2% DDM (Anatrace) for 2 h at 4 °C. The insoluble material was removed by ultracentrifugation. The supernatant was filtered before loading onto a 5 ml His Trap-HP column (GE Healthcare) equilibrated in 20 mM Tris (pH 8.0), 150 mM NaCl, 20 mM imidazole, 10% glycerol and 0.03% DDM. After the clarified supernatant was loaded, the column was initially washed with 50 ml of 20 mM Tris (pH 8.0), 150 mM NaCl, 20 mM imidazole, 10% glycerol and 0.03% DDM and washed again with 50 ml of 20 mM Tris (pH 8.0), 150 mM NaCl, 80 mM imidazole, 10% glycerol and 0.03% DDM. The bound protein was then eluted with 20 mM Tris (pH 8.0), 150 mM NaCl, 300 mM imidazole, 10% glycerol and 0.03% DDM. Peak fractions were pooled, concentrated and injected onto a Superdex 200 GL 10/300 (GE Healthcare) column equilibrated in buffer 20 mM Tris (pH 8.0) and 150 mM NaCl, and 0.5% C_8E_4. Peak fractions were concentrated and was either used immediately or flash-frozen in liquid nitrogen and stored at −80 °C.

Prior to MS analysis, the protein was buffer exchanged into 200 mM ammonium acetate pH 8.0 and 0.5% C_8E_4 using a Biospin-6 (BioRad) column and introduced directly into the mass

spectrometer using gold-coated capillary needles (prepared in-house). Data were collected on a Q-Exactive UHMR mass spectrometer (Thermo Fisher Scientific). The instrument parameters were as follows: capillary voltage 1.1 kV, S-lens RF 100%, quadrupole selection from 1,000 to 20,000 m/z range, collisional activation in the HCD cell 300 V, trapping gas pressure setting 7.5, temperature 200 °C, and resolution of the instrument 12,500. The noise level was set at 3 rather than the default value of 4.64. No in-source dissociation was applied. Data were analysed using Xcalibur 4.2 (Thermo Scientific) and UniDec[56] software packages. Lipids and darobactin were diluted into a buffer containing 200 mM ammonium acetate pH 8.0 and 0.5% (w/v) C_8E_4 and were added in different ratios to solutions of Bam complex in the same buffer. All experiments were repeated three times with similar outcome. Statistical analysis was performed using Prism 8 software (GraphPad). Results are mean ± s.d., n represents three independent experiments. Comparisons for two groups were calculated using unpaired two-tailed Student's t-tests. Exact *P* values are indicated in the figures and in figure legends where possible.

Molecular dynamics simulations

The experimentally solved BamA structure (residues 419-810) was assigned with CHARMM36m[57] parameters at neutral pH. The BamA N-terminus was treated in the neutral, deprotonated form, as part of the BAM complex. The C-terminus was negatively charged. The darobactin parameters were assigned using CgenFF[58], also assuming neutral pH. BamA and BamA-darobactin–membrane systems were built using the CHARMM-GUI membrane builder[59]. The orientation of BamA with respect to the membrane was predicted using the OPM server[60]. Each membrane inner leaflet corresponded to standard *E. coli* lipids[61] in a 0.9:0.05:0.5 ratio for palmitoyl-oleyol-phosphatidyl-ethanolamine (PE), palmitoyl-oleyol-phosphatidyl-glycerol (PG), and cardiolipin (CL), respectively. The outer leaflet was made of R1-type core

LPS, the most frequent among *E. coli* strains. Each simulation box had a dimension of ~10x10x10 nm^3 which resulted in 48 LPS molecules in the outer leaflet and 126 PE, 7 PG and 7 CL lipid molecules in the inner leaflet. Subsequently, the box was filled with ~21,000 TIP3P[62] water molecules. The overall charge of each system was neutralized by the addition of ~150 mM magnesium chloride. Energy minimization using steepest descents was performed for ≤5,000 steps with a 0.1 nm step size. Equilibrations in the NVT and NPT ensembles were performed for ~100 ns in total with position restraints on BamA or BamA-darobactin heavy atoms, while no restraints were applied to the membrane or solvent. In total, 5 membrane systems were prepared: i) BamA in the darobactin free (apo) state; ii) BamA-darobactin complex; iii) Strain 1: G429V, G807V mutations; iv) Strain 2: F394, E435K, G443D mutations; and v) Strain 3: T434A, Q445T, A705T mutations. All in silico point mutations were performed using CHARMM-GUI. All unrestrained production simulations were run in the NPT ensemble in triplicate with different starting velocities for 1 μs each, using GROMACS2018[61]. Equations of motion were integrated through the Verlet leapfrog algorithm with a 2 fs time step, and bonds connected to hydrogens were constrained with the LINCS algorithm. The cutoff distance was 1.2 nm for the short-range neighbor list and van der Waal's interactions with a smooth switching function from 1.0 nm. The Particle Mesh Ewald method was applied for long-range electrostatic interactions with a 1.2 nm real space cutoff[63]. The Nose-Hoover thermostat[64,65] and Parinello-Rahman barostat[66] were used to maintain the temperature and pressure at 310 K and 1 bar, respectively. Simulations were performed on: i) an in-house Linux cluster composed of 7 nodes containing 2 GPUs (Nvidia GeForce RTX 2080 Ti) and 24 CPUs (Intel® Xeon® Gold 5118 CPU @ 2.3 GHz) each; and ii) ASPIRE 1, the petascale cluster at the Singapore National Supercomputing Centre, where each simulation employed 4 nodes each consisting of 1 GPU (Nvidia Tesla K40t) and 24 CPUs (Intel® Xeon® CPU E5-2690 v3 @ 2.6 GHz). Partial densities were calculated using *g_mydensity* over the

entire 3 μs sampling per system[67]. The number of contacts was calculated based on a 0.3 nm cutoff distance for the entire 3 μs of sampling per system. The most dominant conformations were clustered on protein heavy atoms using the Gromos method with a root mean square deviation (RMSD) cutoff distance of 0.35 nm. The percentage of the darobactin molecule exposed to either water, lipids, or protein was calculated based on its solvent accessible surface and averaged over the entire trajectory.

Analysis of β-signal sequences

To identify a set of β-signals we followed a modified version of the procedure established by Paramsivam et al., 2012[68]. A total of ~860,000 sequences of outer membrane β-barrel proteins were obtained from the OMPdb (database release Dec 1, 2020)[69] The prokaryotic sequences were retained and filtered for non-redundancy towards a pairwise identity < 90%. Out of these, only sequences associated with a topology prediction with reliability larger than 80% were kept. For each remaining sequence, the last β-strand in the topology prediction was selected and subjected to the following rules: (1) If its C-terminus corresponded to the protein C-terminus, the 10 C-terminal residues were kept. (2) Otherwise, the sequence was extended towards the protein C-terminus until an aromatic residue was found, but by maximally 4 residues. The 10 C-terminal residues were kept. (3) If no aromatic residue was found, the last β-strand was extended to the protein C-terminus and the first 10-residue long sequence portion that matches the regular expression of Paramsivam et al was selected. Sequences were removed if no matches to the regular expression were found. The final set of "β-signals" contained 263,003 sequences.

To compare the darobactin sequence to the set of β-signals by amino acid type, a position-specific frequency matrix (PSFM) and a position-specific scoring matrix (PSSM) were generated from the β-signals set. Then, the PSSM was used to calculate the log-likelihood score

for each sequence in: (1) the β-signals set ("β-signal sequences", 263,003 entries); (2) an equally large set of 7-residue long, randomly generated amino acid sequences ("random sequences", 263,003 entries); (3) a set of unique permutations of the darobactin sequence ("darobactin sequence permutations", 1,260 entries). The score of darobactin was compared to scores of sets 1–3 using percentiles. The same type of analysis was also done by amino acid chemistry, a position-specific count matrix (PSCM) was generated from the β-signals according to their chemistry: STNQY, polar non-charged; DEKRH, charged; G, Neutral; AVCPLIMWF, Hydrophobic and non-polar; FWY, aromatic. Then, new PSFM and PSSM were generated from the PSCM. Any amino acid sequence was converted to the best scoring chemistry sequence in accordance with the new PSSM. Figure panels were created with matplotlib and logomaker[70,71]. The code was deposited on github (https://github.com/hiller-lab/kaur-jakob-2021).

References

33 Kaur, H., Grahl, A., Hartmann, J. B. & Hiller, S. Sample preparation and technical setup for NMR spectroscopy with integral membrane proteins. *Methods Mol Biol* **2127**, 373-396, (2020).

34 Kabsch, W. Xds. *Acta Crystallogr D Biol Crystallogr* **66**, 125-132, (2010).

35 Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr D Biol Crystallogr* **69**, 1204-1214, (2013).

36 McCoy, A. J. *et al.* Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674, (2007).

37 Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132, (2004).

38 Schuttelkopf, A. W. & van Aalten, D. M. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr D Biol Crystallogr* **60**, 1355-1363, (2004).

39 Adams, P. D. *et al.* PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr D Biol Crystallogr* **58**, 1948-1954, (2002).

40 Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12-21, (2010).

41 DeLano, W. L. J. C. N. O. P. C. Pymol: An open-source molecular graphics tool. **40**, 82-92, (2002).

42 Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *J Struct Biol* **152**, 36-51, (2005).

43 Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* **14**, 290-296, (2017).

44 Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-1612, (2004).

45 Afonine, P. V. *et al.* Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr D Struct Biol* **74**, 531-544, (2018).

46 Vranken, W. F. *et al.* The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687-696, (2005).

47 Pineiro, A. *et al.* AFFINImeter: A software to analyze molecular recognition processes from experimental data. *Anal Biochem* **577**, 117-134, (2019).

48 Kemmer, G. & Keller, S. Nonlinear least-squares data fitting in Excel spreadsheets. *Nat Protoc* **5**, 267-281, (2010).

49 Zhao, H., Piszczek, G. & Schuck, P. SEDPHAT--a platform for global ITC analysis and global multi-method analysis of molecular interactions. *Methods* **76**, 137-148, (2015).

50 Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**, 539, (2011).

51 Pei, J. & Grishin, N. V. AL2CO: calculation of positional conservation in a protein sequence alignment. *Bioinformatics* **17**, 700-712, (2001).

52 Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci USA* **97**, 6640-6645, (2000).

53 Bennion, D., Charlson, E. S., Coon, E. & Misra, R. Dissection of β-barrel outer membrane protein assembly pathways through characterizing BamA POTRA 1 mutants of Escherichia coli. *Mol Microbiol* **77**, 1153-1171, (2010).

54 Ruiz, N., Wu, T., Kahne, D. & Silhavy, T. J. Probing the barrier function of the outer membrane with chemical conditionality. *ACS Chem Biol* **1**, 385-395, (2006).
55 Aoki, S. K. *et al.* Contact-dependent growth inhibition requires the essential outer membrane protein BamA (YaeT) as the receptor and the inner membrane transport protein AcrB. *Mol Microbiol* **70**, 323-340, (2008).
56 Marty, M. T. *et al.* Bayesian deconvolution of mass and ion mobility spectra: from binary interactions to polydisperse ensembles. *Anal Chem* **87**, 4370-4376, (2015).
57 Huang, J. *et al.* CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat Methods* **14**, 71-73, (2017).
58 Vanommeslaeghe, K. *et al.* CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J Comput Chem* **31**, 671-690, (2010).
59 Jo, S., Kim, T., Iyer, V. G. & Im, W. CHARMM-GUI: a web-based graphical user interface for CHARMM. *J Comput Chem* **29**, 1859-1865, (2008).
60 Lomize, M. A., Pogozheva, I. D., Joo, H., Mosberg, H. I. & Lomize, A. L. OPM database and PPM web server: resources for positioning of proteins in membranes. *Nucleic Acids Res* **40**, D370-376, (2012).
61 Van Der Spoel, D. *et al.* GROMACS: fast, flexible, and free. *J Comput Chem* **26**, 1701-1718, (2005).
62 Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J Chem Phys* **79**, 926-935, (1983).
63 Essmann, U. *et al.* A smooth particle mesh Ewald method. *J Chem Phys* **103**, 8577-8593, (1995).
64 Hoover, W. G. Canonical dynamics - equilibrium phase-space distributions. *Phys Rev A* **31**, 1695-1697, (1985).
65 Nose, S. A Unified formulation of the constant temperature molecular-dynamics methods. *J Chem Phys* **81**, 511-519, (1984).
66 Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals - a new molecular-dynamics method. *J Appl Phys* **52**, 7182-7190, (1981).
67 Castillo, N., Monticelli, L., Barnoud, J. & Tieleman, D. P. Free energy of WALP23 dimer association in DMPC, DPPC, and DOPC bilayers. *Chem Phys Lipids* **169**, 95-105, (2013).
68 Paramasivam, N., Habeck, M. & Linke, D. Is the C-terminal insertional signal in Gram-negative bacterial outer membrane proteins species-specific or not? *BMC Genomics* **13**, 510, (2012).
69 Tsirigos, K. D., Bagos, P. G. & Hamodrakas, S. J. OMPdb: a database of β-barrel outer membrane proteins from Gram-negative bacteria. *Nucleic Acids Res* **39**, D324-331, (2011).
70 Hunter, J. D. Matplotlib: A 2D graphics environment. *Comput Sci Eng* **9**, 90-95, (2007).
71 Tareen, A. & Kinney, J. B. Logomaker: beautiful sequence logos in Python. *Bioinformatics* **36**, 2272-2274, (2020).

Data Availability

The data that support the findings of this study are available from the corresponding author upon request. The atomic coordinates have been deposited in the RCSB Protein Data Bank and are available under the accession codes 7NRE and 7NRF. The cryo-EM map has been deposited in the Protein Data Bank under accession code 7NRI and EMDB accession code 12546. Mass spectrometry data has been deposited in figshare with DOI: 10.6084/m9.figshare.12179784. For the study, data was retrieved from the OMPdb (release Dec 1, 2020). The codes developed for β-signal analysis are deposited at https://github.com/hiller-lab/kaur-jakob-2021.

Acknowledgements

We thank Harris Bernstein and Christoph Bieniossek for expression plasmids, Tim Sharpe, Thomas Müntener and the Biozentrum Bio-EM lab for scientific support and discussions. The scientific computing center at University of Basel (sciCORE) and the National Supercomputing Centre Singapore (NSCC) are acknowledged for providing computational resources. This work was supported by the Swiss National Science Foundation (grants 167125, 185388 and 187170 to S.H. and R'Equip 177084 to T.M.), AntiResist: New approaches to combat antibiotic-resistant bacteria (51AU40_180541), the Medical Research Council (program grant MR/N020413/1 awarded to C.V.R), the Bioinformatics Institute (BII) A*STAR and NRF (NRF2017NRF-CRP001-027 to P.J.B. and J.K.M.), and by the NIH (grant P01 AI118687 to K.L.).

Author contributions

CVR, KL, TM and SH designed the study and supervised experiments. RG and YI performed the microbiology experiments. JRB performed the mass-spec experiments. HK and RPJ

performed all other experiments. JKM and PJB ran simulations. EA performed sequence analysis. All authors analyzed data and discussed the findings. HK, KL, TM and SH wrote the manuscript. All authors edited and approved the manuscript.

Declaration of interest

The authors declare that they do not have competing financial interests.

Supplementary Information is available for this paper.

Extended Data Figure Legends

Extended Data Fig. 1. Cryo-electron microscopy structure of the BAM–darobactin complex. (a) Flow chart of data processing to generate the structure, see Methods for details. (b) Purified BAM–darobactin sample used for cryo-EM structure determination analyzed on SDS-PAGE. This experiment was repeated at least three times independently with similar results. (c) Representative electron micrograph of BAM–darobactin. This experiment was repeated at least three times independently with similar results. (d) Selected examples of 2D classes from cryoSPARC. (e) Viewing direction distribution plot for the final three-dimensional reconstruction (f) FSC curves for unmasked, spherically, loose and tight masks and corrected FSC curve for the final reconstruction, yielding a GSFSC resolution of 3.03 Å. (g) Local resolution variations of the EM reconstruction, with coloring as indicated. POTRA domains 1 and 2 are at a local resolution below 4.5 Å and are only visualized at lower contour level where micelle density obscures the view onto the BamA barrel. (h) Plot of directional FSC (red) and histogram of per angle FSC (blue) and ± 1 S.D from mean of the directional FSC. FSC curve indicates a resolution of 3.15 Å. (i) Overview of the cryo-EM reconstruction of the BAM complex. BAM is shown in ribbon representation, the coulomb potential map as blue mesh. Note that the density of POTRA domains P1 and P2 is below the display threshold chosen here due to motional averaging. (j–n) Zoomed local views, showing the map around selected atoms in stick representation from the directions and viewpoints indicated by arrows and letters in i.

Extended Data Fig. 2. Structural details of the cryo-EM and crystal structures of BAM. (a) Superposition of the BAM–darobactin cryo-EM structure (salmon) with the ligand-free BAM crystal structure (green, PDB 5D0O). (b) Superposition of POTRA domains P1–P5 and

the individual components BamB–BamE, as indicated. The dashed horizontal line indicates the pivot between P2 and P3 around which P1 and P2 are rotated by rigid-body movement. (c) Superposition of the BamA-β barrel–darobactin crystal structure (salmon) with a closed gate BamA-β barrel crystal structure (cyan, PDB 4N75). (d) Crystallographic omit map for darobactin bound at the lateral gate region of the BamA β-barrel after refinement of the model without darobactin. The 2mFo-DFc map is shown at 1σ in slate and the mFo –DFc difference map at +/-3σ level in green and red. Top: Overview of an entire BamA barrel, bottom left and right: zoom in without and with overlay of the refined model coordinates. The cyclizations of darobactin can clearly be observed at 2.3 Å resolution. (e) Omit map for strands β1 (top) and β16 (bottom) of the BamA β-barrel visualized as in panel d. (f) Superposition of the cryo-EM and X-ray structures. X-ray structure is colored salmon/blue for BamA/darobactin. Cryo-EM structure bordeaux/white. (g) Same for the ligand darobactin only.

Extended Data Fig. 3. Comparison of BamA β-barrel conformations in aqueous solution. (a) Comparison of 2D [^{15}N,^{1}H]-TROSY fingerprint spectra of different BamA preparations in LDAO. Left: Overlay of BamA-β fingerprint spectra in the absence and presence of darobactin. Middle: Overlay of fingerprint spectra of BamA-β–nanobodyF7 and BamA-β–darobactin. Right: Overlay of fingerprint spectrum of BamA^{+9} and BamA-β-darobactin complex. Bottom panels show zoomed regions of the spectra. (b) Backbone amide chemical shift perturbations between the fingerprint spectrum of BamA-β with and without darobactin (left, black), Bam-β–nanobodyF7 in comparison with BamA–darobactin (middle, blue) and BamA^{+9} in comparison with BamA-darobactin (right, purple). The dotted lines indicate the average CSP, which can be interpreted as a measure of dissimilarity between two spectra. (c) Structures of BamA-β barrel in various conformations of the gate region. Bottom panel shows zoomed in part of the backbone with H-bonds between β1 and β16 or darobactin indicated. From left

to right: open gate (6QGW, red), closed gate (6QGX, blue) and BamA^{+9} (6FSU, purple) and BamA-darobactin complex. (d) *In vivo* functional assay of BamA barrel mutants and C-terminal extensions using JCM166 cells in the absence and presence of arabinose. fl-BamA^{MENVALDFS} and fl-BamA serve as a negative and positive control, respectively.

Extended Data Fig. 4. Isothermal titration calorimetry (ITC) of BamA-β barrel in detergent micelles and its variants titrated with darobactin. Experiments were repeated independently twice with similar results.

Extended Data Fig. 5. MD simulations of BamA-β barrel. (a) Representative snapshots of lipid PE molecule (left) and PG molecule (right) anchored by Ile430 and Leu780 in the gate region. (b) The most dominant conformation of BamA–darobactin complex showing contacts consistently observed between BamA-β barrel and darobactin throughout the simulation sampling. (c) Partial densities of all lipids (top-down view of the membrane); the white arrow highlights the darobactin-binding region. (d) Partial densities of lipid phosphate groups (side-view of the membrane). (e–g) Structural drift and fluctuations of key β-strands around the darobactin binding site. Time-dependent root mean square deviation (RMSD) measured with respect to the initial structure for backbone atoms of β-strands, after performing a least-squares fit. The resulting RMSD is shown for: (e) β16 in ligand-free BamA, (f) β16 in BamA–darobactin complex, (g) hairpin β1/β2 in BamA–darobactin complex.

Extended Data Fig. 6. Interaction of BAM complex with lipids in the absence and presence of bound darobactin. (a) Mass spectra of the BAM complex with different lipids, bottom spectra with PE, middle spectra with PG and top spectra with CL. (b) Deconvolution of the mass spectra in (a), indicates that all the subcomplexes have lipids bound. (c) Relative

intensities of lipid binding peaks from (a) suggest that the negatively charged PG and CL have higher affinity. (d) Mass spectra of BAM complex with lipids and darobactin. Bottom spectra with PE and darobactin, and top spectra with PG and darobactin. Zoom in a section of the 23+ charge state highlights the bound peaks and relative ratio of darobactin binding is shown in bar charts. No significant increase in darobactin binding is observed in these two cases, suggesting that PE and PG lipids do not affect darobactin binding. (e) Mass spectra of BAM complex with lipid mixtures (PE+CL, bottom spectra and PG+CL, top spectra) and darobactin. Lipids, darobactin and their various combination binding peaks are highlighted in the zoom-in section for the charge state 23+. Bar charts of relative peak intensities indicate that darobactin bound with CL is observed to a greater extent than bound alone and PE and PG bound in both cases. This increase is even higher for 2xCL bound species and is slightly lower for PE or PG bound to 1xCL species. However, no change in darobactin binding is observed for PE and PG. The mean relative binding intensities are significantly different in the different lipid-bound forms and are highlighted in bar charts. Bars (**c-e**) represent mean ± s.d., points show data from three independent experiments. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; NS, not significant; two-tailed unpaired Student's t-test. Exact *P* values are indicated in the figures.

Extended Data Fig. 7. MD Simulations of the effect of darobactin-resistance mutations on the BamA–darobactin interaction. (a) Representative snapshots with zoomed views from a simulation with strain 1 bound to darobactin. Mutations G429V and G807V are shown as yellow spheres at the α-carbon position. Time-dependent root mean square deviation (RMSD) of β16 backbone atoms relative to the initial structure. (b) Same for strain 2 (mutations E435K and G443D, cyan) and RMSD of hairpin β1/β2. (c) Same for strain 3 (mutations T434A, Q445P, A705T, green) and RMSD of hairpin β1/β2. In each panel, protein is shown in cartoon

representation, darobactin as sticks (blue = carbons, red = oxygens, white = protons, navy = nitrogens). Hydrogen bonds are shown as black dotted lines.

Extended Data Fig. 8. Relatedness of predicted Gram-negative BamA proteins. (a) Sequence alignment of BamA sequence from Gram-negative bacteria. The region highlighted in yellow is the predicted interaction site of darobactin in sheet β1. Right hand side: This 6 amino acid sequence of *A. baumannii and B. fragilis* was substituted into *E. coli* BamA for *in vivo* assays. (b) Phylogenetic tree of full length BamA sequences from various species of Gram-negative bacteria. Different colours indicate different branches belonging to the species specified next to the branch. Multiple alignments for the tree were carried out using CLUSTAL-W and the phylogenetic tree was derived using SEAVIEW software. (c) Topology plot of BamA from *E. coli* with bound darobactin (blue). For the chimeric mutants the red amino acids were exchanged with the local sequence from either *A. baumannii* or *B. fragilis*.

Extended Data Fig. 9. Comparison of BamA structures involved in molecular interactions and analysis of β-signals. (a) Overlay of the BamA subunit from the BAM–darobactin complex (salmon–blue; this work) with a BamA engaged with a substrate in a late-stage insertion intermediate state (green; PDB 6V05). The substrate has been omitted in this panel. The structures have been globally aligned to the protein backbone. (b) Zoom of the same overlay to the BamA b-barrel. Strand $\beta16^{subs}$ of the substrate is now shown in purple. It is paired to strand $\beta1^{mem}$ of the catalytic BamA. Bold green and red arrows depict the directions of strand β1, forming an ~90° angle. (c) Zoom into the gate region indicating the spatial proximity of the substrate and darobactin interaction sites and their relative rotation of ~90°. (d–f) Comparison of the register of β16 complementation to b1. (d) In BamA–darobactin (salmon), residue I806 pairs with Y432. (e) In the late stage intermediate, I806 of the substrate $BamA^{subs}$ (purple) pairs with F428 in catalyst $BamA^{mem}$ (green), corresponding to a register

shift of 4. (f) Hypothetical position of the four C-terminal residues of substrate BamA, which are not resolved in the available electron density. When paired to $\beta1^{mem}$ in canonical antiparallel β-strand conformation, they locate exactly at the darobactin binding site, with the C-terminal W810 at the position of F^7 of darobactin. (g) Frequency logo of known and putative β-signals from bacterial OMPs, colored by amino acid chemistry. Numbering refers to distance from C-terminus. (h) Distribution of log-likelihood scores in three different sets of sequences, as indicated. The score obtained by the darobactin sequence is indicated by a blue line. The percentile rank of darobactin within each of the three sets is given in brackets. (i,j) Same as g, h, but based on side chain chemistry. "H" = hydrophobic and non-polar residue; "A" = aromatic; "N" = neutral; "C" = charged; "P" = polar non-charged.

Extended Data Fig. 10. Interaction of β-signal peptides with BamA-β barrel in detergent micelles by isothermal titration calorimetry. The first 4 panels show a direct titration of each of the 10aa long β-signal peptides of BamA, BtuB, FhuA, and OmpF to the BamA-β barrel. The next 5 panels show a competition experiment of darobactin titrated to BamA-β barrel in the presence of 10aa long β-signal peptides of OmpT (0.7 mM), BtuB (2.6 mM), OmpF (1.4 mM), FhuA (1.1 mM) and a β-consensus-peptide (1.2 mM). The results from fitting of the data to the competition model is given in Supplementary Table 7.

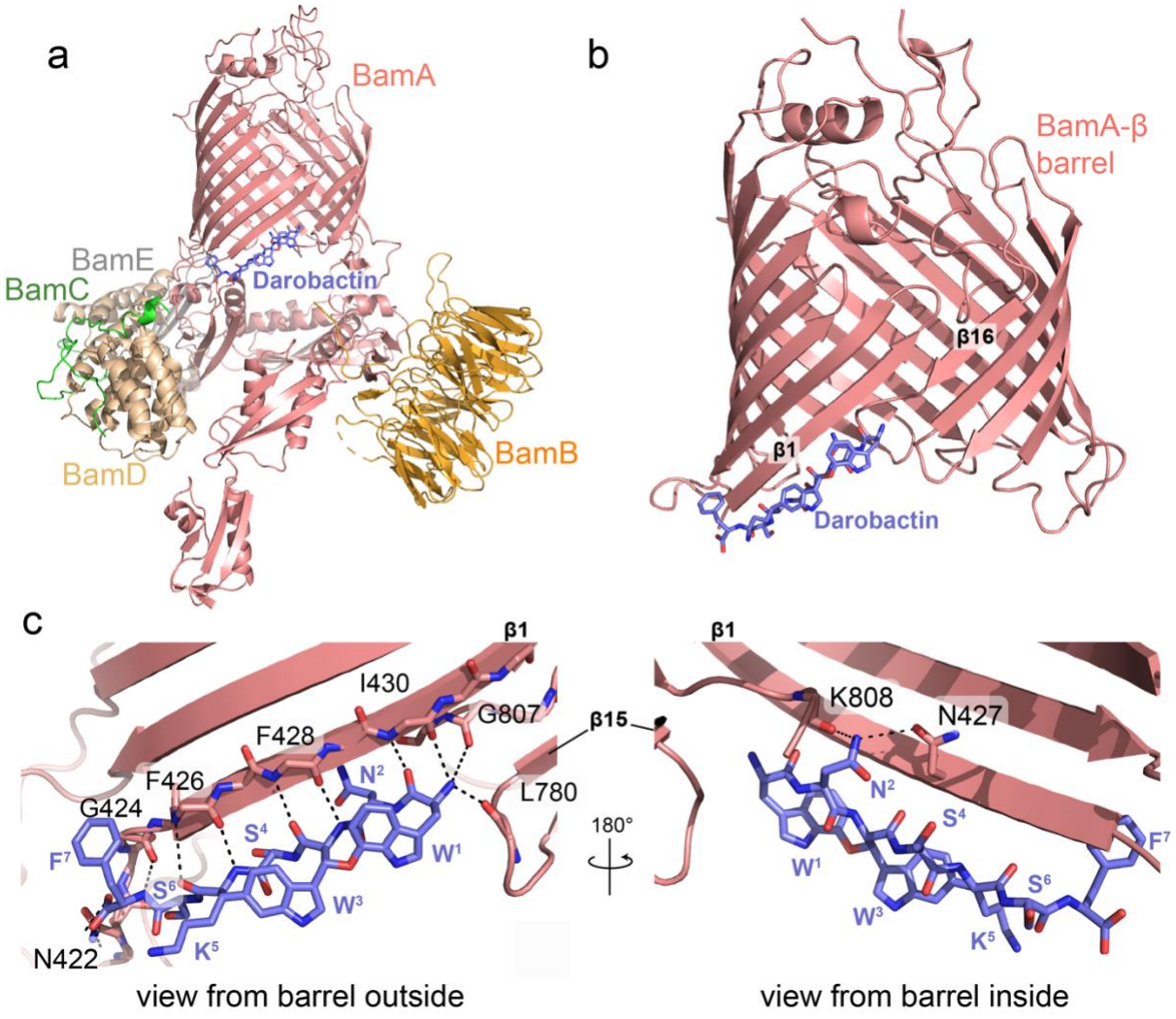

a
BamA
BamE
BamC
Darobactin
BamD
BamB
b
BamA-β barrel
β16
β1
Darobactin
c
β1
I430
G807
F428
F426
G424
N2
S4
W1
F7
S6
W3
K5
N422
L780
β15
180°
β1
K808
N427
N2
S4
F7
W1
S6
W3
K5
view from barrel outside
view from barrel inside

Figure 1

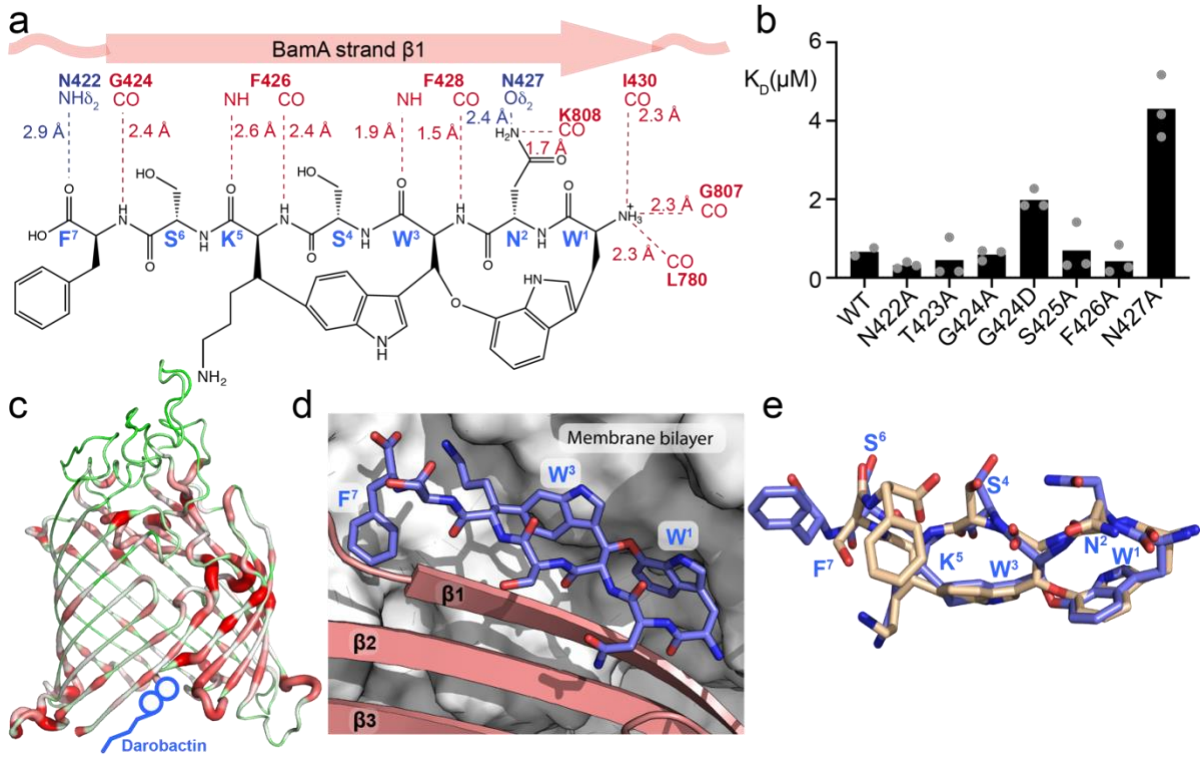

a
BamA strand β1
N422 NHδ2
G424 CO
F426 NH CO
F428 NH CO
N427 Oδ2
K808 CO
I430 CO
2.9 Å
2.4 Å
2.6 Å
2.4 Å
1.9 Å
1.5 Å
2.4 Å
1.7 Å
2.3 Å
G807 CO
2.3 Å
CO L780
F7 S6 K5 S4 W3 N2 W1
b
KD(μM)
6
4
2
0
WT
N422A
T423A
G424A
G424D
S425A
F426A
N427A
c
Darobactin
d
Membrane bilayer
W3
F7
W1
β1
β2
β3
e
S6
S4
F7
K5
W3
N2
W1

Figure 2

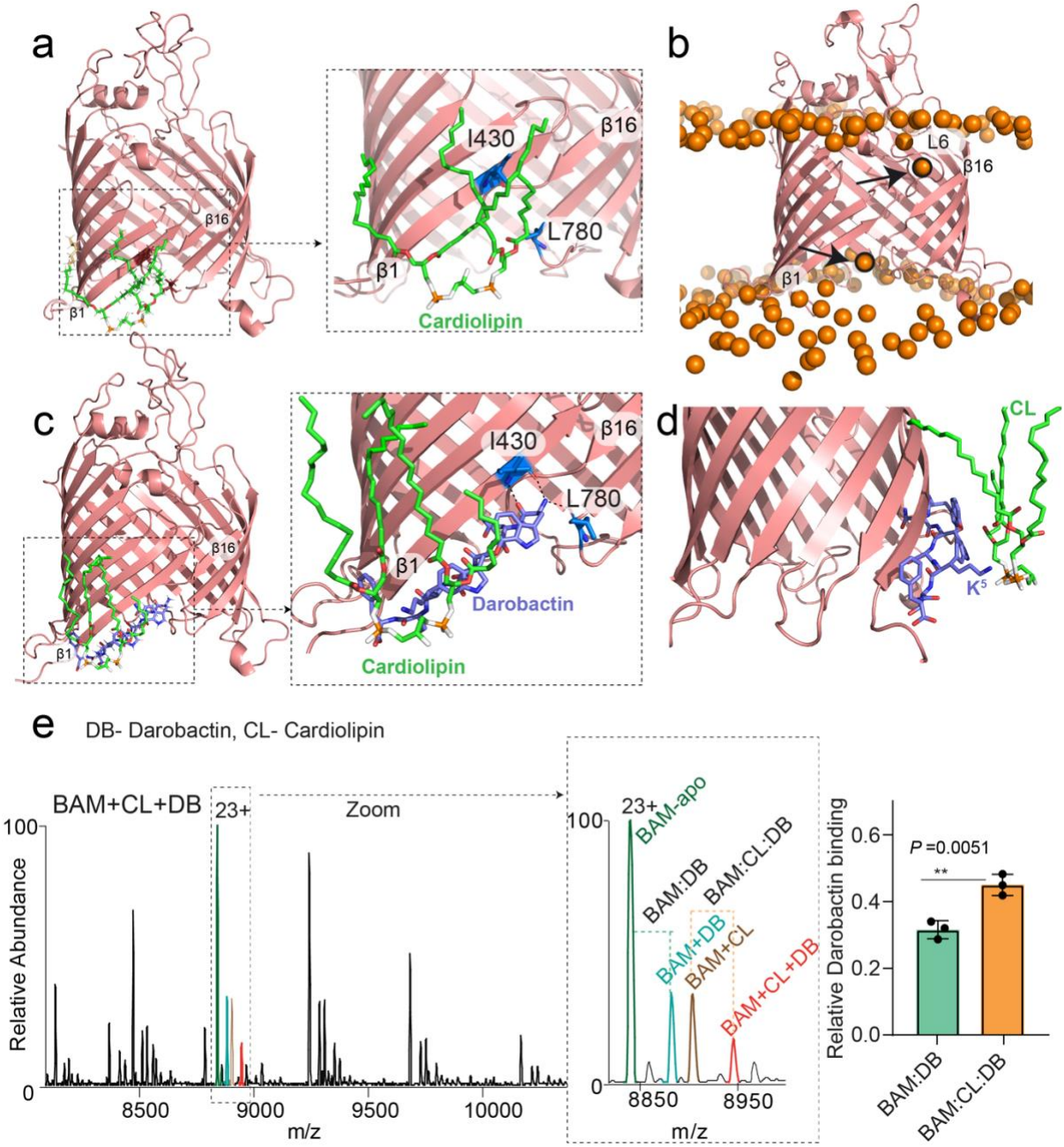

a
β16
β1
I430
β16
L780
β1
Cardiolipin
b
L6
β16
β1
c
β16
β1
I430
β16
L780
β1
Darobactin
Cardiolipin
d
CL
K5
e
DB- Darobactin, CL- Cardiolipin
BAM+CL+DB
23+
Zoom
100
Relative Abundance
0
8500
9000
9500
10000
m/z
23+
BAM-apo
BAM:DB
BAM:CL:DB
BAM+DB
BAM+CL
BAM+CL+DB
100
0
8850
8950
m/z
Relative Darobactin binding
P =0.0051
**
0.6
0.4
0.2
0.0
BAM:DB
BAM:CL:DB

Figure 3

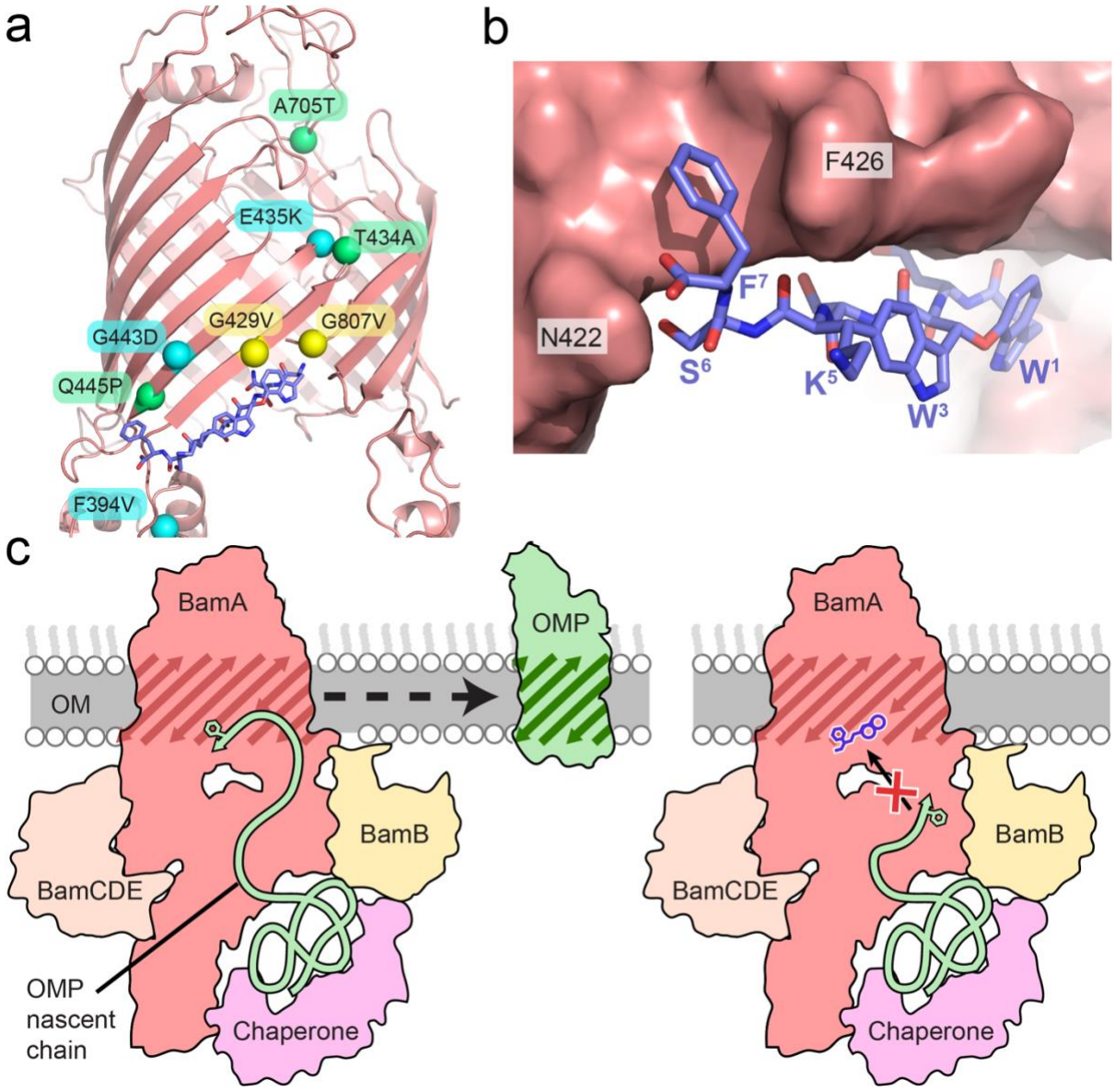
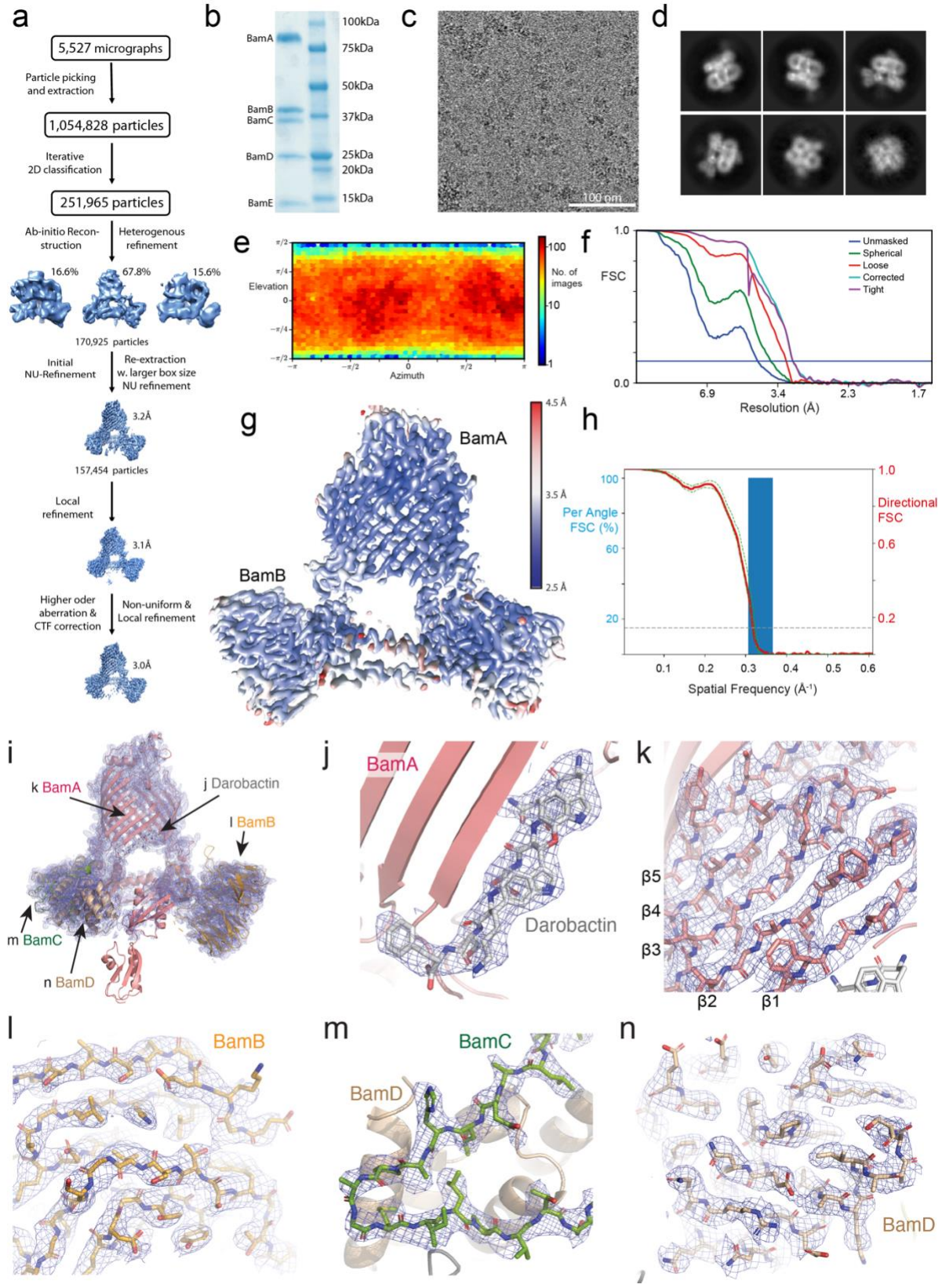
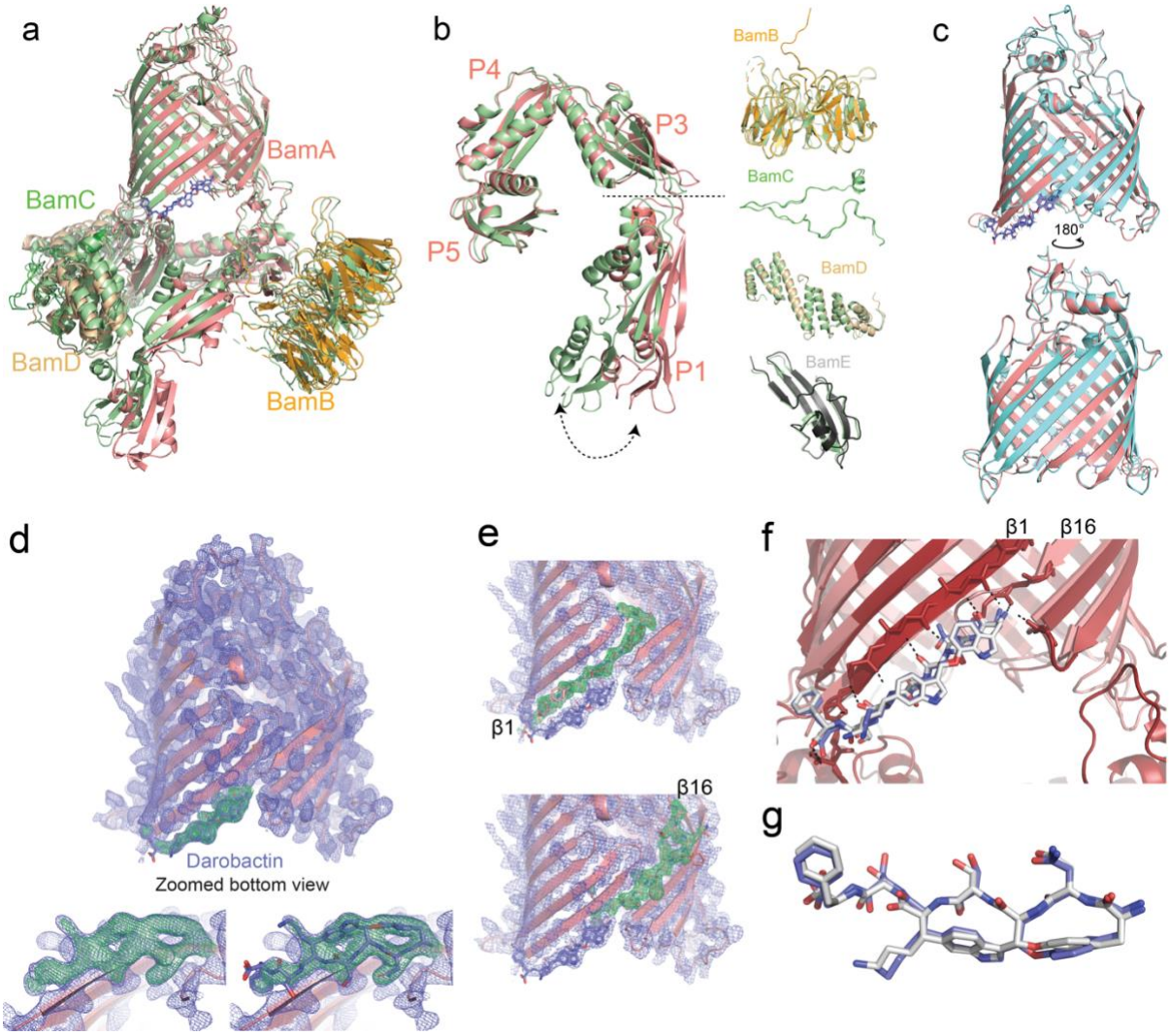

a
A705T
E435K
T434A
G443D
G429V
G807V
Q445P
F394V
b
F426
N422
F7
S6
K5
W3
W1
c
BamA
OMP
OM
BamB
BamCDE
OMP
nascent
chain
Chaperone
BamA
BamB
BamCDE
Chaperone

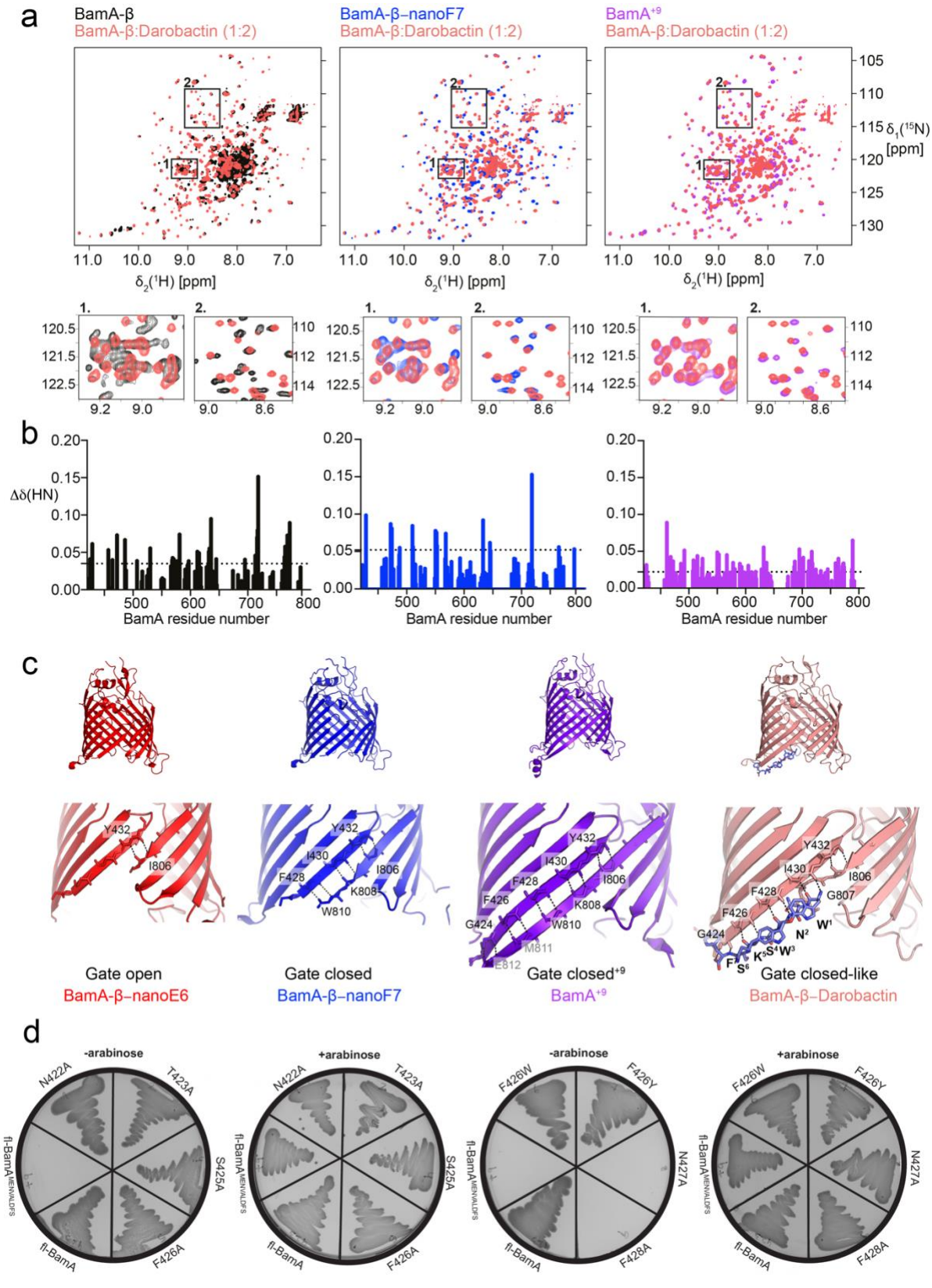
Figure 4



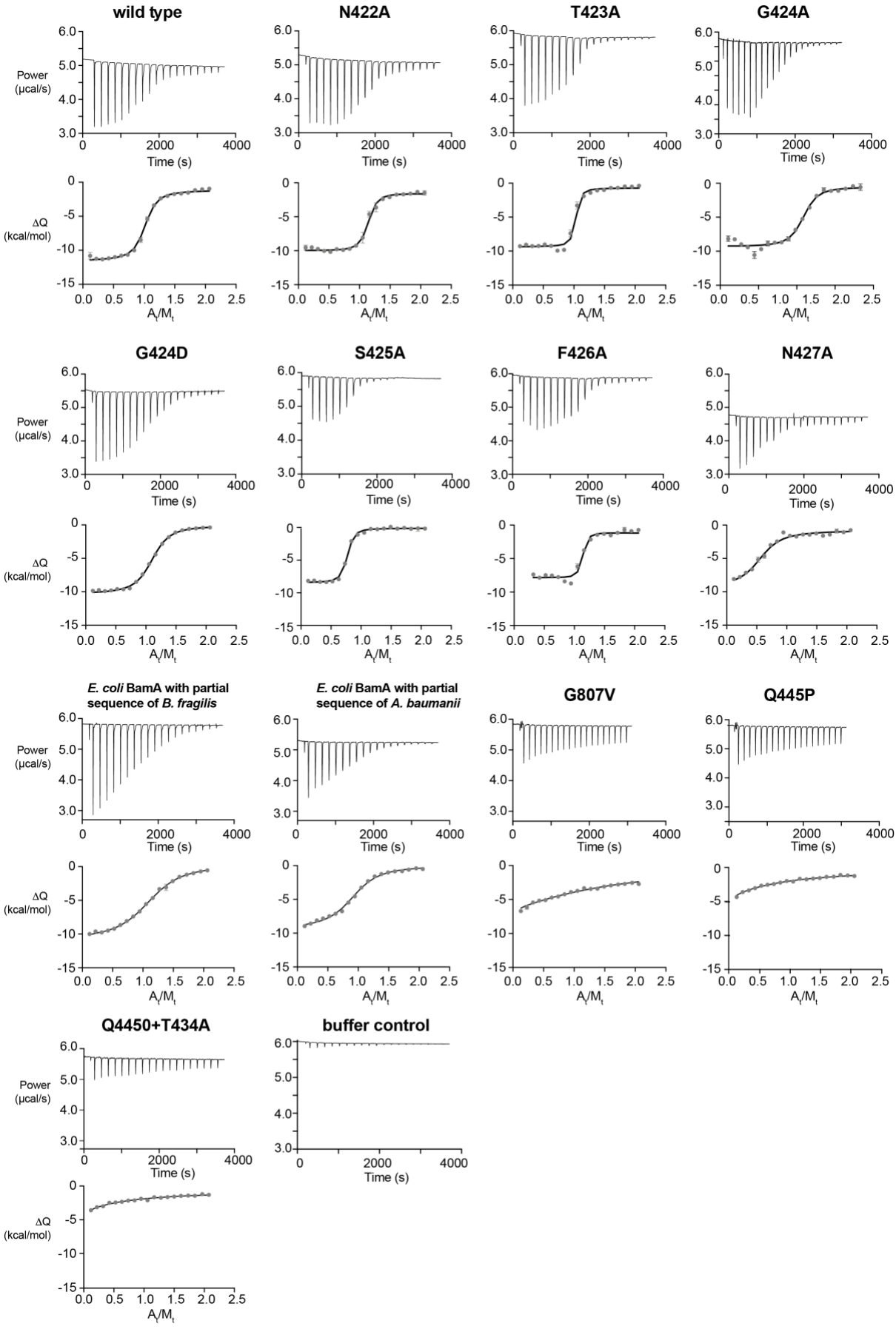
Extended Data Figure 1



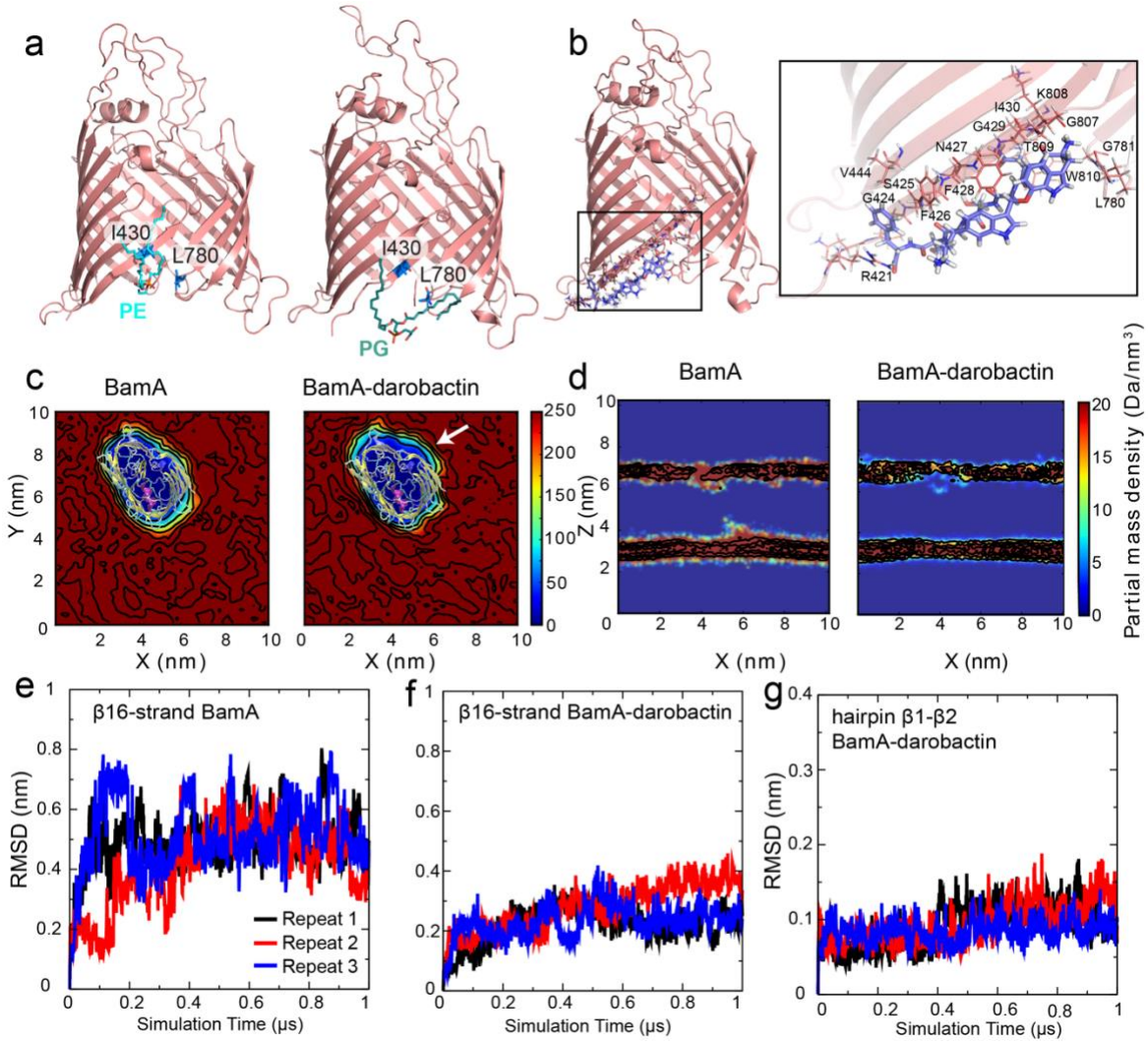
Extended Data Figure 2



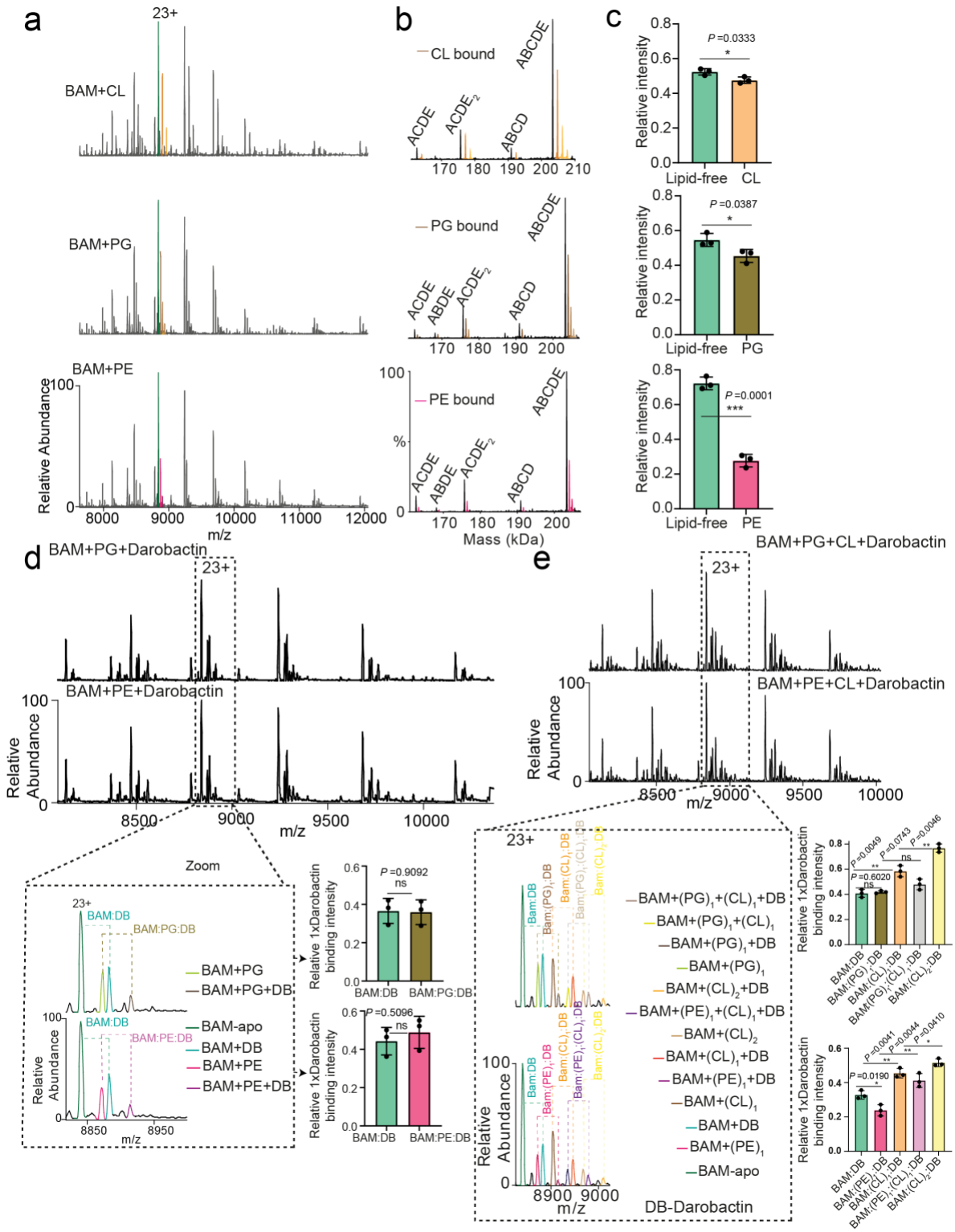
Extended Data Figure 3



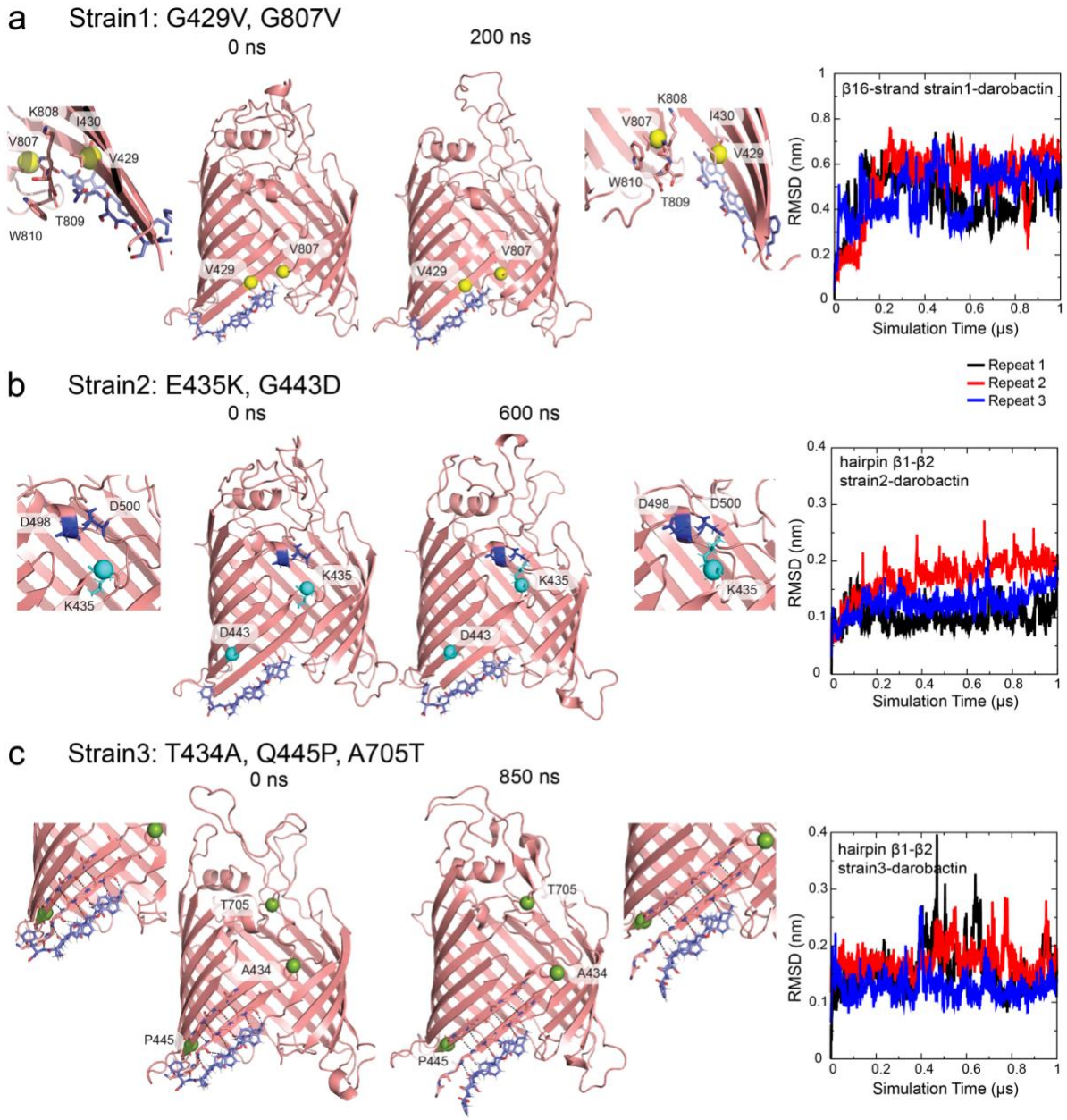
Extended Data Figure 4



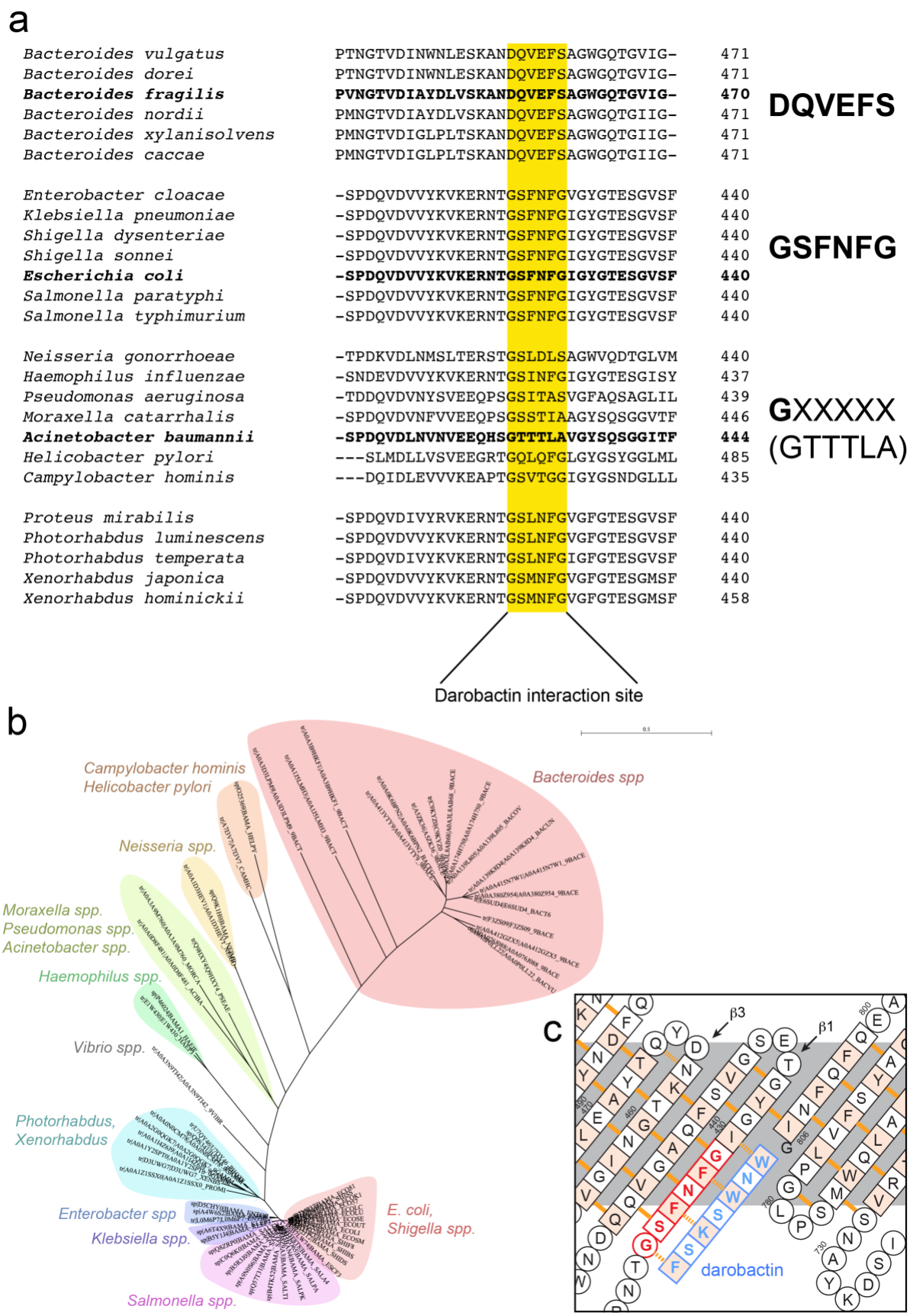
Extended Data Figure 5



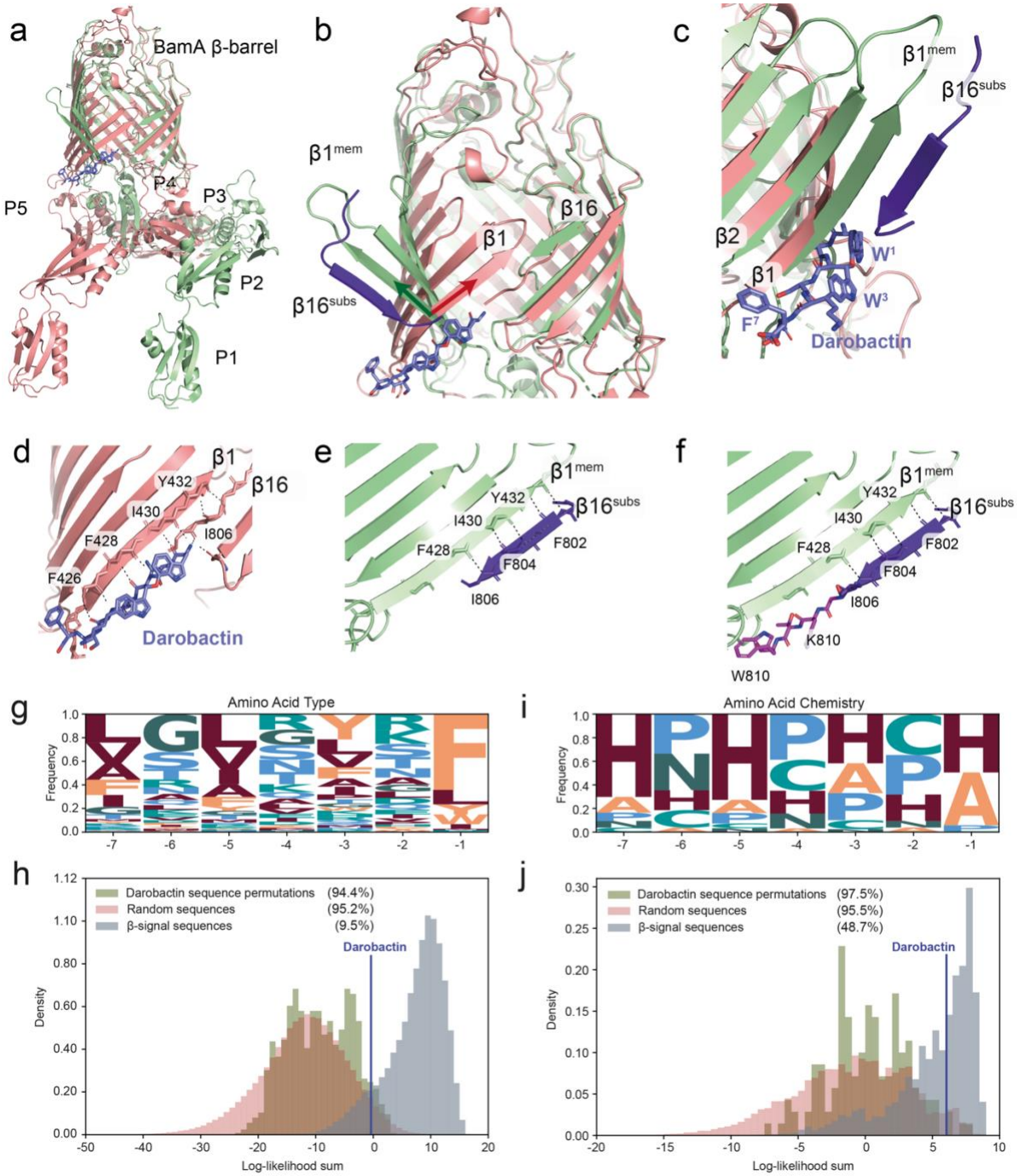
Extended Data Figure 6



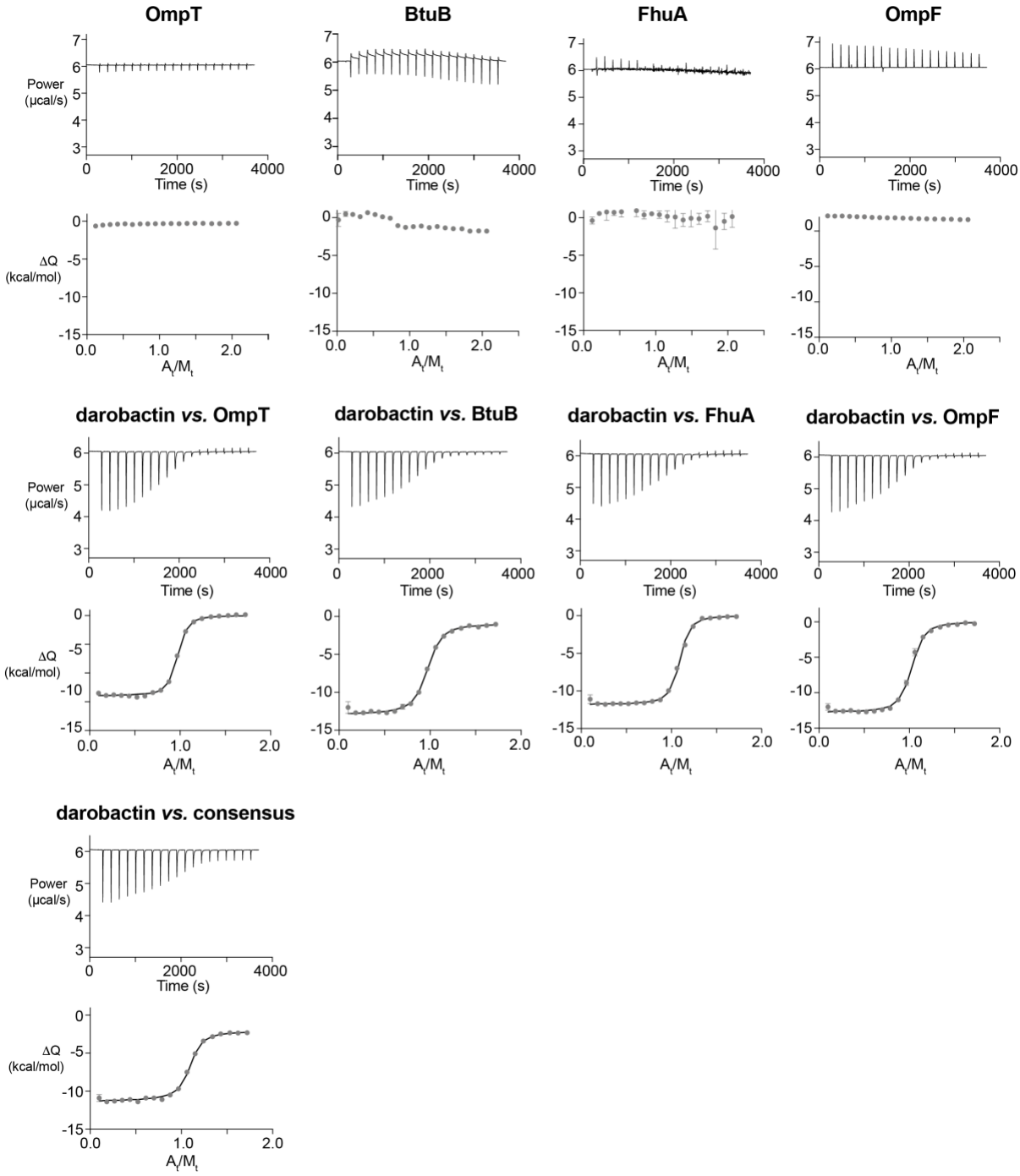
Extended Data Figure 7



Extended Data Figure 8



Extended Data Figure 9



Extended Data Figure 10