

UCLA

UCLA Previously Published Works

Title

CryoEM structure of the Methanospirillum hungatei archaeellum reveals structural features distinct from the bacterial flagellum and type IV pilus

Permalink

<https://escholarship.org/uc/item/2jj1f617>

Journal

Nature Microbiology, 2(3)

ISSN

2058-5276

Authors

Poweleit, Nicole

Ge, Peng

Nguyen, Hong H

et al.

Publication Date

2017

DOI

10.1038/nmicrobiol.2016.222

Peer reviewed



Published in final edited form as:

Nat Microbiol. ; 2: 16222. doi:10.1038/nmicrobiol.2016.222.

CryoEM structure of the *Methanospirillum hungatei* archaellum reveals structural features distinct from the bacterial flagellum and type IV pili

Nicole Poweleit^{1,2}, Peng Ge^{1,2}, Hong H. Nguyen³, Rachel R. Ogorzalek Loo³, Robert P. Gunsalus^{1,4}, and Z. Hong Zhou^{1,2,*}

¹Department of Microbiology, Immunology, and Molecular Genetics, University of California, Los Angeles (UCLA), Los Angeles, California 90095, USA

²Electron Imaging Center for Nanomachines, California Nano Systems Institute, UCLA, Los Angeles (UCLA), Los Angeles, California 90095, USA

³Department of Chemistry and Biochemistry, UCLA, Los Angeles 90095, UCLA, Los Angeles (UCLA), Los Angeles, California 90095, USA

⁴The UCLA-DOE Institute, UCLA, Los Angeles, California 90095, USA

Abstract

Archaea use flagella known as archaella—distinct both in protein composition and structure from bacterial flagella—to drive cell motility, but the structural basis of this function is unknown. Here, we report an atomic model of the archaella, based on the cryo electron microscopy (cryoEM) structure of the *Methanospirillum hungatei* archaellum at 3.4 Å resolution. Each archaellum contains ~61,500 archaellin subunits organized into a curved helix with a diameter of 10 nm and average length of 10,000 nm. The tadpole-shaped archaellin monomer has two domains, a β-barrel domain and a long, mildly kinked α-helix tail. Our structure reveals multiple post-translational modifications to the archaella, including six O-linked glycans and an unusual N-linked modification. The extensive interactions among neighbouring archaellins explain how the long but thin archaellum maintains the structural integrity required for motility-driving rotation. These extensive inter-subunit interactions and the absence of a central pore in the archaellum distinguish it from both the bacterial flagellum and type IV pili.

Prokaryotes use a variety of extracellular protein complexes to interact with their surrounding environment. Both domains of prokaryotes—archaea and bacteria—express extracellular filaments (called archaella and flagella, respectively) that drive cellular motility

Reprints and permissions information is available at www.nature.com/reprints.

*Correspondence and requests for materials should be addressed to Z.H.Z. Hong.Zhou@UCLA.edu.

Author contributions

Z.H.Z. and R.P.G. designed the project. N.P., P.G., R.R.O.L. and H.H.N. performed the experiments and analysed the data. Z.H.Z., R.P.G. and N.P. wrote the paper. All authors contributed to editing the manuscript.

Supplementary information is available for this paper.

Competing interests

The authors declare no competing financial interests.

in the form of ‘swimming’. In addition to motility^{1–3}, the archaella⁴ (also called archaeal flagella) have been shown to participate in cellular adhesion^{5,6}, biofilm formation⁷ and other symbiotic and environmental activities⁵. Archaella have a significantly smaller diameter, on average, than bacterial flagella (10–12 nm and 20–26 nm, respectively), which suggested significant differences between the two analogous motility filaments⁸. Further studies of archaella gene sequences⁹, archaellin post-translational modifications¹⁰, and archaellin export and assembly^{11–14} have shown a greater similarity between bacterial type IV pili and archaella than between archaeal and bacterial flagella¹, although the structural basis was unknown. Additionally, the energy source archaea use in archaella-driven swimming is adenosine triphosphate instead of a proton motive force¹⁵. However, archaella can grow to 10–20 times as long as the cell⁸, much longer than bacterial pili, and the genes in archaella operons are unique to archaella, with a few proteins with homology to proteins in the bacterial pili operons and no proteins with homology to the bacterial flagella operons^{16,17}. The results from these studies indicate that archaella exhibit some similarity to both bacterial flagella (swimming motility and length) and bacterial type IV pili (energy source and biogenesis), but how archaella use the combination of these traits to drive cellular motility is unknown.

Due to the tendency of purified flagellar and pili subunits to form filaments or aggregates in solution, efforts to determine the structure of the archaella by X-ray crystallography and NMR have been unsuccessful. The best available structure of archaella is that from *Halobacterium salinarium*, which was determined at 8 Å resolution from images recorded on photographic films and is insufficient to resolve the two domains of monomers within the filament^{18,19}. Because the domains appeared to be a central α -helix and an exterior globular domain, it was hypothesized that the archaellin had a ‘type IV pili-like’ structure. Here, we have used direct electron-counting cryo electron microscopy (cryoEM) to (1) determine the structure of an endogenous archaellum at 3.4 Å resolution, (2) build an atomic model of the archaellin, (3) identify novel post-translational modifications to the archaellin, and (4) define inter-archaellin interactions essential to the assembly of the archaellar filament. These findings allow us to establish differences between and similarities with the bacterial flagella and type IV pili.

Results and discussion

Atomic structure of the archaella

Our structural determination of purified archaella involved a complex strategy, beginning with helical parameter estimation in high-contrast EM images of negatively stained sample, intermediate-resolution cryoEM structure from film data and a refined structure from high-resolution images by direct electron counting (for details see Methods). Typical frame-aligned images from a Titan Krios with a Gatan K2 Summit direct electron-counting camera (Fig. 1a and Supplementary Fig. 1) shows barely visible filaments, although their Fourier transforms indicate the existence of data better than 3.8 Å resolution (Fig. 1b and Supplementary Fig. 1). Indeed, characteristic high-resolution features can be seen in two-dimensional class averages (Fig. 1c) and Fourier transforms of such class averages show layer lines consistent with the determined helical symmetry parameters of an axial rise of

5.32 Å and a rotation of 108.02° per subunit (Fig. 1d). The filaments were reconstructed to an averaged resolution of 3.4 Å (Fig. 1e, Supplementary Fig. 2 and Supplementary Table 1). This resolution is consistent with visible structural features such as amino-acid side chains (Supplementary Fig. 3). The cryoEM map is characterized with an α -helical core (yellow to green in Fig. 1e,f), surrounded by an all- β -sheet peripheral (blue to red in Fig. 1e,f and Supplementary Video 1). Both the N terminus and C terminus of a single polypeptide chain are clearly visible, permitting *de novo* tracing of the archaellin monomer without any ambiguity (Fig. 1g,h, Supplementary Fig. 3 and Supplementary Video 1). A cryoEM image (Fig. 1i) of intact cells reveals a distinct cell architecture of the cylindrical *M. hungatei* cell: an S-layer enclosing the totality of the cytoplasm of the cell, an external ridged proteinaceous sheath, only permeable to gases and small solutes^{20–22}, and a characteristic plug for larger substrates to pass into and out of the cell at each end of the cell. Multiple long (10,000 nm on average) but thin (10 nm diameter) archaella attach at each end of the cell (Fig. 1i). Unlike its bacterial counterpart, the archaellum exhibits restricted filament flexibility without uniform super-helicity (Fig. 1a,i and Supplementary Fig. 1). Based on the length and identified helical pitch of the *M. hungatei* archaella, we estimate that each archaellar 10 μ m filament contains, on average, ~61,500 archaellin subunits.

The first six amino-acid residues, MRKETA, predicted from all three *flaB* candidate genes from the *M. hungatei* genome²³, are not present in the cryoEM structure. All archaellins studied to date have type III signal peptides that must be cleaved to form the mature protein^{24–27}. To identify which gene codes the observed archaellin and to determine how many amino acids are cleaved to form the mature archaellin or if the full-length protein is present and simply not resolved in the structure due to flexibility, we performed a series of biochemical analyses of the purified archaella. First, mass spectroscopy analysis of the major sodium dodecyl sulfate (SDS) band (Supplementary Fig. 3a) confirms there is a single protein in the purified archaella, and it is the *flaB3* gene product of Mhun_3140 (Supplementary Table 2). Second, we performed Edman N-terminal sequencing to establish that the functional FlaB3 protein's N-terminal sequence is FSGLE, as seen in the atomic model (Supplementary Table 3). Third, we observed that the protein runs anomalously large on the SDS-PAGE gel based on the predicted molecular weight of Mhun_3140 (17.6 kDa), a phenomenon that often indicates that the protein is post-translationally modified. Taken together, our protein analyses identified FlaB3 protein as the *M. hungatei* archaellin subunit, which is cleaved following position six.

No structure is available for any protein of the archaeal motility and adhesion filament family, and we built the atomic model of the above-identified FlaB3 protein *de novo* based on cryoEM data, without reference to any existing crystal or NMR structures (Fig. 1h and Supplementary Fig. 3b). The amino-acid backbone can clearly be traced, and large amino-acid side chains are readily identifiable (Fig. 1h, Supplementary Fig. 3c–f and Supplementary Video 2).

Each archaellin monomer has a tadpole shape and is organized into two domains (Fig. 1h and Supplementary Fig. 3b): an N-terminal domain (NTD) and a C-terminal domain (CTD). The NTD forms a long, hydrophobic α -helix (Fig. 1h and Supplementary Fig. 3c,d) and the CTD an eight-stranded anti-parallel β -barrel (Fig. 1h and Supplementary Fig. 3d,e). The

hydrophobic α -helices of each subunit point into the centre of the filament and form a hydrophobic core down the helical axis of the filament (Fig. 1e–h and Supplementary Fig. 3c,d). About two-thirds of the helix corresponds to the conserved FlaB sequence²⁸ and one-third extends into the variable domain of the protein, making the α -helical domain significantly longer than predicted (Fig. 1j and Supplementary Fig. 4). The CTD β -barrel decorates the outside of the filament (Fig. 1e–h and Supplementary Fig. 3b–f), and is composed entirely of the variable sequence region of the archaellin (Fig. 1j). As detailed below, this domain plays an important role in the archaella as the site of multiple post-translational modifications and maintains ionic interactions between subunits that help to stabilize the long filament.

Post-translational modification

The FlaB3 cryoEM map has eight sites of extra density per subunit (Fig. 2, Supplementary Fig. 5 and Supplementary Video 3) that form dead-end protrusions off the peptide backbone that could not be accounted for by single amino-acid side chains. All of the sites are located on the exterior surface of the β -barrel domain and are therefore exposed to the environment. Seven of the extra densities are probably glycosylation sites based on their size, shape and location (Fig. 2a). Six sites are potential O-linked glycans at positions Ser59, Ser61, Ser65, Thr100, Ser104 (Fig. 2b) and Thr108 (Supplementary Fig. 5a–c). All previously described glycosylation sites for archaella are N-linked glycans, making the primarily O-linked glycans of the *M. hungatei* archaellin unique among archaellins thus far described. The seventh site is a potentially N-linked glycan at position Asn159 (Fig. 2c and Supplementary Fig. 5e). There is one additional amino acid, Gln134, with an unknown modification (Fig. 2d and Supplementary Fig. 5d). Single or multiple sugar residues can be docked into, and fit well with, all of the extra densities except the Gln134 density, which is far too small to hold a glycan based on the previously described sugar moieties decorating the *Methanococcus maripaludis*²⁹ and *Methanococcus voltae*³⁰ archaellins. The N-glycosidic linkage is particularly unusual because the protein sequence in this region does not contain the canonical N-linked glycosylation sequence of N-X-T/S, nor does it occur anywhere in the protein. The amino-acid sequence of the N-linked glycan of the *M. hungatei* archaellin is N-G, which is an atypical glycosylation site³¹.

These post-translational modifications were further confirmed biochemically. First, a mass ion spectrum (Fig. 2e) demonstrated that the mass of the archaellin is ~20 kDa, ~2.4 kDa more than the size predicted from the peptide sequence. The broadness of the matrix-assisted laser desorption/ionization (MALDI) spectrum peak also suggests heterogeneity in the post-translational modification. Second, a glycostain confirmed that the *M. hungatei* archaellin is glycosylated (Fig. 2f, lane 4). Third, deglycosylation with PNGaseF was performed (Fig. 2f, lane 5), and the *M. hungatei* archaellin showed sensitivity to PNGaseF. These SDS–PAGE and mass spectroscopy analyses confirm the presence and heterogeneity of post-translational modifications, as suggested by the cryoEM archaellum structure.

Interactions between archaellins within the archaellum

The cryoEM structure reveals intermolecular interactions between archaellin subunits that can shed light onto how the extremely long, yet thin, filament maintains physical integrity

under stress caused by rotation-driven motility. Individual archaellin subunits within the *M. hungatei* archaellum form a dense network of interactions, with each individual subunit forming intermolecular interactions with eight other subunits (Fig. 3a–c and Supplementary Video 4). There are three types of intermolecular contact identifiable in the protofilament structure: a radial interaction between Phe1 residues at the archaellum helical axis, which create a spokes effect via a π -stacking sandwich (Fig. 3d–f and Supplementary Video 4), hydrophobic interactions between adjacent α -helices, and ionic interactions between α -helices and β -sheets or between two β -sheets of adjacent subunits (Fig. 3b,e,h and Supplementary Videos 5 and 6). Every subunit in the *M. hungatei* archaellum participates in a π -stacking interaction between the Phe1 residues that creates a single, continuous, one-start helix in the centre of the filament with the aromatic side chains projecting into the centre of the filament and stacking on top of one another (Fig. 3d–f and Supplementary Video 4). Through this interaction, every subunit interacts radially with two other subunits, one aromatic side chain above and one below, to which it is not adjacent in any other area of the protein.

The hydrophobic and ionic intermolecular interactions occur between a subunit and the six subunits physically adjacent to it in the archaellar filament (Figs 3g, 4 and Supplementary Videos 5 and 6). First, the N-terminal hydrophobic α -helix is the defining characteristic of the archaellin and this region is highly conserved. Every subunit within the filament interacts hydrophobically through the conserved portion of the α -helix with six adjacent subunits (Figs 3h,i, 4a–d and Supplementary Video 5). In addition to the hydrophobic interactions, each subunit interacts with the same six adjacent subunits through ionic interactions at four positions around the β -barrel (Fig. 4e), the leftmost (Fig. 4f), upper (Fig. 4g), lower (Fig. 4h) and middle interaction (Fig. 4i) areas (Supplementary Video 6). Second, the ionic interaction between adjacent subunits involves a total of eight amino acids—Tyr22, Gln45, Asp53, Lys75, Glu78, Lys141, Lys143 and Asp156—and is confined to discrete pockets where four separate archaellins come together to participate in the ionic interaction network (Fig. 4c–e and Supplementary Video 6). The complex series of interactions explains the intermolecular stability of the archaellar filament.

Although much is known about the structure, biogenesis and functional operation of bacterial flagella, practically nothing is known about the analogous processes in the archaea. Many fundamental questions remain such as how the archaella are anchored to the cell membrane³², how the thin archaella are able to remain intact under the rotating stress of motility, whether subunits are added at the tip as they are in bacterial flagella or at the base of the filament as with bacterial pili³³, and the roles of glycosylation in filament biogenesis and function^{34–37}. The high-resolution structure presented here begins to answer some of these questions. The complex network of archaellin interactions allows the archaellum to remain stable. The lack of a central channel suggests new subunits are added at the base of the filament as they are in bacterial pili. The surface-exposed glycans could be involved in adhesion and proper filament assembly. In several Euryarchaeota, mutations to glycosylation sites result in non-archaellated cells, suggesting that glycan addition is required for filament biogenesis and/or stability^{34,35}.

Comparison to bacterial motility filaments

In bacteria, two types of filamentous structure contribute to cell motility, bacterial flagella and type IV pili, both of which share some characteristics with archaella. At a global level, the bacterial flagella have a uniform super-helical coiling that generates a waveform to drive motility³⁸; however, our data (Fig. 1a,h and Supplementary Fig. 1) show the bacterial flagellar hallmark does not extend to archaellum that have a non-uniform curvature. Before our current 3.4 Å resolution structure of the archaella, two other motility filament structures were known in atomic detail, the R-type bacterial flagellum from *Salmonella enterica*³⁹, solved to 4 Å resolution by cryoEM, and the type IV pili from *Neisseria gonorrhoeae*⁴⁰, solved to atomic resolution by X-ray crystallography and docked into a 12 Å cryoEM map.

At the monomer level, the archaellin, bacterial pilin and bacterial flagellin structures contain two similar structural elements: a long helix domain in the core of the filament and a β -sheet domain lining the periphery (Fig. 5 and Supplementary Video 7). The helix is hydrophobic in all three structures (Fig. 5a–c), creating the non-specific interactions necessary for slight flexibility during filament twisting and bending. The bulk of the β -sheet domain consists of antiparallel strands and is decorated with glycans on the surface (Fig. 5d,e). Beyond these two common structural elements, the structures have a great deal of variation in monomer organization. Within each type of motility filament, the helix domain has a characteristic conserved protein sequence, but the characteristic sequence is not conserved between motility filaments (Fig. 5f). The canonical helix domain of the bacterial pilin is a single helix with two sharp kinks, the archaellin is a single helix with a single very slight kink, and the bacterial flagellin contains four additional helices (ND0, ND1b, CD0 and CD1), which form a coiled-coil domain (Fig. 5d,e). In bacterial pili, the β -sheet domain contains only six β -strands, lacks greater organization, and faces downward (Fig. 5d). The archaellin β -domain forms an eight-stranded β -barrel projecting upwards, creating a taller monomer (Fig. 5d). The bacterial flagellin has two additional β -sheet domains (D2 and D3) that are perpendicular to the helix domain (Fig. 5e). The increasingly complex monomer structures give rise to increasing filament diameter from bacterial pili (6 nm), to archaella (10 nm) to the bacterial flagella (24 nm) (Fig. 6 and Supplementary Video 8).

How these protomers assemble into a filament differs dramatically between the three structures solved to atomic resolution. The helical parameters differ; for example, the number of subunits per turn is 3.6, 3.3 and 5.5 in the bacterial type IV pilus, archaellum and bacterial flagellum, respectively. To achieve a filament length of ~26 nm, 15 bacterial pilins (Fig. 6a), 26 archaellins (Fig. 6b) or 20 bacterial flagellins (Fig. 6c) are needed. Indeed, the axial views of these three filament types bear no similarity. The hollow central pore of the bacterial pilus measures 6–10 Å in diameter (Fig. 6d), the archaellum does not have a pore (Fig. 6e), and the bacterial flagellum has a 20 Å diameter pore (Fig. 6f). This difference is particularly notable for the two flagella, as the bacterial flagellar pore functions as an export channel with subunits travelling through the channel to be assembled at the filament tip. The absence of such a pore in the archaellum supports the view that archaellins are added to the base of the filament as they are in bacterial pili. Central helices of the three filament types have different angles relative to the filament axis: the bacterial pilin helix is doubly kinked and has a variable offset (Fig. 6d); the archaellin is offset 20° (Fig. 6e); and the bacterial

flagellin ND0 and CD0 domains are nearly parallel to the helical axis (Fig. 6f). The number of α -helix domains visible in a slice equal to the axial rise of one subunit is different. In the bacterial pili seven helix domains are visible (Fig. 6g), in the archaella 12 helix domains are visible (Fig. 6h), and in the bacterial flagella 20 helix bundles from different subunits are visible (Fig. 6i). This high level of structural diversity between the three filaments indicates that the previous hypothesis, that archaella are similar to bacterial type IV pili, is oversimplified. This picture of structural diversity is further complicated by other extracellular filaments such as the bacterial fimbria⁴¹, conjugation pilus⁴² and type II secretion system⁴³ and archaeal adhesion filaments such as the *Ignicoccus hospitalis* Iho670 filament⁴⁴ and the *Sulfolobus acidocaldarius* pili⁴⁵. Indeed, we observed distinct pili at a low frequency in our *M. hungatei* archaella data set (Supplementary Fig. 6). The above identified canonical structural elements shared across bacteria and archaea provide a basis for the prediction of other extracellular filaments for which structural information is not yet available.

From a technical view point, cryoEM structures of archaella and bacterial pili have been hitherto limited in resolution to 8 Å due to the thinness and flexibility of the filaments. The recent technological advances in direct electron-counting technology and helical reconstruction incorporating a Bayesian statistics method have made it possible to solve *de novo* atomic structures by cryoEM of thin, curving filaments. Solving the *in situ* archaella structure has also led to the direct observation of post-translational modifications of the *M. hungatei* archaellin. With this high-resolution cryoEM technology and the canonical structure features reported here, microbiologists are now positioned to better understand the biogenesis and operational mechanism of a myriad of bacterial and archaeal filaments involved in diverse functions, such as ‘sex’ conjugation, charge conduction and toxin secretion.

Note added in proof: During the final review stage of this paper, we noted the publication of a paper by Braun *et al.*⁴⁶ that reports an ~4 Å structure of an adhesion filament from *Ignicoccus hospitalis*. This filament is archaeallum-like in appearance, but *I. hospitalis* is not motile and has no archaeallum operon. There are major differences between our structure reported here and the *I. hospitalis* structure. First, our atomic model is complete while the latter has 60 amino acids missing. Second, all post-translational modifications have been resolved in our structure. Third, the helical parameters differ significantly (108.02° rotation per subunit and 5.32 Å axial rise per subunit in our structure versus 106.65° rotation per subunit and 5.45 Å axial rise per subunit in the latter).

Methods

M. hungatei growth and archaella isolation

Methanospirillum hungatei strain JF1 (ATCC# 27890) was cultured in 1 l anaerobic bottles containing 500 ml of Jarrell methanogen acetate medium modified to contain half the normal concentration of phosphate to elevate archaella abundance⁴⁷. Bottles were inoculated (0.5% vol) with exponential phase cells and the vessel headspace was then pressurized (10 p.s.i.) with an 80:20 (vol/vol) mixture of H₂:CO₂. Cultures were incubated at 37 °C with gentle agitation for one week prior to cell harvest⁴⁸.

To isolate archaeella, *M. hungatei* cells were collected by centrifugation at 5 K for 10 min. Cells were resuspended in 2 ml of Tris-buffered saline (TBS: 100 mM NaCl, 50 mM Tris at pH 7.2) before passage through an 18 g syringe needle 30 times. Cells were pelleted at 5 K for 10 min and the supernatant fraction was saved while the cells were resuspended in TBS to be sheared via syringe twice more. All supernatant fractions were pooled and the archaeella were collected by ultracentrifugation in a Beckman TLA100.2 rotor at 60 K for 20 min. The archaeella were resuspended in 50 μ l of TBS for a final 5 K spin for 10 min to remove unwanted cell debris. New cultures were grown, and samples were prepared fresh for each experiment.

SDS-PAGE and glycostaining

M. hungatei archaeella were incubated for 1 h in 0.1% Triton X-100 (ref. 47), and subjected to SDS-PAGE using 4–12% Bis-Tris gradient gels (Invitrogen) according to the manufacturer's instructions. Glycoprotein stains were performed using the Pro-Q Emerald 488 Glycoprotein Stain Kit from Thermo-Fischer (P21855) and incorporating CandyCane glycoprotein molecular weight standards (Molecular Probes) (Supplementary Fig. 7).

Mass spectroscopy and N-terminal identification

Bands excised from an SDS-PAGE gel were digested with trypsin in-gel. The proteolytic peptides were injected for analysis by liquid chromatography tandem-mass spectrometry (LC-MS/MS) on a Thermo-Fisher Q-Exactive. For molecular mass determination, the *M. hungatei* archaeella were first dissociated in Triton X-100 as described above. The dissociated archaeella were mixed with a ferulic acid matrix and deposited on a sample stage for analysis by MALDI mass spectrometry (MS). MALDI-MS was performed in positive ion mode on an Applied Biosystems Voyager-DE-STR MALDI-TOF with 337 nm irradiation. N-terminal identification was performed on archaeellin using Edman sequencing as previously described⁴⁹.

Bioinformatics

Protein sequence alignments were performed with Clustal Omega⁵⁰ using default parameters. Percent identity scores were generated with BLASTP⁵¹. Parameters were changed to compensate for the short length of the protein sequence; specifically, the word size was reduced to 2 and the expected threshold was reduced to 8. Protein secondary structure was predicted *in silico* using psipred⁵².

Negative stain electron microscopy

Purified archaeella (4 μ l) were pipetted onto carbon-coated copper grids and allowed to settle for 1 min before blotting to remove excess sample. A 2 μ l volume of 1% uranyl acetate was added for 1 min and excess stain was blotted away. Data were collected using the FEI TF20 microscope with a 16-megapixel charge-coupled device (CCD) camera and the automated imaging software Leginon⁵³ (Supplementary Fig. 8a–c). For cryoEM samples, the concentration of the *M. hungatei* archaeella was first optimized by negatively staining the sample with 1% uranyl acetate and imaged in an FEI TF12 microscope with a Gatan 2k \times 2k

CCD camera. Samples were diluted as needed to achieve the appropriate sample concentration.

CryoEM

Three successive cryoEM data sets were collected using the same FEI Titan Krios microscope, a data set on photographic film (Supplementary Fig. 8d–f), a test data set collected with the FEI Falcon II direct electron detector (Supplementary Fig. 8g–i), and the final data set used to reconstruct and build the *de novo* atomic model described in the paper collected with the Gatan K2 Summit direct electron detector (Supplementary Fig. 8j–l). For all three data sets, an aliquot of 2–4 μl of the sample was applied to Quantifoil grids and an FEI Vitrobot was used to plunge-freeze the sample in liquid ethane cooled by liquid nitrogen.

For the first cryoEM data set collected on photographic films, an accelerating voltage of 120 kV was used at a magnification of $\times 59,000$ and the defocus values of the images collected were between 2.5 and 4.0 μm . The automated imaging software Legikon was used to collect data. A total of 1,014 images were collected on film over 12 days and digitized using a Nikon Super Coolscan 9000, giving a pixel size of 1.02 \AA per pixel on the specimen (Supplementary Fig. 8d). This data set is very noisy, and the thin filaments are difficult to see in the raw micrographs. However, at the time, we thought the flexibility and lack of contrast provided by the archaeella limited the resolution of the refinement, not the use of film.

The cryoEM data set collected using the Falcon II electron detector in a pre-GIF configuration was a test data set of about 800 videos collected over a day and a half. Each video contains 24 frames with a total exposure of 8 s and a total dose of about 40 electrons per \AA^2 per pixel. Direct electron detectors collect a series of very low-dose images that are combined into a video. The frames in the video are then aligned to correct for sample drift and beam-induced damage using whole frame alignment using the MotionCorr program⁵⁴. The accelerating voltage was 300 kV, the magnification was 97.3k with a pixel size of 1.43 \AA per pixel on the specimen, and the defocus values of the images collected were between 1.5 and 3 μm . The FEI automated imaging software EPU was used to collect data (Supplementary Fig. 8g). Imaging with the Falcon II provided a marked improvement in data quality over the previous film and negative stain data sets.

The final data set was recorded on a Gatan K2 Summit electron detector attached to a Gatan imaging filter (GIF Quanta) in counting mode with the automated imaging software Legikon (Supplementary Fig. 8j). To improve frame alignment, we also added gold fiducials (a 1:50 dilution of 10 nm gold fiducials was added to the optimized samples before freezing). Each video contains 30 frames with a total exposure of 7.5 s and an accumulated dosage of ~ 80 electrons per \AA^2 per video, and frames were aligned using the MotionCorr program. A total of 3,300 videos were collected over four days with an accelerating voltage of 300 kV, with a pixel size of 1.3 \AA per pixel on the specimen, and defocus between 1.0 and 2.5 μm .

Helical parameters

Identification of helical parameters went through multiple iterations as the quality of available data improved. First, we sought to generate a high-resolution layer-line image using cryoEM images from the photographic film data set (Supplementary Fig. 8f). However, the low contrast and high noise in higher-resolution regions prevented sound alignment of filament segments.

Next, to improve the contrast, we used negatively stained micrographs to generate a layer-line image (Supplementary Fig. 8c). However, we were unable to determine the n numbers of many reflections, including the layer-lines $l = 1$ and $l = 2$, leading to major ambiguity of helical parameter assignment. Applying these parameters to the cryoEM data set on photographic films, we generated multiple possible structures, but we were unable to improve these structures beyond 7 Å resolution in any possible helical assignment.

The helical parameters were eventually established with direct electron detector data. Helical parameters were initially guessed from Fourier transforms of two-dimensional class averages (Supplementary Fig. 8i,l) by classical helical cryoEM methods⁵⁵. Briefly, main reflections on the incoherently averaged Fourier transforms of two-dimensional class averages are aligned on a two-dimensional lattice (mesh) and their Bessel orders (n , l) are assigned empirically. The selection rule is therefore deduced from these n , l numbers and is converted to the format of turn and rise per subunit for processing with IHRSR (see section on image processing for details).

More importantly, the final structure shows amino acid side chains, providing a series of solid internal controls for the established helical parameters by matching the densities to the primary sequence.

Image processing packages used

As detailed in the reconstruction process below, we used a combination of different software packages for helical reconstructions. MotionCorr was used for whole frame alignment, EMAN⁵⁶ for image processing, particle boxing and two-dimensional classification, IHRSR (ref. 57) for low-resolution helical reconstruction, a modified EMAN package implemented with the IHRSR method for the 7 Å resolution reconstruction⁵⁸, and lastly a modified RELION⁵⁹ with IHRSR method for both the 4.8 Å resolution reconstruction and the 3.4 Å resolution reconstruction reported in this paper⁶⁰.

Image processing and structure refinement

For both the negatively stained and cryoEM on photographic film data sets, image processing was performed using EMAN 1.8 and IHRSR. CTF parameters were determined using ctffind3 and particles were boxed out with *heliboxer* in Eman 1.8 using a box size of 432 pixels. Two-dimensional classification was performed on both data sets in EMAN. The negative stain data set was processed first to estimate the helical parameters as described above. The helicity determined from the negative stain data and the initial model were used to refine the cryoEM data set collected on photographic films to ~7 Å resolution (Supplementary Fig. 9a).

Frames from the videos collected using the Falcon II electron detector were aligned with the first and last six frames discarded for a total exposure dose of about 30 electrons per \AA^2 . CTF correction was performed using *ctffind3* and particles were boxed using *helixboxer* in EMAN 1.8 using a box size of 256 pixels. Reducing the box size mitigated the effects of filament curvature on the reconstruction. We use RELION for two-dimensional classification. We modified RELION to utilize the IHRSR method for three-dimensional classification and final refinement (Supplementary Fig. 9b).

Frames from the videos collected using the K2 Summit electron detector were first aligned. All frames were used to determine the CTF with *ctffind3* and box particles in EMAN1.8 using a box size of 256 pixels. For the final reconstruction, the first two frames were discarded and the following frames with a total dose of 20 electrons per \AA^2 were used. As described above, particles were boxed out with EMAN and three-dimensional structures were determined with the modified RELION (Fig. 1e and Supplementary Fig. 9c).

Once the final cryoEM density map was obtained, the protein backbone of the archaellin monomer was traced and amino acids assigned in Coot⁶¹ using the baton building method. First, a polyalanine backbone was built starting from the easily identifiable C terminus. Amino-acid side chains were then mutated based on the available sequence. Initial refinement was performed on a single subunit using pseudo crystallographic restraints with the PHENIX⁶² and CNS software packages. Chi angles and rotomers were manual adjusted in Coot, and further refinement was performed in PHENIX using a protofilament of 26 subunits. Figures and videos were created in Chimera⁶³ and Filmora.

Data availability

The data that support the findings of this study are available from the corresponding author upon request. The cryoEM map and the atomic model have been deposited in the Protein Data Bank and EMDB under accession numbers 5TFY and EMD-8405, respectively.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This project received support from National Institutes of Health grants GM071940 and AI094386, NIH/NCRR/NCATS UCLA CTSI grant UL1TR000124, from the UCLA-DOE Institute (DE-FC03-02ER6342) to R.P.G. and R.O.L., and NSF grants DMR-1548924 to Z.H.Z. and 1515843 to R.P.G. N.P. was supported in part by the NIH Biotechnology Training Grant Program (T32GM067555). P.G. was supported in part by an American Heart Association Western States Affiliates Postdoc Fellowship (13POST17340020). The authors acknowledge the use of instruments at the Electron Imaging Center for Nanomachines supported by UCLA and by instrumentation grants from NIH (1S10OD018111) and NSF (DBI-1338135). NIH support for mass spectrometry was provided by grant S10RR025600. The authors acknowledge computer time at the Extreme Science and Engineering Discovery Environment (XSEDE, grant MCB140140 to Z.H.Z.).

References

1. Albers SV, Jarrell KF. The archaellum: how archaea swim. *Front Microbiol.* 2015; 6:23. [PubMed: 25699024]

2. Näther DJ, Rachel R, Wanner G, Wirth R. Flagella of *Pyrococcus furiosus*: multifunctional organelles, made for swimming, adhesion to various surfaces, and cell–cell contacts. *J Bacteriol.* 2006; 188:6915–6923. [PubMed: 16980494]
3. Szabó Z, et al. Flagellar motility and structure in the hyperthermoacidophilic archaeon *Sulfolobus solfataricus*. *J Bacteriol.* 2007; 189:4305–4309. [PubMed: 17416662]
4. Jarrell KF, Albers SV. The archaellum: an old motility structure with a new name. *Trends Microbiol.* 2012; 20:307–312. [PubMed: 22613456]
5. Bellack A, Huber H, Rachel R, Wanner G, Wirth R. *Methanocaldococcus villosus* sp nov., a heavily flagellated archaeon that adheres to surfaces and forms cell–cell contacts. *Int J Syst Evol Micr.* 2011; 61:1239–1245.
6. Jarrell KF, Stark M, Nair DB, Chong JPI. Flagella and pili are both necessary for efficient attachment of *Methanococcus maripaludis* to surfaces. *FEMS Microbiol Lett.* 2011; 319:44–50. [PubMed: 21410509]
7. Schopf S, Wanner G, Rachel R, Wirth R. An archaeal bi-species biofilm formed by *Pyrococcus furiosus* and *Methanopyrus kandleri*. *Arch Microbiol.* 2008; 190:371–377. [PubMed: 18438643]
8. Alam M, Oesterhelt D. Morphology, function and isolation of halobacterial flagella. *J Mol Biol.* 1984; 176:459–475. [PubMed: 6748081]
9. Faguy DM, Jarrell KF, Kuzio J, Kalmokoff ML. Molecular analysis of archaeal flagellins: similarity to the type IV pilin-transport superfamily widespread in bacteria. *Can J Microbiol.* 1994; 40:67–71. [PubMed: 7908603]
10. Correia JD, Jarrell KF. Posttranslational processing of *Methanococcus voltae* preflagellin by preflagellin peptidases of *M. voltae* and other methanogens. *J Bacteriol.* 2000; 182:855–858. [PubMed: 10633127]
11. Patenge N, Berendes A, Engelhardt H, Schuster SC, Oesterhelt D. The *fla* gene cluster is involved in the biogenesis of flagella in *Halobacterium salinarum*. *Mol Microbiol.* 2001; 41:653–663. [PubMed: 11532133]
12. Thomas NA, Pawson CT, Jarrell KF. Insertional inactivation of the *flaH* gene in the archaeon *Methanococcus voltae* results in non-flagellated cells. *Mol Genet Genomics.* 2001; 265:596–603. [PubMed: 11459179]
13. Chaban B, et al. Systematic deletion analyses of the *fla* genes in the flagella operon identify several genes essential for proper assembly and function of flagella in the archaeon, *Methanococcus maripaludis*. *Mol Microbiol.* 2007; 66:596–609. [PubMed: 17887963]
14. Lassak K, et al. Molecular analysis of the crenarchaeal flagellum. *Mol Microbiol.* 2012; 83:110–124. [PubMed: 22081969]
15. Streif S, Staudinger WF, Marwan W, Oesterhelt D. Flagellar rotation in the archaeon *Halobacterium salinarum* depends on ATP. *J Mol Biol.* 2008; 384:1–8. [PubMed: 18786541]
16. Banerjee A, et al. FlaF is a β -sandwich protein that anchors the archaellum in the archaeal cell envelope by binding the S-layer protein. *Structure.* 2015; 23:863–872. [PubMed: 25865246]
17. Chaudhury P, et al. The nucleotide-dependent interaction of FlaH and FlaI is essential for assembly and function of the archaellum motor. *Mol Microbiol.* 2016; 99:674–685. [PubMed: 26508112]
18. Trachtenberg S, Galkin VE, Egelman EH. Refining the structure of the *Halobacterium salinarum* flagellar filament using the iterative helical real space reconstruction method: insights into polymorphism. *J Mol Biol.* 2005; 346:665–676. [PubMed: 15713454]
19. Trachtenberg S, Cohen-Krausz S. The archaeobacterial flagellar filament: a bacterial propeller with a pilus-like structure. *J Mol Microb Biotech.* 2006; 11:208–220.
20. Beveridge TJ, Sprott GD, Whippey P. Ultrastructure, inferred porosity, and Gram-staining character of *Methanospirillum hungatei* filament termini describe a unique cell permeability for this archaeobacterium. *J Bacteriol.* 1991; 173:130–140. [PubMed: 1702777]
21. Xu W, et al. Modeling and measuring the elastic properties of an archaeal surface, the sheath of *Methanospirillum hungatei*, and the implication of methane production. *J Bacteriol.* 1996; 178:3106–3112. [PubMed: 8655487]
22. Toso DB, Henstra AM, Gunsalus RP, Zhou ZH. Structural, mass and elemental analyses of storage granules in methanogenic archaeal cells. *Environ Microbiol.* 2011; 13:2587–2599. [PubMed: 21854518]

23. Gunsalus RP, et al. Complete genome sequence of *Methanospirillum hungatei* type strain JF1. *Standards Genomic Sci.* 2016; 11:2.
24. Bardy SL, Jarrell KF. Flak of the archaeon *Methanococcus maripaludis* possesses preflagellin peptidase activity. *FEMS Microbiol Lett.* 2002; 208:53–59. [PubMed: 11934494]
25. Bardy SL, Jarrell KF. Cleavage of preflagellins by an aspartic acid signal peptidase is essential for flagellation in the archaeon *Methanococcus voltae*. *Mol Microbiol.* 2003; 50:1339–1347. [PubMed: 14622420]
26. Szabó Z, et al. Identification of diverse archaeal proteins with class III signal peptides cleaved by distinct archaeal prepilin peptidases. *J Bacteriol.* 2007; 189:772–778. [PubMed: 17114255]
27. Albers SV, Szabó Z, Driessen AJM. Archaeal homolog of bacterial type IV prepilin signal peptidases with broad substrate specificity. *J Bacteriol.* 2003; 185:3918–3925. [PubMed: 12813086]
28. Ng SYM, Chaban B, Jarrell KF. Archaeal flagella, bacterial flagella and type IV pili: a comparison of genes and posttranslational modifications. *J Mol Microbiol Biotech.* 2006; 11:167–191.
29. Kelly J, Logan SM, Jarrell KF, VanDyke DJ, Vinogradov E. A novel N-linked flagellar glycan from *Methanococcus maripaludis*. *Carbohydr Res.* 2009; 344:648–653.
30. Voisin S, et al. Identification and characterization of the unique N-linked glycan common to the flagellins and S-layer glycoprotein of *Methanococcus voltae*. *J Biol Chem.* 2005; 280:16586–16593. [PubMed: 15723834]
31. Sun S, Zhang H. Identification and validation of atypical N-glycosylation sites. *Anal Chem.* 2015; 87:11948–11951. [PubMed: 26593774]
32. Banerjee A, Neiner T, Tripp P, Albers SV. Insights into subunit interactions in the *Sulfolobus acidocaldarius* archaeallum cytoplasmic complex. *FEBS J.* 2013; 280:6141–6149. [PubMed: 24103130]
33. Jarrell KF, Bayley DP, Kostyukova AS. The archaeal flagellum: a unique motility structure. *J Bacteriol.* 1996; 178:5057–5064. [PubMed: 8752319]
34. Tripepi M, et al. N-glycosylation of *Haloflex volcanii* flagellins requires known Agl proteins and is essential for biosynthesis of stable flagella. *J Bacteriol.* 2012; 194:4876–4887. [PubMed: 22730124]
35. Ding Y, et al. Effects of N-glycosylation site removal in archaeallins on the assembly and function of archaeella in *Methanococcus maripaludis*. *PLoS ONE.* 2015; 10:e0116402. [PubMed: 25700084]
36. Meyer BH, et al. Agl16, a thermophilic glycosyltransferase mediating the last step of N-glycan biosynthesis in the thermoacidophilic crenarchaeon *Sulfolobus acidocaldarius*. *J Bacteriol.* 2013; 195:2177–2186. [PubMed: 23475978]
37. Meyer BH, Birich A, Albers SV. N-glycosylation of the archaeallum filament is not important for archaeella assembly and motility, although N-glycosylation is essential for motility in *Sulfolobus acidocaldarius*. *Biochimie.* 2015; 118:294–301. [PubMed: 25447136]
38. Turner L, Ryu WS, Berg HC. Real-time imaging of fluorescent flagellar filaments. *J Bacteriol.* 2000; 182:2793–2801. [PubMed: 10781548]
39. Yonekura K, Maki-Yonekura S, Namba K. Complete atomic model of the bacterial flagellar filament by electron cryomicroscopy. *Nature.* 2003; 424:643–650. [PubMed: 12904785]
40. Craig L, et al. Type IV pilus structure by cryo-electron microscopy and crystallography: implications for pilus assembly and functions. *Mol Cell.* 2006; 23:651–662. [PubMed: 16949362]
41. Mortezaei N, et al. Structure and function of enterotoxigenic *Escherichia coli* fimbriae from differing assembly pathways. *Mol Microbiol.* 2015; 95:116–126. [PubMed: 25355550]
42. Silverman PM. Towards a structural biology of bacterial conjugation. *Mol Microbiol.* 1997; 23:423–429. [PubMed: 9044277]
43. McLaughlin LS, Haft RJF, Forest KT. Structural insights into the type II secretion nanomachine. *Curr Opin Struc Biol.* 2012; 22:208–216.
44. Yu X, et al. Filaments from *Ignicoccus hospitalis* show diversity of packing in proteins containing N-terminal type IV pilin helices. *J Mol Biol.* 2012; 422:274–281. [PubMed: 22659006]
45. Henche AL, et al. Structure and function of the adhesive type IV pilus of *Sulfolobus acidocaldarius*. *Environ Microbiol.* 2012; 14:3188–3202. [PubMed: 23078543]

46. Braun T, et al. Archaeal flagellin combines a bacterial type IV pilin domain with an Ig-like domain. *Proc Natl Acad Sci USA*. 2016; 113:10352–10357. [PubMed: 27578865]
47. Faguy DM, Koval SF, Jarrell KF. Effect of changes in mineral composition and growth temperature on filament length and flagellation in the Archaeon *Methanospirillum hungatei*. *Arch Microbiol*. 1993; 159:512–520.
48. Patel GB, Roth LA, Berg Lvd, Clark DS. Characterization of a strain of *Methanospirillum hungatii*. *Can J Microbiol*. 1976; 22:1404–1410. [PubMed: 10074]
49. Edman P, Begg G. A protein sequenator. *Eur J Biochem*. 1967; 1:80–91. [PubMed: 6059350]
50. Sievers F, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol*. 2011; 7:539. [PubMed: 21988835]
51. Altschul SF, et al. Protein database searches using compositionally adjusted substitution matrices. *FEBS J*. 2005; 272:5101–5109. [PubMed: 16218944]
52. Jones DT. Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol*. 1999; 292:195–202. [PubMed: 10493868]
53. Suloway C, et al. Fully automated, sequential tilt-series acquisition with Leginon. *J Struct Biol*. 2009; 167:11–18. [PubMed: 19361558]
54. Li X, et al. Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat Methods*. 2013; 10:584–590. [PubMed: 23644547]
55. DeRosier DJ, Moore PB. Reconstruction of three-dimensional images from electron micrographs of structures with helical symmetry. *J Mol Biol*. 1970; 52:355–369. [PubMed: 5485914]
56. Ludtke SJ, Baldwin PR, Chiu W. EMAN: semiautomated software for high-resolution single-particle reconstructions. *J Struct Biol*. 1999; 128:82–97. [PubMed: 10600563]
57. Egelman EH. The iterative helical real space reconstruction method: surmounting the problems posed by real polymers. *J Struct Biol*. 2007; 157:83–94. [PubMed: 16919474]
58. Ge P, et al. Cryo-EM model of the bullet-shaped vesicular stomatitis virus. *Science*. 2010; 327:689–693. [PubMed: 20133572]
59. Scheres SHW. A Bayesian view on cryo-EM structure determination. *J Mol Biol*. 2012; 415:406–418. [PubMed: 22100448]
60. Clemens DL, Ge P, Lee BY, Horwitz MA, Zhou ZH. Atomic structure of T6SS reveals interlaced array essential to function. *Cell*. 2015; 160:940–951. [PubMed: 25723168]
61. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Crystallogr D*. 2010; 66:486–501. [PubMed: 20383002]
62. Adams PD, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D*. 2010; 66:213–221. [PubMed: 20124702]
63. Pettersen EF, et al. UCSF chimera—a visualization system for exploratory research and analysis. *J Comput Chem*. 2004; 25:1605–1612. [PubMed: 15264254]

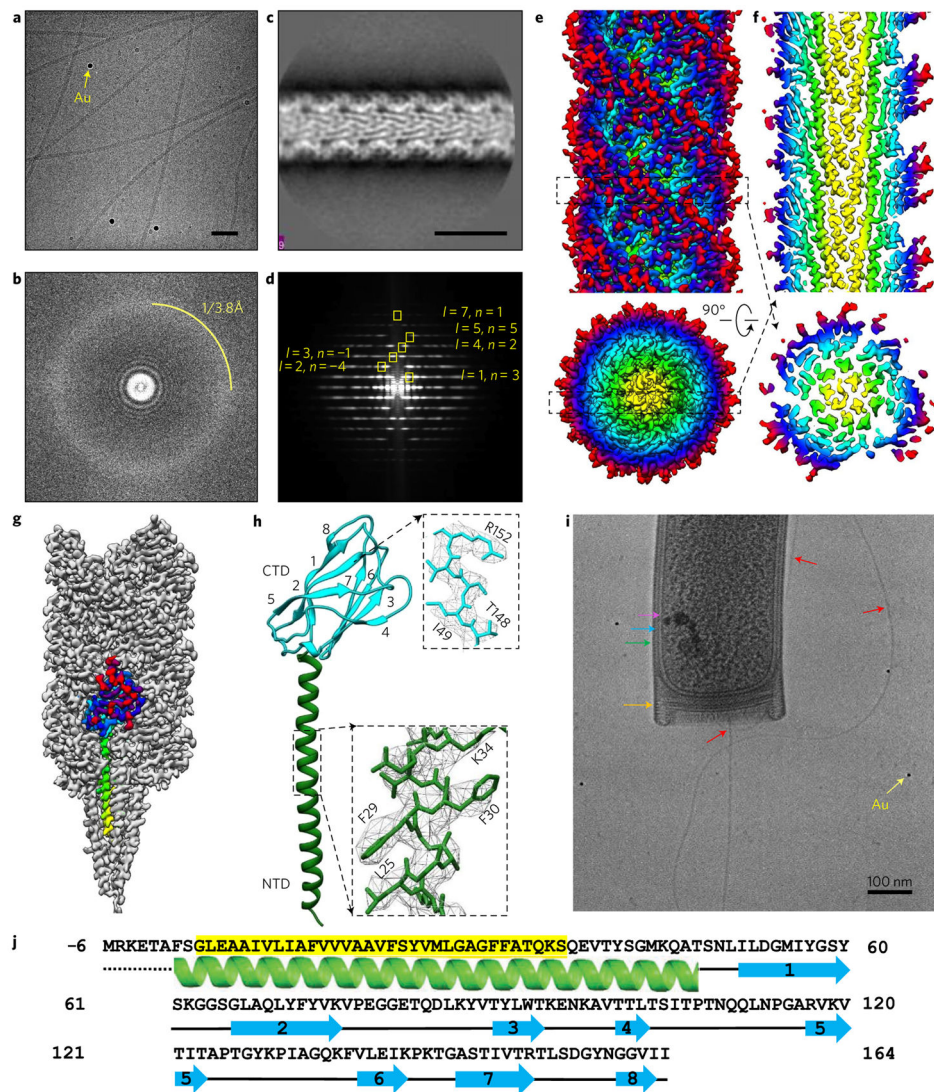


Figure 1. CryoEM data and refinement

a, Representative cryoEM micrograph of *Methanospirillum hungatei* JF1 filaments with 10 nm gold fluidicals (yellow arrow) added. Scale bar, 50 nm. Around 3,300 videos were collected in the data set, of which ~75% had filaments used in two-dimensional classification. **b**, Fourier transform of the micrograph in **a** with the 3.8 Å water ring indicated. **c**, A two-dimensional classification result of CTF corrected images for the *M. hungatei* archaella filaments. Eleven two-dimensional classes (of which the pictured class is one) were used in the final reconstruction. Scale bar, 10 nm. **d**, Fourier transform of the two-dimensional class average with helical indexing. **e**, Reconstruction of the *M. hungatei* archaella at 3.4 Å resolution. **f**, Slab view of the reconstruction showing central α -helices. **g**, Archaella reconstruction with a single subunit highlighted (rainbow). **h**, Reconstruction of an archaellin with two domains, a hydrophobic N-terminal α -helix (green) and the C-terminal eight-stranded β -barrel (cyan) with strands numbered. Insets: close-ups on the α -helix, and on strand 7 of the β -barrel. **i**, A cryoEM image of *M. hungatei* cell end with cytoplasmic membrane (pink), S-layer (blue), sheath (green), plug (orange), 10 nm gold

fiducial (yellow) and protruding archaellum (red) indicated by arrows. Scale bar, 100 nm. **j**, Schematic of the primary and secondary sequences of the *M. hungatei* FlaB3 protein (Mhun_3140). The yellow highlighted portion of the sequence is conserved in archaellins.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

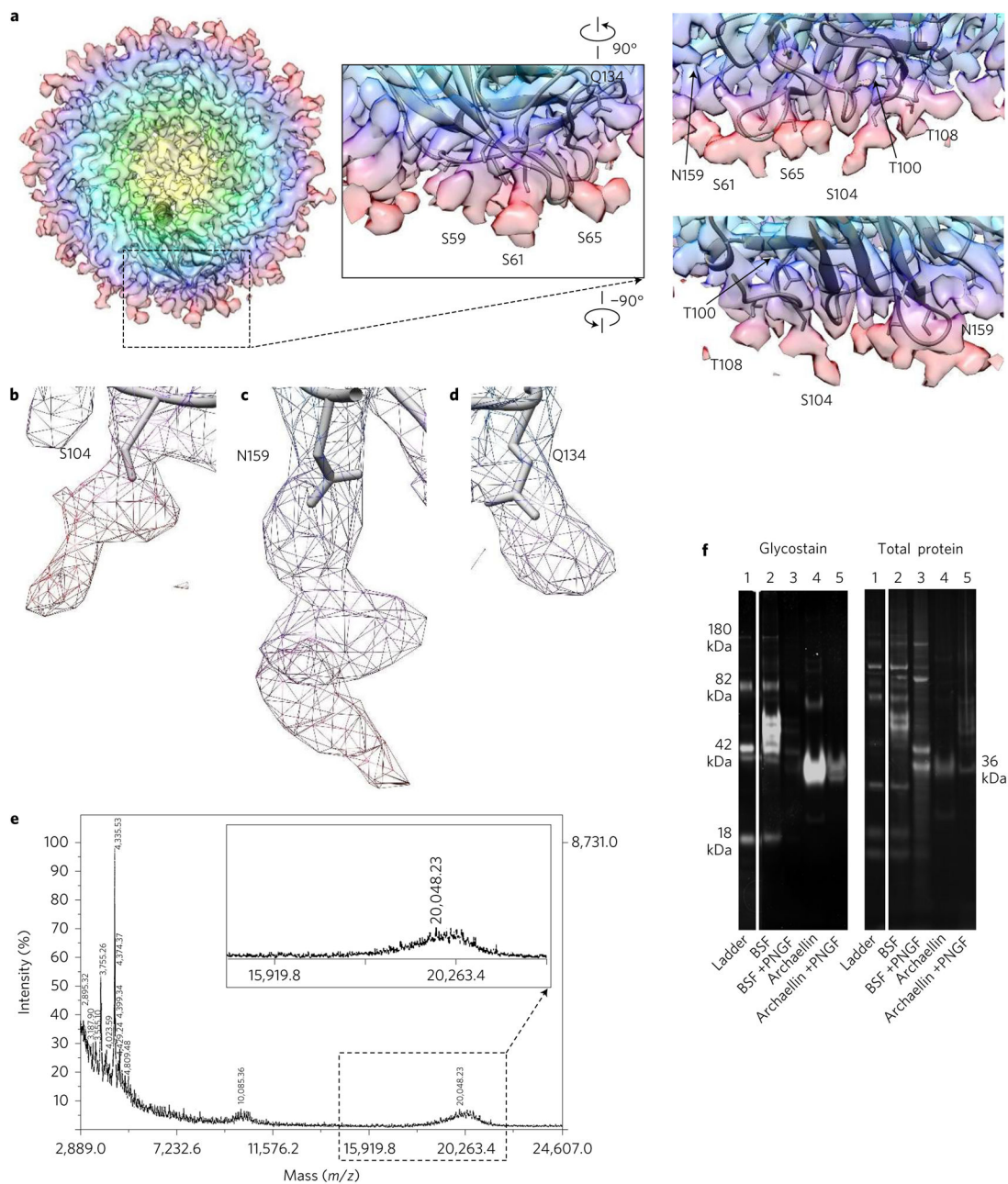


Figure 2. Post-translational modification of the *M. hungatei* flagellin

a, CryoEM reconstruction of the *M. hungatei* archaella with a single archaellin subunit with surface exposed extra densities that do not correspond to the atomic model indicated. **b**, The extra density over Ser104. **c**, The extra density on Asp159. **d**, The extra density on Gln134. **e**, Mass ion spectra of *M. hungatei* archaellin. Inset: enlargement of the broad 20 kDa peak. **f**, Glycostain and total protein stain of *M. hungatei* archaella and treatment with PNGase F. Lane 1, Candy Cane ladder; lane 2, bovine serum fetuin not treated with PNGase F; lane 3, bovine serum fetuin treated with PNGase F; lane 4, *M. hungatei* archaella not treated with PNGase F; lane 5, *M. hungatei* archaella treated with PNGase F.

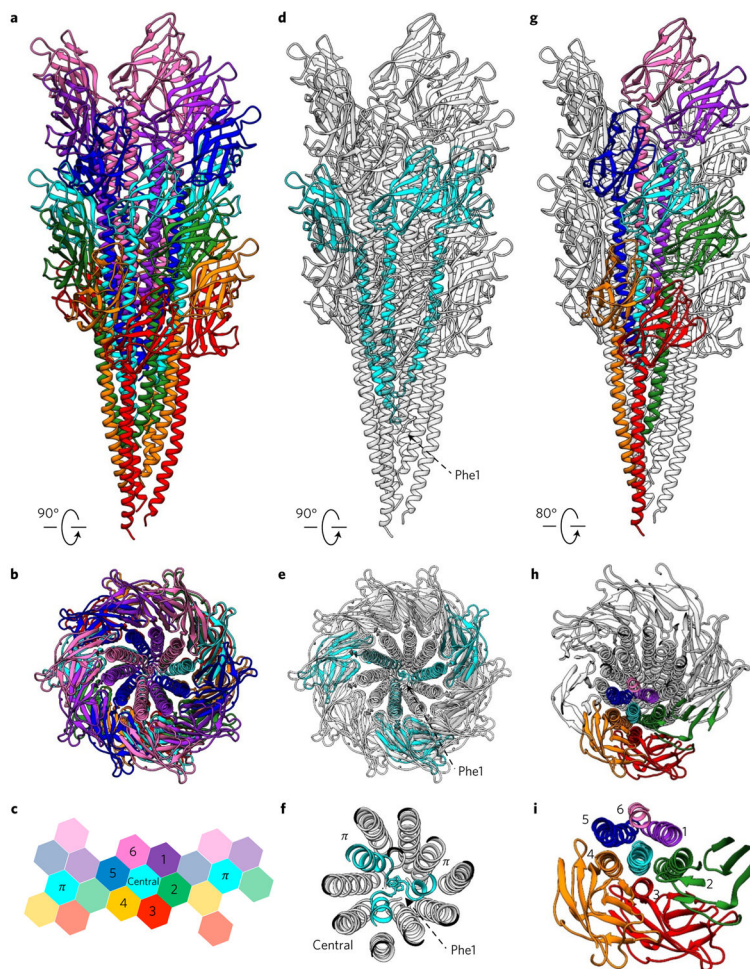


Figure 3. Subunit–subunit interactions

a,b, Network of 21 subunits in the *M. hungatei* archaella filament **c**, Cartoon depicting the interaction unit for a single subunit marked ‘central’. This subunit interacts radially with two subunits marked ‘ π ’ and hydrophobically and ionically with six subunits marked 1 to 6. **d,e**, The two archaellins (cyan) with which the central cyan archaellin interacts radially through π -stacking sandwiches formed by the N-terminal phenylalanine residue shown in the protofilament. **f**, Close-up on the Phe1 residues that form the π -stacking interactions. **g–i**, The six subunits with which the central cyan subunit has adjacent interactions through their hydrophobic α -helices and ionic interactions in the β -barrel domain.

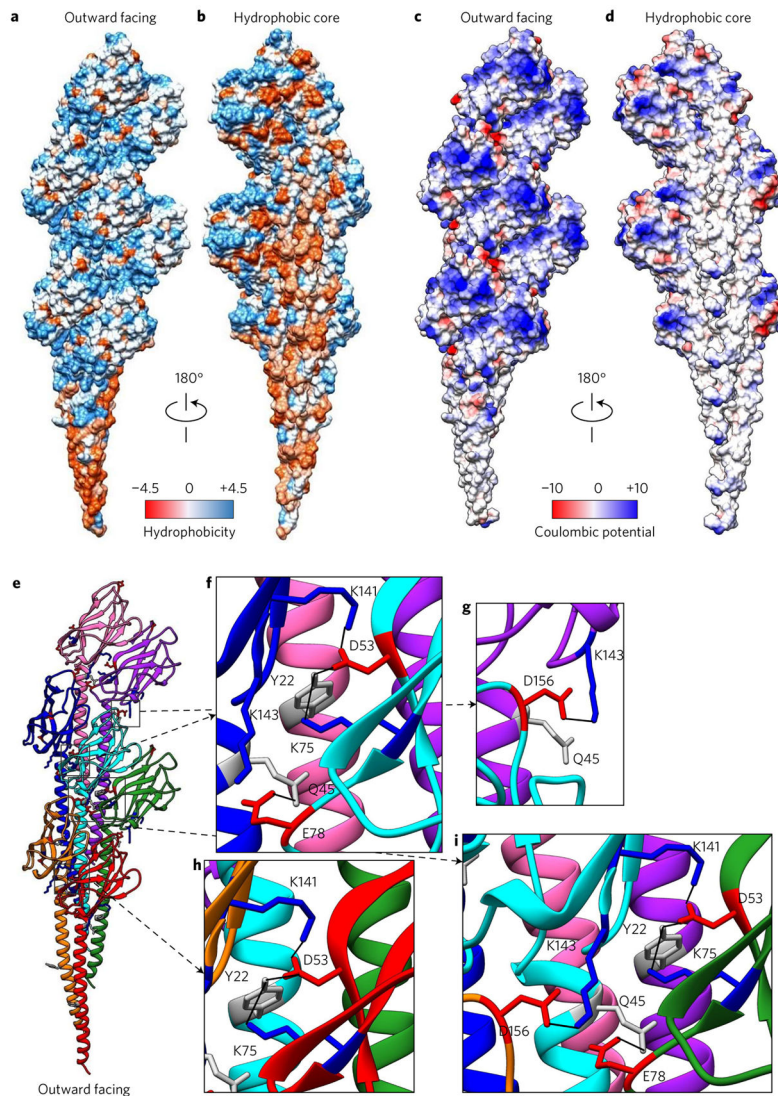


Figure 4. Adjacent subunit interactions

a,b, Hydrophobicity plot of seven subunits involved in adjacent subunit interactions. The inward-facing α -helix domain is hydrophobic, while the outward-facing portion of the β -barrel is hydrophilic. **c,d**, Coulombic potential plot of seven subunits involved in adjacent subunit interactions. The inward-facing α -helix domain has very little charge, and the outside of the filament has patches of positive charge. **e**, An interaction unit of seven subunits. The central (cyan) archaeellin interacts hydrophobically and ionically with the six adjacent archaeellins pictured. The four areas of ionic interaction are indicated with dark grey boxes. The amino acids involved in ionic interactions are coloured in accordance with their charge potential. **f**, Leftmost interaction area of the central (cyan) subunit. The central subunit contributes three amino acids to this interaction: D53, K75 and E78. **g**, Top interaction area of the central (cyan) subunit. The central subunit contributes one amino acid to this interaction: D156. **h**, Lower interaction area of the central (cyan) subunit. The central subunit contributes one amino acid to this interaction: Y22. **i**, Middle interaction area of the

central (cyan) subunit. The central subunit contributes three amino acids to this interaction: Q45, K141 and K143.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

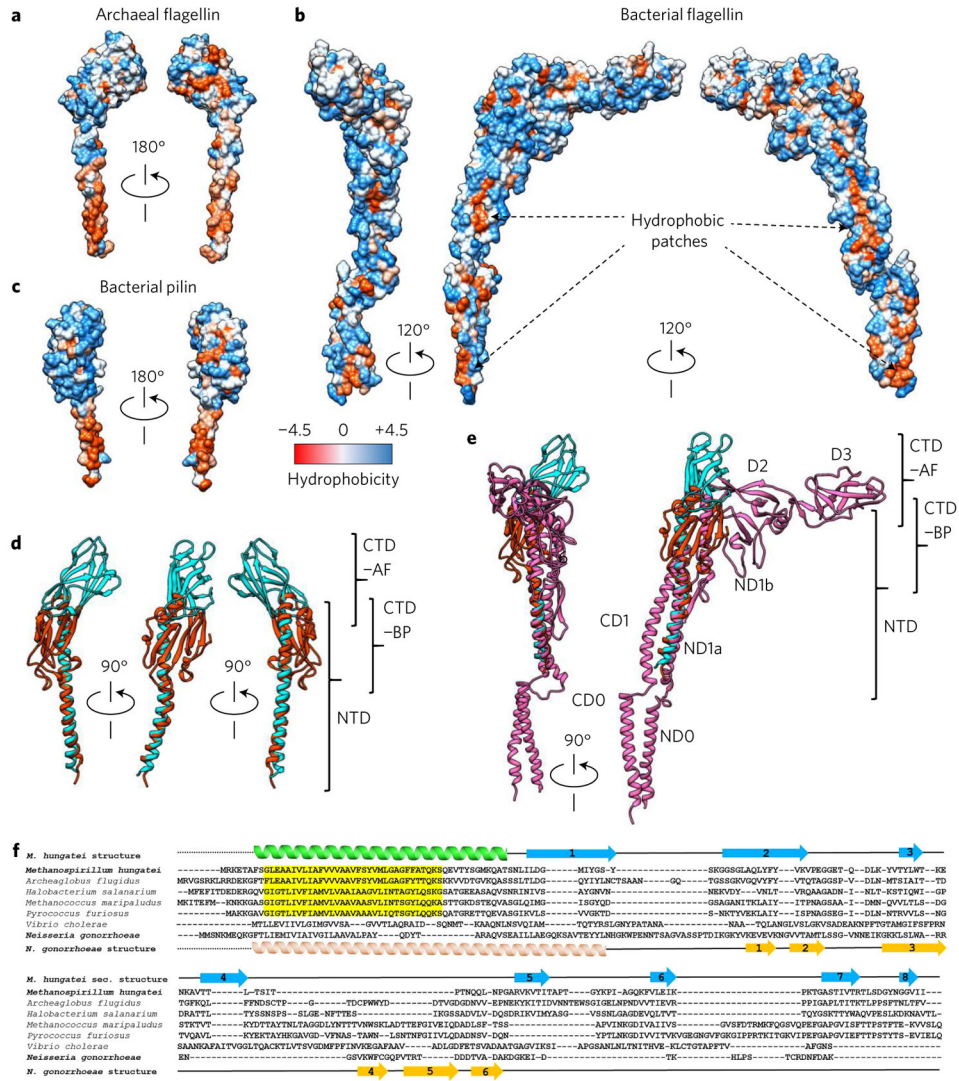


Figure 5. Comparison between the *M. hungatei* archaellin and a bacterial type IV pilin and flagellin
a, Hydrophobicity plot of a single *M. hungatei* archaellin. **b**, Hydrophobicity plot of a single bacterial flagellin from *Salmonella enterica* (PDB: 1UCU). **c**, Hydrophobicity plot of a single bacterial type IV pilin from *Neisseria gonorrhoeae* (PDB: 2HIL). **d**, Comparison of domain organization between the atomic models of the *M. hungatei* archaellin (cyan) and the *N. gonorrhoeae* bacterial type IV pilin (orange). **e**, Comparison of domain organization between the atomic models of the *M. hungatei* archaellin (cyan), the *N. gonorrhoeae* bacterial type IV pilin (orange) and the *S. enterica* bacterial flagellin (pink). **f**, Sequence alignments of five archaellins (top) and two bacterial pilins, *V. cholerae* type IVa and *N. gonorrhoeae* type IVb pilins (bottom). Sequences used are for the major expressed archaellins for each species. The secondary structure of the *M. hungatei* archaellin is shown above the alignments (green and blue), and the secondary structure of the *N. gonorrhoeae* pilin is shown below the alignments (orange). The sequence highlighted in yellow is the consensus sequence for archaeal FlaB proteins. Percent identity scores between the *M.*

hungatei FlaB3 protein and the other sequences in the alignment are as follows: *A. flugidus* 39%, *H. salanarium* 27%, *M. maripaludus* 27%, *P. furiosus* 27%, *V. cholera* 20%, *N. gonorrhoeae* 21%. For comparison, the identity score between the two bacterial type IV pilins is 26%.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

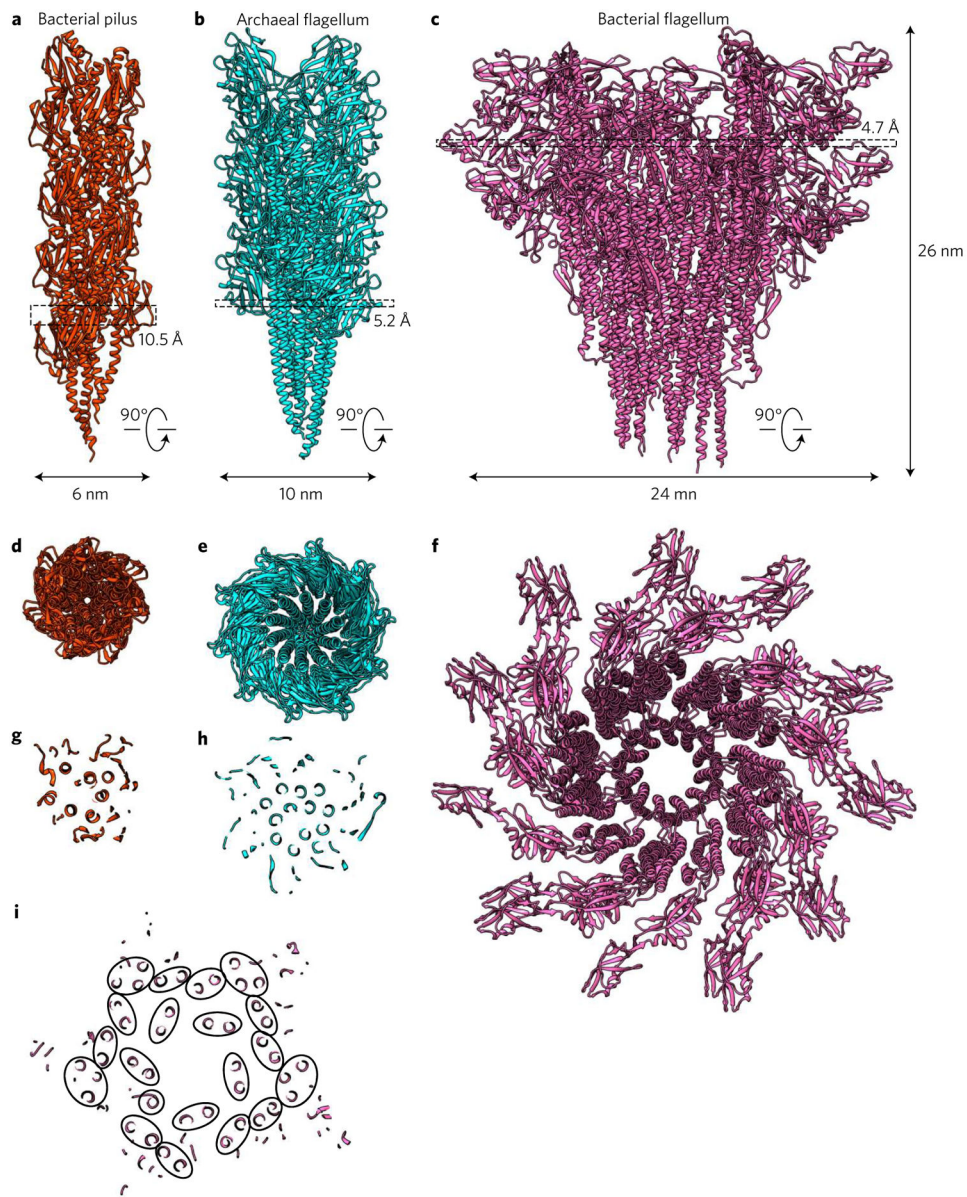


Figure 6. Comparison among prokaryotic motility filaments

a, A 15-subunit protofilament of the *N. gonorrhoeae* bacterial type IV pilus (orange, PDB: 2HIL). **b**, A 26-subunit protofilament of the *M. hungatei* archaeal archaellum (cyan). **c**, A 20-subunit protofilament of the *S. enterica* bacterial flagellum (pink, PDB: 1UCU, protofilament generated based on published helical parameters). **d**, View down the helical axis of the *N. gonorrhoeae* bacterial type IV pilus protofilament. **e**, View down the helical axis of the *M. hungatei* archaellum. **f**, View down the helical axis of the *S. enterica* bacterial flagellum. **g**, Slice one subunit's axial rise (10.5 Å thick) of the *N. gonorrhoeae* bacterial pilus as viewed down the helical axis. **h**, Slice one subunit's axial rise (5.2 Å thick) of the *M. hungatei* archaellum as viewed down the helical axis. **i**, Slice one subunit's axial rise (4.7

Å thick) of the *S. enterica* bacterial flagellum as viewed down the helical axis. Ovals indicate helical bundles belonging to the same subunit.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript