

Supplementary Information

for

Two types of sheathed archaellum structures from *Haloarcula marismortui* differ in their outer layer architectures

Vladimir A. Meshcheryakov¹, Jaekyung Hyun^{1,†}, Satoshi Shibata^{1,††,†††}, Alexey S. Syutkin², Mikhail G. Pyatibratov^{2,*} and Matthias Wolf^{1,*}

¹Molecular Cryo-Electron Microscopy Unit, Okinawa Institute of Science and Technology Graduate University, 1919-1 Tancha, Onna, Kunigami, Okinawa, 904-0495 Japan

²Institute of Protein Research, Russian Academy of Sciences, Institutskaya st. 4, Pushchino, Moscow Region, 142290 Russia

[†]Present Address: School of Pharmacy, Sungkyunkwan University, Suwon-si, Gyeonggi-do, 16419 Republic of Korea

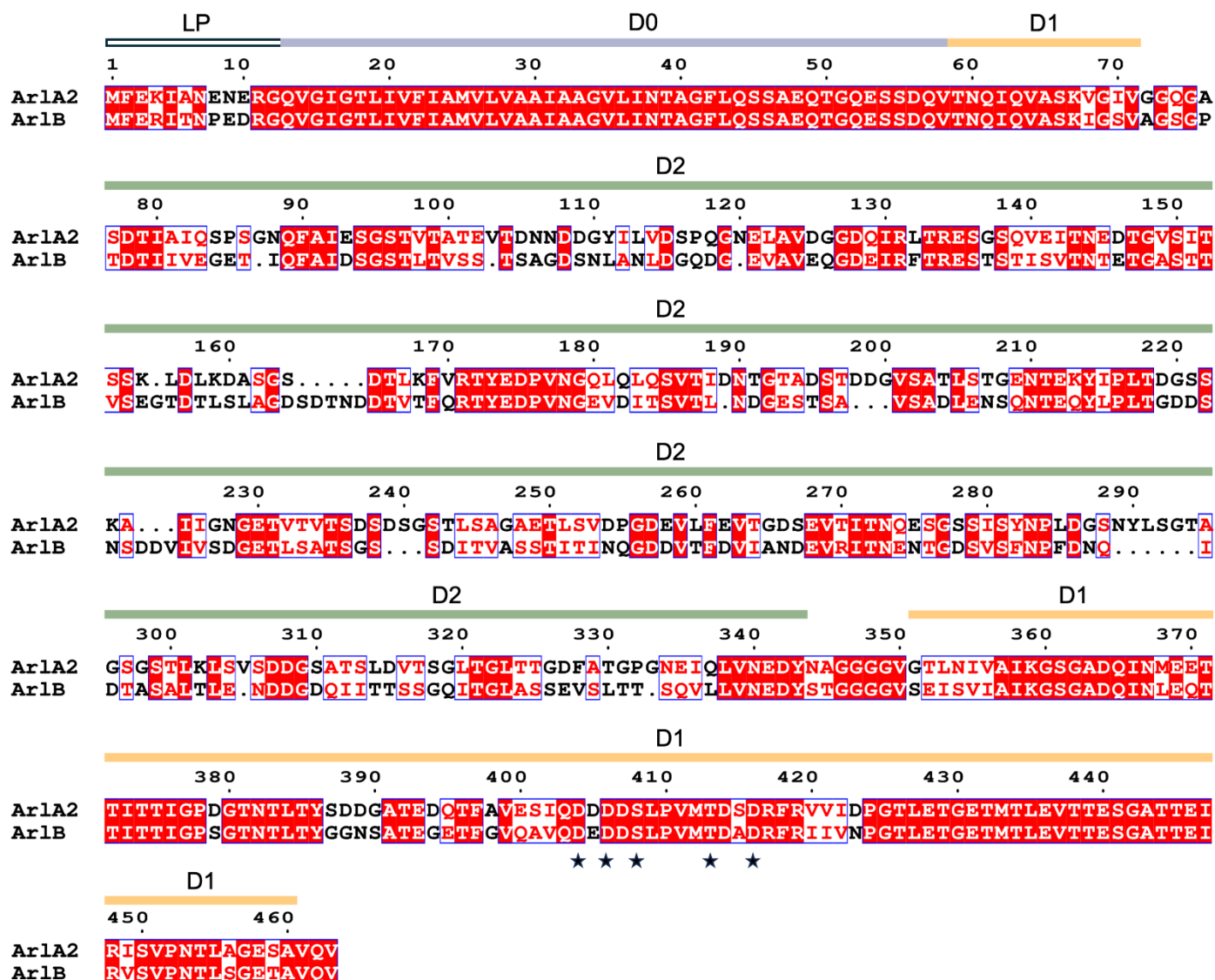
^{††}Present Address: Division of Bacteriology, Department of Microbiology & Immunology, Faculty of Medicine, Tottori University, Tottori, Japan

^{†††}Present Address: Division of Pathogenesis and Disease Control, Avian Infectious Disease and Global Health Research Center, Tottori University, Tottori, Japan

*Correspondence to: Matthias Wolf or Mikhail G. Pyatibratov, E-mail: matthias.wolf@oist.jp, bratov@vega.protres.ru

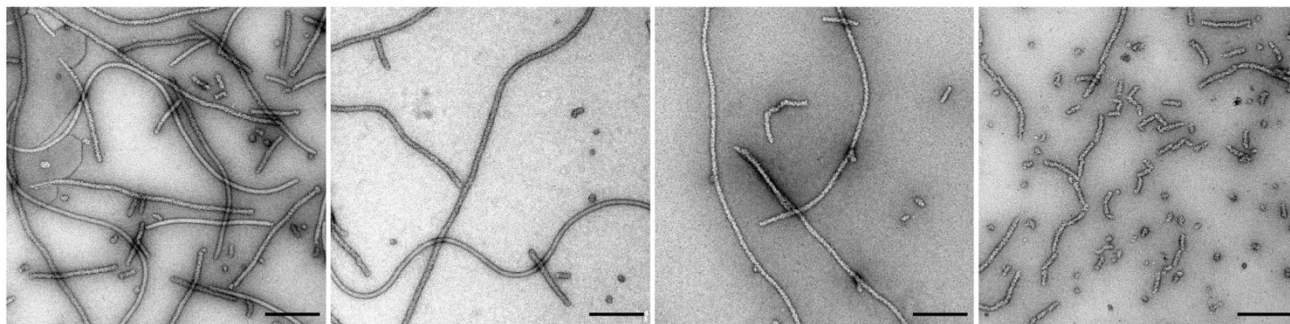


Supplementary Figure 1. The distribution of archaellin lengths based on complete amino acid sequences (including signal peptides). A total of 998 archaellin sequences (n=998) were used to create a histogram. The x-axis represents the length of the archaellin amino acid sequence in terms of the number of amino acids, while the y-axis indicates the number of archaellins with a particular length. It can be observed that the majority of archaellins fall within the range of 180-280 amino acids. Positions of ArlA2 and ArlB *Har. marismortui* are shown by black arrows. Source data are provided as a Source Data file.



Supplementary Figure 2. Amino acid sequence alignment of ArlA2 and ArlB *Har. marismortui*. Identical residues are boxed in red; similar residues are shown in red letters; stars mark metal-binding residues. The figure was prepared with ESPrpt3⁴⁰.

ArlA2



20%

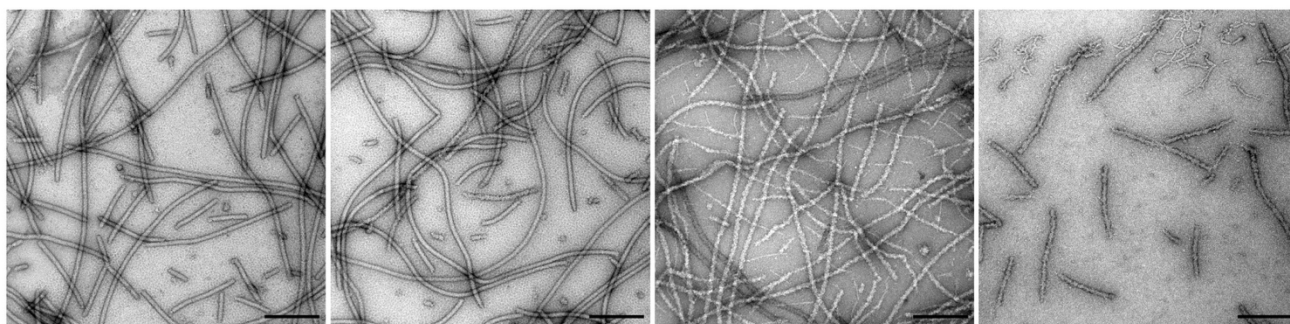
10%

5%

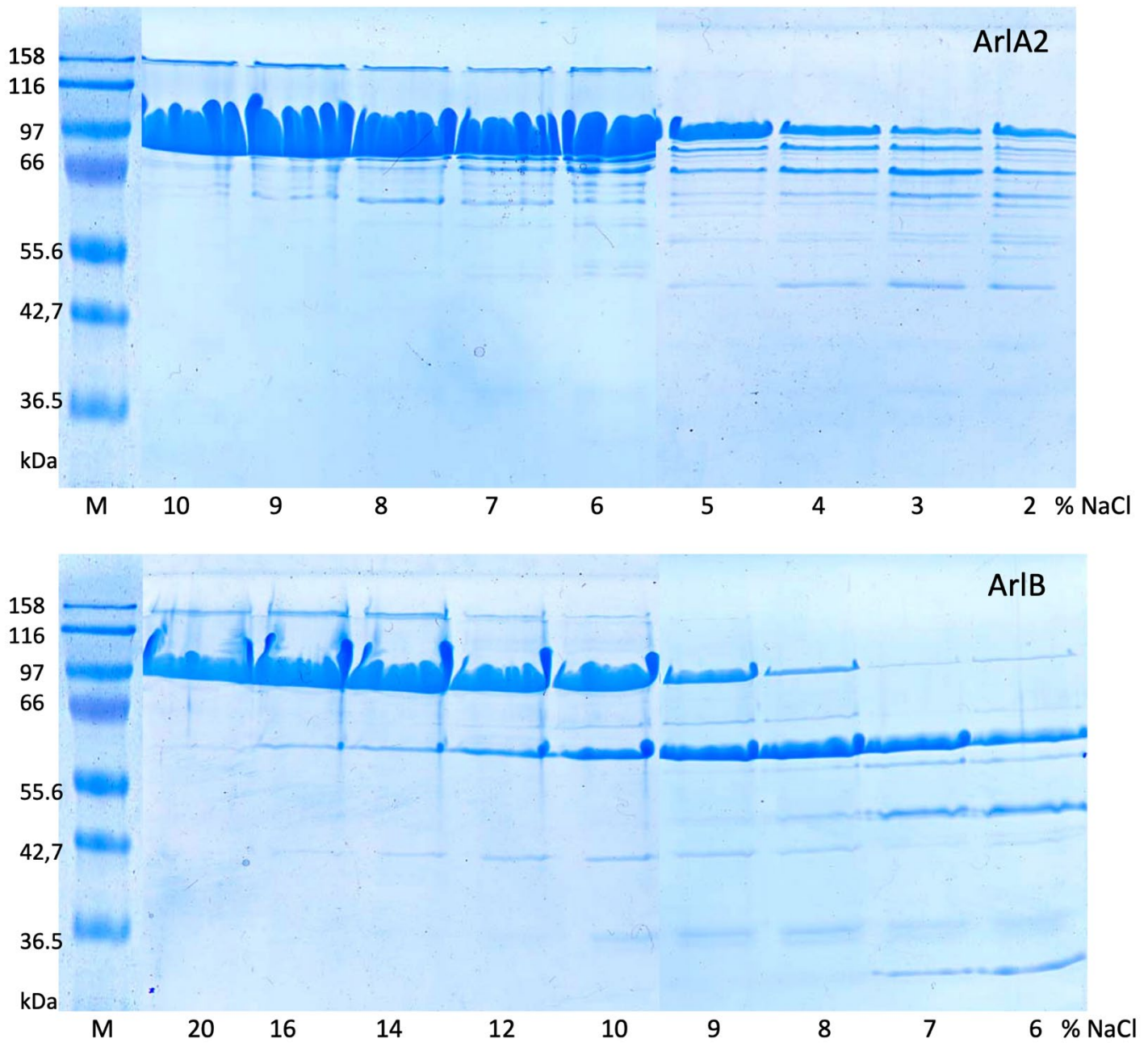
0

NaCl

ArlB

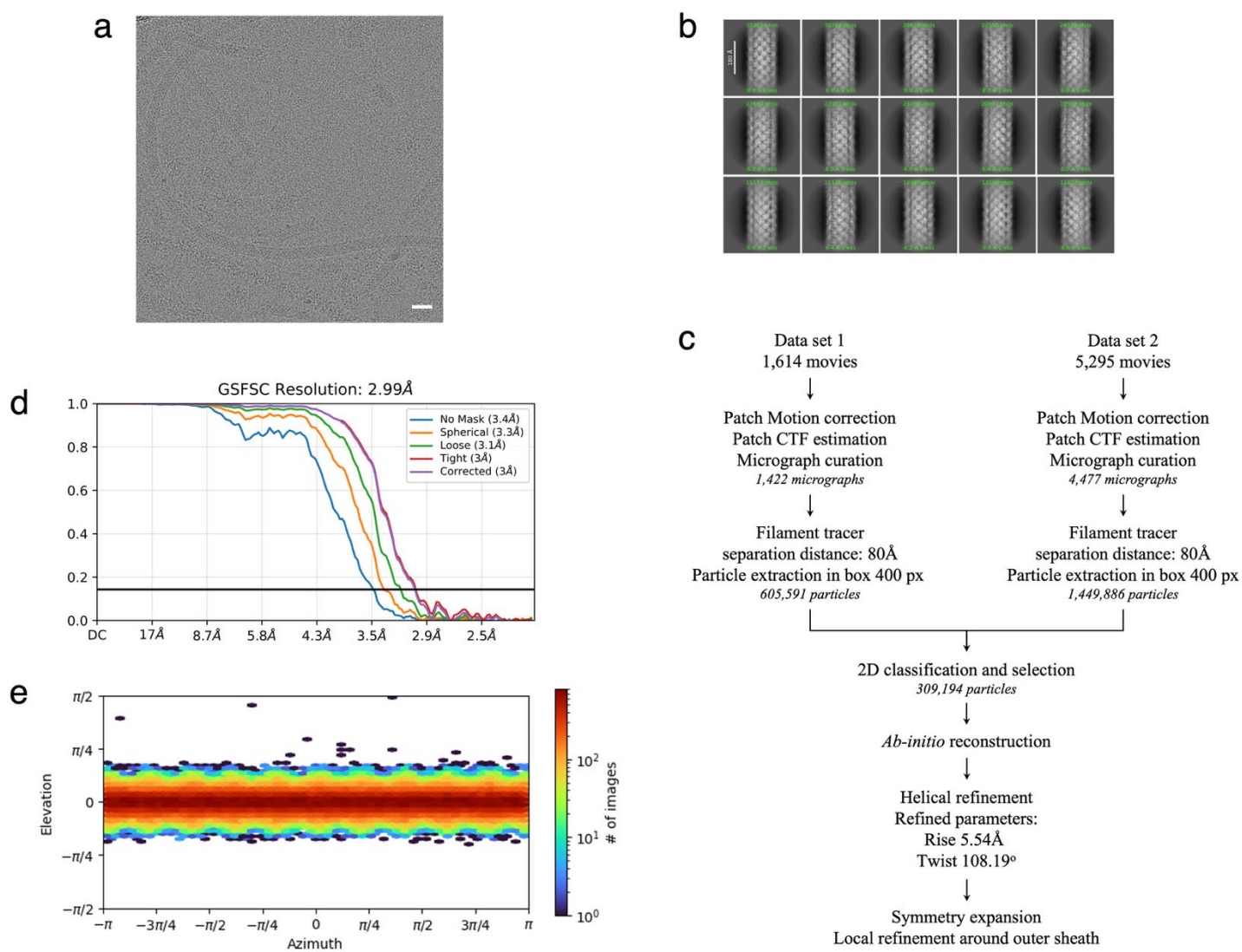


Supplementary Figure 3. Negative-stain EM micrographs of ArlA2 and ArlB filaments at various NaCl concentrations. 5 μ l sample in 10 mM Tris-HCl buffer, pH 8.0 containing appropriate amount of sodium chloride was applied onto glow-discharged carbon grid and stained with 4% uranyl acetate. For 0% NaCl, filaments were absorbed on a carbon grid and washed several times in a drop of 10 mM Tris-HCl buffer, pH 8.0. All samples were imaged using JEOL JEM-1400 Flash transmission electron microscope. Scale bar is 200 nm.

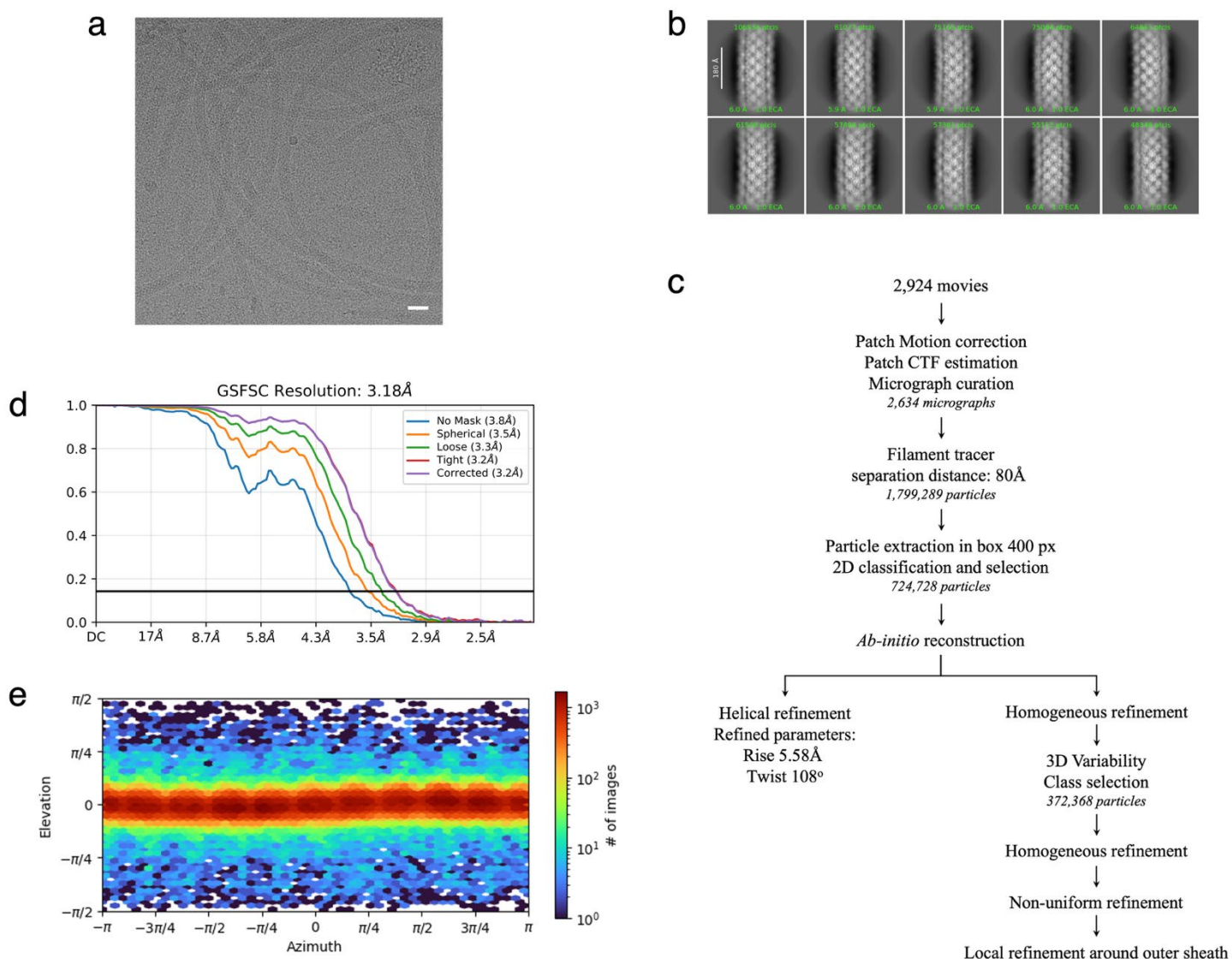


Supplementary Figure 4. Trypsinolysis of ArlA2 and ArlB filaments at various NaCl concentrations.

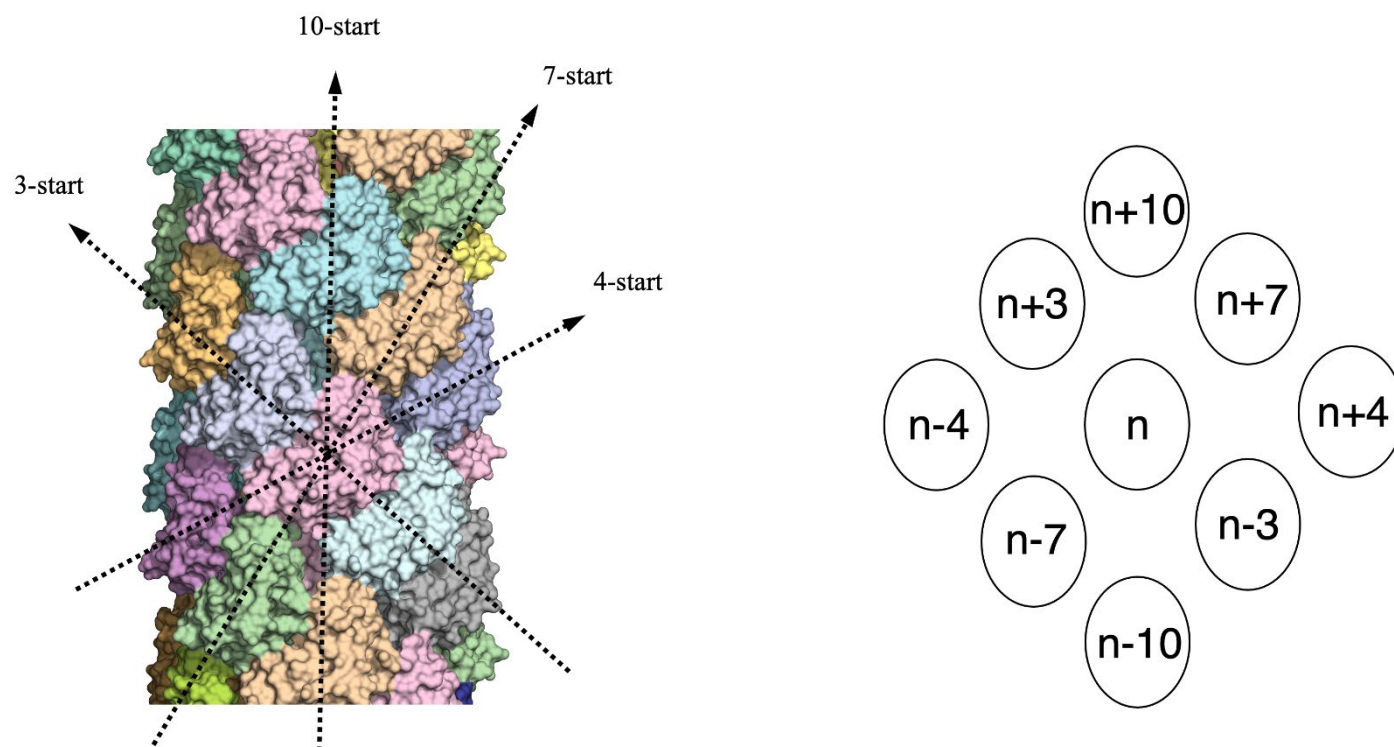
The results of archaella trypsinolysis in 50 mM Tris-HCl, pH 8.0, at various NaCl concentrations are presented. The upper panel shows ArlA2 filaments, while the lower panel shows ArlB filaments. The concentrations of NaCl are indicated on the horizontal axis for each sample in percent. The trypsin treatment time was 30 minutes at room temperature, with a protein/enzyme ratio of 100:1. The results showed that ArlA2 was stable at NaCl concentrations above 5%, and ArlB was stable at concentrations above 10%. However, at 10% NaCl, ArlB archaellin was partially cleaved, but the resulting fragments remained integrated into the filament. At NaCl concentrations below 7.5% NaCl, ArlB filaments disintegrated after trypsin treatment. Original gel images are provided in a Source Data file.



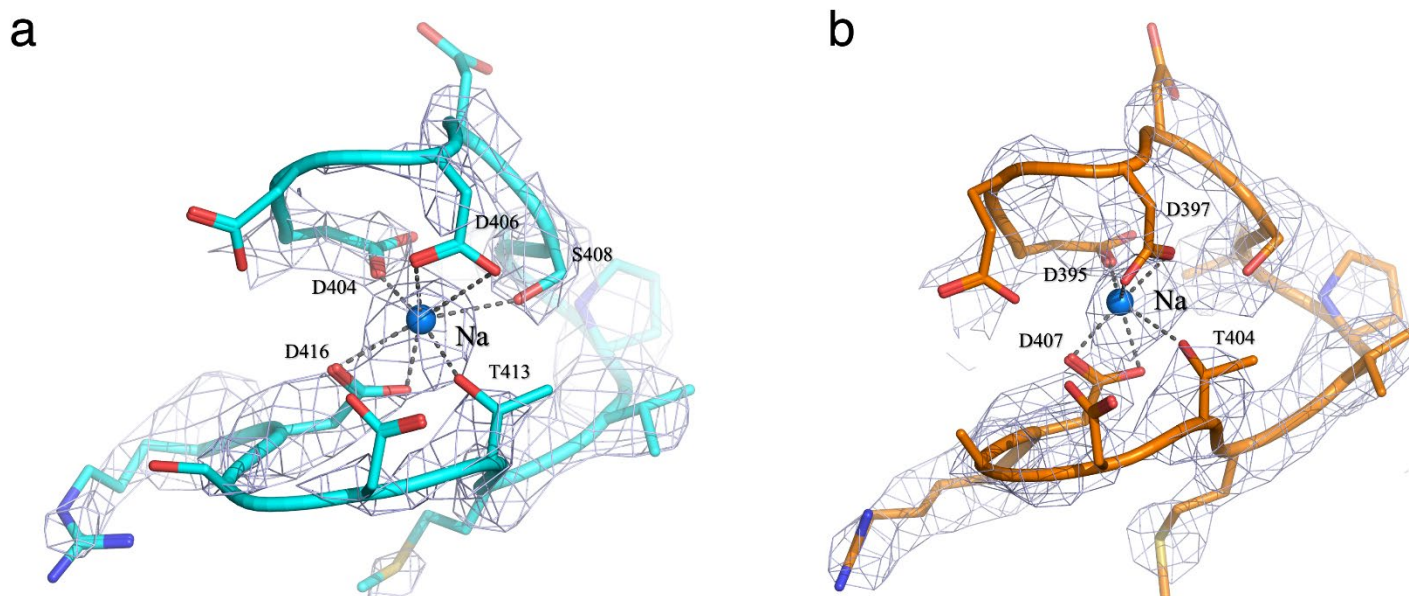
Supplementary Figure 5. Cryo-EM data processing for ArlA2 filaments. a) Representative cryo-EM micrograph. Scale bar is about 400 Å. b) Representative 2D class averages. c) Workflow of cryo-EM data processing. d) Resolution estimation by Fourier Shell Correlation (FSC) of the final cryo-EM reconstruction. e) Angular distribution of particles used in the final reconstruction.



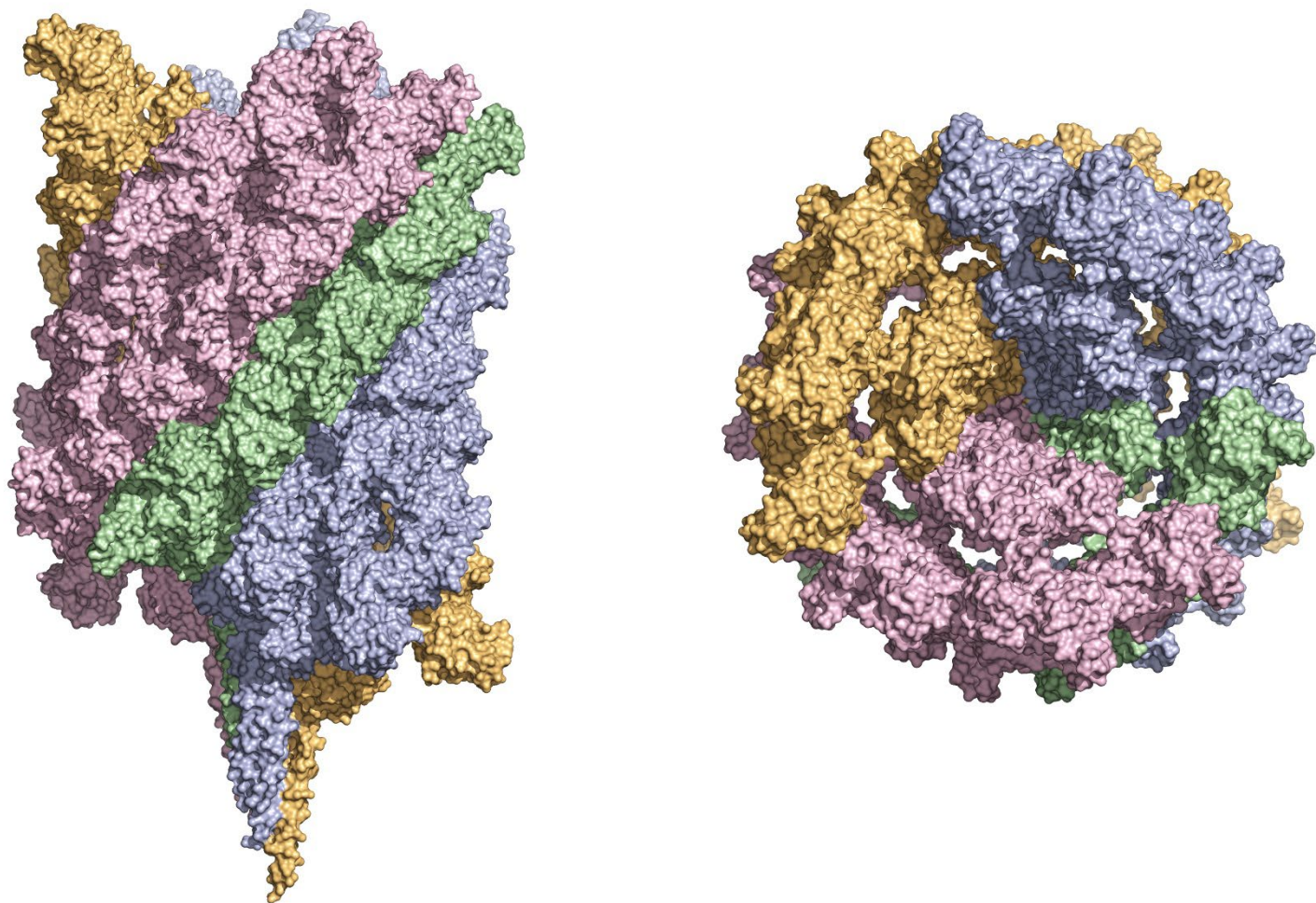
Supplementary Figure 6. Cryo-EM data processing for ArlB filaments. a) Representative cryo-EM micrograph. Scale bar is about 400 Å. b) Representative 2D class averages. c) Workflow of cryo-EM data processing. d) Resolution estimation by Fourier Shell Correlation (FSC) of the final cryo-EM reconstruction. e) Angular distribution of particles used in the final reconstruction.



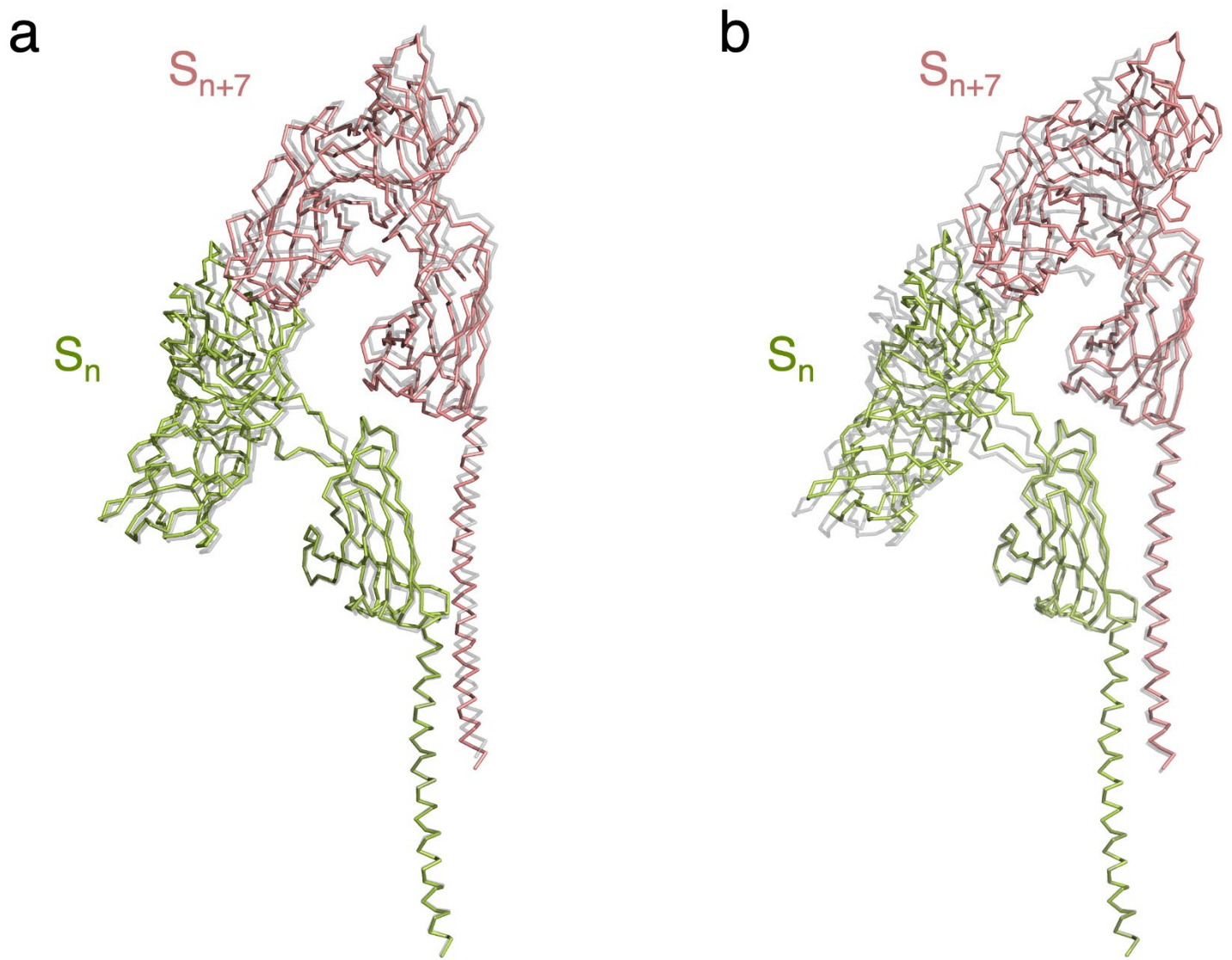
Supplementary Figure 7. Surface representation of the inner core of *Har. marismortui* ArlA2 filament. Different helical directions are depicted.



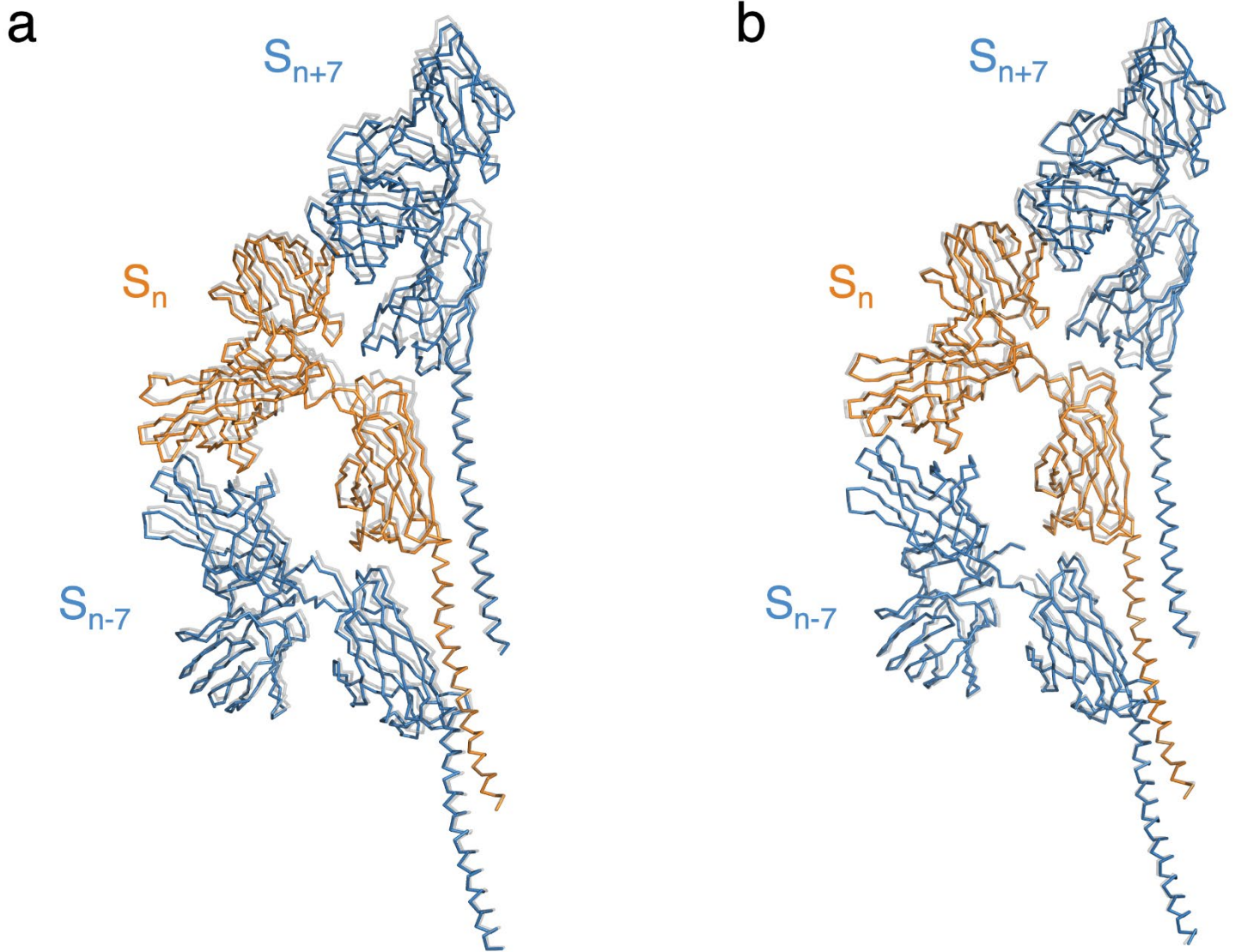
Supplementary Figure 8. Metal-binding sites of ArlA2 (a) and ArlB (b). The map is contoured at 1σ . The metal atom is shown as a blue sphere and coordinating side chains are labeled.



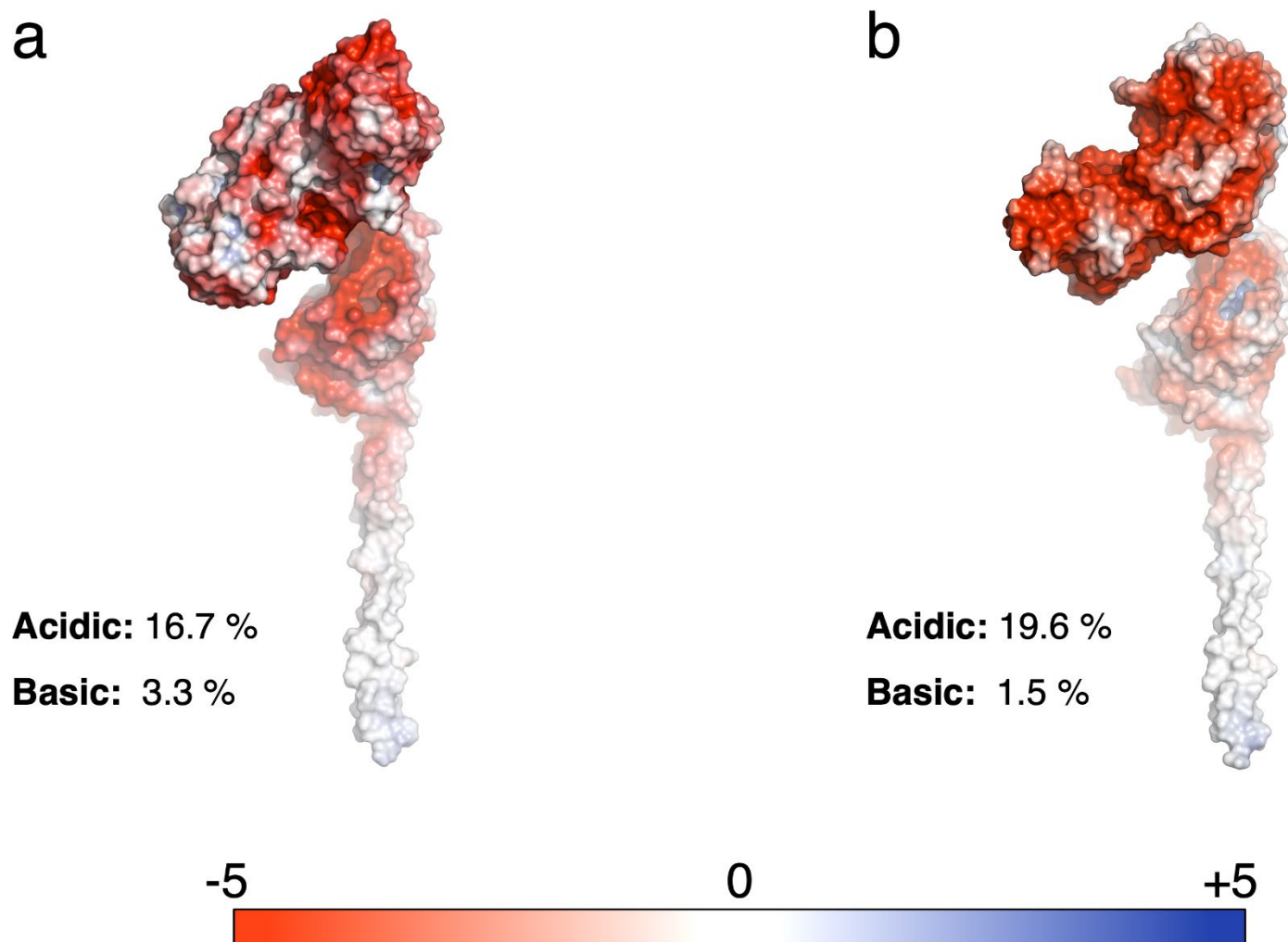
Supplementary Figure 9. Surface representation of the ArlB filament. The outer domain forms strands that wind around the filament surface in the right-handed 7-start direction. Paired strands are colored in pink, blue, and gold. The only unpaired strand is shown in green.



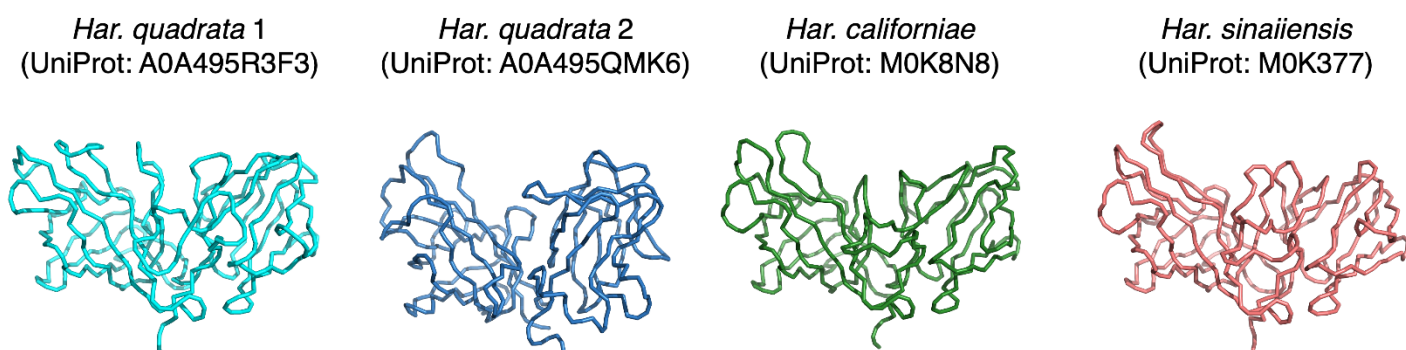
Supplementary Figure 10. 3D variability analysis of ArlA2 filament. a) Movement of the ArlA2 subunits on the curved side of the filament. b) Movement of the subunits positioned 90° from the curved side. Two atomic models along 7-start helical direction were fitted into frame 0 (colored models) and frame19 (gray models) maps from 3DVA. Subunits S_n were aligned with their helical domains D0 using COOT.



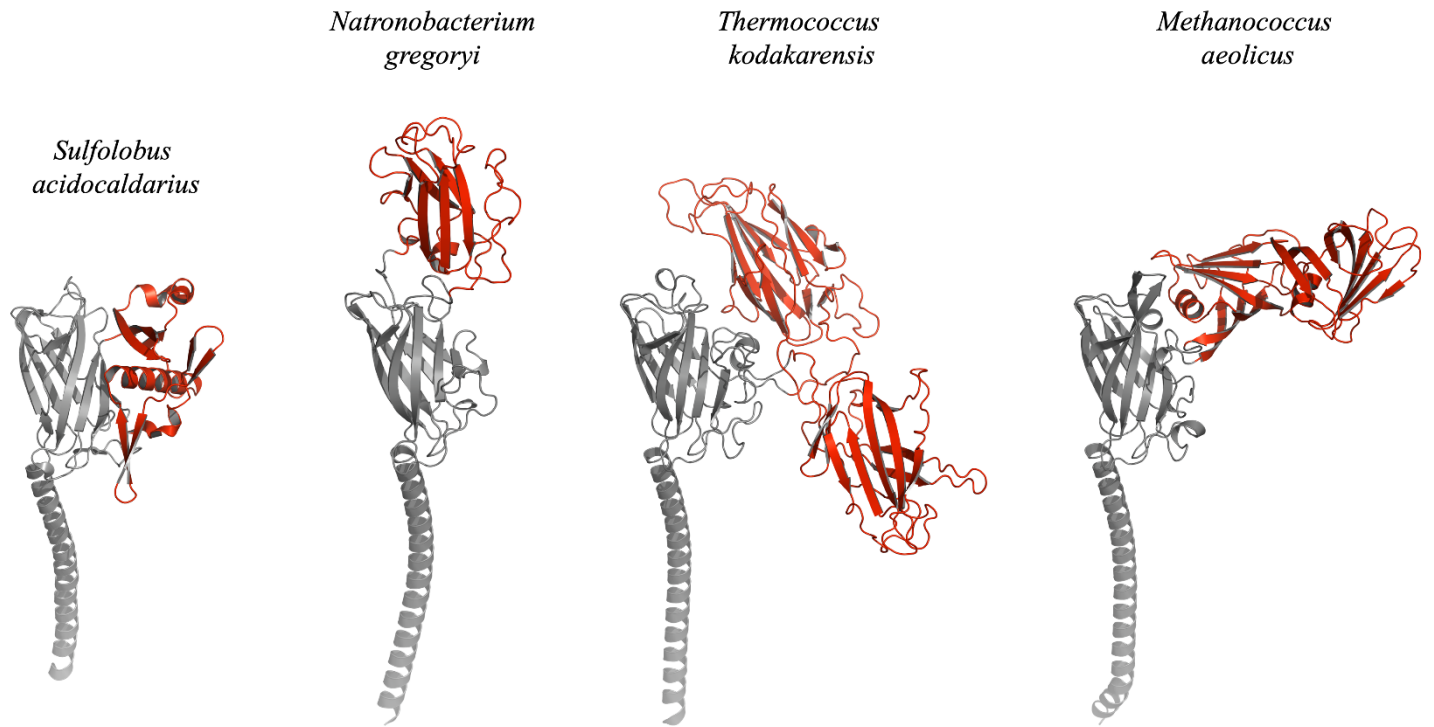
Supplementary Figure 11. 3D variability analysis of ArlB filament. a) Movement of the ArlB subunits on the curved side of the filament. b) Movement of the subunits positioned 90° from the curved side. Three atomic models along 7-start helical direction comprising two different types of interaction within outer domain D2 were fitted into frame 0 (colored models) and frame 19 (gray models) maps from 3DVA. Subunits S_n were aligned with their helical domains D0 using COOT.



Supplementary Figure 12. The electrostatic potential surface of ArlA2 (a) and ArlB (b) monomers was calculated using the APBS software⁴¹ and plotted at $\pm 5 \text{ kT e}^{-1}$. The fractions of acidic and basic amino acids in the outer domain are indicated.



Supplementary Figure 13. AlphaFold3-predicted models of *Har. marismortui* ArlA2/B outer domain homologues of *Haloarcula* species.



Supplementary Figure 14. Examples of the archaeellin outer domains. Experimental model of *Sulfolobus acidocaldarius* archaeellin ArlB (PDB: 8QX4), and AlphaFold3-predicted models of archaeellins from *Natronobacterium gregoryi* (UniProt: L0AJB2), *Thermococcus kodakarensis* KOD1 (UniProt: Q9V2X0), and *Methanococcus aeolicus* (UniProt: A6UTM4). The outer domains are shown in red.