PNAS

Supplementary Information for

Title

The cell envelope of *Thermotogae* suggests a mechanism for outer membrane biogenesis

Authors

Danielle L. Sexton[1], Ameena Hashimi[1], Polina Beskrovnaya[1], Lloyd Sibanda[2], Tao Huan[2], Elitza I. Tocheva[1,*]

[1]Department of Microbiology and Immunology, Life Sciences Institute, 2350 Health Sciences Mall, The University of British Columbia, Vancouver, Canada, V6T1Z3

[2]Department of Chemistry, 2036 Main Mall, University of British Columbia, Vancouver, Canada, V6T1Z1

* **Correspondence to** Elitza I. Tocheva

Email: elitza.tocheva@ubc.ca

Author Contributions: D.L.S and E.I.T designed research; D.L.S, A.H, P.B, L.S. and E.I.T performed research; D.L.S, A.H, P.B, L.S, T.H, E.I.T analyzed data; D.L.S, A.H and E.I.T. wrote the manuscript.

Competing Interest Statement: The authors declare no competing interests.

Classification: BIOLOGICAL SCIENCES; Biochemistry; Microbiology

Keywords: bacterial cell envelope | *Thermotoga maritima* | toga | beta-barrel proteins | cryo-ET | lipidomics | proteomics | OM biogenesis | LUCA | evolution of bacteria

This PDF file includes:

Other supplementary materials for this manuscript include the following:

Movies S1 and S2

Extended Materials and Methods

Protein extraction and analysis

For MS-based protein analysis, the separated IM and toga fractions were resuspended in 6x Laemmli sample buffer diluted in wash buffer and boiled for 10 min at 95 °C. Samples were run 5 mm into a 10% SDS-PAGE gel. The gel was stained with Coomassie and destained in water without exposure to heat. Gel bands were excised in small pieces of about 1 mm^3 using a scalpel blade. The gel bands were washed 3 times for 15 min in 50 mM ammonium bicarbonate/ acetonitrile (50:50, v/v), rinsed in 100% acetonitrile and incubated for 20 min in fresh 100% acetonitrile. After being air dried, proteins were reduced with dithiothreitol (DTT; 10 mM in 100 mM ammonium bicarbonate) for 30 min at 56 °C and alkylated with iodoacetamide (IAA; 50 mM in 100 mM ammonium bicarbonate) for 30 min in the dark at room temperature. The gel bands were then rehydrated with a trypsin solution (Promega; 0.02 µg/µL in 40mM ammonium bicarbonate and 10% acetonitrile) and incubated for 2 hrs on ice. The excess of trypsin solution was removed, and trypsin buffer (40 mM ammonium bicarbonate 10% acetonitrile) was added to cover the gel pieces. Trypsin digestion was performed for 16 h at 37 °C. After the digestion, the supernatant (tryptic peptides) was transferred into a new tube containing 5 µl of extraction solution (acetonitrile/water/10 % trifluoroacetic acid (60:30:10, v/v)). Tryptic peptides were extracted twice from the gel bands by incubation in the extraction solution and vortexing for 10 min. The extracted peptides were combined into the same Eppendorf tube. Samples were then lyophilized and resuspended in 10 µl of 1% formic acid in water. Tryptic peptides were analyzed on an Orbitrap Fusion Lumos Tribrid mass spectrometer (ThermoFisher Scientific) operated with Xcalibur (version 4.0.21.10) and coupled to a ThermoScientific Easy-nLC (nanoflow Liquid Chromatography) 1200 system. Tryptic peptides were loaded onto a C18 trap (75 um x 2 cm; Acclaim PepMap 100, P/N 164946; ThermoFisher Scientific) at a flow rate of 2 µl/min of solvent A (0.1% formic acid in LC-MS grade water). Peptides were eluted using a 45 min gradient from 5 to 40% (5% to 28% in 40 min followed by an increase to 40% B in 5 min) of solvent B (0.1% formic acid in 80% LC-MS grade acetonitrile) at a flow rate of 0.3 µL/min and separated on a C18 analytical column (75 um x 50 cm; PepMap RSLC C18; P/N ES803; ThermoFisher Scientific). Peptides were then electrosprayed using 2.1 kV voltage into the ion transfer tube (300 °C) of the Orbitrap Lumos operating in positive mode. The Orbitrap first performed a full MS scan at a 120,000 full width half maximum (FWHM) resolution to detect the precursor ion having a *m*/*z* between 375 and 1,575 and a +2 to +7 charge. The Orbitrap AGC (Auto Gain Control) and the maximum injection time were set at 4 x 10^5 and 50 ms, respectively. The Orbitrap was

operated using the top speed mode with a 3 sec cycle time for precursor selection. The most intense precursor ions presenting a peptidic isotopic profile and having an intensity threshold of at least 5,000 were isolated using the quadrupole and fragmented with HCD (30% collision energy) in the ion routing multipole. The fragment ions (MS^2) were analyzed in the ion trap at a rapid scan rate. The AGC and the maximum injection time were set at 1 x 10^4 and 35 ms, respectively, for the ion trap. Dynamic exclusion was enabled for 30 sec to avoid of the acquisition of same precursor ion having a similar m/z (plus or minus 10 ppm).

The Lumos raw data files were converted into Mascot Generic Format (MGF) using RawConverter (v1.1.0.18; The Scripps Research Institute) operating in a data dependent mode. Monoisotopic precursors having a charge state of +2 to +7 were selected for conversion. This MGF file was used to search against the *T. maritima* strain MSB8 proteome (Uniprot ID 243274) using Mascot algorithm (Matrix Sciences; version 2.7). Search parameters for MS data included trypsin as enzyme, a maximum number of missed cleavage of 1, a peptide charge equal to 2 or higher, cysteine carbamidomethylation as fixed modification, methionine oxidation as variable modification and a mass error tolerance of 10 ppm. A mass error tolerance of 0.6 Da was selected for the fragment ions. Only peptides identified with a score having a confidence higher than 95% were kept for further analysis. The Mascot data files were imported into Scaffold (v4.3.4, Proteome Software Inc) for comparison of different samples based on their mass spectral counting. Protein localization was predicted using pSORTb [1] and tertiary structures were predicted using RoseTTAFold [2].

Lipid extraction and analysis

Whole membrane, IM, and toga fractions were analyzed for their lipid composition using liquid chromatography-mass spectrometry (LC-MS). Briefly, 1.5 mL of a solvent composed of methanol:acetonitrile:water (2:2:1, vol:vol:vol) was added to each of the three glass tubes corresponding to the individual fractions. The mixture underwent sonication for 45 minutes to homogenize the samples. Technical replicates of each fraction were generated by splitting each fraction into three 400 µL aliquots into 2 mL Eppendorf vials. This was followed by subjecting the samples to three freeze/thaw (F/T) cycles for protein precipitation. One given cycle involved placing the samples in liquid N_2 for 1 minute prior to sonication in an ice-bath for 15 minutes. After completing the F/T cycles the samples were left overnight in a -20 °C freezer. Afterwards, 1.2 mL of MTBE was added to each sample, followed up with vortexing and repetition of the previous

step, except using 0.4 mL H_2O. Subsequently, all samples underwent centrifugation at 14,000 rpm and 4 °C for 15 minutes. After centrifugation, the upper layer, containing the lipids, was transferred into new 1.5-mL Eppendorf vials. To dry the lipid samples, the solvent was evaporated using a vacuum concentrator at 4 °C overnight. A total of 150 µL of isopropanol:acetonitrile (1:1, vol:vol) was added to reconstitute the dried residue. The reconstituted solution was vortexed for 30 s and centrifuged at 14,000 rpm at 4 °C for 15 min. The resulting supernatants were transferred to glass inserts for LC-tandem MS (LC-MS/MS) analysis. Lastly, the QC sample was made by pooling equal aliquots (20 µL) of each sample together.

Lipid profiling of cellular membranes was carried out using an Agilent 1290 Infinity II ultra-high performance liquid chromatography (UHPLC) system (Agilent Technologies) coupled with Bruker Impact II electrospray-ionization quadrupole time-of-flight (ESI-QTOF) mass spectrometer (Bruker Daltonics). LC separation was performed using a reversed phase Acquity UPLC BEH C_{18} column (1.0 × 100 mm, 1.7 µm, Waters), which was maintained at 25 °C. MS detection was performed in both ESI positive (ESI (+)) and ESI negative (ESI (−)) modes in two separate acquisitions. For ESI (+), mobile phase A was acetonitrile/water (6:4, v/v, pH = 4.8 adjusted by 0.1% formic acid) containing 5 mM ammonium acetate, and B was isopropanol/acetonitrile (9:1, v/v). For ESI (−), acetonitrile/water (6:4, v/v, pH = 9.8) and isopropanol/acetonitrile (9:1, v/v) were used as mobile phases A and B, respectively, with only mobile phase A containing 5 mM ammonium acetate. The gradient elution program for ESI (-) was as follows: 0 min, 5% B; 20 min, 95% B; 23 min, 95% B; 24 min, 5% B; 33 min, 5% B. For ESI (+) the gradient elution program was the same as ESI (-), with the exception of the last run at 35 min. The flow rate was set at 0.1 mL/min, and the injection volumes were optimized as 6 µL for ESI (-) and 3 µL for ESI (+). For the MS, capillary voltage was set at 4,500 V for ESI (+) and 3,600 V for ESI (−). The nebulizer gas was set at 1.0 bar, with drying gas flow rate of 6 L/min and source temperature of 220 °C. Data-dependent acquisition (DDA) mode was applied to collect both MS1 and MS2 spectra. A programed injection of 2 µL sodium formate (250 mM) at 25 min was used for internal mass calibration. HPLC-grade methyl tert-butyl ether (MTBE) was purchased from Merck. All the other high purity solvents and chemicals were purchased from ThermoFisher Scientific.

All the raw data were first calibrated using Bruker Compass Data Analysis (version 4.4) and then converted to ABF format with Abf Converter software. Lipid identification was completed on MS-DIAL 4.7 tool [3] based on mass accuracy, isotope ratio, retention time along with MS/MS similarity against publicly available libraries. MS-DIAL 4.7 can be downloaded at the PRIMe website (http://prime.psc.riken.jp/). Raw lipidomics data were collected in positive and negative

ion mode, after intensity correction and normalization (post QC calibration). QC calibration was done to obtain more accurate fold changes. The relative standard deviation (rsd) of the QC injections (at the beginning, middle and end of the injection sequence) were below the threshold of 0.25.

Detection of Ompα homologues in members of the phylum Thermotogae

Prodigal-generated locus tag information for each protein was generated for 49 representative Thermotogae phylum genomes. Proteomes were assessed for the presence of Ompα homologues using PSI-BLAST search similarity to query sequences Ompα1-3 from *T. maritima*, as well as the presence of homologues other toga-associated proteins, as listed in Figure 3. Gene proximity and synteny was assessed to identify putative gene clusters. Top Ompα hits were further assessed for the presence of SLH domains and coiled-coil domains. SLH domain sequences from each putative Ompα is listed. Annotations of proteins were generated using BLASTp searches against the Uniref90 database, and PSI-BLAST search hits to proteins from *T. maritima.*

References

1 Yu, N. Y. *et al.* PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics* **26**, 1608-1615, doi:10.1093/bioinformatics/btq249 (2010).

2 Baek, M. *et al.* Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373**, 871-876, doi:doi:10.1126/science.abj8754 (2021).

3 Lai, Z. *et al.* Identifying metabolites by integrating metabolome databases with mass spectrometry cheminformatics. *Nat Methods* **15**, 53-56, doi:10.1038/nmeth.4512 (2018).

4 Evans, R. *et al.* Protein complex prediction with AlphaFold-Multimer. *bioRxiv*, 2021.2010.2004.463034, doi:10.1101/2021.10.04.463034 (2022).

Supplementary Figures

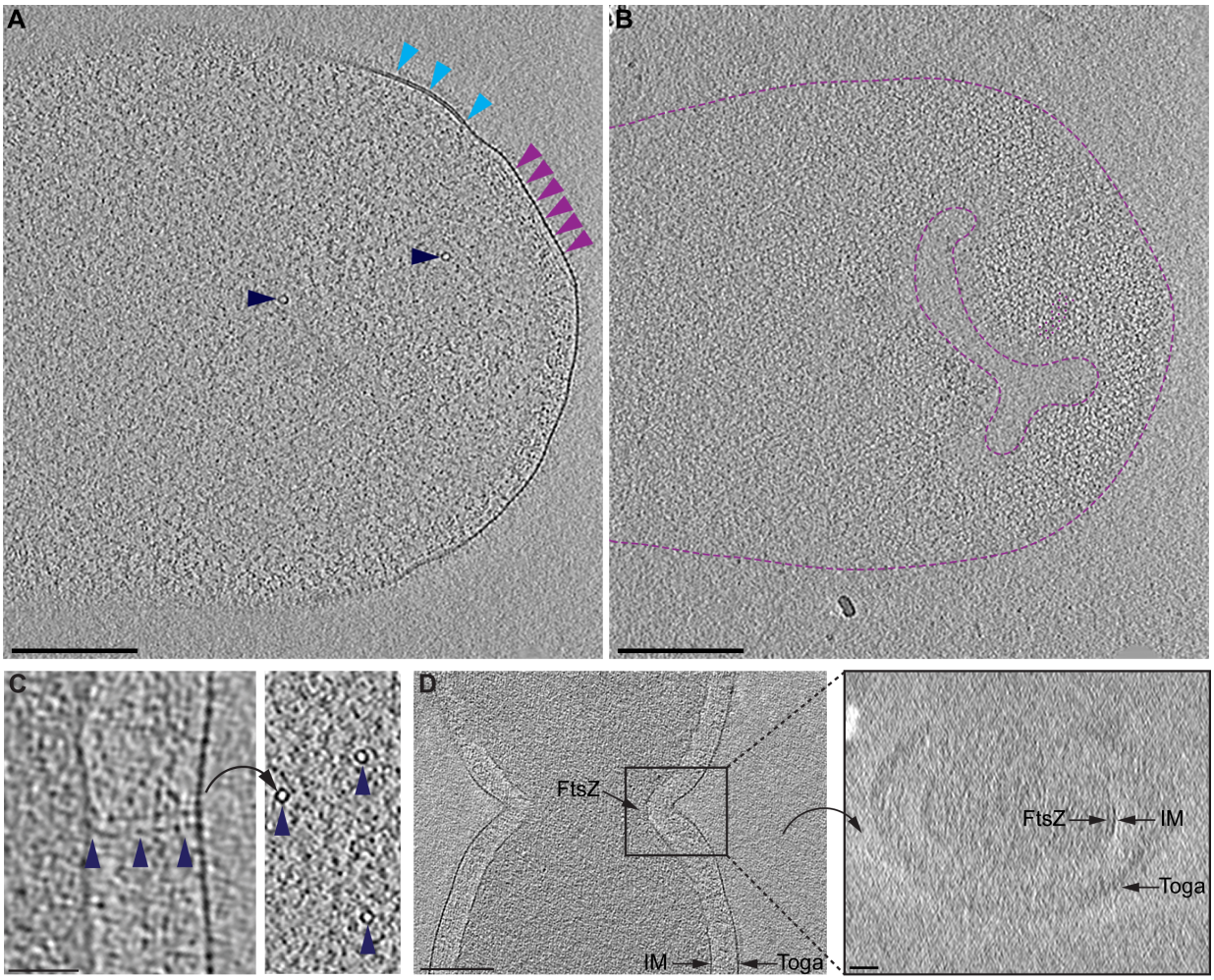

Figure S1. Additional ultrastructural features in *T. maritima.* A) A 10-nm thick tomographic slice shows a lipid bilayer (light blue arrows) and a monolayer of protein array (purple arrows) from Movie S2. B) A top view of the toga revealed the extended lattice of trimers and lipid patches (purple dashed line). C) Side (left) and top (right) views of T4P/T2SS/competence systems (dark blue arrows) traversing the cell envelope. D) A division site revealed that both the inner membrane (IM) and toga invaginate synchronously during cell division. Side views of the division site revealed filamentous structures, likely FtsZ, ~16 nm underneath the IM. Scale bar A, B, D 200 nm, C 50 nm.

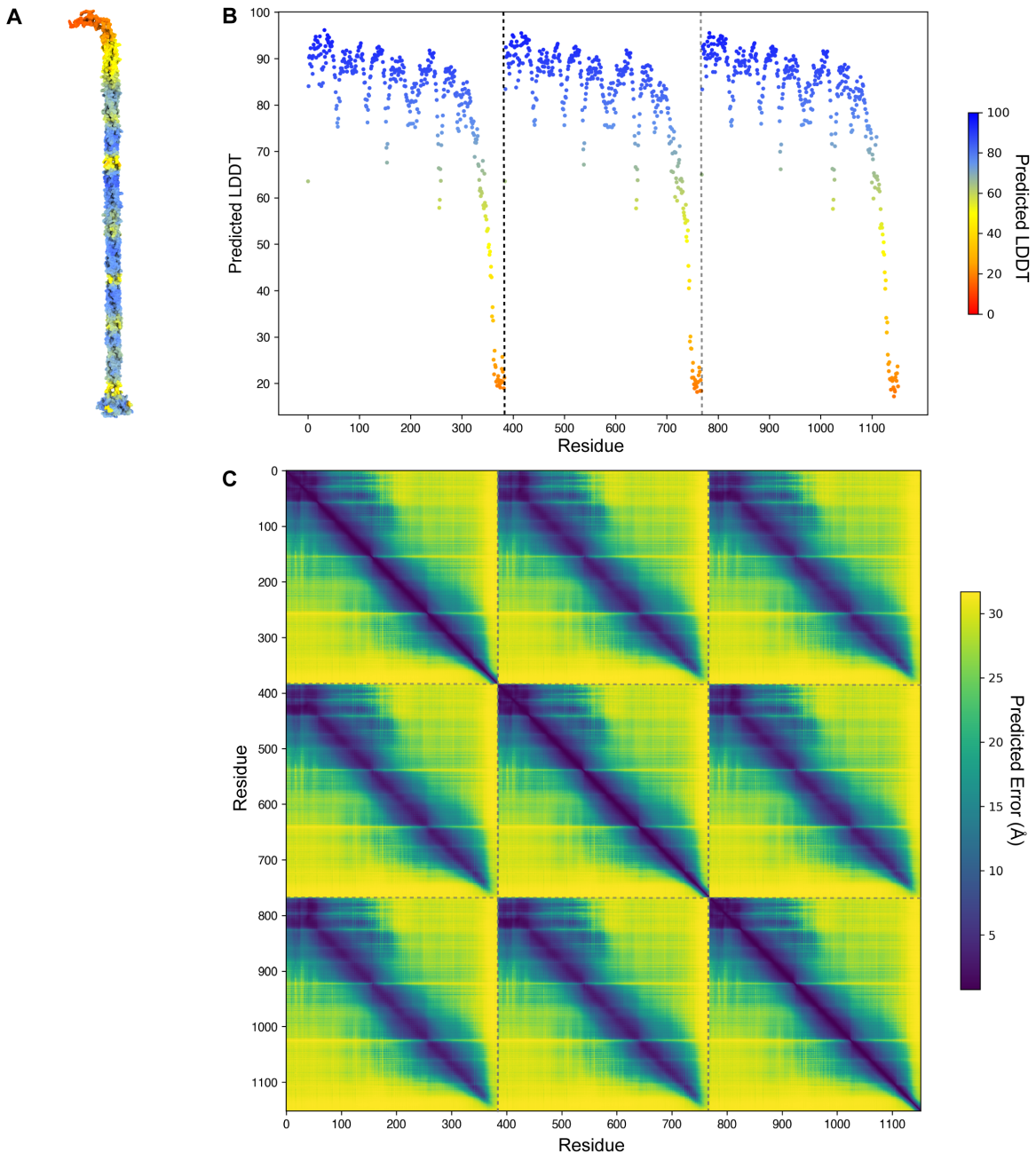

Figure S2. Structural prediction for trimer of Ompα. A) Trimeric structure of Ompα2 was predicted using AlphaFold2 Multimer and colored based on the pLDDT score. SLH domains are located at the bottom of the structure. The predicted B) LDDT and C) error plots for the Ompα2 trimer. The boundary between monomers is indicated with grey dashed lines.

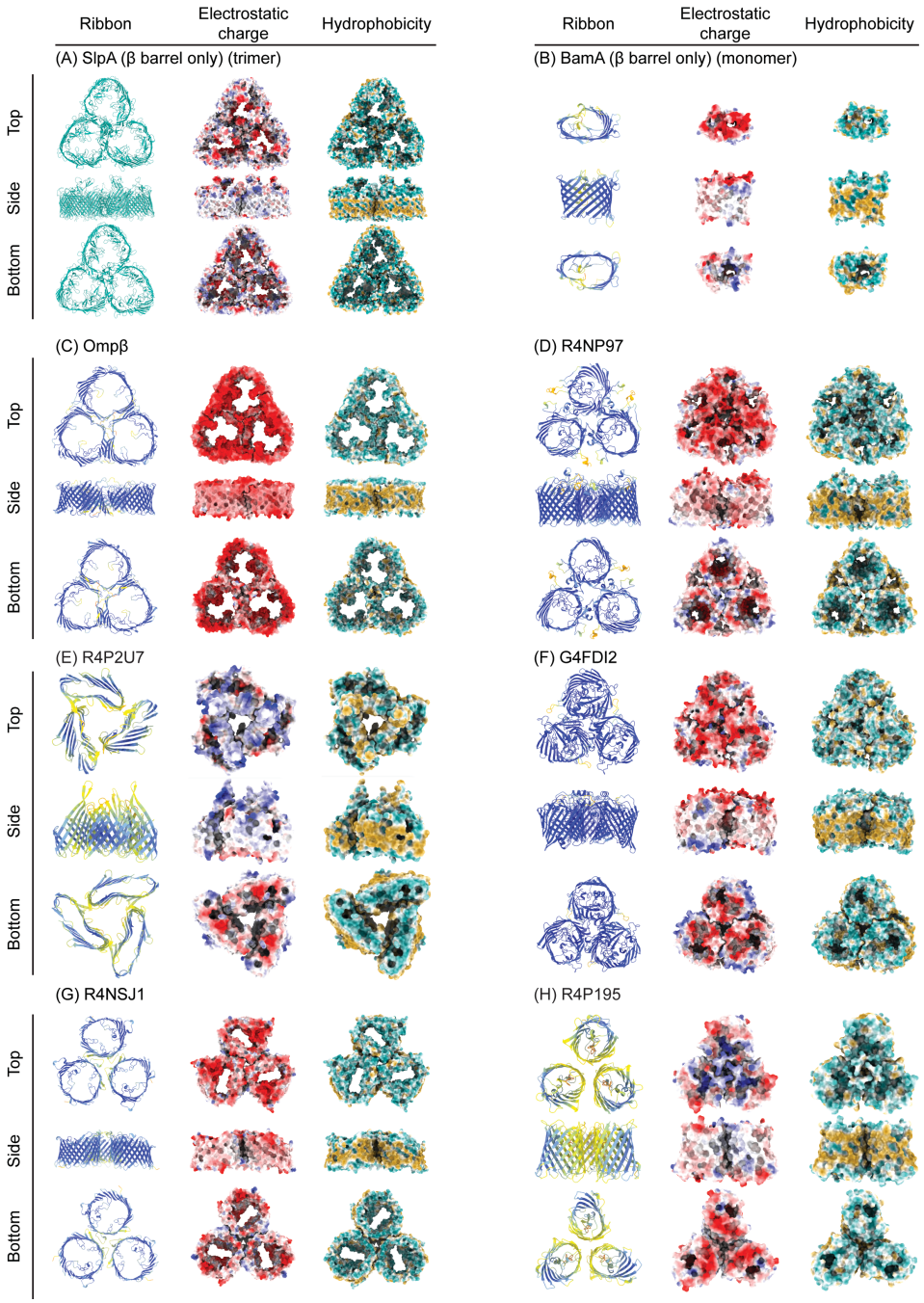

Figure S3. Predicted structures of putative β-barrel proteins in the toga. Trimeric protein structures for the β-barrel portion of all proteins were predicted using AlphaFold2 Multimer. Surface characteristics of A) SlpA from *Deinococcus radiodurans* (PDB 7zgy) and B) β-barrel domain of BamA were compared to predicted β-barrel trimers of C) Ompβ, D) R4NP97, E) R4NP2U7, F) G4FDI2, G) R4NSJ1, and H) R4P195. Structures are colored based on the pLDDTscore from Alphafold2 Multimer [4]. Structures, pLDDT score, electrostatic charge and surface hydrophobicity for each protein were modeled in ChimeraX.

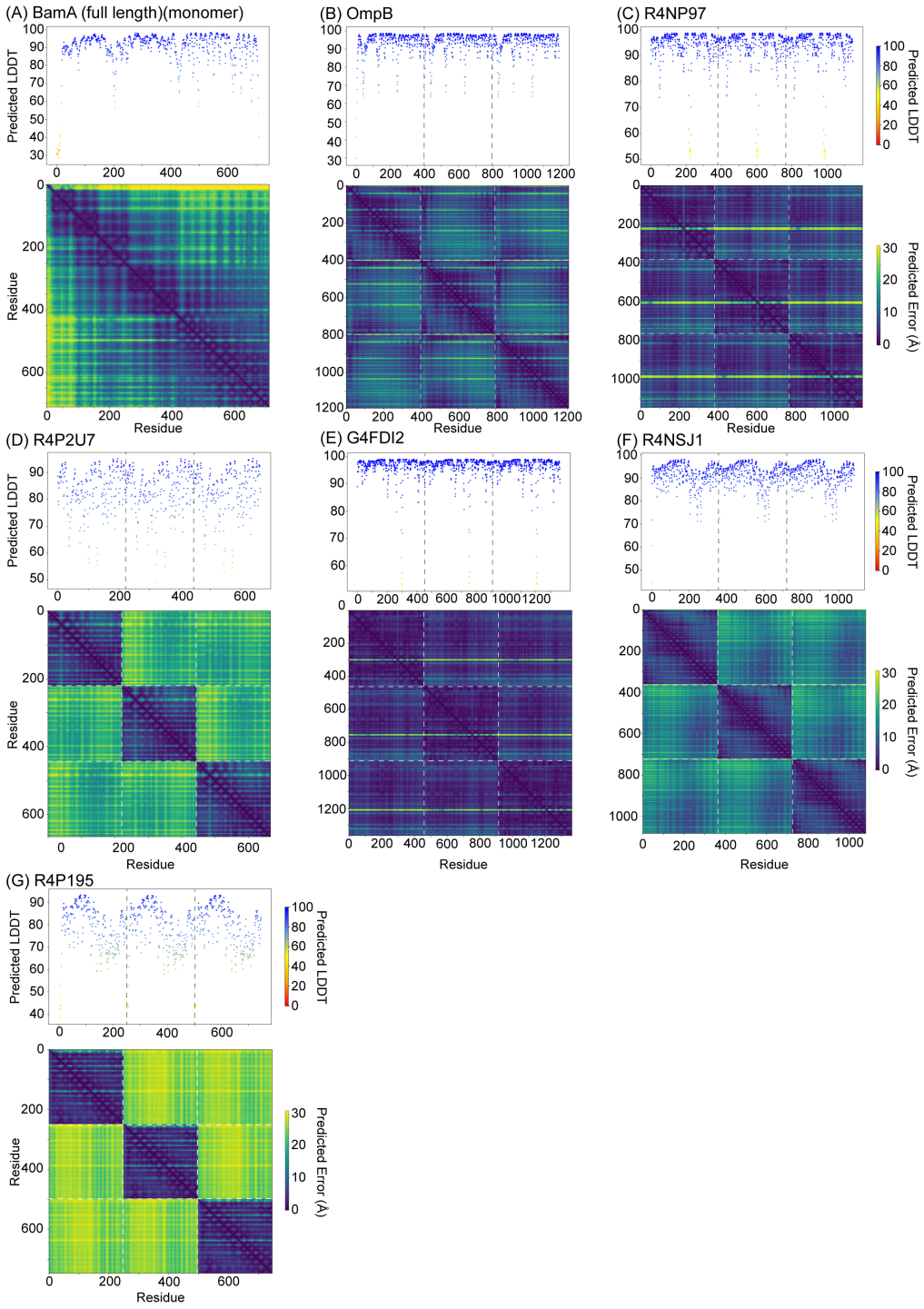

Figure S4. Predicted local distance difference test (LDDT) scores and errors for AlphaFold2 multimer protein structures. The predicted LDDT and error plots for structures depicted in Fig. S3. The boundaries between monomers are indicated with grey dashed lines.

Datasets

Dataset S1: All lipid species and their relative abundances detected in the IM and toga.

Dataset S2: All proteins detected in the IM and toga with 95% confidence.

Dataset S3. Ompα homologues in the phylum Thermotogae.

Dataset S4: All peptides identified in IM and toga fractions.

Dataset S5: Peptide coverage of proteins in IM and toga fractions.

Supplementary Movie Legends

Movie S1. Tomogram of a *T. maritima* cell. 3-dimensional volume shows the cell envelope of *T. maritima* (corresponds to Fig. 1A). The toga is dissociating at the tip of the cell. Side views of monolayers and bilayers of the OM are clearly visible.

Movie S2. Tomogram of the toga of a *T. maritima* cell. 3-dimensional volume of the tip of a *T. maritima* cell shows the extended toga. The monolayer and bilayer side views are highlighted with purple and blue arrows, respectively (Fig. S1). The extended protein arrays and lipid patches are outlined with purple dashed lines.