



An archaellum filament composed of two alternating subunits

Lavinia Gambelli ^{1,2,6}, Michail N. Isupov^{3,6}, Rebecca Conners ^{1,4}, Mathew McLaren^{1,4}, Annett Bellack ⁵, Vicki Gold ^{1,4}, Reinhard Rachel ⁵ & Bertram Daum ^{1,4}✉

Archaea use a molecular machine, called the archaellum, to swim. The archaellum consists of an ATP-powered intracellular motor that drives the rotation of an extracellular filament composed of multiple copies of proteins named archaellins. In many species, several archaellin homologs are encoded in the same operon; however, previous structural studies indicated that archaellum filaments mainly consist of only one protein species. Here, we use electron cryo-microscopy to elucidate the structure of the archaellum from *Methanocaldococcus villosus* at 3.08 Å resolution. The filament is composed of two alternating archaellins, suggesting that the architecture and assembly of archaella is more complex than previously thought. Moreover, we identify structural elements that may contribute to the filament's flexibility.

¹Living Systems Institute, University of Exeter, Exeter EX4 4QD, UK. ²College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter EX4 4QF, UK. ³Henry Wellcome Building for Biocatalysis, Biosciences, College of Life and Environmental Sciences, University of Exeter, Exeter EX4 4QD, UK. ⁴College of Life and Environmental Sciences, University of Exeter, Exeter EX4 4QD, UK. ⁵Institute of Microbiology and Archaea Centre, University of Regensburg, 93053 Regensburg, Germany. ⁶These authors contributed equally: Lavinia Gambelli, Michail N. Isupov. ✉email: b.daum2@exeter.ac.uk

Archaea are ubiquitous microorganisms that successfully colonise diverse environments partially due to their ability to swim by means of archaella^{1–3}. The archaellum is a propulsive nanomachine consisting of an intracellular motor that drives the rotation of an extracellular filament⁴. A clockwise rotation of the archaellum moves the cell forward, whereas a counterclockwise rotation moves the cell backward⁵. While archaella, like bacterial flagella, mediate swimming, the two systems are not homologous, suggesting their independent evolution. Instead, archaella are structurally homologous to filaments of the type-IV filament (T4F) superfamily. The T4F superfamily includes a vast range of filamentous nanomachines, such as bacterial type-IV pili (T4P)⁶, associated with diverse functions⁷. Within the T4F superfamily, archaella are unique in providing swimming motion by means of a rotary propeller. Other functions of T4F in archaea include surface adhesion^{8–11}, cell–cell contacts and biofilm formation^{8,12}.

The archaellum machinery consists of a membrane-embedded motor complex and a ~10 µm long filament, which together are formed by ~10 components, depending on the species¹³. The filament is a helical array of ArlA and/or ArlB archaellins¹⁴ that assemble proximal to the cell surface¹⁵. Filament assembly and rotation is an ATP-dependent process driven by the motor complex, consisting of the platform protein ArlJ, the hexameric AAA + ATPase ArlI, and a putative regulator ArlH^{6,16–18}. The motor is surrounded by a cytosolic ring of ArlX proteins (in Crenarchaeota)¹⁹ or ArlC, D/E (in Euryarchaeota)^{20,21}. In the periplasm, the motor complex is anchored to the S-layer by the ArlF and ArlG stator proteins^{22,23}.

Most genes encoding the archaellum machinery components are organised in one *arl* operon. The first genes in this operon (*arlA* or *arlB*) encode for the archaellin subunits that make up the archaellum filament. The number of archaellin genes is species-specific, usually varying from one to seven^{24,25}. When multiple archaellins are encoded, several or all can be transcribed at the same time^{26,27}. Conventionally, ArlA and ArlB archaellins are divided into major and minor structural components of the filament. Archaellins that are thought to form the bulk of the filament are referred to as major. Biochemical and genetic experiments suggest that in *Methanococcus voltae*²⁸, *Methanococcus maripaludis*²⁹, *Methanothermococcus thermolithotrophicus*³⁰, *Halobacterium salinarum*³¹ and *Haloerubrum lacusprofundi* DL18³² the archaellum filament consists of two major archaellins, hypothesised to be equally distributed along the filament. Nevertheless, the three high-resolution structures of archaella that are available to date show homopolymeric filaments consisting of only a single (major) archaellin^{20,33,34}.

The roles of minor archaellins mostly remain enigmatic. Genetics and molecular biology experiments suggest that in *M. voltae*³⁵, *M. maripaludis*²⁹ and *H. salinarum*³⁶, minor archaellins form a region at the base of the archaellum filament that is reminiscent of the flagellar hook. However, not all archaea show this feature, including those that encode only one archaellin.

Here we employed electron cryo-microscopy (cryoEM) to solve the structure of the archaellum filament from *Methanocaldococcus villosus*, a hyperthermophilic archaeon¹¹ and one of the fastest swimming organisms known³⁷. We report the structure of a heteropolymeric archaellum filament composed of two alternating subunits, ArlB1 and ArlB2 and provide insights into the dynamics of archaella. Our results provide a shift in our understanding of how archaella assemble and function.

Results

High-resolution cryoEM and helical reconstruction of *M. villosus* archaella. The structure of the *M. villosus* archaellum was

determined from cryoEM movies of isolated filaments using Relion 3.1³⁸ and cryoSPARC 3.1.0³⁹ (Supplementary Figs. 1, 2). Helical parameters from previously published archaella structures were applied and optimised by helical search to a rise of 5.57 Å and a twist of 108°. This approach yielded a 3.28 Å resolution map in which the filament appeared to consist of one archaellin homologue (Supplementary Fig. 3a). Surprisingly, further refinement of the cryoEM map without imposing helical symmetry produced two distinct and alternating protein densities (Fig. 1; Supplementary Fig. 4; Supplementary Movie 1; Supplementary Fig. 3b, c). Careful analysis of the filament's heteropolymeric structure revealed that the helical symmetry did not comply with the initially applied rise and twist. Instead, the minimal transformation that superimposes each archaellin subunit onto one of the same types is $n + 6$ (6 subunits along the filament long axis) (Supplementary Fig. 4e). Recalculating the helical symmetry of the archaellum based on this observation resulted in a helical rise of 33.4 Å ($5.57 \text{ Å} \times 6 = 33.42 \text{ Å}$, refined to 33.4 Å) and a twist of -71.8° ($108^\circ \times 6 = 648 = 720 - 72^\circ$). Refining our map with these parameters improved the resolution to 3.08 Å (Fig. 1; Supplementary Figs. 4 and 6).

The *M. villosus* archaellum consists of the subunits ArlB1 and ArlB2.

The archaellum operon of *M. villosus* encodes three archaellin homologs (ArlB1, 2 and 3). To investigate which two constitute the filament, we built atomic models based on their sequences for both distinct protein densities. ArlB1 and ArlB2 could each be modelled unambiguously into alternating positions in the map, guided by aromatic side chains and glycosylation sites (Figs. 2 and 3; Supplementary Fig. 7; Supplementary Movie 2). The amino acid sequence of ArlB3, instead, did not match the map density features.

The proteinaceous part of the *M. villosus* archaellum filament has a ~9.8 nm diameter. ArlB1 and ArlB2 alternate throughout the filament, following left-handed helical strands. The archaellum is composed of three of these helical strands, which by convention are defined as 3-start helices (Figs. 1; 4a, b; Supplementary Fig. 8; Supplementary Movies 1 and 2).

ArlB1 and ArlB2 are homologous by structure and sequence (Fig. 2d–f; Supplementary Figs. 9, 10). Their genes encode proteins of 221 and 225 amino acids, respectively, which share 56.5% sequence identity (75.9% similarity). As is typical for other members of the T4F superfamily, ArlB1 and ArlB2 consist of an N-terminal hydrophobic α -helical “tail” and a globular, C-terminal β -strand rich “head” domain. In the assembled archaellum, the tails constitute the core of the filament, whereas the head domains are exposed to the periphery (Fig. 2a–c). The first 12 amino acids are not present in the mature ArlB1 and ArlB2 proteins due to N-terminal processing by a class-III signal peptidase prior to their insertion into the filament⁴⁰. While the α -helical tail domains are identical in ArlB1 and ArlB2 (I₁₄–A₅₉), the head domains show differences with respect to sequence and structure (Fig. 2f; Supplementary Figs. 9, 10). ArlB1 consists of 12 β -strands and three α -helices, whereas ArlB2 has 13 β -strands and two α -helices. The most striking difference is a glycosylated outward-facing loop (K₁₃₁–A₁₃₇) in ArlB2, which is absent in ArlB1 (Figs. 2f and 3c).

N-glycosylation and metal-binding sites. Refinement of the helical parameters was crucial to clearly resolve the glycosylation pattern of the archaellum filament. Glycan densities were identified as dead-end protrusions from the peptide backbone that were larger than side chains and co-localised with the canonical N-glycosylation consensus sequon (N-X-S/T) (Fig. 3a–c, Supplementary Fig. 7). Including the density occupied by the glycan moieties, the diameter of the archaellum is ~11 nm. All

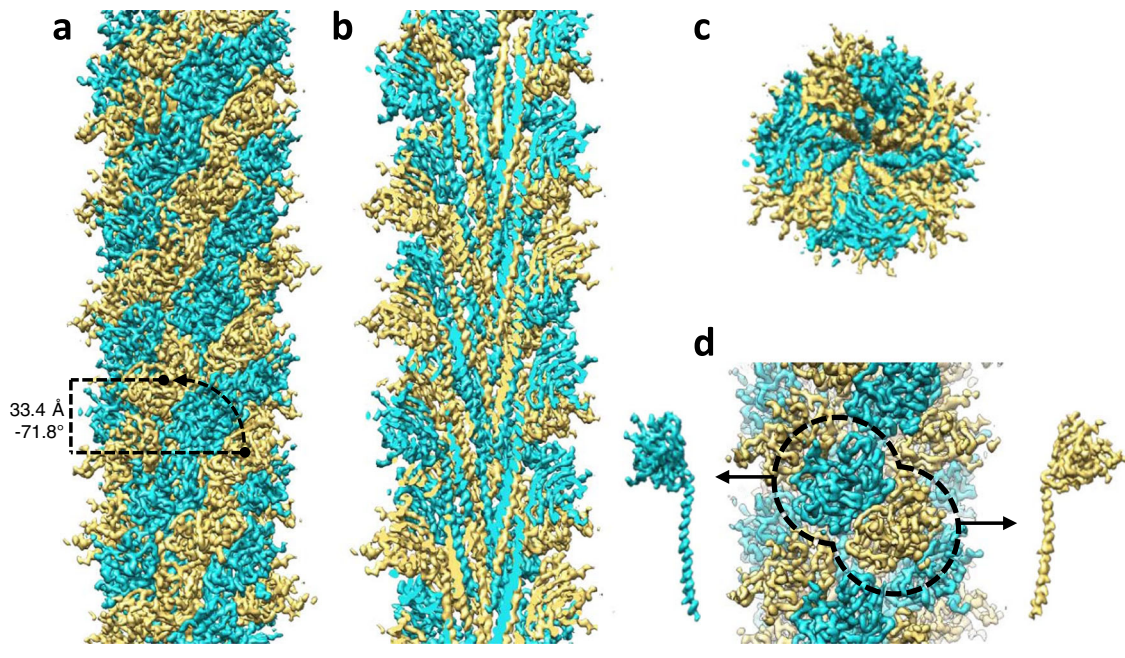


Fig. 1 CryoEM map of the *M. villosus* archaeellum. **a** Surface view; **b**, **c** Cross-sections parallel and perpendicular to the filament's long axis, respectively. The two alternating archaeellin densities are coloured in cyan (ArlB1) and sand (ArlB2). In **a**, the helical parameters (rise, 33.4 Å; twist, -71.8°) are shown. In **d**, the head domains of the two archaeellins are highlighted with a dashed line and the density maps of each subunit are shown left (ArlB1, cyan) and right (ArlB2, sand) of the filament (arrows). Scale bar in **a**, **b** and **c**, 50 Å.

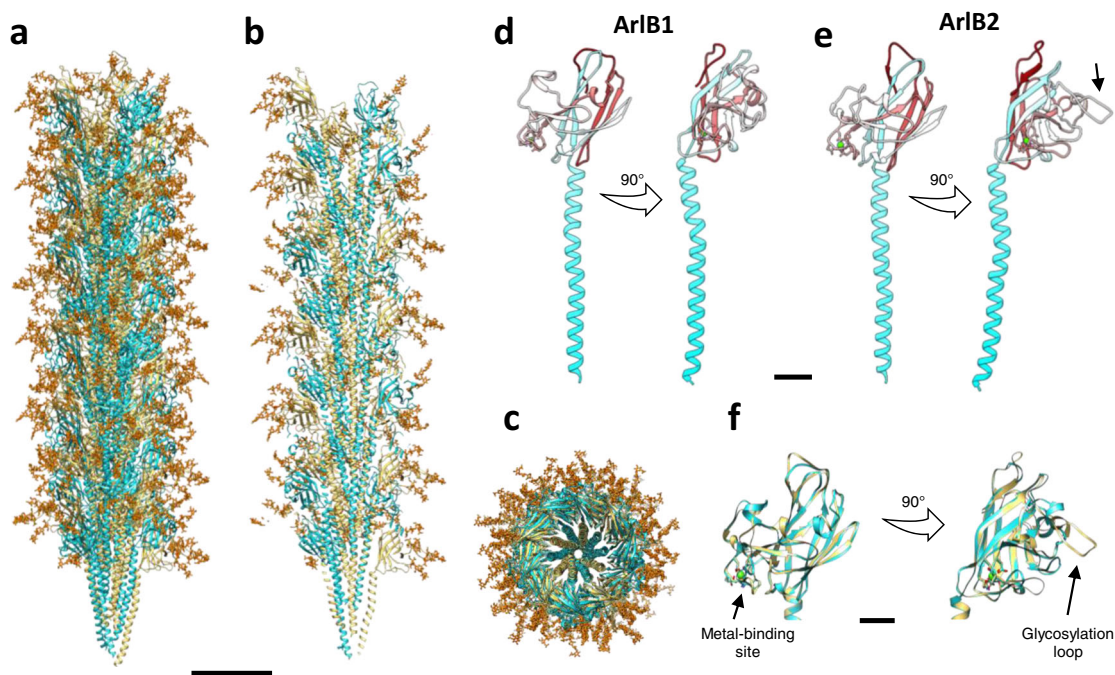


Fig. 2 Atomic model of the *M. villosus* archaeellum filament. **a–c** Atomic model of the archaeellum filament with glycosylation in side view (**a**) and cross-section parallel (**b**) and perpendicular (**c**) to the filament's long axis. ArlB1 is in cyan, ArlB2 in sand, and glycosylation in orange. Scale bar, 50 Å. **d**, **e** Atomic models of ArlB1 (**d**) ArlB2 (**e**) in ribbon representation and cyan-white-maroon colour scheme (N-terminus, cyan; C-terminus, maroon). The arrow in **e** highlights the “glycosylation loop”. **f** Superposition of the ArlB1 (cyan) and ArlB2 (sand) head domain with the metal-binding site and “glycosylation loop” highlighted. Scale bar in **d**, **e** and **f**, 10 Å.

glycosylation sites are located on loops or sites of alternating secondary structure motifs and are equally distributed along the filament's length.

While the head domain of ArlB1 harbours three glycosylation sites (N_{112} , N_{133} and N_{176}) (Fig. 3b), ArlB2 features six (N_{70} , N_{113} , N_{118} ,

N_{133} , N_{134} and N_{180}) (Fig. 3c). Interestingly, two of these glycosylated asparagine residues (N_{133} and N_{134}) are consecutive and located in the NNTT motif of the ArlB2 glycosylation loop (K_{131} – A_{137}).

Since no information on sugar connectivity and identity is available for N-glycans of *M. villosus*, we modelled the branched

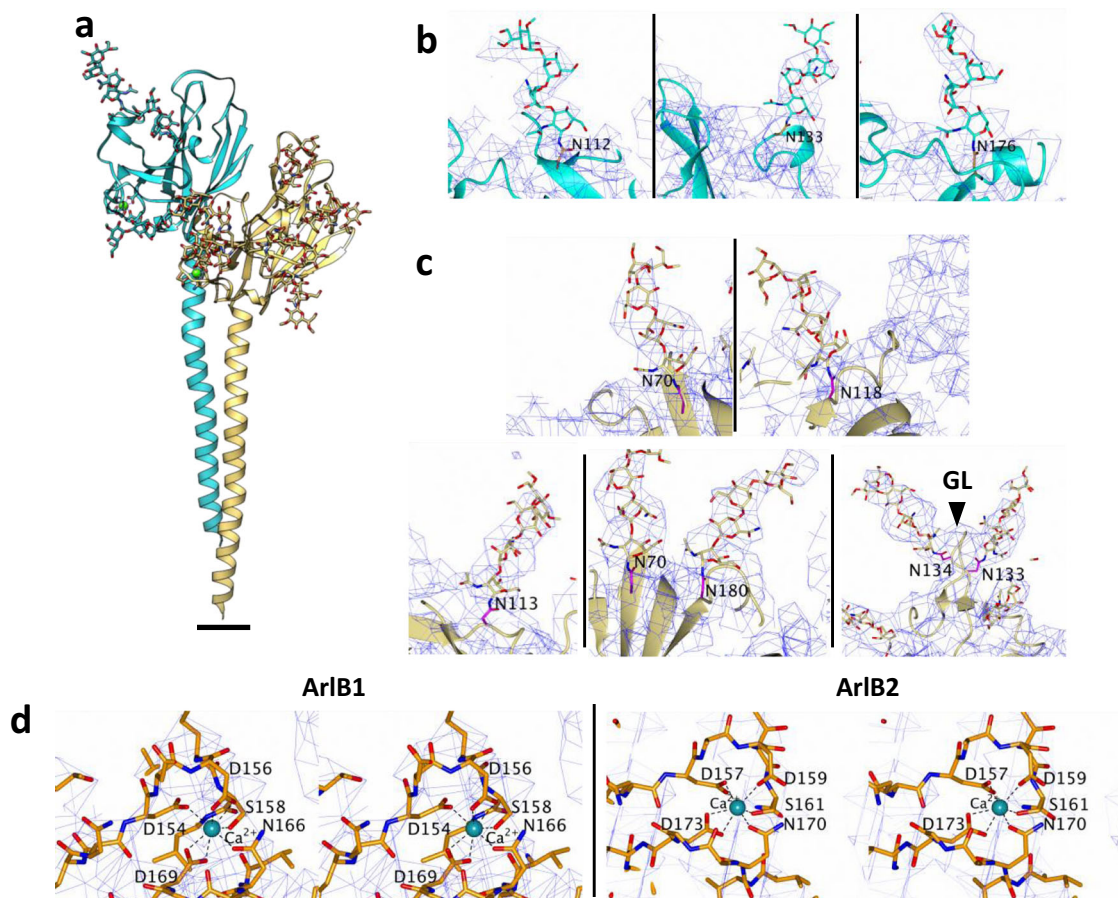


Fig. 3 Glycosylation and metal-binding sites in ArlB1 and ArlB2. **a** Atomic models of ArlB1 (cyan) and ArlB2 (sand) are shown in ribbon representation with modelled glycan molecules, side chains of metal and glycan binding residues shown as stick models and metal ions highlighted as green spheres. Scale bar in **a**, 10 Å. **b** Close-ups of the three glycosylation sites in ArlB1. **c** Close-ups of the six glycosylation sites in ArlB2. GL glycosylation loop. **d** Stereo views of ArlB1 and ArlB2 metal-binding sites. The cryoEM map is represented as a blue mesh and contoured at 2.8 σ . Protein chain is shown as a stick model, Ca^{2+} ion is shown as a cyan sphere with metal-coordinating residues labelled.

heptameric glycan from the related archaeal species *M. thermolithotrophicus*³⁰ in order to aid model building and refinement. For each of the glycan densities, we were able to fit only the first four sugars of the *M. thermolithotrophicus* polysaccharide. No additional saccharides were resolved beyond the fourth unit, which is likely due to their flexibility or a shorter glycan in *M. villosus*. The fitted glycans provide a detailed model of their location and distribution on the surface of the archaellum (Fig. 2a–c; Fig. 3a–c; Supplementary Fig. 11).

As featured by other archaellins³⁴, ArlB1 and ArlB2 also contain a metal-binding site located at the periphery of the head domain, in a groove between $\beta 6$ and $\beta 7$ in ArlB1 and $\beta 7$ and $\beta 8$ in ArlB2 (Fig. 3d). In both archaellins, the metal ion is coordinated by five conserved residues, which differ slightly in their position along the backbone (D₁₅₄, D₁₅₆, S₁₅₈, N₁₆₆, D₁₆₉ in ArlB1 and D₁₅₇, D₁₅₉, S₁₆₀, N₁₇₀, D₁₇₃ in ArlB2). The metal was modelled as a calcium ion. However, due to the prevalence of magnesium over calcium in our growth culture, a mixed population of calcium and magnesium ions cannot be excluded. The metal-binding residues are highly conserved between ArlB1, ArlB2 and the *M. jannaschii* archaellin (Supplementary Fig. 12), for which the site was resolved by X-ray crystallography³⁴.

Heteropolymeric assembly of ArlB1 and ArlB2 introduces a screw asymmetry in the archaellum. To understand the assembly of ArlB1 and ArlB2 in the archaellum structure, we analysed

their organisation and contacts within the filament. Consistent with published structures of homopolymeric archaella^{17,20,33} we find that ArlB1 and ArlB2 undergo hydrophobic interactions via their N-terminal tails in the filament's core. In particular, each tail interacts with neighbours at positions $n + 3$, $n - 3$, $n + 7$, $n - 7$, $n - 1$, $n + 1$, $n - 4$ and $n + 4$ (Supplementary Fig. 13).

In the outer sheath of the archaellum the head domains of ArlB1 and ArlB2 interact via hydrogen bonds and salt bridges. These occur between alternating ArlB1 and ArlB2 subunits along left-handed 3-start helical strands (Fig. 4a–d; Supplementary Table 2) and along right-handed 7-start strands (Supplementary Fig. 14). The alternating sequence in 3-start direction appears to be favoured by high shape and charge complementarity between ArlB1-2 and ArlB2-1 heterodimers. 13 hydrogen bonds and 3 salt bridges with a ΔG_{diss} of 20.4 kcal/mol are formed at the ArlB1-2 contact zone and 8 hydrogen bonds and 2 salt bridges with a ΔG_{diss} of 16.2 kcal/mol generate the ArlB2-1 interface (Fig. 4c; Supplementary Table 2). In contrast, poor charge and shape complementarity would not allow the formation of ArlB1-1 and ArlB2-2 in the 3-start direction. Interestingly, the sequence of alternating subunits is not always in sync in adjacent 3-start helices. One of the helical strands is shifted by one subunit, moving it out of the register (Fig. 4b). This results in an anisotropic distribution of ArlB1 and ArlB2 throughout the archaellum, meaning that ArlB1 and ArlB2 neither alternate nor form perfect homopolymeric strands in any direction but 3-start.

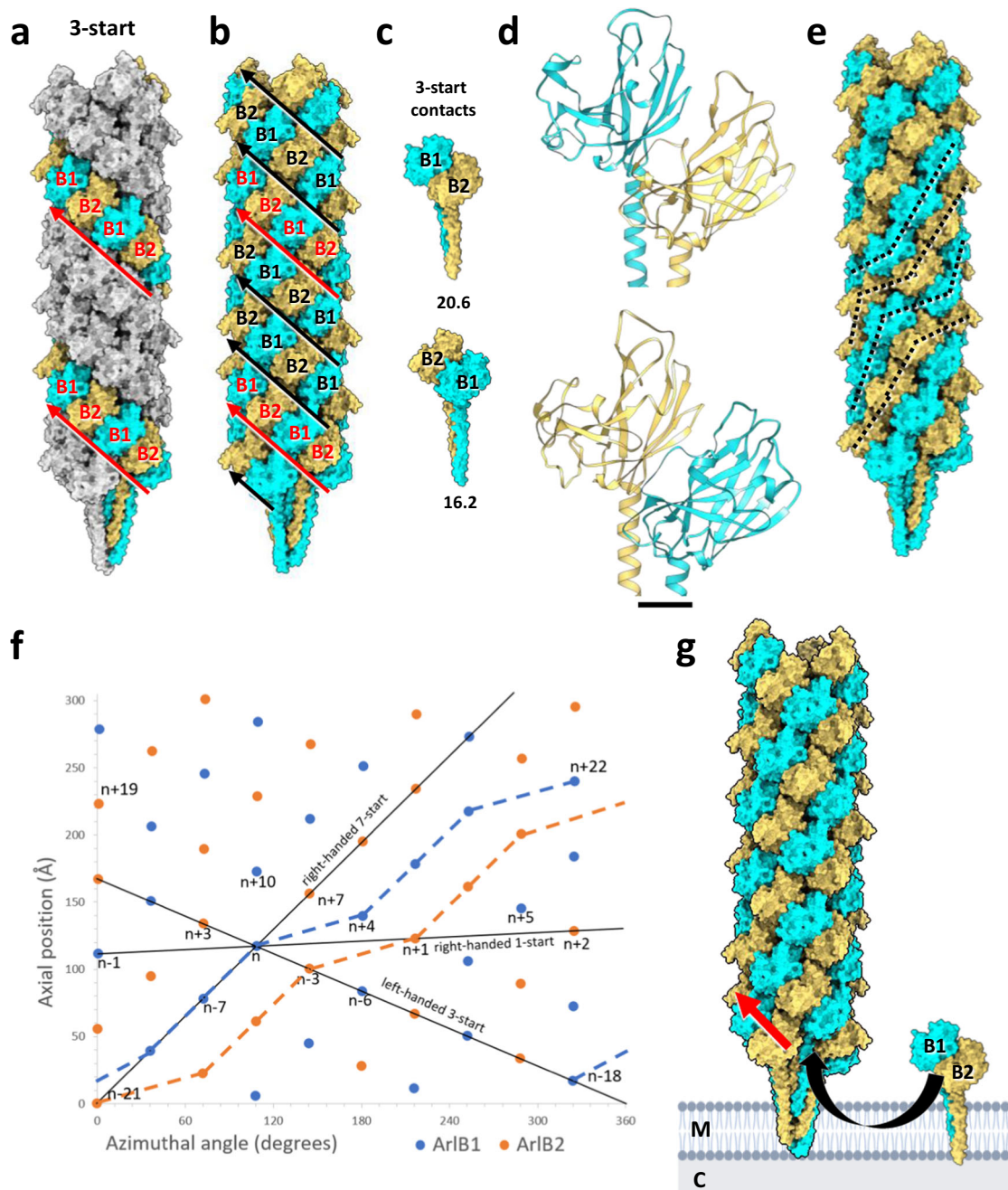


Fig. 4 Screw axis asymmetry in the *M. villosus* archaeellum. **a** The archaeellum atomic model viewed parallel to the filament's axis. ArlB1 (cyan) and ArlB2 (sand) subunits are highlighted along one left-handed 3-start helical strand. ArlB1 and ArlB2 alternate in the 3-start direction (indicated by the red arrows). **b** Subunits within all three 3-start strands are highlighted. Every third 3-start strand (red arrow) is out of register with respect to the other two 3-start strands (black arrows). **c** Intermolecular contacts in 3-start direction between ArlB1-2 (top) and ArlB2-1 (bottom) dimers. The dissociation energies (ΔG_{diss} calculated by PISA) for both interactions are indicated below each dimer (unit: kcal/mol). **d** Close-ups of the atomic models (head domains) of ArlB1-2 and ArlB2-1 as in **c**. **e** Tracing subunits of the same kind through the filament indicates non-helical homopolymeric pseudo-strands (dashed lines). **f** Helical net diagram showing the positions of ArlB1 (blue dots) and ArlB2 (orange dots) in a two-dimensional plot. Solid black lines show various component helices. Blue and orange lines trace non-helical homopolymeric pseudo-strands of ArlB1 and ArlB2. **g** Model of the hypothetical assembly of the heteropolymeric archaeellum from pre-formed ArlB1-2 dimers that are added to the growing filament in 3-start direction. M, membrane; C, cytosol; red arrow, 3-start direction. For simplicity, the archaeellum machinery and the S-layer have not been drawn. Scale bar in **d**, 10 Å.

Along the left-handed 7-start helical strands the two subunits undergo all four possible contacts (ArlB1-2, ArlB2-1, ArlB1-1, ArlB2-2) (Supplementary Fig. 14). The interactions in the right-handed 7-start direction are by an order of magnitude weaker than those in the 3-start direction (ΔG_{diss} of 1.7–4.7 kcal/mol), as they are stabilised by only 2–5 salt bridges per interunit interface

and limited hydrophobic contact between the tails (Supplementary Table 2).

Strikingly, when tracing subunits of the same type (ArlB1 or ArlB2 only) throughout the filament, the archaeellum appears to be defined by homopolymeric right-handed pseudo-strands with broken helical symmetry (Fig. 4e, f). Here, the symmetry of the

pseudo-strands is disrupted at the interface between every third and fourth subunit, giving the pseudo-strands a “stepped” appearance, in contrast to a hypothetical isotropic filament (with $n + 2$ geometry) (Supplementary Fig. 5c, d, g).

The flexibility of the archaellum filament. The archaellum rotates, prompted by an ATP-driven intracellular motor that generates torque. To propagate the motion along its length, the filament transitions from a curved filament in resting cells²⁰ (Supplementary Fig. 15) into a rotating superhelix when cells swim⁴¹. However, due to the nature of helical image processing of cryoEM data, all available structures of archaella suggest idealised helical filaments that are infinitely straight.

In order to obtain a truer representation of the structure of the archaellum and to probe the molecular basis for its flexibility, we performed 3D variability analysis in cryoSPARC, which estimates molecular motion from micrographs⁴². This analysis resulted in a series of 20 maps, displaying a distinctly bent conformation of the archaellum (see Supplementary Movie 3 for morphing between these conformations). Refining our atomic model based on individual conformations, we visualised the dynamic properties of the archaellum at the molecular level (Fig. 5; Supplementary Movies 4–7).

Our analysis reveals that as the filament bends, the subunits are compressed at its concave side and stretched apart along its convex side (Supplementary Movies 3, 4). To allow for this flexibility, the tails in the backbone rotate around and slide along the filament’s axis (Fig. 5a, b; Supplementary Movie 5), whereas the head domains swing diagonally to the filament’s axis (Fig. 5h; Supplementary Fig. 16). Along pseudo-helical ArlB1 and ArlB2 strands (Fig. 5c), the head domains of both archaellins are flexible, and their trajectories differ depending on the subunit location within the filament (Fig. 5d–g). Each head domain is free to swing in all directions within the cylinder surface of the outer sheath and with respect to the tail domain. An example of the head trajectories of two homopolymeric strands is shown in Fig. 5e, g and Supplementary Movies 6, 7. Notably, we do not see substantial flexibility or conformational changes within the head or the tail domains themselves, indicating that they move relative to each other as rigid bodies. Moreover, no significant difference in flexibility between ArlB1 and ArlB2 was observed.

The flexibility of each archaellin is enabled by a hinge of two residues (Ser₆₀, Gly₆₁) linking the head and tail of each subunit. This hinge is highly conserved in ArlB1 and ArlB2 sequences within the archaeal kingdom (Fig. 5i; Supplementary Fig. 17a). Interestingly, the sequence of this region is partially conserved, albeit different from archaella, in archaeal T4P (Supplementary Fig. 17b, d). In bacterial T4P a conserved loop region coincides with the melted central portion of the N-terminal α -helix in the core of the pilus (Supplementary Fig. 17c, d).

Discussion

Many archaea encode for multiple archaellins, which likely originated from gene duplication events⁴³. Studies employing genetics, biochemistry and molecular biology have shown that while some archaella contain a single major archaellin, others are assembled from two homologues^{26,29–32}. The three structures of archaellar filaments that have been published so far suggest homopolymers^{20,33,34}, and until now it was unknown how two or more distinct proteins can assemble into a single archaellum.

Here, we present the structure of a heteropolymeric archaellum. The filament contains two archaellin homologues (ArlB1 and ArlB2), which became distinguishable when the cryoEM map was reconstructed without imposing helical symmetry. This finding highlights the importance of relaxing the symmetry

during the processing of any helical data, as structural heterogeneities may otherwise be overlooked. In this regard, it would be interesting to revisit previously published structures of archaella and use our image processing strategy to investigate if these are indeed homopolymers by nature or perhaps also consist of two alternating subunits.

The evolutionary advantages of a heteropolymeric over homopolymeric archaella are not fully understood. However, accumulating evidence suggests that these include increased cellular motility and filament stability. For example, while individual Δ arlB1 and Δ arlB2 knockout mutants in *H. lacusprofundi* DL18 still assemble archaella, they are less salt-resistant than wild-type³². In addition, micrographs of negatively stained filaments suggested that Δ arlB1 archaella are more flexible and Δ arlB2 archaella are stiffer compared to wild-type. This is in line with the different binding energies that we have calculated for the *M. villosus* archaellum, where ArlB2-2 interfaces in 7-start direction are significantly more stable than the corresponding ArlB1-1 contact (Supplementary Table 2). Provided that *M. villosus* could form homopolymeric archaella, an ArlB2-only filament would thus likely be more stable but stiffer than an ArlB1 homopolymer. Notably, knocking out one of the two archaellar subunits in *H. salinarum* results in reduced motility³¹ and ablating either of the two archaellins in *M. maripaludis* abolishes archaella and swimming motility entirely²⁹. Furthermore, in *Haloarcula marismortui* the two archaellins ArlA2 and ArlB are ecoparalogues, meaning that different archaella are assembled in response to varying environmental situations (in this case the level of salinity in the environment)⁴⁴. This suggests that the ability to form heteropolymeric archaella can, at least in some species, add to the adaptability of the cell to varying habitat conditions.

In the heteropolymeric archaellum of *M. villosus*, ArlB1 and ArlB2 are not isotopically organised. We propose that the resulting screw axis asymmetry, together with the helical variation of rigidity along the 3-start direction, enhances the supercoiling propensity of the filament. This would explain how the archaellum forms the observed stable right-handed supercoil independent of the direction of gyration⁴¹, as a prerequisite for cellular propulsion in opposite directions.

The published structures of T4P in both archaea and bacteria report homopolymeric filaments constituted by one major pilin^{45–49}. In one example, however, two different homopolymeric T4P filaments are assembled simultaneously and by the same machinery⁴⁹. Only one study published to date provides evidence of a heteropolymeric T4P in a Gram-positive bacterium, but the arrangement of the two subunits within the filament is speculative⁵⁰. These findings suggest that the specialised structures of pilins/archaellins may allow for unique patterns of assembly within filaments, and differences within the T4F superfamily may be more widespread than originally thought.

Archaella are assembled by the platform protein ArlJ, powered by the ATPase ArlI⁶. For homopolymeric archaella, the filament is thought to be built from monomers. However, building the complex $n + 6$ symmetry observed in the *M. villosus* archaellum from archaellin monomers would need to be selective for alternating subunits, and thus require a more complex assembly mechanism compared to one that assembles homopolymeric archaella. Therefore, it is conceivable that the filament is instead assembled by pre-formed ArlB1-2 heterodimers, which are then added to the 3-start strands of the growing archaellum (Fig. 4g). This is in line with our observation that ArlB1-2 interfaces show by far the strongest interactions across the filament, which suggests that ArlB1-2 may be the fundamental building block of the *M. villosus* archaellum. Notwithstanding this, knockout mutants of one of two major archaellins still assemble archaella³². This

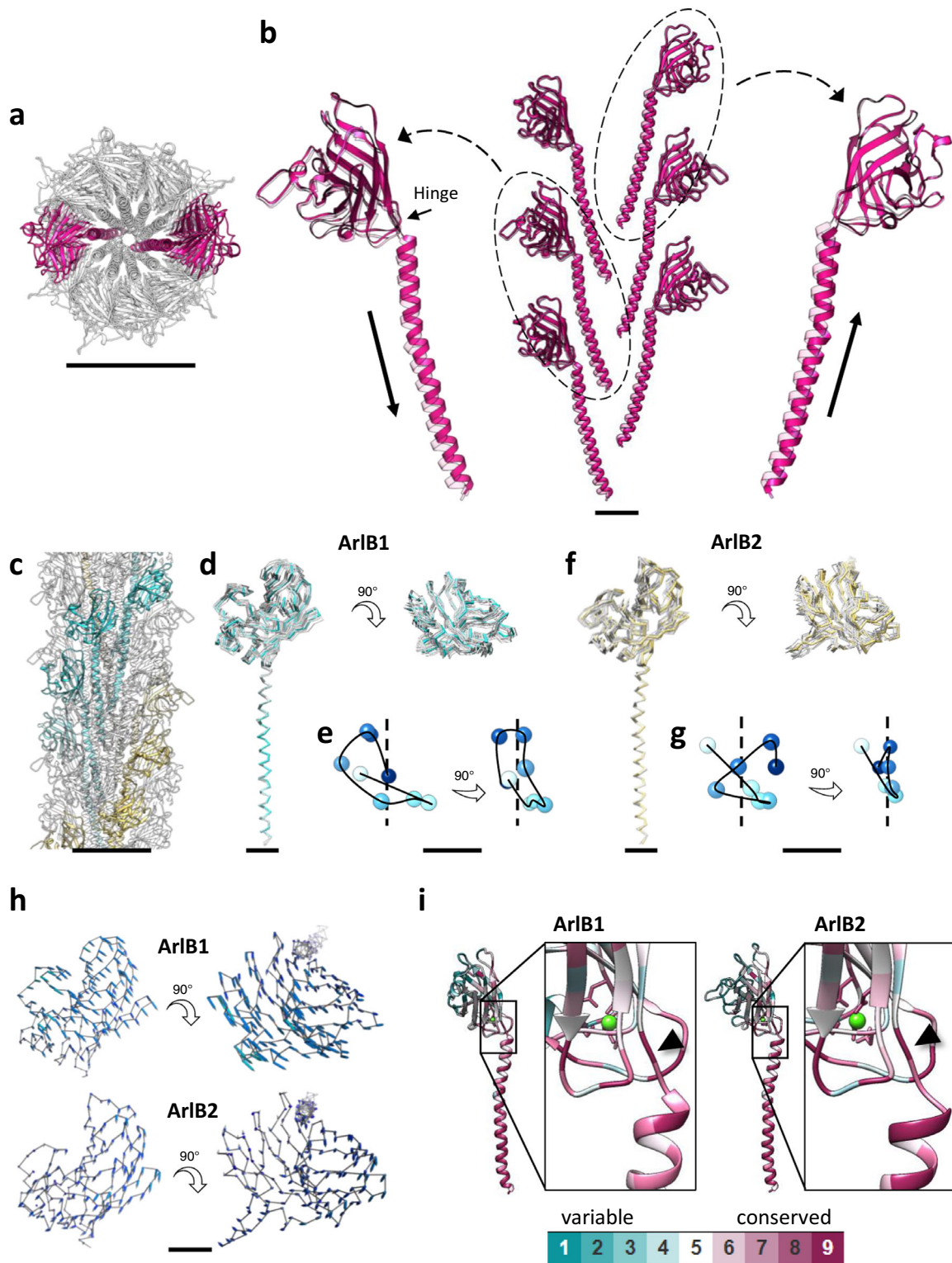


Fig. 5 Flexibility of the *M. villosus* archaeellum. **a, b** Atomic model showing the motion of the tail domains. Two opposite $n + 10$ sets of protein monomers (**a**) were selected and the head domains of the constituting monomers were aligned. The solid and transparent magenta models (**b**) were fitted into the frame 0 and frame 19 maps of the cryoSPARC 3DVA, respectively. The arrows along the tail domains in **b** highlight the up-down motion of the tail domains of opposite $n + 10$ strands. **c-g** Atomic models and motion trajectories showing the flexibility of the head domains along a full filament turn. Eight ArlB1 and eight ArlB2 monomers forming a full helical turn were selected (**c**) and their tail domains aligned. Panels **e, g** represent the motion trajectories of the centroids (shades of blue) of each head domain. The black curve highlights the trajectory, and the dotted black line represents the filament axis. **h** Atomic models of one ArlB1 and one ArlB2 head domains showing their displacement during filament motion. The motion is visualised by vectors linking the backbone atoms of the atomic models fitting into the frame 0 (grey line) and frame 19 (not shown for better clarity) maps of the cryoSPARC 3DVA. **i** Estimation of the evolutionary conservation of ArlB1 and ArlB2. Insets focus on the hinge region. Scale bars in **a** and **c**, 50 Å; in **b, d, f, h**, 10 Å; in **e** and **g**, 2 Å.

were harvested at the early stationary phase (7000–8000 × g) and concentrated (3500 × g, 30 min, 4 °C; Sorvall RC 5 C plus, rotor GS3). A total of 17.5 g of cells (wet weight) were obtained. Cells were resuspended in a small amount of medium and then archaeella were detached by shearing force and cavitation, causing disruption of the cells (ULTRA-TURRAX T25 high-speed homogeniser; 1 min at 13,000 rpm and 10 sec at 22,000 rpm). Cell debris was separated from archaeella by centrifugation (34,500 × g, 4 °C; Sorvall RC 5 C plus, rotor SS34) for 20 min. Archaeella were pelleted from the supernatant by ultracentrifugation (60,000 × g, 90 min, 4 °C; Beckman Optima LE-80K, rotor 70Ti), resuspended in 150 µl 0.1 M HEPES/NAOH pH 7.0 and purified for 48 h (250,000 × g, 4 °C; Beckman Optima LE-80K, rotor SW60-Ti) using a CsCl gradient (0.45 g/ml; Biomol). Fractions were taken by puncturing the ultracentrifuge tubes with sterile syringes, dialyzed against aerobic ½ SME/5 mM HEPES/NaOH pH 7.0, and analysed by SDS-PAGE and transmission electron microscopy¹¹ to identify the archaeella-containing band. After addition of sodium azide to a final concentration of 0.001% (w/v), samples were stored at 4 °C.

CryoEM of archaeella and helical processing. Prior to freezing, a 3 µl drop of undiluted suspension containing purified archaeella was applied to glow-discharged 300 mesh copper R2/2 Quantifoil grids. The grids were blotted with 597 Whatman filter paper for 4 sec, using –1 blot force, in 95% relative humidity, at 21 °C and plunge-frozen in liquid ethane using a Mark IV Vitrobot (FEI). Dose-fractionated movies were collected in linear mode using a 300 kV Titan Krios equipped with a Falcon III detector (Thermo Fisher Scientific) (Supplementary Table 1). 39-fraction movies were recorded using EPU 1.5 (Thermo Fisher Scientific) with an exposure time of 1 sec and a total electron dose of 37 e⁻/Å² at a defocus range of –2.3 to –1.1 µm (0.3 µm increment). A total of 2759 movies were collected and processed using the Relion 3.1 pipeline³⁸ and finalised using cryoSPARC 3.1.0³⁹. Briefly, the movies were motion-corrected using MotionCorr²⁶⁸ and CTF estimation was carried out using CTFFind⁴⁶⁹. Helical segments were manually picked using Helixboxer from the EMAN-2 package⁷⁰. The coordinates of the picked particles were imported into Relion 3.1, where particle extraction, 2D classification, 3D classification, 3D refinement, B-factor sharpening, CTF-refinement and Bayesian polishing were performed. Each filament was subdivided into 256 × 256 pixel size boxes at an offset of 6 Å, resulting in 929,165 segments upon particle extraction. The dataset was then 3D classified (8 classes) and the best class (containing 399,178 segments) was selected for 3D refinement using the helical parameters published for the *P. furiosus* archaeum²⁰. The refined consensus map was improved using CTF-refinement and Bayesian polishing in Relion 3.1. The shiny particles obtained from the Bayesian polishing were further refined using the optimised helical parameters 5.57 Å rise and 108° twist after helical refinement. The resulting map had a resolution of 3.29 Å. The particles of the final Relion 3.1 helical refinement were imported into cryoSPARC 3.1.0 and a new refinement was performed using the helical refinement (BETA) algorithm without imposing helical parameters. This resulted in a reconstruction at a resolution of 3.28 Å by gold-standard FSC in which two distinct subunits could be distinguished. To corroborate these results, another round of 3D refinement with no helical symmetry was performed in Relion 3.1, which provided a comparable result. The two subunits alternate along the left-handed 3-start helices as shown on the archaeum net diagram (Fig. 4f, Supplementary Fig. 5a, e). An attempt to perform a helical refinement with (n + 2) helical parameter 11.14 Å rise and 216° twist resulted in a filament reconstruction with lower resolution and poorly resolved features. Further inspection of the net diagram allowed to establish the minimal transformation of the *M. villosus* archaeum in which each monomer superimposes onto its equivalent, and which proved to be n + 6. A final helical refinement was performed in cryoSPARC 3.1.0, this time imposing a (n + 6) helical parameters 33.4 Å rise and –71.8° twist (see results below). This produced a final map of 3.08 Å resolution by gold-standard FSC 0.143. DeepEMhancer was used for final map sharpening⁷¹.

Model building and validation. Initially, the *P. furiosus* ArlB0 archaeum structure (PDB ID: 5O4U) was positioned into the Relion 3.1 cryoEM map, refined using REFMAC5⁷² and rebuilt to match the *M. villosus* ArlB2 sequence using Coot⁷³. Both candidate sequences (ArlB1 and ArlB3) were attempted to be fitted by Buccaneer⁷⁴ into the density of the second archaeum found in the higher resolution cryoSPARC 3.1.0 cryoEM map. The sequence match the cryoEM map features and glycosylation pattern unambiguously identified ArlB1 as the second protein forming the archaeum. Both ArlB1 and ArlB2 atomic models were then validated using Isolve⁷⁵. Two full turns of the archaeum filament were built in UCSF Chimera⁷⁶ using ArlB1–ArlB2 3-start dimers. Optimised helical parameters with the position of each monomer were rigid-body refined in Coot to adjust for local variations in the non-helically refined map. A script containing a number of CCP4⁷⁷ programs was prepared to propagate any modification to a structure of a single monomer into a full filament. The resulting model was refined using REFMAC5 in CCP-EM interface⁷⁸. Glycans were modelled in Coot using the sugar unit identity and connectivity reported for the glycan structure of *M. thermophilus*. This structure was recently reported as α-d-glycero-d-manno-Hep3OMe6OME-(1–3)-[α-GalNAcA3OME-(1–2)]-β-Man-(1–4)-[β-GalA3OME4OAc6CME-(1–4)-α-GalA-(1–2)]-α-GalAN-(1–3)-β-GalNAc-Asn³⁰. Restraints dictionaries for non-standard sugars were prepared using JLIGAND⁷⁹.

The calcium ion in ArlB1 and ArlB2 was modelled on the basis of the metal coordination and distances observed in our structure and agreed with the metal assignment in³⁴.

3D variability analysis and molecular flexibility. 3D variability analysis (3DVA)⁴² was performed in cryoSPARC 3.1.0 using the 399,178 helical segments from the cryoSPARC 3.1.0 helical refinement. First, the symmetry expansion step was run. Then, the particles stack created by the symmetry expansion was used for 3DVA. Different lowpass filter values were tested (3.5, 5, 8 and 10 Å) in simple mode with 20 intermediate frames and with 10 variability components as output. All components showed the same type of motion.

In order to investigate the filament's flexibility at the molecular level, we selected a 3DVA component obtained at 3.5 Å lowpass filter and used the first frame (number 0) and last frame (number 19) of the component corresponding to two cryoEM maps at the two extremes of the filament motion range. The filament's atomic model was positioned into the two cryoEM maps using phased molecular replacement in MOLREP⁸⁰ and refined using Phenix⁸¹, thereby producing two atomic models fitting the two cryoEM maps. The two atomic models were finally morphed using the morphing function in UCSF Chimera to display the full conformational space of the filament (Supplementary Movie 4), in particular the tail domains (Fig. 5a, b; Supplementary Movie 5). To show the local flexibility of the head domains, one ArlB1 and one ArlB2 subunits were isolated from the two atomic models fitting the two cryoEM maps at frame 0 and frame 19 and r.m.s.d.-coloured vectors were calculated linking the Ca of the head domains (Fig. 5h). To investigate the flexibility of the head domains along the filament, two homopolymeric strands of ArlB1 and ArlB2 were isolated from the atomic model fitting the cryoEM map at frame 0. Then for each strand the tails of its constituting subunits forming a full filament turn were aligned (Fig. 5c–g; Supplementary Movie 6, 7).

STEM tomography. Cells were cryo-immobilised using high-pressure freezing (Leica HPM 100, Munich), then freeze-substituted in ethanol containing 0.5% glutaraldehyde, 1% formaldehyde, 0.5% uranyl acetate, 5% water and finally embedded into Spurr resin. Semithin sections with a thickness of 800–900 nm were cut using a diamond knife (Diatome) on an ultramicrotome (Leica UC6). Sections were mounted on formvar-coated grids (100 mesh Cu grids; Plano, Wetzlar). For STEM tomography at 200 kV⁸², a 3 nm carbon layer (Crossing turbo 208 Carbon) was evaporated onto one side of the sections. Fiducials (protein A-gold, 15 nm; University Medical Center, Utrecht, NL) were added onto both sides of the sections, to aid the alignment and reconstruction process. STEM tomography was done on a JEM-2100F (JEOL GmbH, Freising, Germany) using an electron beam with a small semi-convergence angle, essentially as described⁸². Tiltseries (STEM bright field) were recorded from +66 deg to –66 deg, using the EM-TOOLS software (TVIPS, Gauting, Germany). Reconstruction of 3D volumes was done in IMOD (Univ. of Colorado at Boulder, USA) using the weighted back projection, incl. a SIRT-like filter equivalent to 15 iterations.

Sequence analyses. The evolutionary conservation of amino acids positions in ArlB1 and ArlB2 (Fig. 5i) and bacterial T4P (Supplementary Fig. 17c) was estimated using the ConSurf server⁸³. The sequence alignment and secondary structure assignment of *M. villosus* ArlB1 and ArlB2 were performed with BLASTp and visualised by ESPript⁸⁴ (Supplementary Fig. 9). The archaeal operons (Supplementary Fig. 18) were visualised using ARTEMIS⁸⁵ and the automatic genome annotations were downloaded from NCBI. The sequence comparison and alignment between archaeal and archaeal T4P (Supplementary Fig. 17a,b), and *M. maripaludis* ArlB3, *M. voltae* ArlB3 and *M. villosus* ArlB1, 2, 3 archaeal (Supplementary Fig. 19a) were performed with BLASTp and the Praline server⁸⁶.

Structure analysis and presentation. The structure of the *M. villosus* archaeum filament was visualised using UCSF Chimera, Chimera-X⁸⁷, Coot and CCP4mg⁸⁸. The topology diagrams were prepared using Pro-origami⁸⁹ and TopDraw⁹⁰. Interfaces between different ArlB1 and ArlB2 monomers in the archaeum were analysed using the PISA software⁹¹. The Molprobit score in Supplementary Table 1 was calculated using⁹².

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The archaeum atomic coordinates and electron density map generated in this study have been deposited in the Protein Data Bank (<https://www.rcsb.org/>) with accession number 7OFQ and in the EM DataResource (<https://www.emdataresource.org/>) with the accession number EMDB–12875. The raw movies have been deposited in the EMPIAR database (<https://www.ebi.ac.uk/empair/>) with accession number EMPIAR-10884. Other structural data used in this study are: *P. furiosus* ArlB0 (PDB ID: 5O4U), *M. jannaschii* ArlB1 (5YA6), *T. thermophilus* PilA4 (PDB ID: 6XXD), *P. arsenaticum* PilA (PDB ID: 6W8U).

Received: 6 October 2021; Accepted: 20 January 2022;

Published online: 07 February 2022

References

- Woese, C. R., Magrum, L. J. & Fox, G. E. Archaeobacteria. *J. Mol. Evol.* **11**, 245–251 (1978).
- DeLong, E. F. & Pace, N. R. Environmental diversity of bacteria and archaea. *Syst. Biol.* **50**, 470–478 (2001).
- Schleper, C., Jurgens, G. & Jonuscheit, M. Genomic studies of uncultivated archaea. *Nat. Rev. Microbiol.* **3**, 479–488 (2005).
- Albers, S. V. & Jarrell, K. F. The archaeum: how Archaea swim. *Front. Microbiol.* **6**, 23 (2015).
- Marwan, W., Alam, M. & Oesterheld, D. Rotation and switching of the flagellar motor assembly in *Halobacterium halobium*. *J. Bacteriol.* **173**, 1971–1977 (1991).
- Albers, S. V. & Jarrell, K. F. The Archaeum: an update on the unique archaeal motility structure. *Trends Microbiol.* **26**, 351–362 (2018).
- Berry, J. L. & Pelicic, V. Exceptionally widespread nanomachines composed of type IV pili: the prokaryotic Swiss Army knives. *FEMS Microbiol. Rev.* **39**, 134–154 (2015).
- Jarrell, K. F., Stark, M., Nair, D. B. & Chong, J. P. Flagella and pili are both necessary for efficient attachment of *Methanococcus maripaludis* to surfaces. *FEMS Microbiol. Lett.* **319**, 44–50 (2011).
- Yu, X. et al. Filaments from *Ignicoccus hospitalis* show diversity of packing in proteins containing N-terminal type IV pilin helices. *J. Mol. Biol.* **422**, 274–281 (2012).
- Näther, D. J., 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.* **188**, 6915–6923 (2006).
- 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. Microbiol.* **61**, 1239–1245 (2011).
- Pohlschroder, M. & Esquivel, R. N. Archaeal type IV pili and their involvement in biofilm formation. *Front. Microbiol.* **6**, 190 (2015).
- Beeby, M., Ferreira, J. L., Tripp, P., Albers, S. V. & Mitchell, D. R. Propulsive nanomachines: the convergent evolution of archaella, flagella and cilia. *FEMS Microbiol. Rev.* **44**, 253–304 (2020).
- Cohen-Krausz, S. & Trachtenberg, S. The flagellar filament structure of the extreme acidothermophile *Sulfolobus shibatae* B12 suggests that archaeobacterial flagella have a unique and common symmetry and design. *J. Mol. Biol.* **375**, 1113–1124 (2008).
- Jarrell, K. F., Bayley, D. P. & Kostyukova, A. S. The archaeal flagellum: a unique motility structure. *J. Bacteriol.* **178**, 5057–5064 (1996).
- Reindl, S. et al. Insights into FlaI functions in archaeal motor assembly and motility from structures, conformations, and genetics. *Mol. Cell* **49**, 1069–1082 (2013).
- Meshcheryakov, V. A. & Wolf, M. Crystal structure of the flagellar accessory protein FlaH of *Methanocaldococcus jannaschii* suggests a regulatory role in archaeal flagellum assembly. *Protein Sci.* **25**, 1147–1155 (2016).
- Chaudhury, P. et al. The nucleotide-dependent interaction of FlaH and FlaI is essential for assembly and function of the archaeum motor. *Mol. Microbiol.* **99**, 674–685 (2016).
- Banerjee, A. et al. FlaX, a unique component of the crenarchaeal archaeum, forms oligomeric ring-shaped structures and interacts with the motor ATPase FlaI. *J. Biol. Chem.* **287**, 43322–43330 (2012).
- Daum, B. et al. Structure and in situ organisation of the *Pyrococcus furiosus* archaeum machinery. *Elife* <https://doi.org/10.7554/eLife.27470> (2017).
- Briegel, A. et al. Morphology of the archaeal motor and associated cytoplasmic cone in *Thermococcus kodakaraensis*. *EMBO Rep.* **18**, 1660–1670 (2017).
- Banerjee, A. et al. FlaF is a β -sandwich protein that anchors the archaeum in the archaeal cell envelope by binding the S-layer protein. *Structure* **23**, 863–872 (2015).
- Tsai, C. L. et al. The structure of the periplasmic FlaG-FlaF complex and its essential role for archaeal swimming motility. *Nat. Microbiol.* **5**, 216–225 (2020).
- Tschitschko, B. et al. Genomic variation and biogeography of Antarctic haloarchaea. *Microbiome* **6**, 113 (2018).
- Ng, S. Y., Chaban, B. & Jarrell, K. F. Archaeal flagella, bacterial flagella and type IV pili: a comparison of genes and posttranslational modifications. *J. Mol. Microbiol. Biotechnol.* **11**, 167–191 (2006).
- Kalmokoff, M. L. & Jarrell, K. F. Cloning and sequencing of a multigene family encoding the flagellins of *Methanococcus voltae*. *J. Bacteriol.* **173**, 7113–7125 (1991).
- Näther-Schindler, D. J., Schopf, S., Bellack, A., Rachel, R. & Wirth, R. *Pyrococcus furiosus* flagella: biochemical and transcriptional analyses identify the newly detected flaB0 gene to encode the major flagellin. *Front. Microbiol.* **5**, 695 (2014).
- Kalmokoff, M. L., Jarrell, K. F. & Koval, S. F. Isolation of flagella from the archaeobacterium *Methanococcus voltae* by phase separation with Triton X-114. *J. Bacteriol.* **170**, 1752–1758 (1988).
- 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.* **66**, 596–609 (2007).
- Kelly, J. F. et al. Identification of a novel N-linked glycan on the archaeal and S-layer protein of the thermophilic methanogen, *Methanothermococcus thermolithotrophicus*. *J. Biol. Chem.* **295**, 14618–14629 (2020).
- Tarasov, V. Y., Pyatibratov, M. G., Tang, S. L., Dyal-Smith, M. & Fedorov, O. V. Role of flagellins from A and B loci in flagella formation of *Halobacterium salinarum*. *Mol. Microbiol.* **35**, 69–78 (2000).
- Pyatibratov, M. G. et al. Interaction of two strongly divergent archaeal flagellins stabilizes the structure of the *Halorubrum* archaeum. *Microbiologyopen* **9**, e1047 (2020).
- Poweleit, N. et al. CryoEM structure of the *Methanospirillum hungatei* archaeum reveals structural features distinct from the bacterial flagellum and type IV pilus. *Nat. Microbiol.* **2**, 16222 (2016).
- Meshcheryakov, V. A. et al. High-resolution archaeum structure reveals a conserved metal-binding site. *EMBO Rep.* <https://doi.org/10.15252/embr.201846340> (2019).
- Bardy, S. L., Mori, T., Komoriya, K., Aizawa, S. & Jarrell, K. F. Identification and localization of flagellins FlaA and FlaB3 within flagella of *Methanococcus voltae*. *J. Bacteriol.* **184**, 5223–5233 (2002).
- Bezanosov, S. N., Pyatibratov, M. G. & Fedorov, O. V. On the multicomponent nature of *Halobacterium salinarum* flagella. *Microbiology* **76**, 435–441 (2007).
- Herzog, B. & Wirth, R. Swimming behavior of selected species of Archaea. *Appl. Environ. Microbiol.* **78**, 1670–1674 (2012).
- Scheres, S. H. W. Amyloid structure determination in RELION-3.1. *Acta Crystallogr. D. Struct. Biol.* **76**, 94–101 (2020).
- 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).
- Szabó, Z. et al. Identification of diverse archaeal proteins with class III signal peptides cleaved by distinct archaeal prepilin peptidases. *J. Bacteriol.* **189**, 772–778 (2007).
- Kinosita, Y., Uchida, N., Nakane, D. & Nishizaka, T. Direct observation of rotation and steps of the archaeum in the swimming halophilic archaeon *Halobacterium salinarum*. *Nat. Microbiol.* **1**, 16148 (2016).
- Punjani, A. & Fleet, D. J. 3D variability analysis: resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM. *J. Struct. Biol.* <https://doi.org/10.1016/j.jsb.2021.107702> (2021).
- Desmond, E., Brochier-Armanet, C. & Gribaldo, S. Phylogenomics of the archaeal flagellum: rare horizontal gene transfer in a unique motility structure. *BMC Evol. Biol.* **7**, 106 (2007).
- Syutkin, A. S. et al. Salt-dependent regulation of archaeal flagellins in *Haloarcula marismortui*. *Microbiologyopen* **8**, e00718 (2019).
- Wang, F. et al. The structures of two archaeal type IV pili illuminate evolutionary relationships. *Nat. Commun.* **11**, 3424 (2020).
- Wang, F. et al. An extensively glycosylated archaeal pilus survives extreme conditions. *Nat. Microbiol.* **4**, 1401–1410 (2019).
- Wang, F. et al. Cryoelectron microscopy reconstructions of the *Pseudomonas aeruginosa* and *Neisseria gonorrhoeae* type IV Pili at sub-nanometer resolution. *Structure* **25**, 1423–1435.e1424 (2017).
- Kolappan, S. et al. Structure of the *Neisseria meningitidis* Type IV pilus. *Nat. Commun.* **7**, 13015 (2016).
- Neuhaus, A. et al. Cryo-electron microscopy reveals two distinct type IV pili assembled by the same bacterium. *Nat. Commun.* **11**, 2231 (2020).
- Berry, J. L. et al. Global biochemical and structural analysis of the type IV pilus from the Gram-positive bacterium *Streptococcus sanguinis*. *J. Biol. Chem.* **294**, 6796–6808 (2019).
- Jacobsen, T., Bardiaux, B., Francetic, O., Izadi-Pruneyre, N. & Nilges, M. Structure and function of minor pilins of type IV pili. *Med. Microbiol. Immunol.* **209**, 301–308 (2020).
- Cruden, D., Sparling, R. & Markovetz, A. J. Isolation and ultrastructure of the flagella of *Methanococcus thermolithotrophicus* and *Methanospirillum hungatei*. *Appl. Environ. Microbiol.* **55**, 1414–1419 (1989).
- Karamyshev, A. L. et al. Mapping the interaction of the STT3 subunit of the oligosaccharyl transferase complex with nascent polypeptide chains. *J. Biol. Chem.* **280**, 40489–40493 (2005).
- Matczuk, A. K., Kunec, D. & Veit, M. Co-translational processing of glycoprotein 3 from equine arteritis virus: N-glycosylation adjacent to the signal peptide prevents cleavage. *J. Biol. Chem.* **288**, 35396–35405 (2013).
- Meyer, B. H., Birich, A. & Albers, S. V. N-Glycosylation of the archaeum filament is not important for archaella assembly and motility, although

- N-Glycosylation is essential for motility in *Sulfolobus acidocaldarius*. *Biochimie* **118**, 294–301 (2015).
56. Zaretsky, M., Darnell, C. L., Schmid, A. K. & Eichler, J. N-Glycosylation is important for Halobacterium salinarum archaeal expression, archaeum assembly and cell motility. *Front. Microbiol.* <https://doi.org/10.3389/fmicb.2019.01367> (2019).
 57. Ding, Y. et al. Effects of N-glycosylation site removal in archaealins on the assembly and function of archaea in *Methanococcus maripaludis*. *PLoS ONE* **10**, e0116402 (2015).
 58. Tripepi, M. et al. N-glycosylation of Haloferax volcanii flagellins requires known Agl proteins and is essential for biosynthesis of stable flagella. *J. Bacteriol.* **194**, 4876–4887 (2012).
 59. Chaban, B., Voisin, S., Kelly, J., Logan, S. M. & Jarrell, K. F. Identification of genes involved in the biosynthesis and attachment of *Methanococcus voltae* N-linked glycans: insight into N-linked glycosylation pathways in Archaea. *Mol. Microbiol.* **61**, 259–268 (2006).
 60. Shental-Bechor, D. & Levy, Y. Effect of glycosylation on protein folding: a close look at thermodynamic stabilization. *Proc. Natl Acad. Sci. USA* **105**, 8256–8261 (2008).
 61. Wang, F. et al. A structural model of flagellar filament switching across multiple bacterial species. *Nat. Commun.* **8**, 960 (2017).
 62. Yamaguchi, T. et al. Structure of the molecular bushing of the bacterial flagellar motor. *bioRxiv* <https://doi.org/10.1101/2020.11.12.379347> (2020).
 63. Shibata, S., Matsunami, H., Aizawa, S. I. & Wolf, M. Torque transmission mechanism of the curved bacterial flagellar hook revealed by cryo-EM. *Nat. Struct. Mol. Biol.* **26**, 941–945 (2019).
 64. Kato, T., Makino, F., Miyata, T., Horváth, P. & Namba, K. Structure of the native supercoiled flagellar hook as a universal joint. *Nat. Commun.* **10**, 5295 (2019).
 65. Balch, W. E. & Wolfe, R. S. New approach to the cultivation of methanogenic bacteria: 2-mercaptoethanesulfonic acid (HS-CoM)-dependent growth of *Methanobacterium ruminantium* in a pressurized atmosphere. *Appl. Environ. Microbiol.* **32**, 781–791 (1976).
 66. Fricke, H. et al. Hydrothermal vent communities at the shallow subpolar Mid-Atlantic ridge. *Marine Biology*. **102**, 425–429 (1989).
 67. Huber, H. & Stetter, K. O. Desulfurococcales. *Prokaryotes* **3**, 16 (2006).
 68. Li, X. et al. Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat. Methods* **10**, 584–590 (2013).
 69. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
 70. Tang, G. et al. EMAN2: an extensible image processing suite for electron microscopy. *J. Struct. Biol.* **157**, 38–46 (2007).
 71. Sanchez-Garcia, R. et al. DeepEMhancer: a deep learning solution for cryo-EM volume post-processing. *bioRxiv* <https://doi.org/10.1101/2020.06.12.148296> (2020).
 72. Murshudov, G. N. et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D. Biol. Crystallogr.* **67**, 355–367 (2011).
 73. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 486–501 (2010).
 74. Cowtan, K. The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. D. Biol. Crystallogr.* **62**, 1002–1011 (2006).
 75. Croll, T. I. ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr. D. Struct. Biol.* **74**, 519–530 (2018).
 76. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
 77. Winn, M. D. et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr. D. Biol. Crystallogr.* **67**, 235–242 (2011).
 78. Burnley, T., Palmer, C. M. & Winn, M. Recent developments in the CCP-EM software suite. *Acta Crystallogr. D. Struct. Biol.* **73**, 469–477 (2017).
 79. Lebedev, A. A. et al. Jligand: a graphical tool for the CCP4 template-restraint library. *Acta Crystallogr. D. Biol. Crystallogr.* **68**, 431–440 (2012).
 80. Vagin, A. A. & Isupov, M. N. Spherically averaged phased translation function and its application to the search for molecules and fragments in electron-density maps. *Acta Crystallogr. D. Biol. Crystallogr.* **57**, 1451–1456 (2001).
 81. Afonine, P. V. et al. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D. Biol. Crystallogr.* **68**, 352–367 (2012).
 82. Rachel, R. et al. Dual-axis STEM tomography at 200 kV: Setup, performance, limitations. *J. Struct. Biol.* **211**, 107551 (2020).
 83. Ashkenazy, H. et al. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res.* **44**, W344–W350 (2016).
 84. Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**, W320–W324 (2014).
 85. Carver, T., Harris, S. R., Berriman, M., Parkhill, J. & McQuillan, J. A. Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinformatics* **28**, 464–469 (2012).
 86. Heringa, J. Two strategies for sequence comparison: profile-preprocessed and secondary structure-induced multiple alignment. *Comput. Chem.* **23**, 341–364 (1999).
 87. Pettersen, E. F. et al. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82 (2021).
 88. McNicholas, S., Potterton, E., Wilson, K. S. & Noble, M. E. Presenting your structures: the CCP4mg molecular-graphics software. *Acta Crystallogr. D. Biol. Crystallogr.* **67**, 386–394 (2011).
 89. Stivala, A., Wybrow, M., Wirth, A., Whisstock, J. C. & Stuckey, P. J. Automatic generation of protein structure cartoons with Pro-origami. *Bioinformatics* **27**, 3315–3316 (2011).
 90. Bond, C. S. TopDraw: a sketchpad for protein structure topology cartoons. *Bioinformatics* **19**, 311–312 (2003).
 91. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).
 92. Williams, C. J. et al. MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci.* **27**, 293–315 (2018).

Acknowledgements

We thank the RMS for providing microscope time as part of the Single Particle Cryo-TEM school, University of Leeds. The FEI Titan Krios microscopes were funded by the University of Leeds (UoL ABSL award) and the Wellcome Trust (108466/Z/15/Z). We also acknowledge Alexander Neuhaus for his invaluable assistance with helical reconstruction; Erik Zupa for providing us with the PyMOL script to plot r.m.s.d.-coloured vectors, and Simone Antonio De Rose for help with helical net diagrams. Molecular graphics and analyses were performed with UCSF Chimera, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from NIH P41-GM103311. We are grateful to Yvonne Bilek, Thomas Hader and Konrad Eichinger for their help with mass cultivation of *M. villosus* cells. For this project, L.G., B.D. and M.M. were funded by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 803894). R.C. was funded by the University of Exeter start up funds awarded to V.G., and a Wellcome Trust Seed Award in Science (210363/Z/18/Z) awarded to V.G.; A.B. and R.R. received funding from the Deutsche Forschungsgemeinschaft (WI 731/10-1) awarded to R.R. and Reinhard Wirth. The authors would like to dedicate this study to the memory of Reinhard Wirth, who supervised initial studies on *M. villosus* archaea.

Author contributions

L.G.: Performed the research, provided methodology, wrote the manuscript. M.I.: Performed the research, provided methodology, wrote the manuscript. R.C.: Provided methodology. M.M.: Provided methodology. V.G.: Provided resources and funding for R.C., contributed to writing of the manuscript. A.B.: Provided resources and methodology, contributed to writing of the manuscript. R.R.: Provided resources and methodology. B.D.: Provided funding for L.G., performed research, provided resources, wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-022-28337-1>.

Correspondence and requests for materials should be addressed to Bertram Daum.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022