


















# Donor strand complementation and calcium ion coordination drive the chaperone-free polymerization of archaeal cannulae

Received: 2 March 2025

Accepted: 9 September 2025

Published online: 13 October 2025



Mike Sleutel <sup>1,2,10</sup>, Ravi R. Sonani <sup>3,10</sup>, Jessalyn G. Miller <sup>4,5,10</sup>,  
Fengbin Wang <sup>3,6</sup>, Andres Gonzalez Socorro <sup>4</sup>, Yang Chen <sup>5</sup>,  
Reece Martin <sup>7</sup>, Borries Demeler <sup>7</sup>, Michael J. Rudolph <sup>5</sup>, Vikram Alva <sup>8</sup>,  
Han Remaut <sup>1,2</sup>  , Edward H. Egelman <sup>3</sup>  & Vincent P. Conticello <sup>4,9</sup> 

Cannulae are structurally rigid tubular protein filaments that accumulate on the extracellular surface of archaea within the family *Pyrodictiaceae* during cell growth. These obligate anaerobes propagate under hyperthermophilic conditions in which cannulae form a biomatrix that interconnects and sustains cells. The persistence of cannulae in this environment suggests that these filaments display significant thermostability, which has attracted technological interest in their development as synthetic protein-based biomaterials. Here, we report cryoEM structural analyses of ex vivo and in vitro assembled recombinant cannulae. We demonstrate that the interactions between protomers in native and recombinant cannulae is based on donor strand complementation (DSC), a form of non-covalent polymerization previously observed for bacterial chaperone-usher pili. Unexpectedly, calcium ion coordination at the subunit interfaces reinforces the network of donor strand interactions in the cannulae. This study provides insight into the mechanism of assembly of cannulae and the structural origin of their high stability and rigidity.

Filamentous protein assemblies are common in biological systems in which they are involved in a wide range of functions that are critical for the survival and propagation of the organism<sup>1</sup>. Many of these functions, e.g., locomotion<sup>2</sup>, adhesion<sup>3,4</sup>, tunable mechanical response<sup>3–8</sup>, multicellular organization<sup>9–12</sup>, electrical conductivity<sup>13–16</sup>, directional transport of substrates<sup>17–21</sup>, regulation of enzymatic catalysis<sup>22</sup>, etc., would be desirable to emulate in synthetic protein-based nanomaterials that could be tailored for specific applications<sup>23</sup>. However, this approach requires prior knowledge of the sequence-structure-

function relationships that are responsible for the observed behavior of the corresponding biologically derived assemblies, as well as convenient methods for the ex vivo isolation or in vitro fabrication of the protein filaments. Structural analyses at near-atomic resolution have illuminated the physical principles that underlie the function of many biologically derived protein-based nanomaterials<sup>24,25</sup>. This structural information provides a critical starting point for engineering synthetic analogues that mimic the function of the corresponding native protein assemblies<sup>26</sup>. High-resolution structural information can also enable an

<sup>1</sup>Structural Biology Brussels, Vrije Universiteit Brussel, Brussels, Belgium. <sup>2</sup>Structural and Molecular Microbiology, VIB-VUB Center for Structural Biology, Brussels, Belgium. <sup>3</sup>Department of Biochemistry and Molecular Genetics, University of Virginia, Charlottesville, VA, USA. <sup>4</sup>Department of Chemistry, Emory University, Atlanta, GA, USA. <sup>5</sup>New York Structural Biology Center, New York, NY, USA. <sup>6</sup>Biochemistry and Molecular Genetics Department, University of Alabama at Birmingham, Birmingham, AL, USA. <sup>7</sup>Department of Chemistry and Biochemistry, University of Lethbridge, Lethbridge, AB, Canada. <sup>8</sup>Department of Protein Evolution, Max Planck Institute for Biology Tübingen, Tübingen, Germany. <sup>9</sup>The Robert P. Apkarian Integrated Electron Microscopy Core (IEMC), Emory University, Atlanta, GA, USA. <sup>10</sup>These authors contributed equally: Mike Sleutel, Ravi R. Sonani, Jessalyn G. Miller. ✉ e-mail: [Han.Remaut@vub.be](mailto:Han.Remaut@vub.be); [egelman@virginia.edu](mailto:egelman@virginia.edu); [vcontic@emory.edu](mailto:vcontic@emory.edu)

understanding of the molecular-level interactions that underlie the mechanism of self-assembly, thereby facilitating the development of reliable methods for recombinant production of precursor proteins that can be fabricated into functional filaments that display the desired properties<sup>27,28</sup>.

Here, we employ electron cryomicroscopy (cryoEM) to investigate the structure of cannulae, a structurally unique class of protein-based filaments expressed on the extracellular surface of the hyperthermophilic archaeon *Pyrodicticum abyssi* and related organisms in the family *Pyrodictiaceae*<sup>29–32</sup>. Cannulae form during the life cycle of *P. abyssi*, *occultum*, *brockii* strains under culture conditions that mimic the native growth environment<sup>31–33</sup>. *P. abyssi* is an obligate anaerobe that propagates under hyperthermophilic, deep-sea conditions with optimal growth in the temperature range from 80 °C to 110 °C<sup>34</sup>. The persistence of cannulae under these conditions suggested that the filaments display significant thermostability, which has attracted biotechnological interest for potential use as robust assemblies for the development of synthetic protein-based biomaterials<sup>35</sup>. However, *P. abyssi* strains are challenging to culture at preparative scale and not genetically tractable at present<sup>36</sup>.

The cannulae of *P. abyssi* have been the most thoroughly studied experimental system. Cell cultures of *P. abyssi* display robust growth and achieved high cell densities, albeit under restrictive culture conditions, e.g., high temperature and pressure and corrosive media<sup>32</sup>. Dark field microscopy of growing cultures of *P. abyssi* provided evidence that the flat cells divide through binary fission<sup>36,37</sup>. During cell division, cannulae are synthesized on the cell surface such that the daughter cells remain connected after separation. These extracellular connections were maintained in subsequent rounds of cell division, which resulted in the formation of a dense network of cells that were linked through cannulae. In liquid culture, *P. abyssi* cells were never observed without cannulae, which have been postulated to serve an essential role in cellular propagation. Electron cryotomography (cryoET) studies of *P. abyssi* cultures<sup>38</sup> indicated that the cannulae passed through the S-layer and embedded in the pseudo-periplasmic space. While the biological function of cannulae remains a matter of speculation, cryoET analysis provided evidence for density within the lumen of cannula connecting *P. abyssi* cells, which led to a proposed role for cannulae as a primitive extracellular matrix through which cells could communicate or exchange material.

To gain insight into the structural principles underlying these unique physical properties, cryoEM three-dimensional reconstructions were performed using ex vivo and in vitro recombinant cannulae from *P. abyssi* as substrates<sup>29–32</sup>. These structural analyses revealed that the interactions between protomers in ex vivo and in vitro filaments were based on donor strand complementation (DSC), a form of non-covalent polymerization in which a  $\beta$ -strand from one subunit is inserted into the acceptor groove of a  $\beta$ -sheet in a neighboring subunit. DSC has been observed in structures of bacterial chaperone-usher (CU) pili in which extracellular polymerization occurs through a distinct process<sup>3–7,9–11,39,40</sup>. In cannulae, the donor strand-acceptor groove interaction is reinforced through calcium ion coordination between structurally adjacent protomers. Crystallographic analysis of a recombinant cannula-like protein suggested that calcium ion binding primed polymerization of the monomer into a filament. Bioinformatic analysis provided evidence that structurally similar cannula-like proteins were encoded within the genomes of other hyperthermophilic archaea and were encompassed within the TasA superfamily of biomatrix proteins. CryoEM structural analysis of tubular filaments derived from in vitro assembly of a recombinant cannula-like protein from an uncultured archaeon in phylum *Thermoproteota*, tentatively identified as a *Hyperthermus* species, revealed a common mode of assembly to the *Pyrodicticum* cannulae, in which DSC and calcium ion binding stabilized axial and radial assembly. The facile in vitro assembly of cannula-mimetic tubules represents an attractive

methodology for the fabrication of thermodynamically stable and biochemically addressable filamentous protein nanomaterials. In contrast to CU pili, in which the protein subunits undergo catalytic polymerization requiring a membrane-bound translocon, recombinant cannula proteins self-assemble spontaneously in the presence of calcium ions in aqueous solution under controlled conditions in vitro and displayed a similar structure to the ex vivo *P. abyssi* filament.

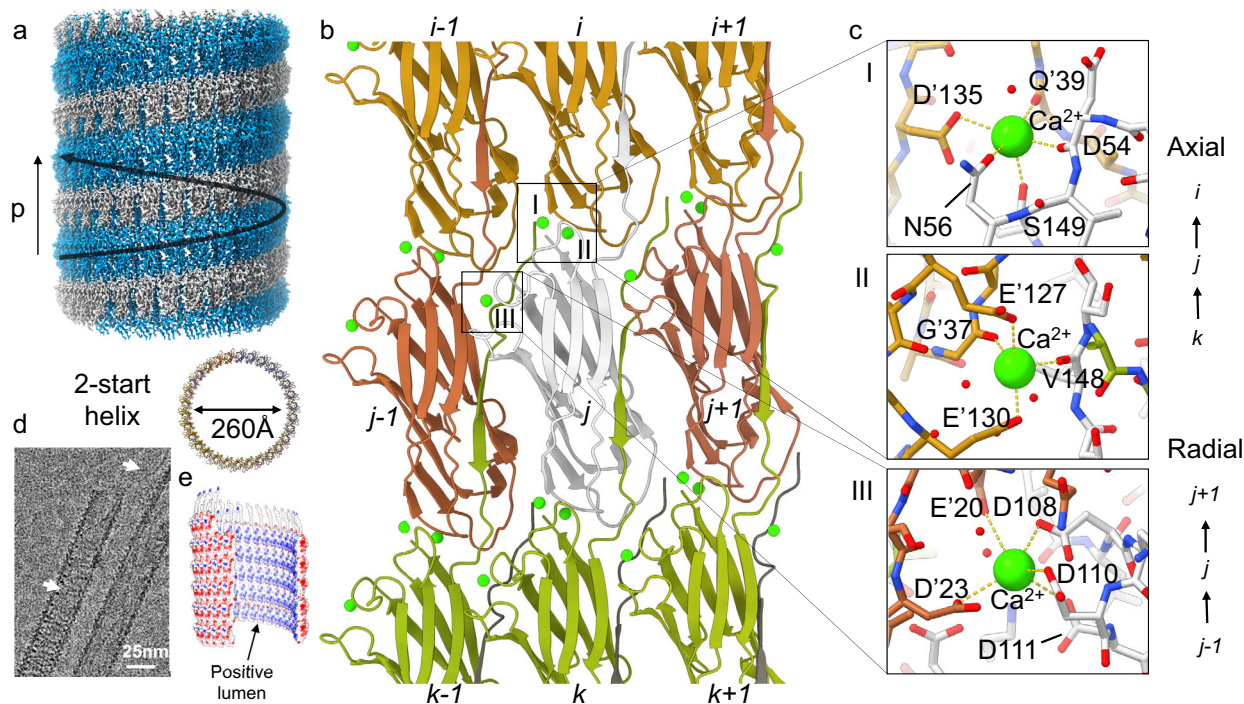
## Results

### CryoEM analysis of ex vivo *P. abyssi* extracellular filaments

We initiated a cryoEM study of ex vivo cannula filaments found on the extracellular surface of *P. abyssi* AV2, the type strain for which the genomic sequence was available (NZ\_AP028907.1)<sup>41</sup>. From the raw micrographs, two main fiber populations could be discerned under the culture conditions (see “Methods”), namely cannulae and an unidentified class of pili (vide infra). Cannulae were present as long (> 5  $\mu$ m) filaments with a diameter of ~26 nm that tended to associate laterally into large rafts (Supplementary Fig. 1a). High magnification cryoEM imaging (Fig. 1d and Supplementary Fig. 1b) revealed the presence of thinner, flexible (~2 nm diameter) fibrils that were either associated with the cannulae or freely floated in suspension. Although these 2 nm fibrils were mostly seen to be randomly bound to cannulae filaments, examples were found in which fibrils were helically wound inside the cannulae scaffold, seemingly following its helical pitch (Fig. 1d), which was suggestive of a specific interaction between both filament types.

After high-resolution data collection, standard helical refinement procedures led to a cryoEM reconstruction of the ex vivo cannulae at 2.3 Å global resolution (FSC 0.143 criterion; Fig. 1a, and Supplementary Fig. 1f, g), which revealed two strands having C2 symmetry, with a helical rise and twist along each strand of 3.18 Å and 12.1°, respectively, and a pitch of 94.6 Å. Given that the molecular identity of the ex vivo cannulae subunits was unknown, Modelangelo<sup>42</sup> was used to build a de novo atomic model without providing an input protein sequence, followed by rounds of alternating manual and automated real-space refinement (see “Methods”). The final model revealed an intricate network of protein subunits that participated in polytypic radial and axial interactions (Fig. 1b and Supplementary Fig. 2). Specifically, the cannula protein subunits interlocked axially via N-terminal DSC (e.g.,  $i \rightarrow j \rightarrow k$ ; Fig. 1b), and axially as well as radially via intermolecular calcium coordination (Fig. 1c and Supplementary Fig. 1h). Three types of calcium binding sites were identified that could be distinguished based on interfacial interactions between subunits. Type I and II binding sites mediated axial interactions between protomers (e.g., protomers  $i$  and  $j$ ), and type III binding sites were located at the interface between two structurally adjacent subunits in the radial plane (e.g., the protomer interface between  $j-1$  and  $j$ ). Each calcium-binding site type was unique and consisted of a combination of Ca<sup>2+</sup> side-chain coordination via Asp or Glu residues, as well as main chain contacts with carbonyl groups and water molecules. We hypothesized that this elaborate network of calcium-binding sites contributed to (i) the remarkable stability of the cannulae (*P. abyssi* was cultured at 98 °C), but (ii) could also represent an environmental trigger for cannula protein assembly (vide infra).

To identify the major subunit of the ex vivo cannulae, we first performed a hidden Markov search using the previously investigated cannula protein CanA (UniParc UPI0024181436)<sup>35,43</sup> from *P. abyssi* TAG11<sup>32,33</sup> as query sequence against the *P. abyssi* AV2 genome (NZ\_AP028907.1)<sup>41</sup>. This search led to 19 putative CanA-like candidate proteins, 15 of which aligned onto a common consensus sequence (Supplementary Fig. 3). Of those 15 sequences, 6 sequences were retained as plausible candidates based on sequence length arguments (i.e., primary mature sequence between 140 aa and 160 aa; Filter 1 in Supplementary Fig. 4). The final assessment of the major subunit sequence was based on sidechain identification through manual inspection of the cryoEM map (Filters 2 and 3 in Supplementary Fig. 4).



**Fig. 1 | CryoEM structure of ex vivo cannulae from the hyperthermophilic archaeon *Pyrodicticum abyssi* AV2.** **a** cryoEM volume revealing a two-start (C2) helical filament with a pitch  $p$  and diameter of 94.6 Å and 260 Å, and a rise and twist of 3.18 Å and 12.1°, respectively; **b** local interaction network between neighboring cannula protofilaments in which subunits interlock through donor strand complementation along the axial (e.g.,  $i \rightarrow j \rightarrow k$ ) direction and partake in further inter-molecular contacts mediated by calcium coordination along the axial and radial (e.g.,  $j-1 \rightarrow j \rightarrow j+1$ ) direction; **c** stick representation of three types of calcium pockets:

(i) axial stacking of subunits via type I and II calcium binding centers, (ii) radial contacts are mediated via type III  $\text{Ca}^{2+}$  binding sites; **d** raw cryoEM image of a cannulae filaments loaded with a helically winding thinner filament (2 nm; cargo). White arrows denote winding cargo and dangling cargo associated with the cannulae filaments; **e** surface electrostatics showing a predominantly negatively charged surface and a positively charged lumen for the cannula model. The representative image in this figure (**d**) is derived from a total of 2365 cryo-EM movies from two technical replicates.

WP\_338251948.1 (annotated as “hypothetical protein”, InterProScan: no identified Pfam or PANTHER domains) was retained as the major subunit. Given the similarity (49.0% sequence identity) to previously identified CanA, and the fact that this sequence was derived from ex vivo analysis, we propose “CanX” as the naming convention. Interestingly, we found clear additional hexose-like density linked to residue Asn51 (Supplementary Fig. 4). N-glycosylation in *Thermoproteota* was found to comprise a eukaryote-like GlcNAc $\beta$ (1-4)GlcNAc core structure<sup>44</sup>. The excess density at Asn51 and the adjacent hydrogen-bond network fit well with a  $\beta$ -linked N-acetylglucosamine (GlcNAc; albeit with some additional density at the 3' position), including additional density for a partially resolved hexose bound to its 4' position (Supplementary Fig. 4). We concluded that Asn51 is an N-linked glycosylation site, likely with GlcNAc core, which resulted in a band of glycans that helically wound around the cannula outer surface (Supplementary Fig. 1i).

Intrigued by the number of detected cannula-like proteins in the *P. abyssi* genome, we inspected the genomic loci and found that 11 out of 15 sequences were grouped into two extensive gene clusters, with the remaining 4 scattered across the genome (i.e., orphan sequences; Supplementary Fig. 5). Interestingly, CanX (WP\_338251948.1) was flanked by WP\_338251946.1 (37.4% sequence identity) and WP\_338251950.1 (29.5% sequence identity), which suggested that these three genes constituted an operon. Although we can only speculate on the function of the CanX homologues present in the genome at this point, we wondered if one or more of these CanX-like proteins could be sporadically integrated into cannulae (i.e., as minor subunits). The cryoEM volume shows no clear signs of subunit heterogeneity, but if one or more CanX-like proteins were randomly integrated into the

cannula lattice at low occupancy, then their impact on and contribution to the final 3D reconstruction would be limited. After inspecting the AlphaFold2<sup>45</sup> structural predictions (Supplementary Fig. 5), numerous candidates were found with substantial domain insertions, and we reasoned that such domains could act as fiducials to guide a particle classification process. To test that hypothesis, we tuned the filament tracer algorithm of cryoSPARC to selectively pick the edges of the cannula filaments and extracted the particles using a smaller box size covering a length equivalent to 2.5 helical pitches to allow for greater sensitivity to local changes in downstream 2D classification jobs. Despite this approach, the resulting 2D class averages showed no signs of heterogeneity, nor could the presence of any additional domains be discerned as outward projections (Supplementary Fig. 1j). We recognize that this approach cannot provide definitive proof that the cannula filaments were solely composed of CanX, but it suggested that if any other CanX homologues were incorporated into the cannulae, it would most likely occur at low sub-stoichiometric levels.

Next, we set out to determine the identity of the 2 nm fibrils that were associated with the cannula filaments (Fig. 1d and Supplementary Fig. 1b). Given the clear interaction between the two types of filaments, we inferred that the 2 nm fibrils might be putative cannula-cargo. The reconstructed cryoEM volume showed no signs of the presence of the cargo, but that would be expected due to the helical reconstruction method that was employed, i.e., the cargo locally breaks helical symmetry, and, as a result, any cargo density would be averaged out in the reconstruction. To account for such a symmetry mismatch, we performed C1 reconstructions followed by 3D classification, but did not identify any 3D classes in which the cargo was resolved, which was most likely due to the absence of a specific cannula-cargo epitope



(as suggested by the raw micrographs). In the absence of a cryoEM resolved cannula-cargo interaction, we hypothesized that the cargo could be nucleic acid and collected a second cryoEM dataset after extended DNase I treatment. The resulting raw micrographs showed no signs of the presence of any cargo after DNase I treatment. To formalize that observation, we adopted the following workflow. Starting from high-resolution helical reconstructions of both DNase I-treated and non-treated particles, we used the respective consensus volumes as input in particle subtraction jobs—the rationale being that particle alignment and classification would be dominated by the cannula signal for the non-subtracted particles. Next, the subtracted particles were 2D classified and inspected for the presence of fibrillar classes that would match the dimensions of the cargo in raw micrographs. For the non-treated cannulae dataset, such cargo 2D class averages were produced as evidenced from the apparent diameter (i.e., 2 nm) and the mapping of the particles-corresponding to the cargo 2D class averages-back onto the raw micrographs, which confirmed that these corresponded to boxed regions with cargo clearly present. In contrast, no cargo 2D class averages could be obtained following DNase I treatment, which allowed us to conclude that the cargo fibrils were DNA. Although the cryoEM images were suggestive of a luminal location, these did not unambiguously exclude the possibility that the cargo could be bound to the cannulae surface. However, indirect evidence further suggested a luminal interaction based on the surface electrostatics of the cannula filaments. Cannulae displayed a highly polar structure that has a predominantly negative outer surface and a highly positive lumen (Fig. 1e). The presence of a positively charged lumen for a negative cargo might seem contradictory from the perspective of DNA transport but could be rationalized through a mechanism of (i) counterion-mediated charge neutralization or (ii) the involvement of effector molecules bound to the DNA to facilitate transport through the cannulae fibers<sup>17–19,46</sup>. Taken together, these data suggested a possible function for cannulae as molecular conduits between cells for the exchange of DNA, although we can't exclude the possibility that other types of cargo exist as well.

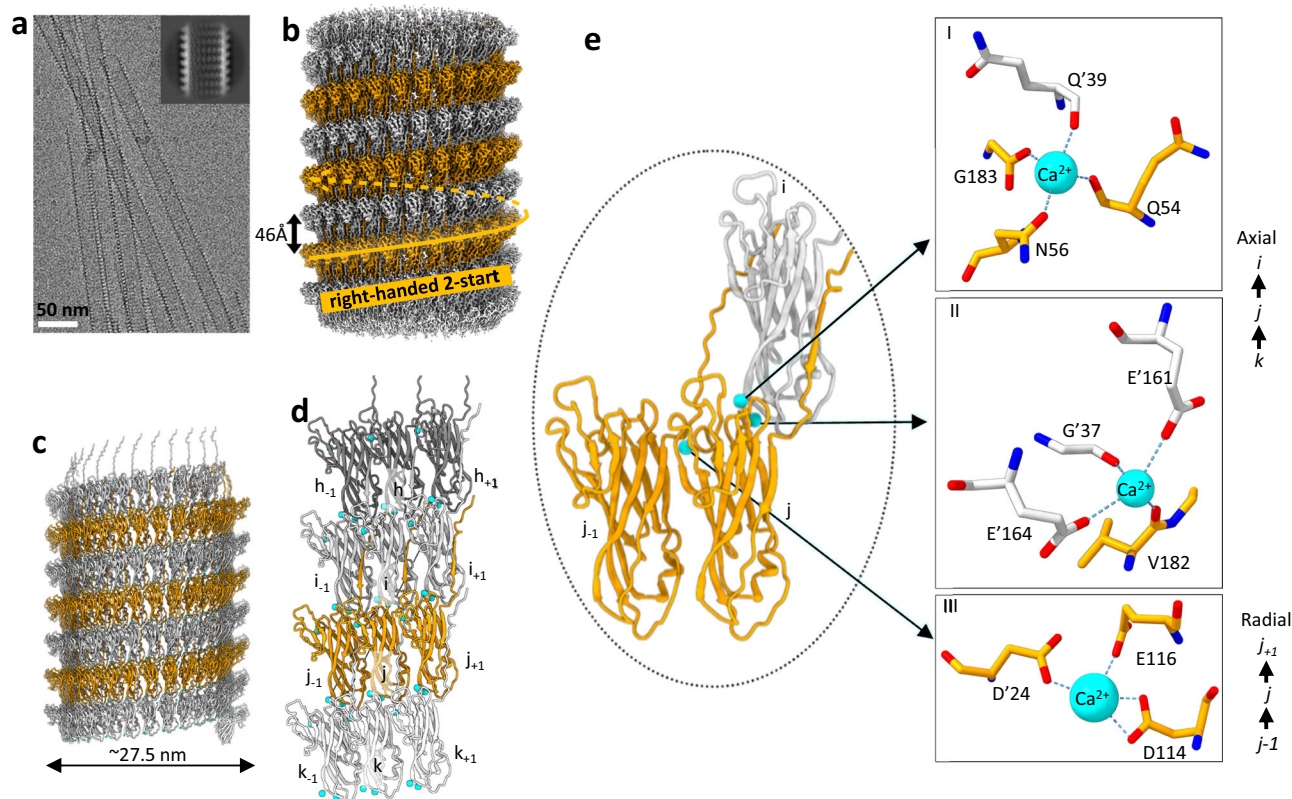
In addition to the cannulae filaments, a second class of pili was also found to be present (Supplementary Fig. 1d). These pili existed as large super bundles with a local undulating character that was reminiscent of the structure of the archaeal bundling pili (ABP) of *Pyrobaculum calidifontis* (PDB: 7UEG)<sup>40</sup>. Although the 2D class averages of these putative *P. abyssi* bundling pili do not show homogenous packing of the filaments, the resulting power spectrum showed a pronounced maximum at 1/47 Å, which we presume corresponded to the axial rise of the subunits in the individual pili. In contrast, an axial rise of 33 Å was observed for individual filaments of *P. calidifontis* ABP, which were derived from DSC of the cellular protein AbpA (WP\_011849593.1). Using HMMER<sup>47</sup> and Foldseek<sup>48</sup>, we found no putative homologues of AbpA encoded in the genome of *P. abyssi* AV2. AbpA was distantly related to TasA (UniProt: P54507), the major biofilm matrix component of Gram-positive *Bacillus subtilis*, which was shown to have an axial rise of 48 Å (PDB: 8AUR)<sup>9</sup>. Using the TasA structure as a Foldseek<sup>48</sup> query, we identified the hypothetical protein WP\_338249486.1 (AAA988\_RS09245) in a structural similarity search (Supplementary Fig. 7). Although the sequence identity to TasA is low (16%), the AlphaFold3<sup>49</sup> prediction of a WP\_338249486.1 dimer revealed notable structural similarities, i.e., WP\_338249486.1 is predicted to form a donor strand complemented dimer with a monomer-to-monomer distance of 47 Å, which agreed with the experimentally determined axial rise of the putative bundling pili of *P. abyssi* AV2 (Supplementary Fig. 1d, e). Looking at the local genomic context of WP\_338249486.1, we identified the protein WP\_338249487.1 (AAA988\_RS09250), presumed to form an operon with AAA988\_RS09245 and predicted to encode a dedicated class I signal peptidase. Based on these data, we hypothesized that WP\_338249486.1 formed the structural basis of the bundling pili found in the extracellular milieu of *P. abyssi* AV2. In analogy to CanX, we

propose AbpX as the naming convention for WP\_338249486.1. While never previously identified in EM analyses of *Pyrodicticum* extracellular filaments, the AbpX filaments were more similar in structure to the taxonomically more distant TasA filaments than to previously characterized AbpA pili from *P. calidifontis*. Remarkably, two different precursor proteins, CanX and AbpX, undergo polymerization into structurally distinct extracellular filaments through a common process of DSC on the surface of *P. abyssi*.

### Structural analysis of recombinant *P. abyssi* cannulae

While the cryoEM structural analysis of the CanX filament provided insight into the origin of the thermodynamic stability and mechanical rigidity of *P. abyssi* cannulae, the challenging culture conditions hinder preparation of cannulae in sufficient quantities to evaluate their potential in biomaterials applications. Previous experimental investigations demonstrated that recombinant expression of *P. abyssi* cannula-like proteins in heterologous bacterial hosts represented a viable preparative route to soluble precursors of structurally similar filaments<sup>35,43</sup>. CanA, a cannula-like protein from *P. abyssi* strain TAG11, has been the most thoroughly studied of these recombinant precursors<sup>32</sup>. Previous investigators reported that recombinant CanA could assemble in vitro into tubular filaments in the presence of calcium ions. The resultant filaments displayed a similar diameter and morphology to ex vivo cannulae from *P. abyssi* TAG11 under low-resolution electron microscopy imaging<sup>35</sup>. In addition, a recombinantly expressed truncation of CanA, K<sub>1</sub>-CanA, corresponding to a deletion of the first ten amino acid residues of the putative donor strand of the mature protein, displayed no propensity for polymerization under similar conditions<sup>43</sup>. The induction of self-assembly of recombinant CanA in the presence of calcium ions and its inhibition upon partial deletion of the N-terminal donor strand suggested that recombinant cannula-like proteins could polymerize in vitro through a mechanism that mimicked the assembly of the ex vivo filaments.

A structural determination of in vitro assemblies of recombinant CanA was undertaken to compare the corresponding structure to that of the ex vivo filament. A codon-optimized version of the coding sequence of the mature CanA protein, i.e., after deletion of N-terminal twenty-five amino acids of the putative signal peptide sequence<sup>50</sup>, was designed for bacterial cell expression (Supplementary Figs. 8a and 9a). Previous research demonstrated that this CanA protein variant could be expressed and purified as a soluble protein in *Escherichia coli* host systems in good yield<sup>35</sup>. Following the published procedure, recombinant CanA was obtained in an unoptimized yield of 31 mg of purified protein per L of culture from the cell lysate after induction of expression in *E. coli* strain BL21Gold(DE3) (see “Methods”). The purity and identity of the recombinant CanA protein were confirmed using SDS-PAGE analysis and electro-spray ionization (ESI) mass spectrometry, respectively (Supplementary Figs. 10a and 11a). Sedimentation velocity analytical ultracentrifugation indicated that CanA migrated as a monomer in low-salt buffer (50 mM Tris-HCl, pH 7.5, 80 mM NaCl, 0.1 mM EDTA, 9% glycerol) at an apparent molecular mass that agreed with mass spectrometric analysis and the monoisotopic molecular mass calculated from the amino acid sequence of mature CanA (Supplementary Fig. 12a). Thermal treatment (50°–80 °C) of solutions of purified CanA in a low-salt buffer led to rapid polymerization in the presence of CaCl<sub>2</sub> (20 mM)<sup>35</sup>. The experimental conditions for CanA assembly differed significantly from the native growth conditions of *P. abyssi* (seawater, 80–110 °C, 20–35 MPa hydrostatic pressure). The facile in vitro assembly of CanA under non-physiological conditions for *P. abyssi* growth suggests that cannula biosynthesis under native conditions may be regulated to prevent premature assembly. Negative-stain TEM images of the in vitro assemblies revealed the presence of a structurally uniform population of high aspect-ratio tubular assemblies of ~25 nm in diameter (Supplementary Fig. 13a), which agreed with previous reports from EM imaging of ex vivo



**Fig. 2 | CryoEM structure of in vitro cannula-like tubules formed by the self-assembly of recombinant CanA.** **a** representative raw cryoEM micrograph and 2D-class average (inset) of the recombinant CanA filament; **b** cryoEM volume depicting right-handed 2-start helical packing of the helical reconstruction (rise and twist of 1.59 Å and  $-173.75^\circ$ , respectively); **c** helical model of CanA filament; **d** local

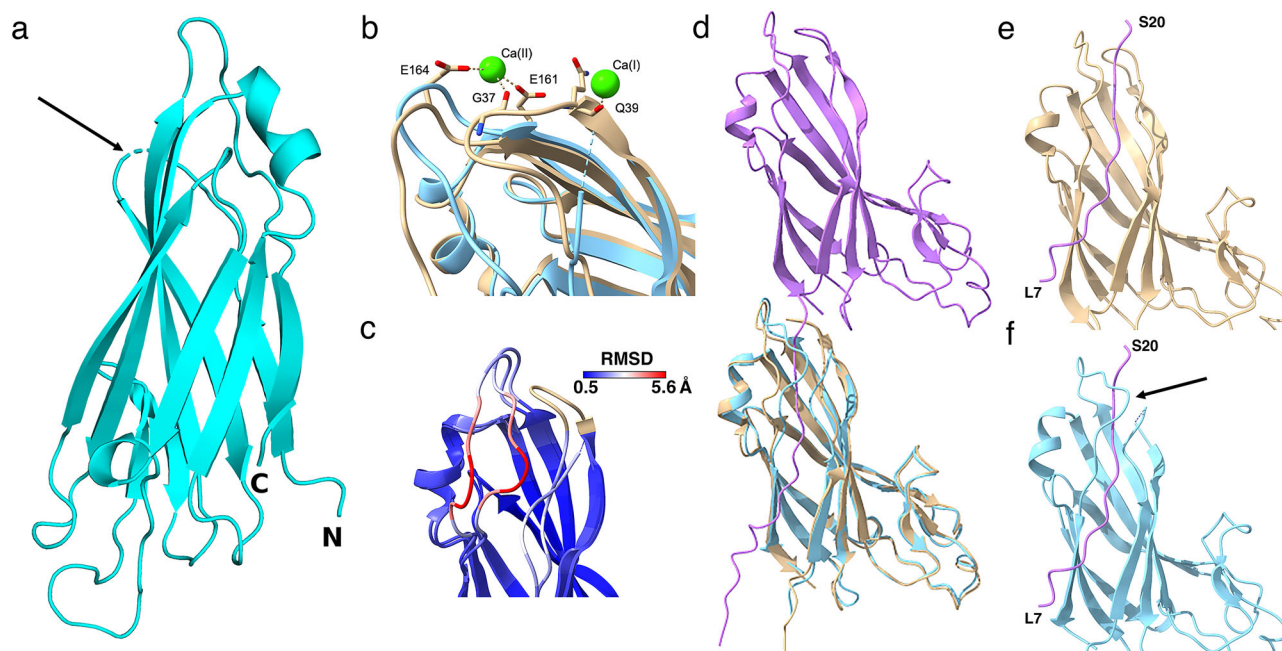
interaction network between neighboring CanA subunits with three  $\text{Ca}^{2+}$  ions bridging protomers in a 12-mer segment of the cannulae; **e** close-up view of three  $\text{Ca}^{2+}$  ions binding sites in stick representation. The representative image in this figure (a) is derived from a total of 11143 cryoEM micrographs from a single technical replicate.

*P. abyssi* TAG11 cannulae and recombinant polymers of CanA<sup>32,35,38</sup>. The cannula-like filaments of CanA laterally associated into large rafts like those observed in ex vivo preparations of *P. abyssi* AV2 cannulae (vide supra). Similar behavior was observed for cannulae under native growth conditions in cultures of *P. abyssi* TAG11, in which fast freeze/deep etch electron microscopy imaging revealed the presence of bundles containing up to a hundred cannulae at the cell surface<sup>32</sup>.

CryoEM analysis was employed to determine the structure of the recombinant CanA assemblies (Fig. 2). Despite extensive lateral association between filaments observed in the raw micrographs, a sufficient number of isolated particles could be classified as the basis for helical reconstruction (Fig. 2a). The 2D class averages confirmed that the tubular filaments were uniform in inner and outer diameter,  $\sim 18$  nm and  $\sim 27$  nm, respectively, despite in vitro assembly conditions that differed significantly from the native growth environment of *P. abyssi*. A 3D volume was reconstructed at 2.6 Å global resolution (FSC 0.143 criterion) from the 2D projection images using iterative helical real-space reconstruction<sup>51,52</sup> after determination of helical symmetry (Supplementary Fig. 14a)<sup>25</sup>. An unambiguous atomic model could be built into the reconstructed density (Fig. 2b and Supplementary Fig. 14b, c). In contrast to the C2 symmetry observed for the ex vivo cannulae, the CanA filament was a C1 assembly of protomers along a left-handed 1-start helix with a rise of 1.59 Å and a twist of  $-173.75^\circ$ . The most prominent structural feature of the CanA filament, as for the CanX filament, was the presence of a pair of right-handed protofilaments with a  $a \sim 4.6$ – $4.7$  nm periodicity (Figs. 1a and 2b, c), i.e., half the pitch of the 2-start helices, which was consistent with previously reported TEM images of native cannulae and recombinant cannula-like tubules<sup>32,35,38</sup>.

As observed for the ex vivo filament, the atomic model of the CanA tubules suggested that DSC was the primary mechanism that drove polymerization of CanA into cannula-like filaments and stabilized the resultant assemblies. The relative orientation of protomers in the polymer, i.e., along the  $h \rightarrow i \rightarrow j \rightarrow k$  direction (Fig. 2d) was defined by a rotation of  $1.25^\circ$  and an axial rise of 46 Å ( $1.5^\circ$  and 48 Å, respectively, for the corresponding interaction in CanX). Successive protomers were aligned nearly parallel to the central axis of the helical assembly (Figs. 1b and 2d). Donor-strand complementation linked protomers located on structurally adjacent 2-start helices (Fig. 2b, c) and reinforced the structural interfaces within the CanA filament. The structure of the CanA cannulae displayed an extensive network of inter-protomer contacts that involved subunits located as far away as  $j \pm 60$  on either side of a central protomer (Supplementary Fig. 14d). Protein, Interfaces, Surfaces, and Assemblies analysis<sup>53</sup> indicated that, while DSC between protomers buried the greatest amount of surface area, significant interactions also occurred between subunits located within the same protofilament (e.g., subunits  $j-1$  and  $j$ ) along a 2-start helix (Fig. 2d and Supplementary Fig. 14). In addition, as observed for the ex vivo filament, three calcium ions could be fit into the cryoEM density map of the in vitro filament (Fig. 2d). The calcium ion binding sites were structurally homologous to those of the ex vivo filament, mediating similar polytypic radial and axial interactions between structurally adjacent protomers (Fig. 2e). Other than the observed helical symmetry, the most significant difference between the ex vivo and in vitro cannula filaments was the structure of the protomers. While the core structures of CanX and CanA were based on homologous jelly-roll folds (49.0% sequence identity), extended loops were observed in the CanA protomer and the AlphaFold3<sup>49</sup> predicted





**Fig. 3 | Crystal structure analysis of the CanA monomer.** **a** crystal structure of the CanA monomer (cyan) depicted as a ribbon diagram. The arrow indicates a rendering of the disordered loop region 34–40; **b** close-up of the superposed CanA monomer from the crystal structure (blue) with the cryoEM CanA protomer (tan) depicting the conformational change within the loop region 161–173; **c** close-up of the RMSD rendering from a Matchmaker backbone alignment between the CanA protomer and monomer that highlights the difference in loop region 161–173; **d** alignment of the CanA monomer from the crystal structure (blue) to an

interacting axial pair of CanA protomers from the cryoEM structure. The N-terminal donor strand from protomer *i* (violet) inserts into the acceptor groove of protomer *j* (tan); **e**, **f** zoomed-in views of loop region 161–173 in the CanA cryoEM structure (**e**) and the superposed CanA monomer (**f**). The latter view revealed a potential steric clash (arrow) between this region of the CanA crystal structure with donor strand residues Tyr11 to Ser20 that potentially precludes CanA oligomerization in the absence of calcium ion binding.

monomer structure, which were not present in CanX (Supplementary Fig. 15a). In the recombinant CanA structure, these loops projected from the outer surface of the filament, which resulted in a pair of ridges that coincided with the 2-start helices (Fig. 2b, c). The CanX protomer lacked these loop insertions, although a homologous protein (WP\_338251966.1, 51.1% sequence identity to CanA) was identified in the *P. abyssi* AV2 proteome that presented loop protrusions from the core fold (Supplementary Fig. 5). However, outer surface projections that could account for the presence of this CanA-like protein were not detected in the ex vivo filament (vide supra).

#### Potential role of calcium ion binding in assembly of cannulae

The conservation of the calcium-binding sites in the structures of the ex vivo and in vitro cannula filaments, as well as the calcium ion induction of CanA assembly, raised a question regarding the relationship between calcium ion binding and DSC. The polymerization of the cannula-like proteins resulted from progressive DSC in combination with calcium ion binding. However, for CanA, calcium ion binding was a necessary but not sufficient criterion for polymerization under the conditions that we investigated. While the monomer was competent for polymerization in the presence of calcium ion<sup>54</sup>, the polymerization process was kinetically slow at ambient temperature and required heating to promote more rapid assembly. Under native growth conditions for *P. abyssi* (~100 °C), the temperature would be sufficient to initiate polymerization of cannula-like proteins at the extracellular surface since the calcium concentration in seawater (400 ppm, ~10 mM) would be within the permissible concentration range. Hypothetically, other divalent metals ions present in seawater could promote self-assembly of cannula-like proteins into filaments at the cellular surface of *P. abyssi*. The Mg<sup>2+</sup> ion is the most abundant divalent cation in seawater (1300 ppm, ~54 mM). However, under the assembly conditions that were employed (50–100 μM protein, 20 mM

divalent ion, 50–80 °C, low-salt buffer, pH 7.5), filamentation of recombinant CanA was not observed in vitro, although a recent report suggested that polymerization occurred at higher protein concentration (1–2 mM)<sup>54</sup>. The difference in ionic radius between Mg<sup>2+</sup> and Ca<sup>2+</sup>, 72 pm versus 100 pm, can lead to different binding affinities between the two divalent metal ions within coordination sites in metalloproteins. Notably, the addition of terbium ion (10 mM), which is often employed as a calcium ion surrogate in biophysical studies of calcium-binding proteins (Tb<sup>3+</sup> ionic radius ~106 pm)<sup>55</sup>, rapidly initiated polymerization of recombinant CanA into cannula-like filaments at ambient temperature. EDTA (10 mM), a strong chelator of calcium ion (K<sub>b</sub> ~10<sup>8</sup> M<sup>-1</sup> in 10 mM Tris-HCl, pH 7.5)<sup>56</sup>, inhibited the Ca<sup>2+</sup> induced polymerization of CanA into cannula-like filaments. However, treatment of pre-formed cannulae with EDTA did not induce dis-assembly unless the solution was heated to ≥ 60 °C. SV-AUC measurements indicated that CanA remained a monomer in solution over the course of the experiments in the presence of sub-critical concentrations of Ca<sup>2+</sup> ion below the polymerization temperature (Supplementary Fig. 12a).

To gain insight into the potential role of calcium ions in driving filament formation, single-crystal X-ray diffraction was employed to determine the structure of the monomeric form of CanA in the absence of calcium ions (Fig. 3a). A comparison between the crystal structure of the CanA monomer and the corresponding CanA protomer in the assembly revealed significant conformational differences between the two structures that localized to protein segments associated with subunit-subunit interfaces in the filament. The most obvious difference between the two structures lies in the conformation of the N-terminal nineteen amino acids in CanA. These residues were not observed in the crystal structure of the CanA monomer, which suggested that the N-terminal segment was disordered and did not contribute to the diffraction. In contrast, these N-terminal amino acids adopted a β-strand conformation (strand A) in the cryoEM

structure of the CanA protomer, which mediated the donor strand-acceptor groove interaction between axially adjacent protomers (e.g.,  $i,j$ ) in the filament (Fig. 2b, c). Deletion of the first 10 amino acids of mature CanA resulted in a construct, K<sub>1</sub>-CanA<sup>43</sup>, that was incapable of polymerization, which further argued for the importance of an intact donor strand in cannulae formation.

Significant structural differences were also observed in the conformation of two loops, corresponding to residues 34–40 and 161–173, which contributed to the axial Ca<sup>2+</sup> binding sites (Fig. 3b) within the cannula filament. In the crystal structure, the electron density corresponding to residues 35–39 was absent, which implied that the loop conformation was dynamic in the absence of calcium ions. In addition, a backbone alignment of the CanA monomer crystal structure and the CanA cryoEM protomer structure indicated that the loop region corresponding to residues 161–173 underwent a significant displacement upon formation of the filament (Fig. 3b, c). We attributed the observed conformational changes in these two loops to the influence of calcium ion binding. Alignment of the monomeric CanA structure to an axially interacting pair of protomers ( $i,j$ ) in the filament suggested a mechanism whereby calcium ion binding could promote DSC (Fig. 3d). In the cryoEM structure of the CanA filament (Fig. 3e), calcium ion binding stabilized the conformation of the two loop regions and promoted a local interaction that exposed the acceptor groove of the monomer. In the crystal structure of the monomer, loop region 161–173 adopted a conformation that occluded the acceptor groove, which prevented it from accepting the donor strand of another monomer and precluded DSC potentially preventing filament formation (Fig. 3f). We propose that calcium ion binding to the monomer triggers a structural rearrangement that renders the monomer competent for subsequent donor-strand complementation and polymerization into a cannula-like filament. However, the rate at which polymerization occurred was observed to depend on temperature, perhaps mediated through a thermally activated process that was not apparent from the available biophysical data. The crystal structure of CanA showed that N-terminal strand A was not self-complemented in the monomeric form and, consequently, available to interact with another subunit upon calcium coordination through donor-strand complementation, thereby initiating polymerization into a filament. Structural comparison between the CanA protomer in the cryoEM structure and a recently reported NMR structure of N-terminal deletion K<sub>1</sub>-CanA (PDB: 9HDI) similarly indicated that structural differences between monomer and protomer were localized to the N-terminal donor strand and loop regions associated with divalent ion binding<sup>54</sup>. The most pronounced structural difference between the two structures, outside of the donor strand conformation, occurred within the loop region 34–40, which undergoes a significant conformational rearrangement to accommodate calcium ion coordination at the axial binding sites (Supplementary Fig. 16).

### CryoEM of recombinant cannula-like protein homologues

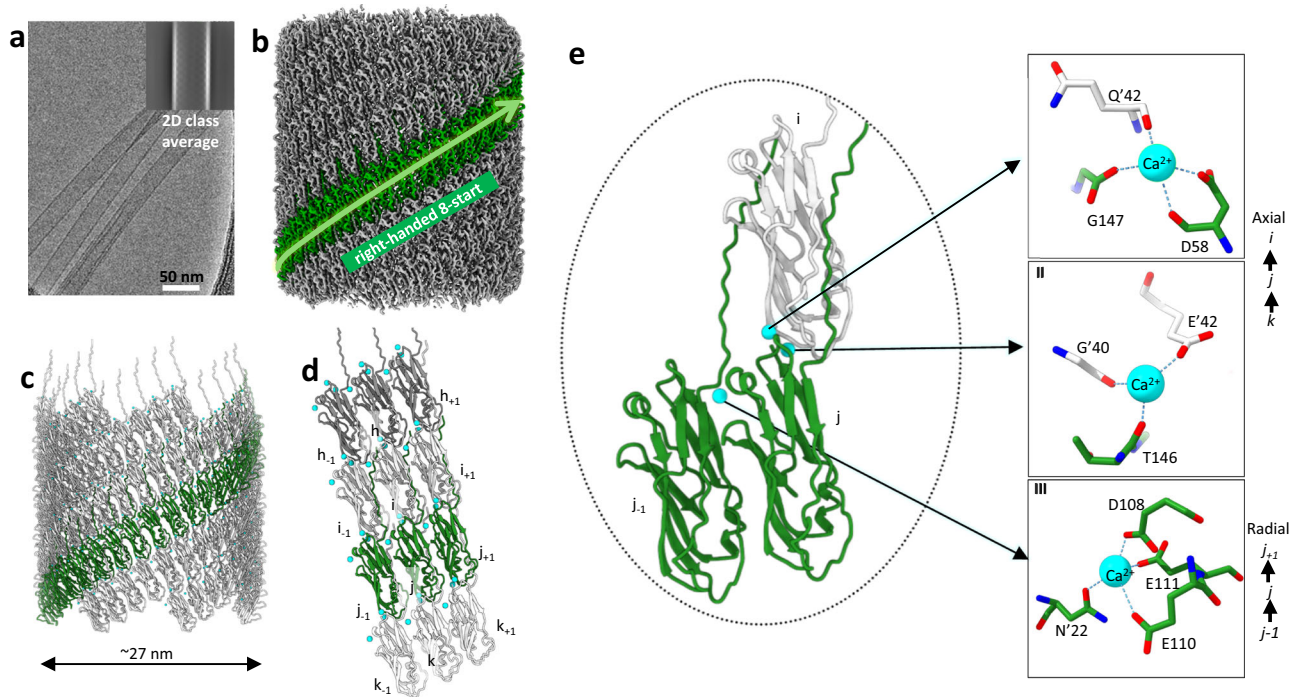
To identify other potential cannula-mimetic proteins in the genomes of related archaeal species, Foldseek was employed using the protomer structure of CanA (PDB: 7UII) as a query<sup>48</sup>. Several putative cannula-like sequences were identified within the proteomes of species within the family *Pyrodictiaceae*, including five proteins from the related species *Pyrodictium occultum* and four proteins derived from metagenome-assembled genomic (MAG) data reported for an uncultured archaeon in phylum *Thermoproteota*, tentatively assigned as a species within the genus *Hyperthermus* (Supplementary Table 1). AlphaFold2 structural predictions were available for the respective proteins through the UniProt webserver ([www.uniprot.org](http://www.uniprot.org))<sup>57</sup>. In each case, the presence of a jelly-roll fold was predicted for the respective monomeric structures with high pLDDT and pTM, although poorly predicted insertions and N-terminal extensions were also observed for several of the larger proteins as previously observed for putative

cannula proteins encoded within *P. abyssi* AV2 genome (Supplementary Fig. 5). In most cases, AlphaFold2 predicted an unstructured N-terminal domain for the mature protein sequences. As determined for CanA and CanX, most of these cannula-mimetic protein sequences contained potential *N*-glycosylation sequons (NXS/T) that localized to surface loops in the AlphaFold2 structural predictions.

Two of these cannula-like proteins, Hyper1 (hypothetical protein DSY37\_00495, UniProt: A0A432RA20) and Hyper2 (hypothetical protein DSY37\_03180, UniProt: A0A432R7L7), were chosen as substrates for a more detailed analysis. The mature sequences<sup>50</sup> of these two proteins were similar in length to CanX and shared a pairwise sequence identity of 38.9% to each other and of 36.4% and 26.2%, respectively, to CanX. Codon-optimized coding sequences (Supplementary Fig. 9) of the putative mature proteins were cloned into the plasmid pD451-SR and transformed into *E. coli* strain BL21Gold(DE3). Expression cultures were screened for the presence of the respective proteins under IPTG induction using the experimental procedure optimized for recombinant CanA (see “Methods”). While SDS-PAGE analysis indicated that both proteins were successfully expressed, the Hyper1 protein accumulated in inclusion bodies that could not be successfully refolded after denaturation and was not further studied. The Hyper2 protein accumulated in the soluble fraction of the cell lysate and could be purified using the procedure developed for recombinant CanA. Recombinant Hyper2 was obtained in an unoptimized yield of 80 mg of purified protein per L of culture from the cell lysate of *E. coli* expression cultures. The purity and identity of the Hyper2 protein were confirmed using SDS-PAGE analysis and ESI mass spectrometry (Supplementary Figs. 10b and 11b). As observed for CanA, sedimentation velocity AUC analysis in low-salt buffer indicated that Hyper2 behaved as a monomer in the absence of calcium ion (Supplementary Fig. 12b). The apparent molecular mass of the Hyper2 protein from SV-AUC analysis agreed with that observed from mass spectrometry and the monoisotopic molecular mass calculated from the amino acid sequence of the recombinant Hyper2 protein.

In contrast to CanA, Hyper2 assembled rapidly at ambient temperature in the presence of calcium ion ( $\geq 10$  mM) without heat treatment. However, negative-stain TEM analysis (Supplementary Fig. 13b) indicated a range of filamentous assemblies, including single-walled tubular filaments of different diameters and multi-walled tubes. The structural polymorphism of the Hyper2 assemblies was confirmed in the raw cryoEM micrographs and 2D class averages (Fig. 4a and Supplementary Fig. 17a, b). Helical reconstruction was employed to structurally analyze two of the single-walled tubes corresponding to diameters of  $\sim 240$  Å and  $\sim 270$  Å. After independent 2D classification (Supplementary Fig. 17b) and manual assignment of helical symmetry (Supplementary Fig. 18a), standard helical refinement procedures led to cryoEM reconstructions (Fig. 4b and Supplementary Fig. 17c) at 3.5 Å global resolution for both classes of filaments (FSC 0.143 criterion; Supplementary Fig. 18b, c and Supplementary Table 2). Atomic models could be reliably built into the respective EM density maps for the two classes of Hyper2 filaments, which revealed that they were structurally related but based on different helical symmetries. The structure of the thicker tube displayed C8 symmetry with a helical rise of 11.16 Å and twist of 9.59°, while the structure of the thinner tube revealed C1 symmetry with a helical rise of 1.42 Å and a twist of 46.23° (Fig. 4b, c and Supplementary Fig. 17c). As observed for the *Pyrodictium* cannulae reconstructions, the protomers in the Hyper2 filaments were integrated through a combination of DSC and calcium ion binding interactions (Fig. 4d, e and Supplementary Fig. 17d).

The most significant structural difference observed between the *Pyrodictium* and *Hyperthermus* cannulae lay in the number and geometrical arrangement of the protofilaments. For the putative *Hyperthermus* cannulae, eight protofilaments propagated along right-handed 8-start helices having a periodicity of  $\sim 5.2$  nm (one-eighth of the 8-start helical pitch) and twists of 9.59° (C8) and 9.84° (C1),



**Fig. 4 | CryoEM structure of cannula-like tubules (Hyper2 thick tube) formed by the in vitro self-assembly of recombinant Hyper2.** **a** representative cryoEM micrograph and 2D-class average of recombinant Hyper2 tubes; **b** cryoEM map showing right-handed 8-start helical packing with C8 rotational symmetry (rise and twist of 11.16 Å and 9.59°, respectively); **c** helical model of Hyper2 thick tube; **d** local

interaction network between neighboring Hyper2 subunits with three  $\text{Ca}^{2+}$  ions bridging protomers in a 12-mer segment of the cannulae; **e** close-up view of the three  $\text{Ca}^{2+}$  ions binding sites in stick representation. The representative image in this figure (**a**) is derived from a total of 5502 cryoEM micrographs from a single technical replicate.

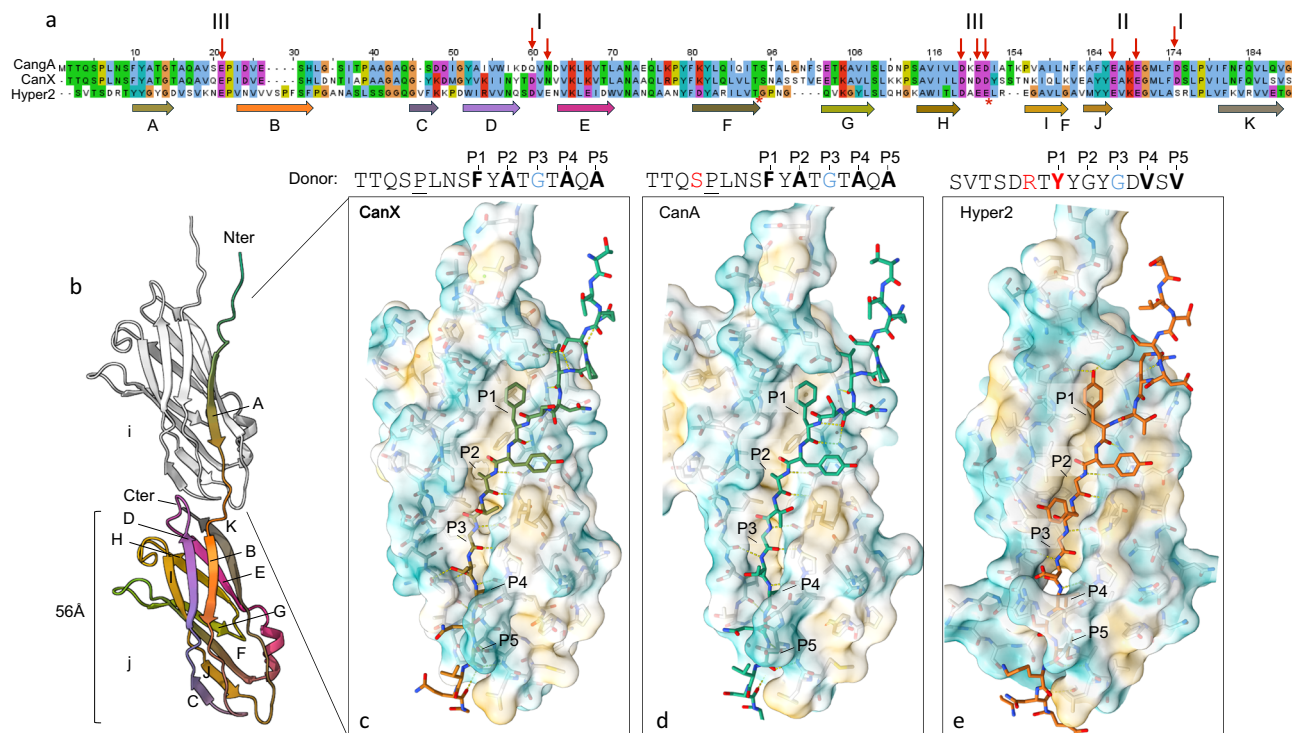
respectively (Fig. 4b and Supplementary Fig. 17c). For the *Pyrodicticum* cannulae, two protofilaments propagated coincidentally with a pair of right-handed 2-start helices having a periodicity of 4.6–4.7 nm (Figs. 1a and 2b, c). However, the distance between the protofilaments along the direction of donor strand polymerization was comparable in the structures of the *Hyperthermus* cannulae (~4.6–4.7 nm) and the *Pyrodicticum* cannulae (~4.6–4.8 nm). In all cases, DSC occurred between structurally adjacent protofilaments and reinforced the supramolecular structure of the respective cannulae. The donor strand-acceptor groove interaction occurred along a series of left-handed 31-start (C1) or 32-start (C8) helices for *Hyperthermus* cannula-like filaments (Fig. 4d and Supplementary Fig. 17c). In contrast, the donor strand polymerization in the *Pyrodicticum* cannulae occurred along a set of right-handed 30-start helices for CanX (Fig. 1c) or 29-start helices for CanA (Fig. 2d). The propagation axis of the DSC interaction was nearly parallel to the central helical axis for the *Pyrodicticum* cannulae (~4° tilt), whereas donor strand propagation was tilted at an angle (~18°) from the central axis in the Hyper2 assemblies.

Despite these differences in helical symmetry, the protomer structure and local structural interfaces were remarkably conserved between the putative *Hyperthermus* and *Pyrodicticum* cannulae. The core structure of the protomer was based on a donor-strand complemented jelly-roll fold in which calcium ions were observed to bind at the radial and axial interfaces between protomers. Structural alignment of the C $\alpha$  backbone atoms of the CanX, CanA, and Hyper2 protomers suggested a close correspondence between the respective jelly-roll folds despite the presence of two large loops in the CanA structure (Supplementary Fig. 15a). Two distinct dimeric structural interfaces were identified in the local interaction network of the respective filaments (Supplementary Fig. 2), which were associated with radial interactions between laterally adjacent protomers within the same protofilament (e.g., subunits  $j-1$ ,  $j$ ,  $j+1$ ) and axial interactions

between protomers located in different protofilaments (e.g., subunits  $i$ ,  $j$ ,  $k$ ). Pairwise backbone alignments of the dimeric radial and axial interfaces between protomers in the different cannulae revealed that the structures were strongly conserved (RMSD  $\approx$  0.8–1.2 Å, Supplementary Fig. 15b). This strong conservation of interfacial interactions among the cannulae can be contrasted to the weaker sequence identity between the *Pyrodicticum* and *Hyperthermus* cannula-like proteins (Supplementary Fig. 15c).

As in the ex vivo and in vitro *P. abyssi* cannulae, three distinct calcium ion binding sites were observed in the Hyper2 filaments at structurally homologous positions. A multiple sequence alignment (MSA) of the mature CanX, CanA, and Hyper2 proteins indicated that the residues in the calcium ion binding sites were conserved (Fig. 5a), which suggested a similar role for calcium ions in inducing assembly and maintaining the structural integrity of the two classes of cannulae (Figs. 1b, c, 2d, e, and 4d, e). Likewise, the structural conservation of the donor strand-acceptor groove interaction was observed between the *Pyrodicticum* and *Hyperthermus* cannulae. In each case, donor strand A of subunit  $j$  forms a series of hydrogen-bonded interactions with strands B and K of the acceptor groove on an axially adjacent subunit  $i$  (shown in Fig. 5b for the CanX filament). Strands B and K define a hydrophobic groove consisting of five pockets, P1–P5, which accommodate residues on one face of the donor strand (Fig. 5c–e). The donor strands of the cannula-like proteins displayed an amphipathic sequence pattern in which hydrophobic residues alternated with hydrophilic residues. The sidechains of the former residues were buried in the acceptor groove, while the polar sidechains of the latter residues were exposed on the filament surface. The donor strands of the cannula-like proteins displayed significant sequence homology, in which an aromatic residue (F/Y) occurred at the site that occupied the P1 pocket, while smaller sidechain hydrophobic residues (G/A/V) occupied the sites on strand A that interacted with P2–P5 pockets





**Fig. 5 | Conserved donor strand complementation across cannulae forming subunits.** **a** multiple sequence alignment (MSA) of mature CanA, CanX, and Hyper2 sequences (Jalview, Clustal coloring). Red stars denote regions where CanA loop/domain insertions were removed for visual clarity. Red arrows highlight conserved residues that interact with calcium in type I, II, and III coordination centers. Arrows below the MSA denote secondary structure elements color coded according to the monomer cartoon representation in **(b)**; **b** Illustration of the cannula donor strand complementation mechanism shown here for CanX: the

N-terminal strand A of subunit *j* docks into the surface exposed hydrophobic groove defined by the region between strands B and K of subunit *i*; **c–e** zoomed-in renditions of the donor strand (stick representation) docking into the respective B/K grooves for CanX, CanA and Hyper2, respectively. Donor sequences above **(c–e)**: hydrophobic residues docked into complementary pockets highlighted in bold; residues forming sidechain-mediated intermolecular hydrogen bonds shown in red; proline directing the N-terminus towards the “lattice contact” underlined.

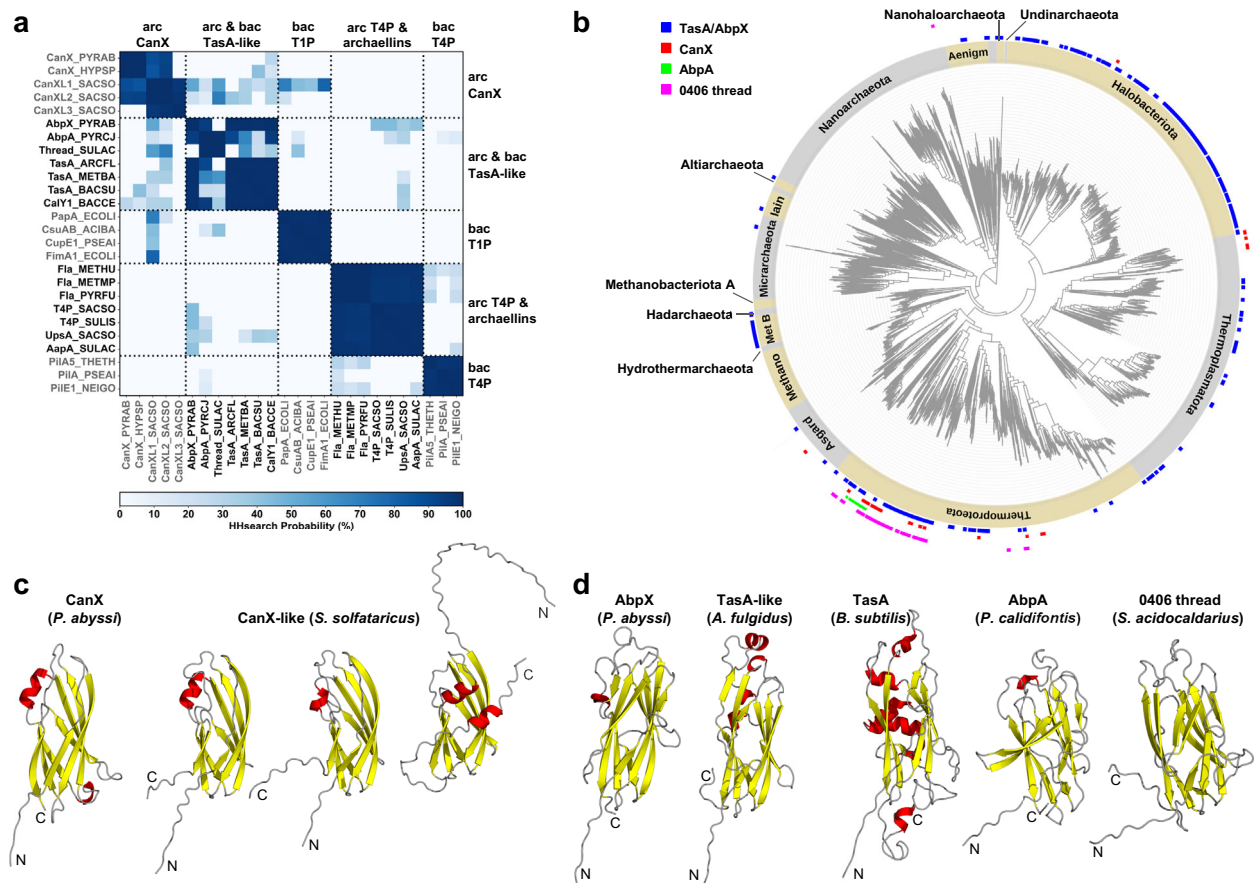
(Fig. 5c–e). Notably, a central Gly residue (cyan, Fig. 5c–e) was conserved in all three proteins, as well as for all putative cannula-like proteins tentatively identified in *Pyrodicticum* and *Hyperthermus* species. We hypothesize that the conserved glycine residue may guide the alignment between the donor strand and acceptor groove, ensuring the donor strand adopts the registry that promotes structural complementation. To test the importance of the glycine residue, a recombinant protein corresponding to the G13A mutant of Hyper2 was prepared and tested for its ability to undergo polymerization in the presence of calcium ions. Under identical conditions to those successfully employed for the *wild-type* Hyper2 protein, the Hyper2G13A mutant was severely compromised in its ability to form filaments, which implied that donor-strand complementation was partially to completely inhibited through this conservative substitution. In bacterial CU pili, a similar process of molecular recognition between the side chains of residues on the donor strand and pockets in the acceptor groove promotes correct registry during donor-strand complementation and governs the order of incorporation of different pilins into the filament during biosynthesis<sup>3,4,8,58,59</sup>.

### Distribution of cannula-like proteins in archaea

This similarity in donor-strand/acceptor groove recognition prompted us to systematically evaluate the evolutionary relationships of CanX and AbpX with other bacterial and archaeal cell surface filaments formed through donor-strand complementation<sup>3–7,9–11,39</sup>. Based on HMMER searches and signal peptide prediction, all CanX homologs are exclusively found in the *Pyrodictiaceae* family and possess a Sec/SPI signal sequence, which suggested that they undergo transport through

the Sec translocon and are subsequently cleaved by Signal Peptidase I, similar to other characterized donor strand complemented filaments. However, the genomic neighborhood of these putative cannula-like proteins lacked genes encoding putative chaperones or dedicated signal peptidases, supporting the hypothesis that CanX could assemble in the presence of calcium without external assistance. Unlike CanX, HMMER searches revealed that AbpX homologs were more widely distributed across archaea<sup>10</sup>, including in *Pyrodictiaceae* that lacked CanX homologs, such as *P. delaneyi* Su06<sup>34</sup> (WP\_055409108.1). The sequence homology searches with AbpX and its archaeal homologs also identified sequence matches to bacterial TasA family proteins, with pairwise sequence identity levels of 30%–35%, which supported the classification of AbpX within the TasA superfamily (Fig. 6a, b). Like bacterial TasA proteins<sup>9</sup>, AbpX homologs in archaea were frequently found adjacent to genes encoding a signal peptidase I or other AbpX/TasA-like proteins, underscoring a potential evolutionary and functional connection between these systems.

To uncover potentially distant CanX homologs that HMMER searches might have missed, sensitive profile HMM-based sequence searches were employed using HHsearch. We searched the profile HMM databases of proteomes from several representative archaea for homologs of CanX, which identified three distant homologs in *Saccharolobus solfataricus* (WP\_009990924.1, WP\_009991716.1, WP\_029552507.1) with HHsearch probability values of greater than 90% (Supplementary Fig. 19a). Despite sharing only ~18% pairwise sequence identity with *P. abyssi* CanX, the structures of these putative proteins, predicted by AlphaFold2, retained the jelly-roll fold, showing a striking structural resemblance to CanX (Fig. 6c).



**Fig. 6 | CanX and AbpX belong to the TasA superfamily.** **a** pairwise profile HMM comparison of representative archaeal and bacterial cell-surface filament subunits. The heatmap shows HHsearch probabilities (%) for each pairwise comparison. The proteins are grouped into five categories: archaeal CanX proteins (arc CanX), archaeal and bacterial TasA-like proteins (arc & bac TasA-like), bacterial type I pilins (bac T1P), archaeal type IV pilins and archaeellins (arc T4P & archaeellins), and bacterial type IV pilins (bac T4P), all of which are  $\beta$ -sandwich/jelly-roll fold proteins that assemble into filaments. The NCBI/UniProt accession numbers of the proteins are provided in the “Methods” section; **b** distribution of the TasA superfamily across archaea, visualized on a species-level tree from the Genome Taxonomy

Database. Colored dots indicate the presence of specific protein families: TasA/AbpX (blue), CanX (red), AbpA (green), and *S. acidocaldarius* 0406 thread (magenta). Phyla are indicated on the outer ring of the tree. **c**, **d** structural representations of representative TasA superfamily proteins. AlphaFold3 models are shown for selected proteins, while cryoEM structures are displayed for *P. calidifontis* AbpA (PDB: 7UEG) and the *S. acidocaldarius* 0406 thread subunit (PDB: 7PNB).  $\beta$ -helices are colored red, and  $\beta$ -strands are colored yellow.

Homo-oligomers of these proteins, as predicted by AlphaFold2-multimer<sup>60</sup>, suggested their potential to polymerize through DSC (Supplementary Fig. 19b). These proteins were employed in subsequent HMMER searches that identified over 100 CanX-like proteins in archaea, mostly within the phylum *Thermoproteota*, specifically in the *Sulfolobaceae* and *Thermocladaceae* families (Fig. 6b). For instance, we identified a homolog each in *Vulcanisaeta distributa* DSM 14429 (WP\_013336327.1) and *Caldivirga maquilensis* IC 167 (WP\_156769888.1), two in *Sulfolobus acidocaldarius* DG1 (WP\_011279025.1, WP\_011279026.1), and four in *Metallosphaera yellowstonensis* MK1 (WP\_009071828.1, WP\_009072785.1, WP\_009072754.1, WP\_009071324.1). AlphaFold3 models for these homologs suggested that they could also adopt the jelly-roll fold (Supplementary Fig. 19b) and could potentially undergo DSC. However, experimental validation will be required to confirm whether these predicted homologs indeed assemble into extracellular filaments that have physiological function in the corresponding organisms in their native environment.

The jelly-roll fold of cannula-like proteins was a structural feature that was shared with the bacterial TasA protein from *B. subtilis* and its recently identified archaeal homologs, AbpA, from *P. calidifontis* and 0406 thread subunit from *S. acidocaldarius*. To explore the

evolutionary relationships among these proteins, HHsearch was employed for pairwise sequence comparisons and searches against the PDB and Pfam profile HMM databases. These searches identified statistically significant matches (HHsearch probability values between 80% and 95%) between CanX-like proteins and AbpA (PDB: 7UEG), the *S. acidocaldarius* 0406 thread subunit (PDB: 7PNB), as well as the Pfam Peptidase\_M73 family (PF12389), which encompassed bacterial TasA homologs. The pairwise sequence identity across these matches was low (below 20%), indicating that CanX and CanX-like proteins form a highly divergent family within the TasA superfamily. The *S. acidocaldarius* thread subunit, which was recently described as representative of a newly discovered class of archaeal surface filaments, also showed statistically significant matches to CanX-like proteins, AbpA, and TasA family proteins in our HHsearch searches. However, the low sequence identity among these proteins (below 20%) suggested that both the thread subunit and AbpA constituted distinct families within the TasA superfamily (Fig. 6b). Additionally, our searches uncovered subdomain-level matches to archaeal flagellin C-terminal globular subunits and bacterial type I filament-forming proteins, such as CupE1, FimA, and PapA (Fig. 6a), all of which adopted the immunoglobulin-like (Ig) fold. Despite the low sequence similarity, the core structures of these proteins displayed a remarkably conserved topology.

While type IV filament (T4F) superfamily assemblies, including the archaeal flagellins, are widespread across archaea<sup>61</sup>, the distribution of TasA-like proteins in archaea remains less well understood<sup>10</sup>. To address this, we used HMMER to search for homologs of bacterial TasA, AbpX, AbpA, CanX, and the O406 thread subunit in representative archaeal genomes from the Genome Taxonomy Database (GTDB). Our searches revealed that homologs of AbpA, CanX, and the O406 filament are primarily restricted to genomes in the *Thermoproteota* phylum, whereas putative homologs of AbpX and bacterial TasA are found more extensively across the genomes of several archaeal phyla (Fig. 6b). Interestingly, the genomes of many archaeal species encode two or more of these putative TasA-like proteins. For example, the *V. distributa* DSM 14429 genome encodes potential homologs of CanX (WP\_013336327.1), AbpA (WP\_148678262.1), and the thread subunit (WP\_013336193.1, WP\_013336196.1), highlighting the potential diversity and functional versatility of these putative filament-forming proteins. These findings demonstrate that genes encoding TasA-like proteins may be far more widespread in the genomes of archaea than previously recognized. The observation of bundling pili in *P. calidifontis* and *S. acidocaldarius*, as well as the presence of the AbpX filaments in the ex vivo *P. abyssi* culture, suggest that many archaeal species may have the potential capacity to form biofilms under conditions in which these TasA-like proteins are expressed as filaments at the extracellular surface.

## Discussion

DSC underlies the polymerization of several classes of bacterial<sup>3–7,9,39</sup> and archaeal<sup>10,11</sup> pili that are involved in cellular adhesion and biofilm formation. Proper assembly of bacterial pili often requires the involvement of specialized cellular machinery such as the CU secretion system<sup>62,63</sup>. In absence of the usher catalyst, in vitro polymerization of recombinant CU-type pilin monomers is kinetically slow and results in polymers that have shorter lengths and lower stability than the native pili, which has been attributed to the introduction of structural defects in the assemblies<sup>63,64</sup>. In contrast, in vitro polymerization of recombinant cannula-like protein monomers afforded filaments that mimicked the quaternary structure of ex vivo cannulae and preserved the critical interfacial interactions between protomers. Accordingly, unlike most bacterial protein filaments derived from DSC, the polymerization of the cannulae does not require the presence of a dedicated membrane-bound translocation/assembly platform<sup>65</sup> or accessory proteins such as chaperones<sup>62</sup> or proteases<sup>39,66</sup> to mediate polymerization. In agreement with this hypothesis, neither dedicated signal proteases nor putative chaperone proteins were identified within the genetic loci encoding cannula-like proteins. The presence of calcium ion binding sites distinguished cannulae from other extracellular filaments that arose from DSC. Our results suggested that calcium ion coordination primed DSC and promoted the in vitro polymerization of recombinant cannula proteins into cannula-like filaments (Fig. 3). In addition, calcium ion coordination stabilized the quaternary structure of the filaments through formation of a network of axial and radial interactions between structural subunits in both the ex vivo and in vitro cannulae.

Coordination of calcium has been previously reported to induce structural transitions in proteins secreted at prokaryotic cell surfaces. These transitions have been attributed to the structural influence of calcium ion coordination, which arises from the higher  $\text{Ca}^{2+}$  ion concentration between the extracellular and intracellular environment ( $[\text{Ca}]_{\text{extra}}/[\text{Ca}]_{\text{intra}} \sim 10^4$ )<sup>67</sup>. Calcium-dependent structural transitions have been studied extensively in extracellular filaments such as the repeats-in-toxin adhesion proteins (RTX adhesins)<sup>68</sup>, which are megadalton proteins that are translocated through the outer membrane (OM) of Gram-negative bacteria via the Type I secretion system (TISS)<sup>69</sup>. RTX adhesins can consist of >100 copies of tandemly repeated bacterial Ig-like (Blg) protein domains capped with a C-terminal RTX  $\beta$ -roll domain. Since the secretion signal is located at the

C-terminus of the protein, translocation cannot proceed until completion of protein synthesis, which requires that the nascent polypeptide chains are maintained in an unfolded state that is permissive for translocation through the TISS translocon. Biophysical analyses of pathogenic<sup>70</sup> and non-pathogenic<sup>71,72</sup> RTX adhesins have demonstrated that calcium ion binding plays several critical roles for this class of proteins, including facilitation of transport across the OM and induction of folding in the extracellular environment<sup>73,74</sup>. RTX adhesin proteins are proposed to be maintained in a fully or partially unfolded state at the low intracellular calcium ion concentration prior to secretion through the translocon<sup>74–76</sup>. In the extracellular environment, the higher calcium ion concentration initiates folding of the Blg protein domains, essentially serving as a “chemical foldase”<sup>69</sup>. Crystallographic analyses of single or tandemly repeated RTX Blg domains provided evidence that calcium binding not only stabilized the internal structure of individual proteins but also structured flexible loops between domains, causing a rigidification of the structure that enabled the filament to extend outward from the cellular surface<sup>70,72,74,75,77</sup>. We envision a similar dual functional role for calcium ion coordination in the assembly of cannulae. At the low intracellular calcium ion concentration, cannula precursor proteins remain in a monomeric or low oligomeric state. Upon secretion into the extracellular environment, the higher calcium ion concentration triggers self-assembly of the monomers into cannulae by facilitating DSC. Calcium ions coordinate at binding sites at the axial and radial interfaces between protomers, further rigidifying the structure of the cannulae. Likewise, a calcium ion gradient has been demonstrated to induce self-assembly of bacterial S-layers<sup>78–80</sup>. Binding of calcium ion to interfacial loops of these cell surface structures induced a structural transition into their assembly-competent conformation and promoted adoption of the correct quaternary structure. Like the RTX adhesins and cannulae, calcium-induced structural transitions were proposed to provide a control mechanism to prevent premature assembly of nascent subunits in the cytoplasmic environment, where calcium is held at low nanomolar concentrations<sup>78,79</sup>.

The large lumen (~18 nm) of the cannula filaments provides an attractive target for the development of functional nanomaterials that can encapsulate and release large substrates, but also begs the question of the native function of cannulae. CryoET analysis of *P. abyssi* TAGII cannulae provided evidence for the presence of density within the lumen, which led to a proposed role in substrate transport<sup>38</sup>. Our structural analysis of the ex vivo cannulae indicated that they could accommodate filamentous cargo within the lumen, which in several instances was observed to be helically wound at the interior surface (Fig. 1d). The thin diameter of the filamentous cargo (~2 nm) and its susceptibility to DNase I treatment suggested the presence of DNA in the lumen, although the structural analysis was inconclusive regarding the cargo identity. Extracellular protein filaments have been previously demonstrated to mediate horizontal gene transfer (HGT) in prokaryotes through transport of DNA between cells. The proposed transport mechanism in archaea involves the synthesis of a thin conjugative pilus (~7–9 nm outer diameter) through which DNA uptake can occur from a donor cell. We wondered if cannulae could serve as an alternative conduit that could mediate HGT through cross-cellular DNA transport. Conjugative DNA uptake in archaea involves the Ced/Ted DNA import systems<sup>19,81</sup>, which are encoded on the archaeal chromosome and are homologous to the prokaryotic type IV secretion system (T4SS)<sup>82</sup>. Could cannulae replace the more widespread conjugative DNA uptake process in *Pyrodictiaceae*? To answer this question, we searched for the presence of genes encoding Ced/Ted-related proteins within the genomes of organisms in *Pyrodictiaceae* family. Pilin protein sequences *Aeropyrum pernix* CedA1 (WP\_010865579) and *P. calidifontis* TedF (WP\_011849449) were employed to query the genomes of the family *Pyrodictiaceae* in a PSI-BLAST search. The TedF query resulted in no positive hits, which was not surprising given that the two families



reside in different orders. In contrast, the *A. pernix* CedA1 sequence, belonging to the same order *Desulfurococcales*, returned multiple hits at E-values ( $10^{-24}$ – $10^{-27}$ ) that were suggestive of high sequence identity. The recovered CedA1 homologs encompassed all members of the family *Pyrodictiaceae* for which genomic information was available as well as MAG data from uncultured *Pyrodictium* and *Hyperthermus* species. While cannulae may play a role in HGT through DNA transport, the conservation of an endogenous Ced DNA import machinery suggested that a putative cannula-mediated mechanism must be complementary to and independent from HGT through the Ced conjugative uptake pathway.

The narrow distribution of putative cannula-like proteins in archaeal genomes raises a question regarding the evolutionary and structural relationship between cannulae and the more widely distributed ABP of the TasA superfamily. The co-expression of structurally distinct AbpX and CanX extracellular filaments in *P. abyssi* cultures implies different functional roles despite similar jelly-roll folds and a common process of polymerization through donor-strand complementation. A structural homology search of the proteome of *Pyrodictium occultum* PL-19<sup>83</sup>, a shallow-sea hyperthermophilic archaeon in which cannulae have been physically observed and for which a genome sequence (NZ\_LNTB01000001.1) is available<sup>31</sup>, identified a gene encoding a putative ABP (WP\_058370724.1). Previous EM analyses of ex vivo samples in *Pyrodictium occultum* cultures did not detect Abp-like structures<sup>29,31</sup>, although these thin filaments would have been difficult to detect under the experimental imaging conditions. For comparison, a gene encoding a putative AbpX-like protein (WP\_055409108.1, Uniprot: A0A0N7JD56) was identified within the genome of *Pyrodictium delaneyi* Su06 (NZ\_CP013011.1), a motile deep-sea archaeon phylogenetically most closely related to *P. abyssi* AV2<sup>34,41</sup>. Neither sequence nor structural homology searches identified genes encoding putative cannula-like proteins in the genome of *P. delaneyi* or in the complete genomes available for other archaea within the *Pyrodictiaceae* family.

We speculate that cannulae may have evolved from the structurally similar and more widely distributed ABP as a unique evolutionary adaptation within this sub-class of archaea to accommodate a sessile life cycle under hyperthermophilic environmental conditions. These *Pyrodictium* species, thriving at elevated temperature (~100 °C) in deep-sea hydrothermal vents and submarine solfataric fields, form intricate colonies with individual non-motile, disk-shaped cells linked through cannulae. Biofilm integrity may be further supported by the presence of filamentous networks derived from assembly of other matrix proteins, e.g., AbpX, that independently undergo polymerization through donor-strand complementation. These cellular colonies may display a clonal form of multicellular organization that is distinct from that observed in most microbial communities. Cannulae can potentially enable physical and chemical communication (including possible DNA exchange) between cells, creating a structure reminiscent of biological tissues. Interconnected cells form a mesh-like network that, while not truly multicellular in the sense of differentiated cells in animals or plants, represents a cooperative and coordinated community. This network potentially enhances resilience under extreme conditions, allowing for resource sharing (e.g., DNA repair) and robustness to environmental stress—characteristics often associated with the early stages of multicellularity.

## Methods

### Materials

Luria-Bertani (LB) medium was purchased from IBI Scientific (Dubuque, IA). Isopropyl β-D-1-thiogalactopyranoside (IPTG) was purchased from GoldBio (St. Louis, MO). Chicken egg white lysozyme was purchased from Research Products International (Prospect, IL). Benzoylase nuclease was purchased from Merck KGaA (Darmstadt, Germany). Protease Inhibitor Cocktail Set V was purchased from

Calbiochem (San Diego, CA). Competent cells of *E. coli* strain BL21Gold(DE3) were purchased from Agilent Technologies (Cedar Creek, TX). Precision Plus Protein™ Kaleidoscope™ pre-stained protein standards were purchased from Bio-Rad (Hercules, CA). Custom gene synthesis was performed by ATUM (Newark, CA). Carbon-coated copper grids (200 mesh) and C-flat cryoEM grids were purchased from Electron Microscopy Sciences (Hatfield, PA). Lacey carbon cryo grids and 2% (w/w) solutions of methylamine vanadate (Nano-Van) and methylamine tungstate (Nano-W) negative stains were purchased from Ted Pella (Redding, CA). All other chemical reagents were purchased from either VWR (Radnor, PA) or Sigma-Aldrich Chemical Co. (St. Louis, MO) and used without further purification.

### Protein expression and purification

Codon-optimized genes corresponding to the expression cassettes of CanA, Hyper1, and Hyper2 (Fig. S9) were purchased from ATUM (Newark, CA) as sequence-verified clones in plasmid pD451-SR (Kan<sup>R</sup>). The synthetic genes were cloned as transcriptional fusions under the control of a T7 promoter with a strong ribosome binding site in the high copy number plasmid pD451-SR (pUC *ori*). Purified plasmids encoding the respective proteins were transformed into *E. coli* strain BL21Gold(DE3). Single colonies of the respective transformants were inoculated into sterile LB medium (4 mL) supplemented with kanamycin (100 g/mL) and incubated at 37 °C for 12 h. The pre-incubated cultures were used to inoculate sterile LB media (200 mL) in a shake flask culture supplemented with kanamycin to a final concentration of 100 µg/mL. Bacterial cultures were incubated on a rotary shaker at 37 °C until the OD<sub>600</sub> reached a value of ca. 0.4 absorption units (AU). Protein expression was induced through addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM and the bacterial cultures were incubated with shaking for an additional 4 h at 30 °C. Cells were isolated from the expression culture through centrifugation at 9100 × *g* for 3 min and resuspended in 2 mL of low-salt buffer (50 mM Tris-HCl pH 7.5, 80 mM NaCl, 9% glycerol). The following reagents were added to the cell lysate: hen egg-white lysozyme (final concentration of 1.25 mg/mL), Benzonase nuclease (250 units), MgCl<sub>2</sub> (final concentration of 1 mM), and 100× protease inhibitor cocktail (to 1×). The cell lysate was incubated with shaking at 30 °C for 1 h. The clarified cell lysate was dialyzed against low-salt buffer containing EDTA (final concentration of 5 mM). After several changes in this high EDTA buffer, the protein was dialyzed in a low-EDTA version of the same buffer (0.01 mM EDTA) with several changes. The dialysate was heated to 80 °C for 20 min to denature the heat-labile proteins of the host bacterium and was subsequently cooled to ambient temperature. The resultant solution was centrifuged at 14,000 × *g* for 10 min at ambient temperature. The clarified supernatant fraction was dialyzed overnight against low-salt buffer in a membrane having MWCO of 10,000 Da for CanA and 7000 Da for Hyper 2. The protein concentration of the dialysate was determined spectrophotometrically using a Nanodrop spectrophotometer. Purified protein samples (4 mL) in low-salt buffer achieved a concentration of ca. 400 µM after purification. Protein purity was assessed using SDS-PAGE gel electrophoresis and electrospray-ionization (ESI) mass spectrometry.

### Mass spectrometry

Recombinant protein samples were dialyzed extensively against deionized water to remove salt prior to mass spectrometric analysis. Mass spectra were acquired on a Thermo Exactive Plus using a nano-flex source with off-line adapter. Solutions of peptide (10 µL) were deposited into a New Objective econo picotip and placed in the nanosource using the off-line adapter. A backing pressure was applied using an off-line syringe to assure flow to the tip. A voltage of 1.0–2.5 kV was applied to the picotip. Typical settings employed for the analysis were a capillary temperature of 320 °C and an S-lens RF level between 30 and 80 with an AGC setting of 1 × 10<sup>6</sup>. The maximum

injection time was set to 50 ms. Spectra were taken at 140,000-resolution at a mass/charge ratio ( $m/z$ ) of 200 using Tune software and analyzed with Thermo Freestyle software. Representative mass spectra for the protein samples were acquired from analysis involving two technical replicates.

### Analytical ultracentrifugation

Sedimentation velocity experiments were performed on a Beckman Coulter Optima AUC at the Canadian Center for Hydrodynamics at the University of Lethbridge. Recombinant CanA (8.3 and 35  $\mu\text{M}$ ) and Hyper 2 (10.0 and 34  $\mu\text{M}$ ) samples were measured in low-salt buffer (50 mM Tris-HCl pH 7.5, 80 mM NaCl, 0.01 mM EDTA, 9% glycerol); either alone or supplemented with aqueous  $\text{CaCl}_2$  to a final concentration of 1–20 mM. All measurements were conducted in epon-charcoal centerpieces fitted with quartz windows and measured in intensity mode at 45,000 RPM and 20 °C. Data were collected at 220 nm for the low concentration protein samples and 295 nm for the high concentration samples. All data were analyzed with UltraScan III version 4.0<sup>84</sup>. The data were fitted with an iterative two-dimensional spectrum analysis<sup>85</sup> to fit time- and radially invariant noise and meniscus positions. Sample analysis was further refined using Monte Carlo analysis<sup>86</sup>, and diffusion-corrected integral sedimentation coefficient distributions were generated using the enhanced van Holde–Weischet analysis methods<sup>87</sup>. UltraScan calculated the buffer density and viscosity to be 1.023810 g/cm<sup>3</sup> and 1.27097 cP, respectively. Representative sedimentation coefficient distribution curves were acquired from datasets derived from two independent technical replicates.

### In vitro cannulae assembly

Aliquots (100  $\mu\text{L}$ ) of the recombinant CanA and Hyper2 proteins (50–100  $\mu\text{M}$ ) in low-salt buffer were incubated in the presence of different concentrations of  $\text{CaCl}_2$  and  $\text{MgCl}_2$  (final concentration from 5 to 20 mM in the respective divalent ions). The temperature of the polymerization mixture for CanA was heated in the temperature range from 50 to 80 °C for 20 min and cooled to ambient temperature. The polymerization mixture of Hyper2 was incubated for 24 h at ambient temperature after addition of  $\text{Ca}^{2+}$  ion to the desired concentration. Polymerizations samples were screened using negative-stain transmission electron microscopy (nsTEM) for the presence of fibrillar assemblies.

### Negative-stain transmission electron microscopy

Aliquots (5  $\mu\text{L}$ ) of the polymerization mixtures under different assembly conditions were deposited onto 200-mesh carbon-coated copper grids. After 1 min of incubation, excess solution was removed by wicking the sample. An aliquot (5  $\mu\text{L}$ ) of negative stain (1% mixture of 50% NanoVan/50% NanoW) was placed on the grid and incubated for 1 min. Excess stain was wicked away and the grids were air-dried. TEM images were collected with a JEOL JEM-1400 transmission electron microscope at an accelerating voltage of 80 kV.

### CryoEM structural analysis of ex vivo cannulae

An active culture of *P. abyssi* strain AV2 (DSMZ 6158) was purchased from DSMZ-German Collection of Microorganisms and Cell Cultures GmbH. *P. abyssi* was cultured in medium 283 (*Pyrodictum* medium) under anaerobic conditions at 98 °C in a pressurized vessel with 2 bar over-pressure of sterile 80%  $\text{H}_2$  and 20%  $\text{CO}_2$  gas mixture. Aliquots (100  $\mu\text{L}$ ) of pressurized culture were taken using a 1 mL Injekt-F syringe (Braun) fitted with a 23-gauge sterile needle (BD Microlance) by puncturing the rubber fitting. Harvested aliquots were transferred to 1.5 mL Eppendorf tubes, effectively breaking the anaerobe storage conditions. To minimize the effects of exposure to oxygen, care was taken to prepare cryoEM grids shortly after sample harvesting. For that, aliquots (3  $\mu\text{L}$ ) of the culture were used as is during grid

preparation with no further processing to avoid sample processing artefacts.

A high-resolution cryoEM dataset was collected using Quantifoil™ R2/1 300 copper mesh holey carbon grids. Grids were glow-discharged at 5 mA plasma current for 30 s in an ELMO (Agar Scientific) glow-discharger. A Gatan CP3 cryo-plunger set at −176 °C and relative humidity of 90% was used to prepare the cryo-samples. A sample (3  $\mu\text{L}$ ) of the culture was applied on the holey grid and incubated for 60 s. The sample was back-blotted using Whatman type 2 paper for 3 s and plunge-frozen into precooled liquid ethane at −176 °C. High-resolution movies were recorded at the VIB-VUB Facility for Bio Electron Cryogenic Microscopy (BECM) at 300 kV on a JEOL Cryoarm300 microscope equipped with an in-column  $\Omega$  energy filter (operated at slit width of 14 eV), automated with SerialEM v4.1.8. The movies were captured with a K3 direct electron detector run in counting mode at a magnification of 60 K with a calibrated pixel size of 0.71 Å/pixel, and a total exposure of 60 e/Å<sup>2</sup> over 60 frames. A total of 2365 movies were collected and imported into cryoSPARC v4.5.3<sup>88</sup> for further processing. Movies were motion-corrected using Patch Motion Correction and defocus values were determined using Patch CTF. Exposures were curated and segments were picked using the filament tracer and extracted (750k particles) with a box size of 300 × 300 pixels at a pixel resolution of 1.42 Å/pixel. After several rounds of 2D classification (310 k particles retained), Fourier-Bessel indexing was used to calculate initial estimates of the helical rise and twist values, which were used in a subsequent helical refinement job. Next, particles were re-extracted at 0.71 Å/pixel (600 × 600 pixel) and used as input for a second round of helical refinement, followed by local and global CTF refinement, and helical refinement. The resulting high-resolution volume and particle stack were used for reference-based motion correction after particle duplicate removal (310k particles), followed by a final helical refinement (2.30 Å global resolution, FSC 0.143 criterion) after mask application in cryoSPARC. An initial atomic model was built using ModelAngelo<sup>42</sup> v1.0.12 without providing an input sequence (build\_no\_seq), after which the model was manually adjusted to the CanX sequence (WP\_338251948.1) and real-space refined in Coot<sup>89</sup>. A final round of refinement was performed using Phenix<sup>90,91</sup> and figures were created using ChimeraX<sup>92</sup>. Map and model statistics are found in Supplementary Table 2.

### CryoEM imaging of in vitro assembled cannulae

The CanA and Hyper2 samples (ca. 3–3.5  $\mu\text{L}$ ) were applied to glow-discharged lacey carbon grids and plunge frozen using an EM GP Plunge Freezer (Leica). Grids were positively glow-discharged using amylamine for the Hyper2 sample. The cryoEM movies were collected on a 300 keV Titan Krios with a K3 camera (University of Virginia) at 1.08 Å/pixel and a total dose of 50 e/Å<sup>2</sup> using software EPU v2.5.0.4799REL. First, motion corrections and CTF estimations were done in cryoSPARC<sup>88,93,94</sup> v3.2 (CanA) or v3.3.2 (Hyper2). Next, particles were automatically picked by “Filament Tracer” with a shift of 13 and 24 pixels for CanA and Hyper2, respectively. All auto-picked particles were subsequently 2D classified through multiple rounds of selection, and all particles in bad 2D averages were discarded. After initial processing, 659,059 particles remained for CanA. In contrast to the CanA tubes, the Hyper2 assemblies displayed significant polymorphism, including the presence of single-walled and multi-walled tubes. The single-walled tubes constituted populations of tubes that displayed different diameters. 2D classification of the single-walled tubes gave enough particles for 3D reconstruction for two distinguishable classes (95,418 and 140,207 particles) representing tubes having diameters of 240 Å (Hyper2 thin tube) and 270 Å (Hyper2 thick tube), respectively. A range of possible helical symmetries was calculated from the respective averaged power spectra derived from aligned raw particles. To speed up the reconstruction, a smaller subset of particles was first used to test all possible helical symmetries by trial and error until amino acid

side chains, and the hand of a short  $\alpha$ -helix was seen, which was followed by final reconstructions with all good particles<sup>51,95</sup>. The final volumes were then sharpened with a negative B-factor,  $-85$ ,  $-88$ , and  $-89$  Å<sup>2</sup> for CanA, Hyper2 thick tube, and Hyper2 thin tube, respectively, which was automatically estimated in cryoSPARC. After final helical refinement, resolution was estimated from GS-FSC calculations (0.143 criterion)<sup>96</sup> after mask autogeneration in cryoSPARC. The resolutions of the reconstructions are reported in Supplementary Figs. 14 and 17 and Supplementary Table 2. EMReady was employed to perform post-processing on raw non-sharpened maps<sup>97</sup>.

### Model building of recombinant cannulae

The absolute helical hand of the maps of CanA and Hyper2 tubes was determined from the hand of an  $\alpha$ -helix. The hand assignment also agreed with the AlphaFold2<sup>45</sup> prediction of CanA and Hyper2. The cryoEM map corresponding to a single subunit of CanA and Hyper2 was segmented in Chimera<sup>98</sup>. The AlphaFold2 predictions of CanA and Hyper2 were used as the starting model and docked into the cryoEM maps. The docked models were refined by iterative cycles of manual refinement using Coot<sup>89</sup> and automatic refinement using PHENIX<sup>99</sup> and ISOLDE<sup>100</sup> as needed to reach the reasonable map-to-model agreement and model geometry. Densities for eight (1–8) N-terminal residues and four (149–152) C-terminal residues of Hyper2 were not resolved in the map and thus these residues were left unmodeled. MolProbity<sup>101</sup> was used to evaluate the quality of the filament model. The refinement statistics are shown in Supplementary Table 2.

### CanA crystallization and data collection

The recombinant CanA protein was concentrated to 10 mg/ml and put into crystallization trials. All crystals were grown by sitting drop vapor diffusion at 20 °C using a protein to reservoir volume ratio of 1:1 with total drop volumes of 0.2  $\mu$ L. Crystals of CanA were grown using a crystallization solution containing 100 mM MgCl<sub>2</sub>, 30% PEG 4 K, 3% Xylitol, and 100 mM Tris pH 8.5. All crystals were flash frozen in liquid nitrogen after a short soak in the appropriate crystallization buffers supplemented with 20–25% ethylene glycol. Data for CanA were collected at the beamline 19-ID NYX at the National Synchrotron Light source II (NSLS II), Brookhaven National Labs. All data was indexed, merged, and scaled using HKL2000, then converted to structure factors using CCP4.

### Crystal structure determination and refinement

The CanA crystal structure was solved by molecular replacement using the program Phaser<sup>102</sup>. Molecular replacement calculations were performed using the coordinates of the CanA monomeric subunit from the CanA cryoEM structure (PDB: 7UII as the search model. The resulting phase information from molecular replacement was used for manual model fitting of the CanA structure using the graphics program COOT<sup>89</sup>. Structural refinement of the CanA coordinates was performed using the PHENIX package<sup>90</sup>. During refinement, a cross-validation test set was created from a random 5% of the reflections. Data collection and refinement statistics are listed in Supplementary Table 3. Molecular graphics of the CanA crystal structure were prepared using the PyMOL Molecular Graphics System (Schrodinger) (DeLano Scientific LLC, Palo Alto, CA) or ChimeraX<sup>92</sup>. A representative image of a portion of the CanA electron density map is provided in Supplementary Fig. 20.

### Bioinformatic analyses of CanX and AbpX

All sequence similarity searches were performed using the HMMER webserver<sup>47</sup> against the UniProtKB database and a local installation of HMMER v3.4<sup>103</sup> against the non-redundant (nr) protein sequence database. Additionally, HHsearch within the MPI Bioinformatics Toolkit<sup>104</sup> was utilized to detect distant homologs. For pairwise sequence comparisons of representative archaeal and bacterial cell-

surface filament-forming proteins (Fig. 6a), we constructed MSAs by running three iterations of HHblits<sup>105</sup> against the UniClust30 database<sup>106</sup>. Secondary structure information was incorporated into these alignments using the addss.pl script from HH-suite3. Profile hidden Markov models (HMMs) were then generated using hmmake, and context-specific transitions were calculated using cstranslate. A custom-built profile HMM database was built as outlined in the HH-suite3 user guide. To obtain pairwise HHsearch probabilities, individual alignments were searched against this custom-built profile HMM database using default parameters.

The query sequences correspond to previously characterized archaeal and bacterial surface proteins that assemble into filaments based on  $\beta$ -sandwich/jelly-roll folds, employing either donor-strand complementation (type I pili, TasA-like proteins) or helical packing of N-terminal  $\alpha$ -helices (type IV pili and archaeal flagellins/archaeellins). For clarity, the proteins analyzed in Fig. 6a were grouped into five categories: archaeal CanX proteins, archaeal and bacterial TasA-like proteins, bacterial type I pilins, archaeal type IV pilins and archaeellins, and bacterial type IV pilins. Figure 6a includes the following proteins: CanX from *P. abyssi* (WP\_338251948.1) and *Hyperthermus* sp. (RUM47266.1); CanX-like from *S. solfataricus* (WP\_009990924.1, WP\_009991716.1, WP\_029552507.1); AbpX from *P. abyssi* (WP\_338249486.1); AbpA from *P. calidifontis* (A3MUL8); 0406 thread subunit from *S. acidocaldarius* (Q4JBK8); TasA from *Archaeoglobus fulgidus* (AOA075WE38), *Methanosarcina barkeri* (AOA0E3QZX2), and *B. subtilis* (P54507); CalY1 from *B. subtilis* (AKB91166); bacterial type I pilins from *E. coli* (PapA–P04127, FimA1–P04128), *Acinetobacter baumannii* (CsuA/B–A0A6F8TDQ5), and *Pseudomonas aeruginosa* (CupE1–Q9HVE4); archaeal flagellins from *Methanospirillum hungatei* (Q2FUM4), *Methanococcus maripaludis* (Q6LWP3), and *Pyrococcus furiosus* (AOA0B4ZYM1); archaeal type IV pilins from *S. solfataricus* (AOA157T322, UpsA–Q7LXX9), *Sulfolobus islandicus* (M9UD72), and *S. acidocaldarius* (AapA–Q4J6I2); and bacterial type IV pilins from *Thermus thermophilus* (PilA5–Q72GL2), *P. aeruginosa* (PilA–P02973), and *Neisseria gonorrhoeae* (Pile1–P02974).

To investigate the prevalence of TasA-like filaments in archaea (Fig. 6b), we searched for homologs of TasA, AbpX, AbpA, the 0406 thread subunit, and CanX-like proteins in archaeal proteomes within the GTDB (release R220)<sup>107</sup> using HMMER. First, we built MSAs by running three iterations of jackhmmmer against the nr70\_arc database, a filtered version of the nr protein sequence database containing only archaeal sequences and with a maximum pairwise sequence identity of 70%. For TasA, we used the pre-existing alignment (PF12389) from the Pfam database<sup>108</sup>. The resulting alignments were then converted into profile HMMs using hmmbuild. To detect homologs of each query protein in the GTDB archaeal proteomes, we performed searches with hmmsearch using a stringent *E*-value threshold ( $-E$ ) of  $1e-10$ . We visualized the presence of these homologs in a GTDB archaeal phylogenetic tree using iTOL<sup>109</sup>. For this, we converted the GTDB archaeal trees into iTOL-compatible format using the convert\_to\_itol method from GTDB-tk v2.4.0+<sup>110</sup>. For signal peptide prediction, we used DeepTMHMM<sup>111</sup>, and for obtaining structural models, we used a local installation of AlphaFold2<sup>45</sup>, the AlphaFold3 webserver<sup>49</sup>, as well as the AlphaFold Protein Structure Database<sup>112</sup>.

### Reporting summary

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

### Data availability

The ex vivo CanX cryo-EM map and the corresponding atomic model were deposited in the EMDB and the PDB under accession numbers EMD-51935 and 9H8B, respectively. The in vitro CanA cryo-EM map and the corresponding atomic model were deposited in the EMDB and the PDB under accession numbers EMD-26546 and 7UII, respectively. The



Hyper2 thick tube cryo-EM map and the corresponding atomic model were deposited in the EMDB and the PDB under accession numbers [EMD-44403](#) and [9BAB](#), respectively. The Hyper2-thintube cryo-EM map and the corresponding atomic model were deposited in the EMDB and the PDB under accession numbers [EMD-44404](#) and [9BAC](#), respectively. The coordinate file and diffraction data for the CanA crystal structure were deposited into the PDB under accession code [9DLO](#). Unless otherwise stated, all data supporting the results of this study can be found in the article, supplementary, and source data files. Source data are provided with this paper.

## References

- Egelman, E. H. Cryo-EM of bacterial pili and archaeal flagellar filaments. *Curr. Opin. Struct. Biol.* **46**, 31–37 (2017).
- Kreutzberger, M. A. B. et al. Convergent evolution in the supercoiling of prokaryotic flagellar filaments. *Cell* **185**, 3487–3500.e3414 (2022).
- Hospenthal, M. K. et al. Structure of a chaperone-usher pilus reveals the molecular basis of rod uncoiling. *Cell* **164**, 269–278 (2016).
- Spaulding, C. N. et al. Functional role of the type 1 pilus rod structure in mediating host-pathogen interactions. *Elife* **7**, e31662 (2018).
- Pakharukova, N. et al. Archaic chaperone-usher pili self-secrete into superelastic zigzag springs. *Nature* **609**, 335–340 (2022).
- Doran, M. H., Baker, J. L., Dahlberg, T., Andersson, M. & Bullitt, E. Three structural solutions for bacterial adhesion pilus stability and superelasticity. *Structure* **31**, 529–540.e527 (2023).
- Bohning, J. et al. Architecture of the biofilm-associated archaic chaperone-usher pilus CupE from *Pseudomonas aeruginosa*. *PLoS Pathog.* **19**, e1011177 (2023).
- Hospenthal, M. K. et al. The cryoelectron microscopy structure of the type 1 chaperone-usher pilus rod. *Structure* **25**, 1829–1838.e1824 (2017).
- Bohning, J. et al. Donor-strand exchange drives assembly of the TasA scaffold in *Bacillus subtilis* biofilms. *Nat. Commun.* **13**, 7082 (2022).
- Wang, F., Cvirkaitė-Krupovic, V., Krupovic, M. & Egelman, E. H. Archaeal bundling pili of *Pyrobaculum caldifontis* reveal similarities between archaeal and bacterial biofilms. *Proc. Natl. Acad. Sci. USA* **119**, e2207037119 (2022).
- Gaines, M. C. et al. Electron cryo-microscopy reveals the structure of the archaeal thread filament. *Nat. Commun.* **13**, 7411 (2022).
- Sleutel, M., Pradhan, B., Volkov, A. N. & Remaut, H. Structural analysis and architectural principles of the bacterial amyloid curli. *Nat. Commun.* **14**, 2822 (2023).
- Wang, F. et al. Structure of microbial nanowires reveals stacked hemes that transport electrons over micrometers. *Cell* **177**, 361–369.e310 (2019).
- Wang, F. et al. Cryo-EM structure of an extracellular *Geobacter* OmcE cytochrome filament reveals tetrahaem packing. *Nat. Microbiol.* **7**, 1291–1300 (2022).
- Wang, F. et al. Structure of *Geobacter* OmcZ filaments suggests extracellular cytochrome polymers evolved independently multiple times. *Elife* **11**, e81551 (2022).
- Baquero, D. P. et al. Extracellular cytochrome nanowires appear to be ubiquitous in prokaryotes. *Cell* **186**, 2853–2864.e2858 (2023).
- Costa, T. R. D. et al. Structure of the bacterial sex F pilus reveals an assembly of a stoichiometric protein-phospholipid complex. *Cell* **166**, 1436–1444.e1410 (2016).
- Zheng, W. et al. Cryoelectron-microscopic structure of the pKpQIL conjugative pili from carbapenem-resistant *Klebsiella pneumoniae*. *Structure* **28**, 1321–1328.e1322 (2020).
- Beltran, L. C. et al. Archaeal DNA-import apparatus is homologous to bacterial conjugation machinery. *Nat. Commun.* **14**, 666 (2023).
- Loquet, A. et al. Atomic model of the type III secretion system needle. *Nature* **486**, 276–279 (2012).
- Flacht, L. et al. Integrative structural analysis of the type III secretion system needle complex from *Shigella flexneri*. *Protein Sci.* **32**, e4595 (2023).
- Hvorecny, K. L. & Kollman, J. M. Greater than the sum of parts: mechanisms of metabolic regulation by enzyme filaments. *Curr. Opin. Struct. Biol.* **79**, 102530 (2023).
- Zhu, J. et al. Protein assembly by design. *Chem. Rev.* **121**, 13701–13796 (2021).
- Egelman, E. H. Three-dimensional reconstruction of helical polymers. *Arch. Biochem. Biophys.* **581**, 54–58 (2015).
- Wang, F., Gnewou, O., Solemanifar, A., Conticello, V. P. & Egelman, E. H. Cryo-EM of helical polymers. *Chem. Rev.* <https://doi.org/10.1021/acs.chemrev.1c00753> (2022).
- Miller, J. G., Hughes, S. A., Modlin, C. & Conticello, V. P. Structures of synthetic helical filaments and tubes based on peptide and peptido-mimetic polymers. *Q. Rev. Biophys.* 1–103 <https://doi.org/10.1017/S0033583522000014> (2022).
- Fulton, D. A., Dura, G. & Peters, D. T. The polymer and materials science of the bacterial fimbriae Caf1. *Biomater. Sci.* **11**, 7229–7246 (2023).
- Nguyen, P. Q., Courchesne, N. D., Duraj-Thatte, A., Pra-veschotinunt, P. & Joshi, N. S. Engineered living materials: prospects and challenges for using biological systems to direct the assembly of smart materials. *Adv. Mater.* **30**, e1704847 (2018).
- Stetter, K. O. Ultrathin mycelia-forming organisms from submarine volcanic areas having an optimum growth temperature of 105 °C. *Nature* **300**, 258–260 (1982).
- Stetter, K. O., König, H. & Stackebrandt, E. *Pyrodictium* gen. nov., a new genus of submarine disc-shaped sulphur reducing archaeobacteria growing optimally at 105 degrees C. *Syst. Appl. Microbiol.* **4**, 535–551 (1983).
- König, H., Messner, P. & Stetter, K. O. The fine structure of the fibers of *Pyrodictium occultum*. *FEMS Microbiol. Lett.* **49**, 207–212 (1988).
- Rieger, G., Rachel, R., Hermann, R. & Stetter, K. O. Ultrastructure of the hyperthermophilic archaeon *Pyrodictium abyssi*. *J. Struct. Biol.* **115**, 78–87 (1995).
- Pley, U. et al. *Pyrodictium abyssi* sp. nov. represents a novel heterotrophic marine archaeal hyperthermophile growing at 110 °C. *Syst. Appl. Microbiol.* **14**, 245–253 (1991).
- Lin, T. J. et al. *Pyrodictium delaneyi* sp. nov., a hyperthermophilic autotrophic archaeon that reduces Fe(III) oxide and nitrate. *Int. J. Syst. Evol. Microbiol.* **66**, 3372–3376 (2016).
- Barton, N. R. et al. *Chimeric Cannulae Proteins, Nucleic Acids Encoding them and Methods for Making and Using them* (United States Patent, 2008).
- Rieger, G., Müller, K., Hermann, R., Stetter, K. O. & Rachel, R. Cultivation of hyperthermophilic archaea in capillary tubes resulting in improved preservation of fine structures. *Arch. Microbiol.* **168**, 373–379 (1997).
- Horn, C., Paulmann, B., Kerlen, G., Junker, N. & Huber, H. In vivo observation of cell division of anaerobic hyperthermophiles by using a high-intensity dark-field microscope. *J. Bacteriol.* **181**, 5114–5118 (1999).
- Nickell, S., Hegerl, R., Baumeister, W. & Rachel, R. *Pyrodictium* cannulae enter the periplasmic space but do not enter the cytoplasm, as revealed by cryo-electron tomography. *J. Struct. Biol.* **141**, 34–42 (2003).

39. Shibata, S. et al. Structure of polymerized type V pilin reveals assembly mechanism involving protease-mediated strand exchange. *Nat. Microbiol.* **5**, 830–837 (2020).
40. Wang, T. et al. CryoSeek: A strategy for bioentity discovery using cryoelectron microscopy. *Proc. Natl. Acad. Sci. USA* **121**, e2417046121 (2024).
41. Miyazaki, U. et al. *Pyrofollius japonicus* gen. nov. sp. nov., a novel member of the family *Pyrodictiaceae* isolated from the Iheya North hydrothermal field. *Extremophiles* **27**, 28 (2023).
42. Jamali, K. et al. Automated model building and protein identification in cryo-EM maps. *Nature* **628**, 450–457 (2024).
43. Kreitner, R. et al. Complete sequential assignment and secondary structure prediction of the cannulae forming protein CanA from the hyperthermophilic archaeon *Pyrodicticum abyssi*. *Biomol. NMR Assign.* **14**, 141–146 (2020).
44. Jarrell, K. F. et al. N-linked glycosylation in Archaea: a structural, functional, and genetic analysis. *Microbiol. Mol. Biol. Rev.* **78**, 304–341 (2014).
45. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
46. Beltran, L. et al. The mating pilus of *E. coli* pED208 acts as a conduit for ssDNA during horizontal gene transfer. *mBio* **15**, e0285723 (2024).
47. Potter, S. C. et al. HMMER web server: 2018 update. *Nucleic Acids Res.* **46**, W200–W204 (2018).
48. van Kempen, M. et al. Fast and accurate protein structure search with Foldseek. *Nat. Biotechnol.* **42**, 243–246 (2024).
49. Abramson, J. et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500 (2024).
50. Almagro Armenteros, J. J. et al. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* **37**, 420–423 (2019).
51. Egelman, E. H. Reconstruction of helical filaments and tubes. *Methods Enzymol.* **482**, 167–183 (2010).
52. Egelman, E. H. The iterative helical real space reconstruction method: surmounting the problems posed by real polymers. *J. Struct. Biol.* **157**, 83–94 (2007).
53. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).
54. Munte, C. E. et al. Biophysical characterization and solution structure of the cannulae-forming protein CanA from the hyperthermophilic archaeon *Pyrodicticum abyssi*. *Sci. Rep.* **15**, 28563 (2025).
55. Wang, C. L., Leavis, P. C. & Gergely, J. Kinetic studies show that Ca<sup>2+</sup> and Tb<sup>3+</sup> have different binding preferences toward the four Ca<sup>2+</sup>-binding sites of calmodulin. *Biochemistry* **23**, 6410–6415 (1984).
56. Griko, Y. V. Energetics of Ca(2+)-EDTA interactions: calorimetric study. *Biophys. Chem.* **79**, 117–127 (1999).
57. UniProt, C. UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Res.* **51**, D523–D531 (2023).
58. Puorger, C., Vetsch, M., Wider, G. & Glockshuber, R. Structure, folding and stability of FimA, the main structural subunit of type 1 pili from uropathogenic *Escherichia coli* strains. *J. Mol. Biol.* **412**, 520–535 (2011).
59. Zyla, D., Echeverria, B. & Glockshuber, R. Donor strand sequence, rather than donor strand orientation, determines the stability and non-equilibrium folding of the type 1 pilus subunit FimA. *J. Biol. Chem.* **295**, 12437–12448 (2020).
60. Evans, R. et al. Protein complex prediction with AlphaFold-Multimer. Preprint at *bioRxiv* <https://doi.org/10.1101/2021.10.04.463034> (2022).
61. Denise, R., Abby, S. S. & Rocha, E. P. C. Diversification of the type IV filament superfamily into machines for adhesion, protein secretion, DNA uptake, and motility. *PLoS Biol.* **17**, e3000390 (2019).
62. Waksman, G. Structural and molecular biology of a protein-polymerizing nanomachine for pilus biogenesis. *J. Mol. Biol.* **429**, 2654–2666 (2017).
63. Zyla, D. S. et al. The assembly platform FimD is required to obtain the most stable quaternary structure of type 1 pili. *Nat. Commun.* **15**, 3032 (2024).
64. Nishiyama, M., Ishikawa, T., Rechsteiner, H. & Glockshuber, R. Reconstitution of pilus assembly reveals a bacterial outer membrane catalyst. *Science* **320**, 376–379 (2008).
65. Costa, T. R. D. et al. Secretion systems in Gram-negative bacteria: structural and mechanistic insights. *Nat. Rev. Microbiol.* **13**, 343–359 (2015).
66. Xu, Q. et al. A distinct type of pilus from the human microbiome. *Cell* **165**, 690–703 (2016).
67. Bronner, F. Extracellular and intracellular regulation of calcium homeostasis. *ScientificWorldJournal* **1**, 919–925 (2001).
68. Guo, S., Vance, T. D. R., Stevens, C. A., Voets, I. & Davies, P. L. RTX Adhesins are Key Bacterial Surface Megaproteins in the Formation of Biofilms. *Trends Microbiol.* **27**, 453–467 (2019).
69. Spitz, O. et al. Type I secretion systems—one mechanism for all? *Microbiol. Spectr.* **7** <https://doi.org/10.1128/microbiolspec.PSIB-0003-2018> (2019).
70. Griessl, M. H. et al. Structural insight into the giant Ca(2+)-binding adhesin SiiE: implications for the adhesion of *Salmonella enterica* to polarized epithelial cells. *Structure* **21**, 741–752 (2013).
71. Guo, S. et al. Role of Ca(2+)-binding in folding the tandem beta-sandwich extender domains of a bacterial ice-binding adhesin. *FEBS J.* **280**, 5919–5932 (2013).
72. Vance, T. D. et al. Ca<sup>2+</sup>-stabilized adhesin helps an Antarctic bacterium reach out and bind ice. *Biosci. Rep.* **34**, e00121 (2014).
73. Bumba, L. et al. Calcium-driven folding of RTX domain beta-rolls ratchets translocation of RTX proteins through type I secretion ducts. *Mol. Cell* **62**, 47–62 (2016).
74. Peters, B. et al. Structural and functional dissection reveals distinct roles of Ca<sup>2+</sup>-binding sites in the giant adhesin SiiE of *Salmonella enterica*. *PLoS Pathog.* **13**, e1006418 (2017).
75. Vance, T. D. R., Ye, Q., Conroy, B. & Davies, P. L. Essential role of calcium in extending RTX adhesins to their target. *J. Struct. Biol. X* **4**, 100036 (2020).
76. Vance, T. D. R., Guo, S., Assaie-Ardakany, S., Conroy, B. & Davies, P. L. Structure and functional analysis of a bacterial adhesin sugar-binding domain. *PLoS ONE* **14**, e0220045 (2019).
77. Oude Vrielink, A. S., Vance, T. D., de Jong, A. M., Davies, P. L. & Voets, I. K. Unusually high mechanical stability of bacterial adhesin extender domains having calcium clamps. *PLoS ONE* **12**, e0174682 (2017).
78. Baranova, E. et al. SbsB structure and lattice reconstruction unveil Ca<sup>2+</sup> triggered S-layer assembly. *Nature* **487**, 119–122 (2012).
79. Herdman, M. et al. High-resolution mapping of metal ions reveals principles of surface layer assembly in *Caulobacter crescentus* cells. *Structure* **30**, 215–228.e215 (2022).
80. Sogues, A. et al. Structure and function of the EA1 surface layer of *Bacillus anthracis*. *Nat. Commun.* **14**, 7051 (2023).
81. van Wolferen, M., Wagner, A., van der Does, C. & Albers, S. V. The archaeal Ced system imports DNA. *Proc. Natl. Acad. Sci. USA* **113**, 2496–2501 (2016).
82. Costa, T. R. D. et al. Type IV secretion systems: advances in structure, function, and activation. *Mol. Microbiol.* **115**, 436–452 (2021).
83. Utturkar, S. M. et al. Draft genome sequence of *Pyrodicticum occultum* PL19T, a marine hyperthermophilic species of archaea that grows optimally at 105 degrees C. *Genome Announc.* **4**, e00016 (2016).
84. Demeler, B. & Gorbet, G. E. Analytical Ultracentrifugation Data Analysis with UltraScan-III. In *Analytical Ultracentrifugation* (eds.

- Uchiyama, S., Arisaka, F., Stafford, W.E. & Laue, T.) 119–143 (Springer, 2016).
85. Brookes, E., Cao, W. & Demeler, B. A two-dimensional spectrum analysis for sedimentation velocity experiments of mixtures with heterogeneity in molecular weight and shape. *Eur. Biophys. J.* **39**, 405–414 (2010).
  86. Demeler, B. & Brookes, E. Monte Carlo analysis of sedimentation experiments. *Colloid Polym. Sci.* **286**, 129–137 (2008).
  87. Demeler, B. & Van Holde, K. E. Sedimentation velocity analysis of highly heterogeneous systems. *Anal. Biochem.* **335**, 279–288 (2004).
  88. 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).
  89. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D. Biol. Crystallogr.* **60**, 2126–2132 (2004).
  90. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 213–221 (2010).
  91. Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D. Struct. Biol.* **75**, 861–877 (2019).
  92. Goddard, T. D. et al. UCSF chimeraX: meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018).
  93. Rohou, A. & Grigorieff, N. CTFFIND4: fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
  94. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
  95. Egelman, E. H. A robust algorithm for the reconstruction of helical filaments using single-particle methods. *Ultramicroscopy* **85**, 225–234 (2000).
  96. Afonine, P. V. et al. New tools for the analysis and validation of cryo-EM maps and atomic models. *Acta Crystallogr. D. Struct. Biol.* **74**, 814–840 (2018).
  97. He, J., Li, T. & Huang, S. Y. Improvement of cryo-EM maps by simultaneous local and non-local deep learning. *Nat. Commun.* **14**, 3217 (2023).
  98. Pettersen, E. F. et al. UCSF Chimera-a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
  99. Afonine, P. V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. D. Struct. Biol.* **74**, 531–544 (2018).
  100. 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).
  101. Williams, C. J. et al. MolProbity: more and better reference data for improved all-atom structure validation. *Protein Sci.* **27**, 293–315 (2018).
  102. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
  103. Eddy, S. R. Accelerated profile HMM searches. *PLoS Comput. Biol.* **7**, e1002195 (2011).
  104. Zimmermann, L. et al. A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. *J. Mol. Biol.* **430**, 2237–2243 (2018).
  105. Steinegger, M. et al. HH-suite3 for fast remote homology detection and deep protein annotation. *BMC Bioinforma.* **20**, 473 (2019).
  106. Mirdita, M. et al. Uniclust databases of clustered and deeply annotated protein sequences and alignments. *Nucleic Acids Res.* **45**, D170–D176 (2017).
  107. Parks, D. H. et al. GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic Acids Res.* **50**, D785–D794 (2022).
  108. El-Gebali, S. et al. The Pfam protein families database in 2019. *Nucleic Acids Res.* **47**, D427–D432 (2019).
  109. Letunic, I. & Bork, P. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Res.* **52**, W78–W82 (2024).
  110. Chaumeil, P. A., Mussig, A. J., Hugenholtz, P. & Parks, D. H. GTDB-Tk v2: memory friendly classification with the genome taxonomy database. *Bioinformatics* **38**, 5315–5316 (2022).
  111. Hallgren, J. et al. DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks. Preprint at *bioRxiv* <https://doi.org/10.1101/2022.04.08.487609> (2022).
  112. Varadi, M. et al. AlphaFold protein structure database in 2024: providing structure coverage for over 214 million protein sequences. *Nucleic Acids Res.* **52**, D368–D375 (2024).

## Acknowledgements

High-resolution cryoEM imaging on the recombinant cannula-like filaments was conducted at the Molecular Electron Microscopy Core facility at the University of Virginia, which is supported by the School of Medicine and built with NIH grant G20-RR31199. In addition, the Titan Krios (SIG S10-RR025067) and K3/GIF (U24-GM116790) were purchased in part or in full using the designated NIH grants. High-resolution cryoEM imaging on the ex vivo cannulae was performed at the VIB-VUB Facility for Bio Electron Cryogenic Microscopy (BECM) and we thank Marcus Fislage and Dirk Reiter for assistance in data collection. This study was supported by the Robert P. Apkarian Integrated Electron Microscopy Core (RRID: SCR\_023537) at Emory University, which is subsidized by the School of Medicine and Emory College of Arts and Sciences. Additional support was provided by the Georgia Clinical & Translational Science Alliance of the National Institutes of Health under award number UL1TR000454. Negative stain TEM images were collected on the JEOL JEM-1400, 120 kV TEM supported by the National Institutes of Health Grant S10 RR025679. Use of the NYX beamline 19-ID at the National Synchrotron Light Source II was supported by the New York Structural Biology Center. This research used resources of the National Synchrotron Light Source II, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Brookhaven National Laboratory under Contract No. DE-SC0012704. NYX detector instrumentation was supported by grant S10OD030394 through the Office of the Director, National Institutes of Health. The AUC studies were supported by the Canada 150 Research Chairs program C150-2017-00015, the National Institutes of Health grant 1R01GM120600, and the Canadian Natural Science and Engineering Research Council Discovery Grant DG-RGPIN-2019-05637 to B.D. The Canadian Center for Hydrodynamics is funded by the Canada Foundation for Innovation grant CFI-37589 (B.D.). UltraScan supercomputer calculations were supported through NSF/XSEDE grant TG-MCB070039N to B.D. This research was supported by grants from NSF (2003962) to V.P.C., NIH (GM138756) to F.W., the Human Frontier Science Program (RGY0074/2021) to V.A., FWO (G043021N) to M.S., and NIH (GM122510) to E.H.E. The content is solely the responsibility of the authors and does not necessarily reflect the official views of the National Institutes of Health. V.P.C. thanks Elizabeth Trembath-Reichert (Arizona State University) for discussions on the metagenomic data that served as the basis for the design of the Hyper2 protein. V.A. thanks Andrei Lupas (MPI Tübingen) for continued support. This work was partly supported by institutional funds of the Max Planck Society. This material is based in part upon work supported by the National Science Foundation Graduate Research Fellowship Program to A.G.S. under grant number (2439564). Any opinions, findings, and conclusions or recommendations expressed in this



material are those of the authors) and do not necessarily reflect the views of the National Science Foundation.

## Author contributions

M.S., H.R., E.H.E., and V.P.C. conceptualized and designed the project. M.S. performed cryogenic freezing, nsTEM, and cryoEM imaging and data processing of the ex vivo *P. abyssi* filaments. J.G.M. and A.G.S. performed protein expression and purification, optimized in vitro assembly conditions, and acquired the initial EM images of recombinant cannulae. F.W. performed cryogenic freezing and cryoEM imaging and data processing of the in vitro *P. abyssi* filaments. R.R.S. performed cryogenic freezing and cryoEM imaging and data processing of the in vitro *Hyperthermus* filaments. R.M. acquired the analytical ultracentrifugation data and R.M. and B.D. performed the AUC data analysis. Y.C. performed crystallization trials. M.R. collected the diffraction data and performed the crystallographic analysis. V.A. performed bioinformatic analyses. M.S., F.W., R.R.S., J.G.M., A.G.S., M.R., R.M., V.A., and V.P.C. made figures. M.S., V.A., and V.P.C. wrote the manuscript with contributions from all authors.

## 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/s41467-025-64120-8>.

**Correspondence** and requests for materials should be addressed to Han Remaut, Edward H. Egelman or Vincent P. Conticello.

**Peer review information** *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.

**Reprints and permissions 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-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025