



**HAL**  
open science

## Archaeal bundling pili of *Pyrobaculum calidifontis* reveal similarities between archaeal and bacterial biofilms

Fengbin Wang, Virginija Cvirkaite-Krupovic, Mart Krupovic, Edward H. Egelman

### ► To cite this version:

Fengbin Wang, Virginija Cvirkaite-Krupovic, Mart Krupovic, Edward H. Egelman. Archaeal bundling pili of *Pyrobaculum calidifontis* reveal similarities between archaeal and bacterial biofilms. *Proceedings of the National Academy of Sciences of the United States of America*, 2022, 119 (26), pp.e2207037119. 10.1073/pnas.2207037119 . pasteur-03701529

**HAL Id: pasteur-03701529**

**<https://pasteur.hal.science/pasteur-03701529>**

Submitted on 22 Jun 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License



# Archaeal bundling pili of *Pyrobaculum calidifontis* reveal similarities between archaeal and bacterial biofilms

Fengbin Wang<sup>a</sup>, Virginija Cvirkaite-Krupovic<sup>b</sup>, Mart Krupovic<sup>b,1</sup> , and Edward H. Egelman<sup>a,1</sup> 

Contributed by Edward Egelman; received April 22, 2022; accepted May 17, 2022; reviewed by Katrina Forest and Lisa Craig

While biofilms formed by bacteria have received great attention due to their importance in pathogenesis, much less research has been focused on the biofilms formed by archaea. It has been known that extracellular filaments in archaea, such as type IV pili, hami, and cannulae, play a part in the formation of archaeal biofilms. We have used cryo-electron microscopy to determine the atomic structure of a previously uncharacterized class of archaeal surface filaments from hyperthermophilic *Pyrobaculum calidifontis*. These filaments, which we call archaeal bundling pili (ABP), assemble into highly ordered bipolar bundles. The bipolar nature of these bundles most likely arises from the association of filaments from at least two different cells. The component protein, AbpA, shows homology, both at the sequence and structural level, to the bacterial protein TasA, a major component of the extracellular matrix in bacterial biofilms, contributing to biofilm stability. We show that AbpA forms very stable filaments in a manner similar to the donor-strand exchange of bacterial TasA fibers and chaperone-usher pathway pili where a  $\beta$ -strand from one subunit is incorporated into a  $\beta$ -sheet of the next subunit. Our results reveal likely mechanistic similarities and evolutionary connection between bacterial and archaeal biofilms, and suggest that there could be many other archaeal surface filaments that are as yet uncharacterized.

cryo-EM | helical polymers | archaea | TapA

In the environment, most microbial species alternate between free-living, planktonic lifestyle, whereby single cells float or swim in liquid medium, and sessile lifestyle, characterized by formation of complex microbial consortia, known as biofilms, in which cells attach to each other as well as to various biotic or abiotic substrates (1–4). It has even been suggested that biofilms represent a default mode of growth for many species, whereas planktonic lifestyle is an artifact of in vitro propagation (5). The cells within a biofilm undergo functional specialization toward a common goal of optimization of the maintenance and viability of the community (1, 6). Thus, different subpopulations within a biofilm are physiologically distinct from each other as well as from planktonic cells, although pinpointing the underlying gene expression pattern remains challenging (7, 8). One important manifestation of these physiological differences is that microbial communities within biofilms display increased resistance to viruses (9) and various antimicrobials, including antibiotics (10, 11). Indeed, studies on biofilms have profound biomedical relevance (12) and hence, most of our understanding of biofilm formation stems from studies on bacterial species.

A common feature of bacterial biofilms is formation of an extracellular polymeric substance, which consists of various polymers, including exopolysaccharides, proteins, and extracellular (e)DNA, that are excreted by the cells in the biofilm and encase the bacterial community within a protective matrix (13). In a soil-dwelling, monoderm bacterium *Bacillus subtilis*, one of the most extensively studied biofilm models (14, 15), maintenance of the structural biofilm integrity and the timing of biofilm development depend on a secreted protein TasA (16, 17), which has a globular jelly-roll domain (18) and forms fibrous assemblies (19). TasA is encoded in an operon with *tapA* and *sipW*. A minor matrix component TapA is also a secreted protein and has been described as a TasA assembly and anchoring protein, but its exact function remains unclear (20). The crystal structure of the central TapA domain revealed a  $\beta$ -sandwich fold, most closely related to that of bacterial macroglobulins (21). The third gene in the operon encodes a signal peptidase SipW, which is responsible for the processing of the N-terminal signal peptides in both TasA and TapA (22, 23).

Although biofilms have been most extensively studied in bacteria, archaea are also known to widely practice this lifestyle (24, 25). Indeed, biofilm formation has been documented for multiple metabolically different archaeal species inhabiting highly diverse ecological niches (24, 25). Archaeal biofilms are thought to provide protection against various environmental stresses (26–28), promote horizontal gene exchange (29, 30) and

## Significance

Biofilms are communities of microbes where cells attach to each other as well as to various surfaces. Bacterial biofilms have been intensively studied due to their importance in many infections, whereas much less is known about biofilms in Archaea, the third domain of life. Using cryo-electron microscopy, we have determined the atomic structure of a surface filament that forms bi-polar bundles connecting archaeal cells. We show that the major protein of these bundling pili has common ancestry with a protein known to be an important component of bacterial biofilms. This adds to our understanding of the evolutionary relationship between bacteria and archaea and may provide new insights into bacterial biofilms.

Author affiliations: <sup>a</sup>Department of Biochemistry and Molecular Genetics, University of Virginia, Charlottesville, VA 22903; and <sup>b</sup>Institut Pasteur, Université Paris Cité, CNRS UMR6047, Archaeal Virology Unit, 75015 Paris, France

Author contributions: M.K. and E.H.E. designed research; F.W., V.C.-K., and M.K. performed research; F.W., M.K., and E.H.E. analyzed data; and F.W., M.K., and E.H.E. wrote the paper.

Reviewers: K.T.F., University of Wisconsin–Madison; and L.C., Simon Fraser University

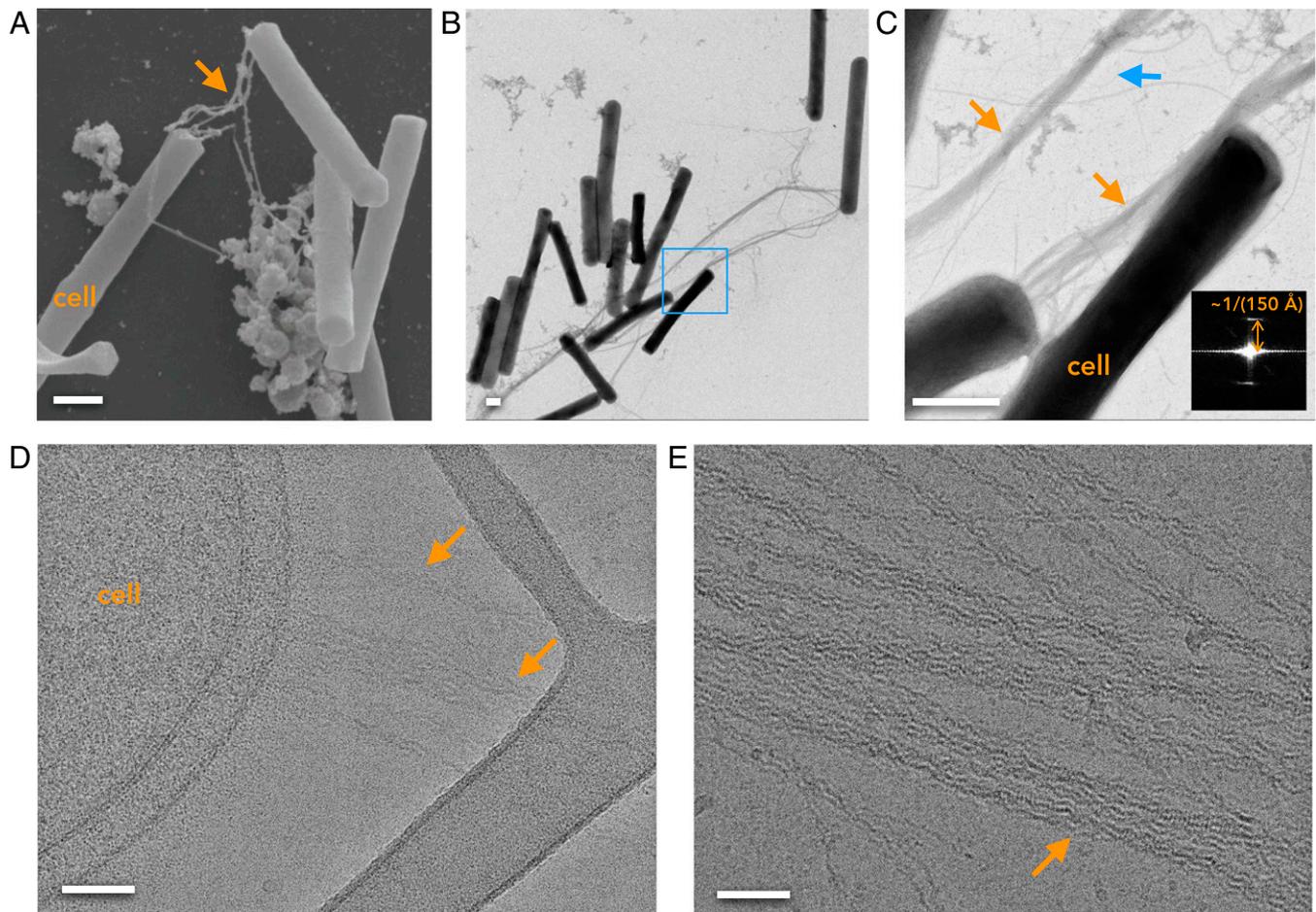
The authors declare no competing interest.

Copyright © 2022 the Author(s). Published by PNAS. This article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

<sup>1</sup>To whom correspondence may be addressed. Email: mart.krupovic@pasteur.fr or egelman@virginia.edu.

This article contains supporting information online at <http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2207037119/-/DCSupplemental>.

Published June 21, 2022.



**Fig. 1.** ABP at different scales. (A) SEM image of *P. calidifontis* ABP fibers (orange arrow) connecting cells. (Scale bar: 500 nm.) (B) Negative staining image of ABP fibers connecting *P. calidifontis* cells. (Scale bar: 500 nm.) (C) Higher magnification image of the boxed area in (B). The orange arrows indicate the dense mesh of ABP fibers seen at this higher magnification. The blue arrow indicates a single thin filament that forms the mesh. The averaged power spectrum from several regions of this mesh (bottom right) shows a single strong periodicity of  $\sim 150 \text{ \AA}$ . (Scale bar: 500 nm.) (D) Cryo-EM image of ABP (orange arrow) attached to a *P. calidifontis* cell. (Scale bar: 50 nm.) (E) Cryo-EM image of ABP bundles (orange arrow) that have been sheared off cells. The cryo-EM images of the bundles also show a strong periodicity of  $\sim 150 \text{ \AA}$  as seen in negative stain (C). (Scale bar: 50 nm.)

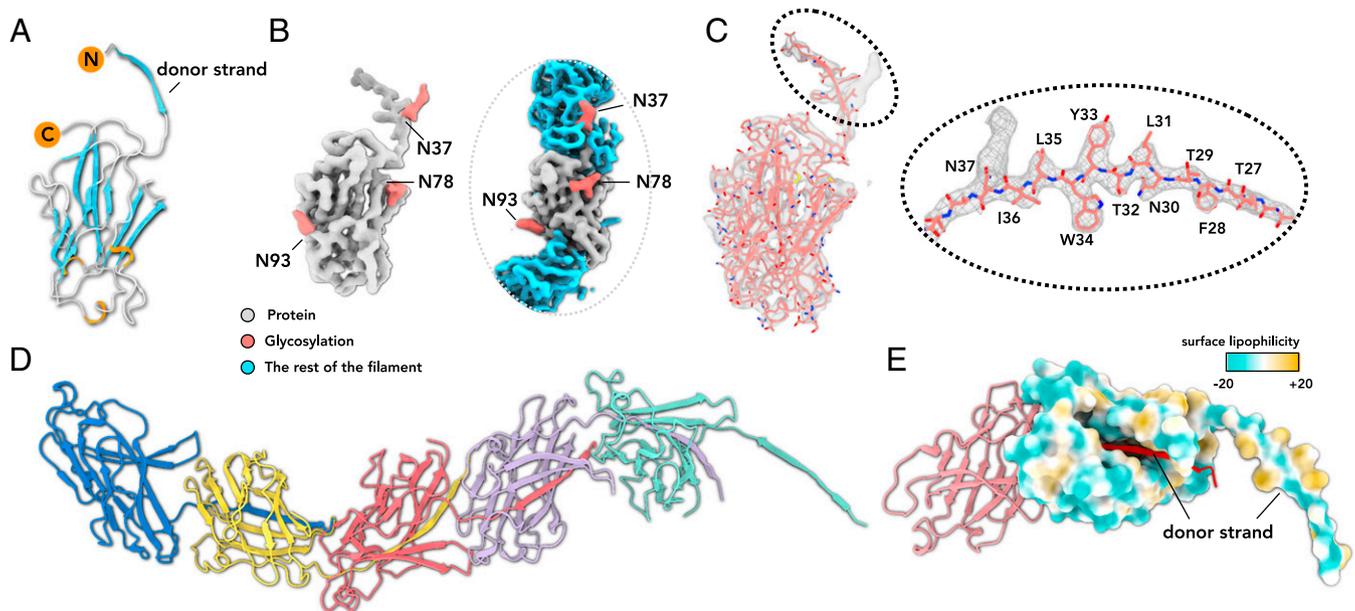
enable syntrophy with other archaea and bacteria (31–34). Molecular studies on archaeal biofilm formation are still in their infancy and have largely focused on genetically tractable hyperthermophilic members of the order Sulfolobales and hyperhalophilic archaea of the class Halobacteria (24). Both Sulfolobales and halophilic archaea deploy different type 4 pili (T4P) during early stages of biofilm formation, whereby cells adhere to various substrates (35, 36). Glycosylation was found to play an important role in this process (37), consistent with the fact that archaeal T4P are typically heavily glycosylated (37–39). The molecular details of archaeal biofilm maturation, involving formation of a complex 3D architecture, are less well understood. In halophilic archaea, exopolysaccharides and eDNA are important components of the extracellular matrix (30, 40), whereas in Sulfolobales biofilms, only small amounts of eDNA were detected (41, 42). Whether fibrous protein assemblies, similar to *B. subtilis* TasA, are involved in the structuring of archaeal biofilms remains unknown.

Here we present a structure of an extracellular filament, which we named archaeal bundling pilus (ABP), produced by a hyperthermophilic (optimal growth temperature  $\sim 90^\circ\text{C}$ ) and neutrophilic (optimal growth pH  $\sim 7$ ) archaeon *Pyrobaculum calidifontis* (order Thermoproteales). We show that the ABP subunit, AbpA, is homologous to the major biofilm matrix protein TasA of *B. subtilis*. Our structural data explains not only how AbpA

subunits polymerize through donor-strand exchange, but also how ABP filaments from different cells form antiparallel bundles, which could promote biofilm matrix formation in archaea. These insights might be also relevant to understanding TasA-mediated biofilm formation in bacteria.

## Results

**Bundling Filaments at Different Microscopic Scales.** *P. calidifontis* is a hyperthermophilic and metabolically versatile archaeon belonging to the order Thermoproteales (phylum Thermoproteota, formerly known as Crenarchaeota) (43). Under scanning electron microscopy (SEM), rod-shaped *P. calidifontis* cells are often seen connected to each other with fibers ranging from 200 to 500  $\text{\AA}$  in diameter (Fig. 1A). These filaments appear to have a polar distribution based upon examining 70 cells by SEM, where 63 were clearly polar. Under negative staining electron microscopy (Fig. 1B and C), these fibers can be resolved to be a mesh formed from thin filaments  $\sim 50 \text{ \AA}$  in diameter, which we name archaeal bundling pili (ABP). The mesh between cells has a single strong axial periodicity of  $\sim 150 \text{ \AA}$  (Fig. 1C, *Inset*) generating the transverse striations that are clearly seen in the micrographs. ABP are seen to directly extend from the cell surface using cryo-electron microscopy (cryo-EM) (Fig. 1D). Multiple forms of the filaments/bundles



**Fig. 2.** The archaeal bundling pilin inserts a donor strand into the adjacent pilin. (A) A ribbon representation of a single *P. calidifontis* AbpA subunit, with N and C termini labeled. The  $\beta$ -strands are in cyan, and three very short  $\alpha$ -helices are in orange. (B) The density accounted for by the AbpA pilin atomic model is in gray, and the extra density, due to post-translational modifications, is in red. This is also shown in a filament map, where the rest of the filament is colored blue. Three putative glycosylation sites are N37, N78, and N93. The amino acids are numbered from the preprotein to be consistent with the Uniprot database, so that the first amino acid in the mature pilin is T25. (C) The cryo-EM map of a single pilin subunit. The close-up view of the donor strand is shown on the right. (D) A filament model of ABP. Each subunit is in a different color. (E) Two adjacent subunits in the ABP. The left one is shown in a ribbon representation, and the right one is shown in an atomic surface view colored by lipophilicity. The donor strand from the subunit on the left is colored red in the ribbon representation, and it can be seen residing in the deep groove in the subunit on the right.

are found, ranging from rare single filaments to bundles up to a couple of hundred Angstroms in diameter (Fig. 1E).

#### Donor Strands Inserted into the Polymerized ABP Subunits.

Since there were too few single filaments found for three-dimensional reconstruction, we used cryo-EM to determine the structure of ordered helical bundles of ABP containing six filaments with a diameter of  $\sim 170$  Å: a central filament surrounded by five outer filaments. We chose these bundles for further analysis as the helical order was not apparent in the larger bundles which may be aggregates of these six-filament bundles. We were able to generate an overall 4.0 Å resolution reconstruction as judged by a map:map Fourier shell correlation (FSC) (SI Appendix, Fig. S1). Each of the filaments in the bundle shares the same helical symmetry, and the subunits are related to each other by a left-handed rotation of  $-77.4^\circ$  and a translation of 32.8 Å along the helical axis, and this same symmetry also relates each of the five outer filaments to the other four. This symmetry, with a helical pitch of 152.5 Å arising from  $4.65$  units/turn  $\times 32.8$  Å rise per subunit, generates the periodicity of  $\sim 150$  Å seen in the mesh between cells (Fig. 1C), establishing that the ABP must be the dominant component of this mesh.

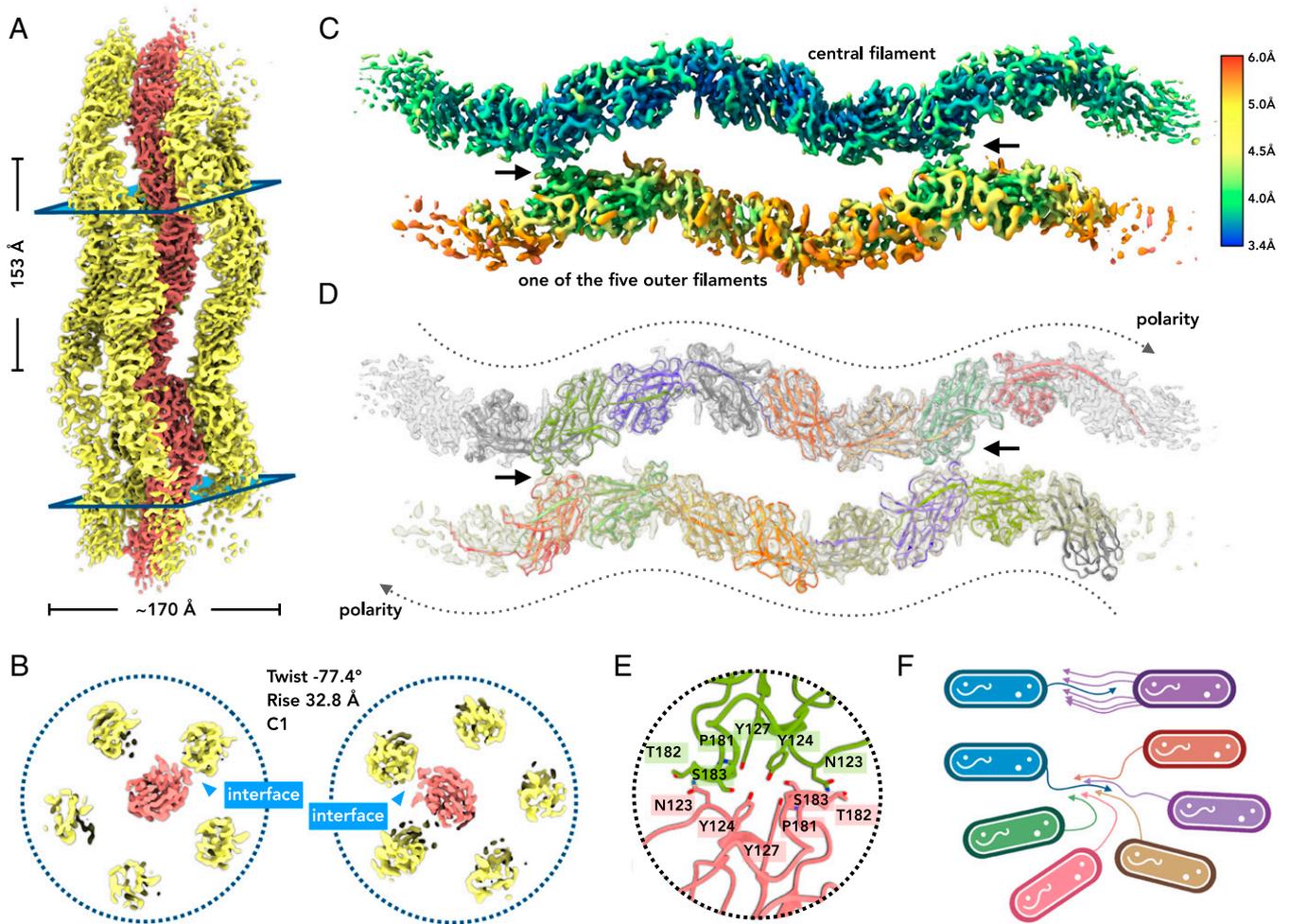
We will first focus on the structure of the component filament, and then discuss the structure of the bundle. At this resolution, we were able to manually generate a C $\alpha$  trace for the protein backbone, with an estimated length of  $\sim 175$  residues. We were then able to determine the pilin sequence by comparing the cryo-EM map with 148 AlphaFold predictions for proteins encoded by *P. calidifontis* that contained between 175 and 300 residues (44). We used a range that included much larger sequences in case there was a cleaved signal sequence that would not be found in the processed protein. There was only one good match to the map from these 148 predictions,

Pcal\_0910 (WP\_011849593), a 204 aa-long protein, which we name AbpA. The first residue of AbpA in the ABP filament is T25, suggesting that the preprotein is N-terminally processed prior to the ABP assembly. Indeed, SignalP analysis (45) predicts that AbpA carries a cleavable signal peptide, with the predicted signal peptidase I cleavage site, 22-AMA↓TL-26, coinciding perfectly with the determined structure. The structure of the ABP subunit AbpA is  $\beta$ -strand-rich and forms a variant of the jelly-roll fold (Fig. 2A). Similar to chaperone-usher (46–48) and type V pili (49), AbpA inserts a donor strand into a groove in the  $\beta$ -sheet of an adjacent subunit (Fig. 2). The donor-strand corresponds to the B strand of the canonical jelly-roll fold, creating a  $\beta$ -sheet with strands I, D, and G (BIDG) and further extending this sheet by connecting it with a  $\beta$ -hairpin (see below), which is an additional structural element compared to the typical jelly-roll folds.

Extra densities were observed on three asparagines, including one on the donor strand (Fig. 2B and C), all of which are part of an Asn-X-Ser/Thr motif (where X is any amino acid) for N-linked glycosylation (50). The interface between two adjacent subunits (Fig. 2D) is  $2,100$  Å<sup>2</sup>, as calculated by PDB-PISA, which is quite large compared to the size of the pilin. The surface of the pilin is predominantly hydrophilic, while part of the donor strand and the groove in which it binds is more hydrophobic (Fig. 2E).

#### Filaments within the Ordered ABP Bundles Have Opposing Polarities.

We asked how filaments within the six-filament ABP bundle (Fig. 3A) interact with each other. Within any cross-section perpendicular to the helical axis, the central filament only makes contact with one outer filament, consistent with the helical twist of the bundle with 4.65 subunits per turn (Fig. 3B), which is the same twist as in the component filaments. The outer filaments make no contacts with each other,



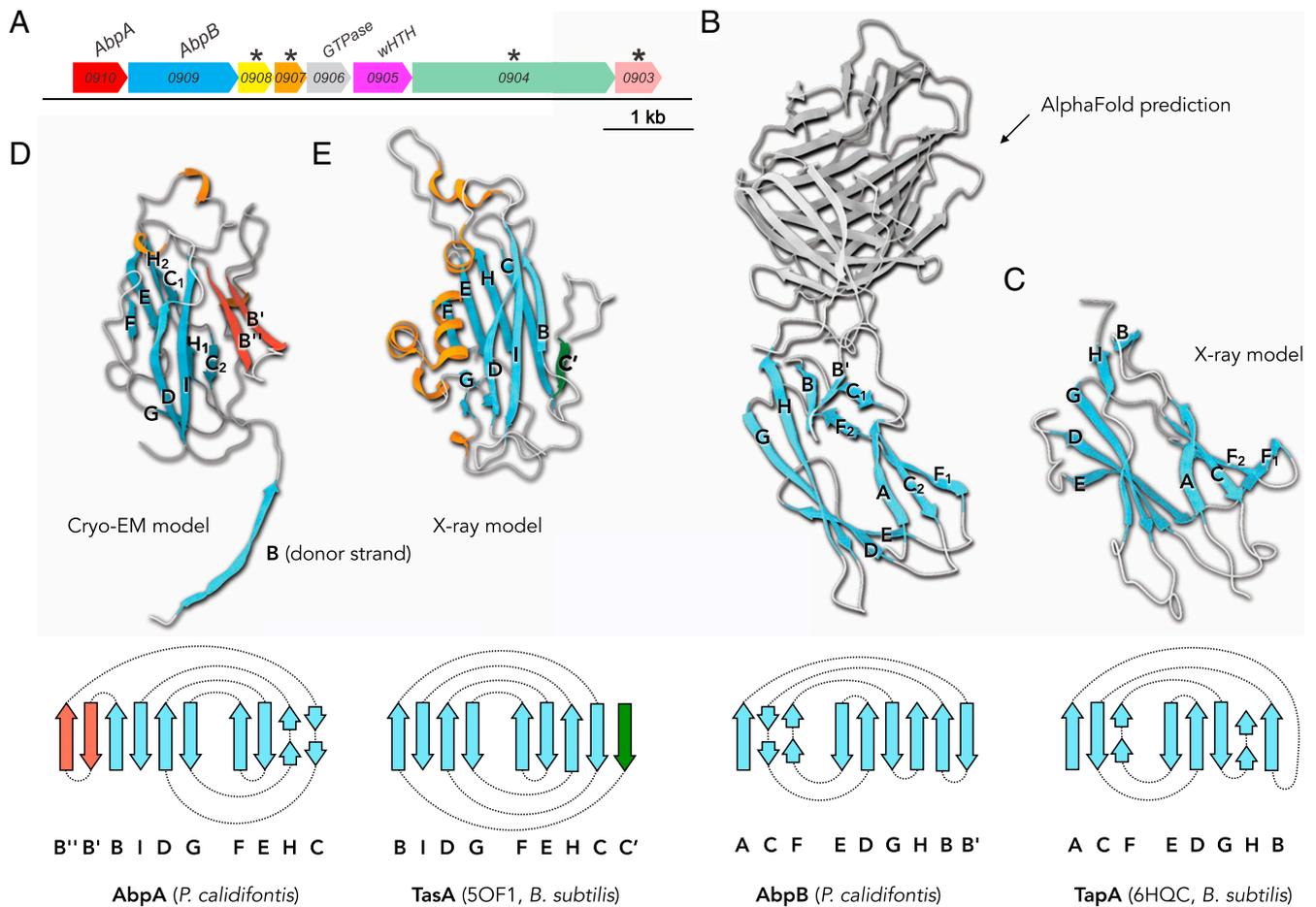
**Fig. 3.** Filaments in the ABP bundle have opposing polarities, elucidating biofilm formation. (A) Cryo-EM map of ABP. A global helical symmetry exists in the six-filament bundle, with the central filament in red and the other five filaments in yellow. The diameter of the six-filament bundle is  $\sim 170$  Å, and the pitch of this bundle is 154 Å. (B) The filament-filament interface between the central filament and the outer filaments, shown at the level of the two planes indicated in (A). (C) The local resolution estimation of the central filament and one of the five outer filaments. The filament-filament interfaces are indicated by black arrows. (D) The same view as (C), with atomic models of these two filaments shown in the map. The map shows that these two filaments have opposite polarities. (E) A close-up view of the filament-filament interface, with the residues involved in the interface labeled. The central filament is in green, and the outer filament is in pink. (F) A schematic model for the role of the bundling pili in biofilm formation. Such a bundle can be formed between as few as two cells (top), but it can also connect, in principle, up to six cells within one bundle (bottom).

so the bundle is held together exclusively by contacts with the central filament. A local resolution estimation (Fig. 3C) shows that the central filament has a higher resolution than determined by the overall FSC, at better than 4 Å. In contrast, for the outer filaments, only the subunits making contact with the central filament have a resolution close to 4 Å, and the remainders of the outer filaments are at a lower resolution, presumably due to flexibility.

What is most striking in the bundle is that the outer filaments have the opposite polarity compared to the central filament, suggesting that they probably come from different cells (Fig. 3D), consistent with the electron microscopy observations showing that ABP bundles are anchored within the envelopes of interconnected cells (Fig. 1B and C). Each interface between an outer filament and the central filament is mediated mainly by hydrogen bonds, with a small buried interfacial area of 400 Å<sup>2</sup> (Fig. 3E). The interface is between one subunit in the central filament and one subunit in an outer filament. While this buried surface area is quite small compared to that found in many complexes or in the interfaces between subunits in the same filament (in each ABP filament the interface between adjacent subunits buries  $\sim 2,100$  Å<sup>2</sup> of surface area), it would be

repeated  $\sim 67$  times in every  $\mu\text{m}$  length of the bundle. Thus, the stability of the interaction between an outer filament and the central filament would be a function of these many interactions. Due to the bipolar nature of the bundle, this interface has a local pseudo-C2 symmetry (Fig. 3E) and repeats every five subunits. With this bundling mechanism, one *P. calidifontis* cell can form a robust bridge with another cell and have the potential to connect up to five different cells simultaneously (Fig. 3F).

**AbpA Is a Homolog of Bacterial TasA.** We next assessed the evolutionary provenance of AbpA by performing sensitive profile-profile comparisons using HHpred (51), whereby a profile hidden Markov model (HMM) of a query sequence was compared against the profile HMMs of the protein family database (Pfam). The search yielded a highly significant (97% probability) hit to a profile HMM of Peptidase\_M73 protein family (PF12389), which includes camelysin, a secreted metallo-endopeptidase (52), and TasA, a key component of bacterial biofilms (18, 53–58). Indeed, it has been suggested that TasA has evolved from an inactivated camelysin-like protease (18). The match encompassed the N-terminal signal sequence and the C-terminal  $\beta$ -strand rich ectodomain of both proteins (5F)



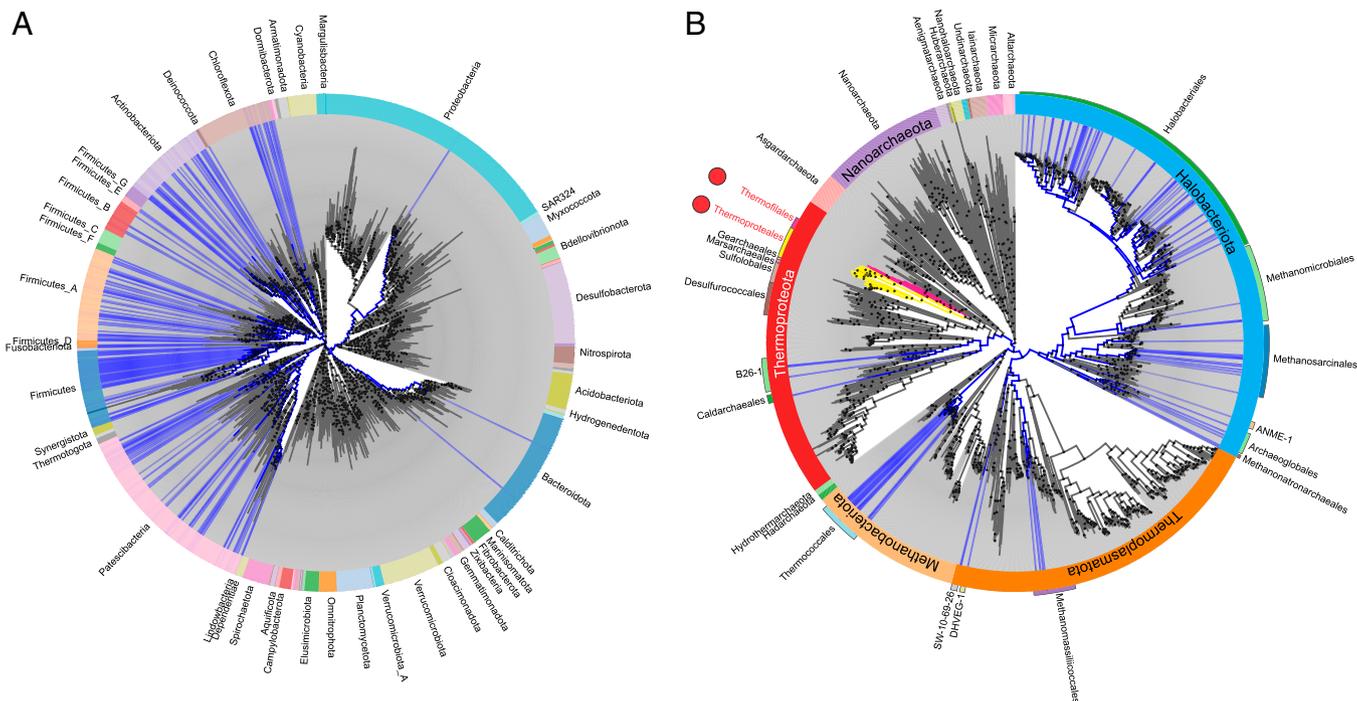
**Fig. 4.** AbpA is a homolog of bacterial TasA. (A) Genes in the conserved *Pyrobaculum* operon, including *abpA* (TasA-like) and *abpB* (TapA-like) genes. Asterisks denote ORFs with predicted transmembrane domains. (B) The AlphaFold prediction of *P. calidifontis* TapA-like protein AbpB. The domain homologous to bacterial TapA is highlighted and the schematic representation of the fold is shown below. (C) The TapA protein of *B. subtilis* determined by X-ray crystallography (PDB: 6HQ3), and its corresponding schematic representation. (D) Cryo-EM model of *P. calidifontis* AbpA has a jelly-roll fold, with an additional  $\beta$ -hairpin (red). (E) The X-ray model of *B. subtilis* TasA (PDB: 5OF1), which also has a jelly-roll fold.

*Appendix, Fig. S2*). While the Dali server (59) failed to find significant structural similarity between AbpA and TasA, possibly, at least in part, due to donor-strand insertion in the ABP subunit but not in the crystallized monomeric TasA (PDB: 5OF1) (18), manual comparison of the two proteins revealed a high degree of structural similarity in the jelly-roll domain (Fig. 4). Furthermore, a very recent cryo-EM study reveals globular TasA subunits arranged into a filament with a donor strand from one subunit inserted into a sheet in the next subunit (60), very similar to what we report herein for ABP filaments. Collectively, these sequence and structural comparisons establish an evolutionary relationship between AbpA and TasA.

We further explored the genomic neighborhood similarities of TasA and AbpA-encoding genes. As mentioned above, in *B. subtilis*, TasA is encoded within an operon together with a putative anchoring protein TapA and a signal peptidase SipW (15). AbpA is conserved in hyperthermophilic archaea of the orders Thermoproteales and Thermofilales. In Thermoproteales, the pilin is encoded within a potential operon including eight genes (Fig. 4A and *SI Appendix, Fig. S3*). In *P. calidifontis*, immediately downstream of Pcal\_0910 is a gene encoding a 410 aa-long protein (Pcal\_0909) with a predicted signal peptide. The protein does not have recognizable homologs outside of Thermoproteales and Thermofilales. However, AlphaFold modeling of this protein revealed three  $\beta$ -strand rich domains, with the first, signal peptide-proximal  $\beta$ -sandwich domain showing remarkable

structural similarity to the ectodomain of the *B. subtilis* protein TapA (Fig. 4 A–C), the putative anchoring component for the extracellular TasA filaments (21, 53). Both Pcal\_0909 and TapA homologs have a distinct  $\beta$ -sandwich fold, different from the jelly-roll domain of TasA and AbpA (Fig. 4 D and E). We therefore suggest that Pcal\_0909 is a homolog of TapA and name it AbpB. Analysis of the six other proteins encoded within the *P. calidifontis* ABP operon yielded only generic functional predictions. Namely, Pcal\_0906 is a predicted GTPase (not conserved in the operons of other Thermoproteales species; *SI Appendix, Fig. S3*) and Pcal\_0905 is a winged helix-turn-helix domain containing transcription regulator. Notably, Pcal\_0908 (WP\_193322921), Pcal\_0907 (WP\_011849590), and Pcal\_0903 (WP\_193322920) each have four predicted transmembrane domains, with an N-terminal signal sequence, whereas Pcal\_0904 (WP\_011849587) has a single N-terminal transmembrane domain (*SI Appendix, Fig. S4*) and a large  $\beta$ -strand-rich ectodomain. These proteins might also participate in biofilm formation and/or anchoring of the ABP filaments. In members of the Thermofilales, ABP-encoding loci differ considerably from those of Thermoproteales, with only ABP pilin (*abpA*) and TapA-like *abpB* genes being shared between the two types of loci.

**Archaeal TasA Homologs beyond Thermoproteales and Thermofilales.** The relationship between the AbpA and TasA is recognizable using sensitive profile-profile and structural comparisons. We inquired whether archaea encode TasA



**Fig. 5.** Phylogenetic distribution of TasA-like proteins in bacterial (A) and archaeal (B) genomes. Branches corresponding to bacterial or archaeal species which encode TasA-like proteins matching the Pfam profile PF12389 (E-value cutoff of 0.005) are shown in blue, whereas those lacking it are in black. Branches of Thermoproteales and Thermofilales which encode more divergent TasA homologs are shown in yellow and magenta, respectively. The trees are scaled to show family and species level diversities for bacteria and archaea, respectively. The outer circle indicates the phyla, which are labeled. Relevant archaeal orders are also shown as bars. The figure was prepared using Annotree ([annotree.uwaterloo.ca](http://annotree.uwaterloo.ca)). Sequences of bacterial and archaeal TasA-like proteins are provided in *SI Appendix, Supplementary Data 1*.

homologs which would be more readily recognizable at the sequence level. To this end, we assessed the distribution of the Pfam PF12389 family in bacterial and archaeal genomes available in the Genome Taxonomy Database (GTDB) using Annotree (61, 62) (*SI Appendix, Supplementary Data 1*). In bacteria, TasA homologs are widespread but not universally conserved. They are readily identifiable in different Firmicutes lineages as well as in several other large phyla, including Actinobacteriota, Chloroflexota, Fusobacteriota, and Patescibacteria, whereas diderm bacteria of the phylum Proteobacteria nearly completely lack TasA homologs (Fig. 5A). Notably, we also found TasA homologs to be relatively widespread in archaea, especially, in several lineages previously constituting the phylum Euryarchaeota (recently redistributed into phyla Halobacteriota, Methanobacteriota, and Thermoplasmatota) (63). These include members of the orders Halobacteriales, Methanosarcinales, Archaeoglobales, Methanonaeronarchaeales, Methanomassiliococcales, and Thermococcales as well as several unclassified lineages (Fig. 5B). In Thermoproteota (former Crenarchaeota), only Caldarchaeales and an unclassified lineage B26-1 encode recognizable TasA homologs. Archaeal proteins showing closer similarity to bacterial TasA homologs than AbpA display a patchy distribution suggestive of extensive horizontal gene transfer. Although the function of these archaeal proteins remains to be investigated experimentally, it is likely that they are also implicated in biofilm formation, further suggesting that the mechanism of biofilm structuring in bacteria and, at least, some archaea depends on homologous proteins.

## Discussion

There is extensive literature detailing how pili in the chaperone-usher pathway use a mechanism of strand exchange to form filaments with interlocking  $\beta$ -sheets (64). For these pilin subunits, a chaperone provides a strand to complete a  $\beta$ -sheet in the pilin

prior to assembly. The usher in the outer membrane has the catalytic activity to exchange the strand from the chaperone with a donor strand from another subunit, forming the filament. Since the structure of ABP filaments has a formal resemblance to that of chaperone-usher pili (*SI Appendix, Fig. S5*), with a donor strand from one subunit inserted into a  $\beta$ -sheet in the next subunit, we must ask whether a chaperone and usher are also involved. Analysis of the *P. calidifontis* genome did not reveal any obvious candidates for either a chaperone or a catalytic usher. Indeed, comparison of the AbpA pilin encoding operons in Thermoproteales and Thermofilales showed that only two genes, namely, the pilin gene itself and a gene encoding AbpB (the TapA-like protein), are conserved between the two groups of archaea (*SI Appendix, Fig. S3*).

An AlphaFold prediction for the structure of the AbpA monomer (*SI Appendix, Fig. S6*) has the donor strand folding back into the subunit, to self-complement. Since the AlphaFold model is only a prediction, we cannot experimentally address the question of how the conformation of a monomer may change when it is incorporated into a filament. Nevertheless, it is possible that such self-complementation as shown in the AlphaFold model does take place, eliminating the need for a chaperone to maintain the fold of the subunit prior to assembly. But it is hard to imagine how such a self-complemented subunit could remove the donor strand and insert it into another subunit, displacing the strand there, in the absence of some complicated catalytic pathway. Since the donor strand is rather hydrophobic (Fig. 2E), it is possible that the strand partially folds back on the subunit, but is not inserted into the sheet as in the AlphaFold model, rather than projecting out as in the diagrams of Fig. 2 A–C. If that were the case, glycosylation at N37 on the donor strand and N78 near the gate of the groove might shield the hydrophobic residues and keep the donor strand soluble without self-inserting. The latter case

would lead to a much smaller energy barrier when the pilin assembles into a filament, as the strand would not need to be removed from the sheet in the donor subunit to insert into the next subunit. Notably, the donor-strand (strand B) not only reconstitutes the BIDG  $\beta$ -sheet of the canonical jelly-roll fold but also extends this  $\beta$ -sheet to include a  $\beta$ -hairpin formed by B' and B''  $\beta$ -strands (Fig. 4D), further stabilizing the ABP filament. The archaeal protein CanA (65) might also provide some insights into assembly of ABP filaments. It has been shown using a purely in vitro system that CanA self-assembles into filaments (PDB: 7UII) in the absence of any chaperone or catalyst. In the filament, a donor strand from one subunit is inserted into a  $\beta$ -sheet of another subunit. Thus, CanA might serve as a model for a self-assembly process that does not require a chaperone or catalytic usher as is required in the bacterial chaperone-usher pathway.

Our results have uncovered an unexpected evolutionary relationship between the loci encoding ABP and TasA, a major matrix protein orchestrating biofilm formation in bacteria. We show that not only are the major pilins homologous at the sequence and structural level, but the minor component of the bacterial biofilm matrix, protein TapA, also has a counterpart in archaea. A potential role for the archaeal AbpB could be to initiate the filament assembly by providing the complementing strand to the first ABP subunit. If so, the distal  $\beta$ -stranded domains of AbpB could serve as adhesins, similar to bacterial FimH (66, 67). Whereas in bacteria the signal peptides of both TasA and TapA are processed by a dedicated signal peptidase, SipW, encoded in the *tapA-sipW-tasA* operon, proteolytic processing of the archaeal counterparts appears to be performed by the generic signal peptidase as no dedicated gene was identified in the vicinity of the ABP operon. The AbpA and AbpB proteins conserved in archaeal orders Thermoproteales and Thermofilales display considerable divergence compared to their bacterial homologs, with the relationship being detectable only using some of the most sensitive profile-profile or structural comparisons. However, we found that proteins more closely related to the bacterial TasA homologs display rather broad, even if patchy, distribution in archaea, suggesting that TasA-mediated biofilm formation might be more widespread across both prokaryotic domains of life than currently appreciated. Future functional and structural studies on these TasA-like proteins (SI Appendix, Supplementary Data 1) are bound to uncover further parallels between bacterial and archaeal biofilms. It has been previously considered that T4P are the only cell surface filaments shared between bacteria and archaea (24, 38, 39, 68, 69). Our current results extend the similarities between bacterial and archaeal surface structures to TasA/ABP filaments. It is almost certain that further research on archaeal cell envelopes will uncover additional surface filaments, both archaea-specific and those shared with cellular organisms from the other domains of life. Notably, none of the species in the order Thermoproteales, including members of the genus *Pyrobaculum*, are genetically tractable. Thus, our present study further highlights the utility of cryo-EM in gaining important insights into the biology of nonmodel organisms.

The structure of the *P. caldifontis* ABP bundle provides unique insights into how ABP filaments emanating from different cells can promote cellular aggregation and biofilm formation. In particular, bipolar organization of the bundle can entangle up to six individual cells (Fig. 3F). Notably, negative staining electron microscopy observation of ABP bundles (Fig. 1B) suggests that the bundles may include more than six filaments and hence more cells can be envisioned to participate

in such interactions. Bundles composed of multiple TasA fibers have been observed in *B. subtilis* biofilms. However, high-resolution structural data for such bundles is not available at the moment, with only doublets of TasA filaments having been investigated using cryo-electron tomography-based modeling and molecular dynamics simulation (60). The bipolar ABP structure presented in this work might be relevant for understanding the formation of larger bacterial TasA bundles. Collectively, our results uncover previously unsuspected mechanistic similarities and evolutionary connections underlying the formation of bacterial and archaeal biofilms.

## Materials and Methods

**Cell Growth and Pili Purification.** *Pyrobaculum caldifontis* DSM 21063 cells (43) were purchased from the DSMZ culture collection and grown in 1090 medium (0.1% yeast extract, 1.0% tryptone, 0.3% sodium thiosulfate, pH7) at 90 °C without agitation. Preculture (30 mL) was started from a 200  $\mu$ L cryo-stock, grown for 2 d and then diluted into 200 mL of fresh medium. When OD<sub>600</sub> reached ~0.2, the cells were collected by centrifugation (Sorval SLA1500 rotor, 7,000 rpm, 10 min, 20 °C). The resultant pellet was resuspended in 4 mL of phosphate-buffered saline (PBS) buffer, and the cell suspension was vortexed for 15 min to shear off the extracellular filaments. The cells were removed by centrifugation (Eppendorf F-35-6-30 rotor, 7,830 rpm, 20 min, 20 °C). The supernatant was collected and the filaments were pelleted by ultracentrifugation (Beckman SW60Ti rotor, 38,000 rpm, 2 h, 15 °C). After the run, the supernatant was removed and the pellet was resuspended in 200  $\mu$ L of PBS.

**Scanning Electron Microscopy.** The *P. caldifontis* cells were grown in 1090 medium at 90 °C without shaking until OD<sub>600</sub> reached ~0.2. Then, 18 mL of cell culture was fixed by adding 2 mL of 25% glutaraldehyde solution, incubated at room temperature for 2 h and kept overnight at 4 °C. The fixed culture was then centrifuged (Eppendorf F-35-6-30 rotor, 7,000 rpm, 10 min, 20 °C) and the cell pellet was resuspended in 2 mL of 0.1 M Hepes, 2.5% glutaraldehyde, pH 7.2 buffer. Cell suspension (1 mL) was placed on poly-L-lysine-coated coverslip and incubated overnight at 4 °C. Then the cells were washed three times with 0.1 M Hepes buffer pH 7.2, postfixed for 1 h 30 in 1% osmium tetroxide in 0.1 M Hepes pH 7.2, and rinsed with distilled water. Samples were then dehydrated by gradually washing in ethanol solutions (25%, 50%, 75%, 95%, and 100%), followed by critical point drying with CO<sub>2</sub>. Dried specimens were sputtered with 20 nm gold palladium using a GATAN Ion Beam Coater and imaged with a JEOL JSM 6700F field emission scanning electron microscope.

**Transmission Electron Microscopy.** For negative staining transmission electron microscopy, 10  $\mu$ L of cell suspension were adsorbed onto glow-discharged copper grids with carbon-coated Formvar film and negatively stained with 2.0% (wt/vol) uranyl acetate. The samples were observed under FEI Tecnai Spirit Bio-Twin 120 microscope operated at 120 kV.

**Cryo-EM and Image Processing.** The *P. caldifontis* sample was applied to glow-discharged lacey carbon grids and vitrified using a Leica EM GP plunge freezer (Leica). Grids were imaged on a Titan Krios (300 keV, Thermo Fisher) with a K3 camera (Gatan). A total of 2,027 micrographs were collected under electron counting mode at 1.08 Å per pixel, using a defocus range of 1–2  $\mu$ m with ~50 electrons/Å<sup>2</sup> distributed into 40 fractions. Motion correction and contrast transfer function estimation were done in cryoSPARC (70–72). Initial 2D references were generated from ~200 manually picked particles and these were used for auto-picking 380,371 segments. From these segments, 2D class averages of bundles with high-resolution details were found, and the segments classified by these averages were used for a helical reconstruction (SI Appendix, Fig. S6). The two possible helical symmetries were calculated from an averaged power spectrum generated from the raw particles and one was found to be correct (73, 74). Typically, the hand of a cryo-EM protein reconstruction is determined by inspecting the hand of  $\alpha$ -helices in the map (75). However, this map only has two tiny one-turn helices that were not obvious without an atomic model. So, the map hand was initially determined by comparing with the AlphaFold prediction, and later validated by the hand of two tiny  $\alpha$ -helices. The resolution of the reconstruction

**Table 1. Cryo-EM of *P. calidifontis* bundling pili**

Parameter	Bundling pili
Data collection and processing	
Voltage (kV)	300
Electron exposure (e <sup>-</sup> Å <sup>-2</sup> )	50
Pixel size (Å)	1.08
Particle images (n)	204,565
Shift (pixel)	40
Helical symmetry	
Point group	C1
Helical rise (Å)	32.79
Helical twist (°)	-77.40
Map resolution (Å)	
Map:map FSC (0.143)	4.0
Model:map FSC (0.38)	4.1
d <sub>99</sub>	4.4
Refinement and model validation	
Ramachandran favored (%)	91.2
Ramachandran outliers (%)	0.2
RSCC	0.79
Clashscore	18
Bonds rmsd, length (Å)	0.028
Bonds rmsd, angles (°)	0.731
Deposition ID	
PDB (model)	7UEG
EMDB (map)	EMD-26474

was estimated by both Map:Map FSC, Model:Map FSC, and local resolution estimation in cryoSPARC. The statistics are listed in *SI Appendix, Table S2*.

**Protein Identification and Model Building.** The density corresponding to a single bundling pilin was segmented from the experimental cryo-EM density using Chimera (76). Manual C $\alpha$  tracing estimated the subunit to have ~175

amino acids. A total of 148 protein structures with lengths between 175 and 300 in *P. calidifontis* were predicted by AlphaFold (44). All predictions were then compared with the C $\alpha$  trace for the best (lowest) score using two measures combined into a single metric: (a) the rmsd to the DeepTracer model, and (b) the reciprocal of the percentage of residues in the DeepTracer model that can be aligned to the corresponding AlphaFold prediction. The AbpA protein was identified as the correct pilin with a score of 1.38, while the next best fit had a score of 4.96. This protein identification approach has been developed into an online tool, DeepTracer-ID (<https://deepttracer.uw.edu/deepttracerid-new-job>) (77). The full-length sequence was threaded into the map using DeepTracer (78), manually adjusted in Coot (79), and real-space refined in PHENIX (80). MolProbity was used to evaluate the quality of the filament model (81). The refinement statistics are shown in Table 1.

**Sequence Analyses.** Profile-profile comparisons were performed using HHpred against the Pfam database (51). Signal peptides and transmembrane domains were predicted using SignalP v5 (45) and TMHMM (82), respectively. Genomic neighborhoods were analyzed using the enzyme function initiative-genome neighborhood tool (EFI-GNT) (83) and SyntTax (84). The distribution of TasA-like proteins (Pfam PF12389) in the bacterial and archaeal genomes available in the Genome Taxonomy Database (GTDB) was assessed using AnnoTree (61, 62).

**Data Availability.** The cryo-EM atomic model and map have been deposited in the Protein Data Bank (7UEG) (85), and Electron Microscopy Database (EMD-26474) (86), respectively.

**ACKNOWLEDGMENTS.** The cryo-EM imaging was done at the Molecular Electron Microscopy Core Facility at the University of Virginia, which is supported by the School of Medicine. This work was supported by NIH (GM122510 E.H.E. and K99GM138756 to F.W.). The work in the M.K. laboratory was supported by grants from l'Agence Nationale de la Recherche (ANR-20-CE20-009-02 and ANR-21-CE11-0001-01) and Ville de Paris (Emergence(s) project MEMREMA). We are grateful to Ultrastructural Bioluminescence Core Facility of Institut Pasteur for the access to electron microscopes and Christine Schmitt for help with sample preparation for SEM.

- P. Stoodley, K. Sauer, D. G. Davies, J. W. Costerton, Biofilms as complex differentiated communities. *Annu. Rev. Microbiol.* **56**, 187-209 (2002).
- H. C. Flemming *et al.*, Biofilms: An emergent form of bacterial life. *Nat. Rev. Microbiol.* **14**, 563-575 (2016).
- M. Berlanga, R. Guerrero, Living together in biofilms: The microbial cell factory and its biotechnological implications. *Microb. Cell Fact.* **15**, 165 (2016).
- R. Kolter, E. P. Greenberg, Microbial sciences: The superficial life of microbes. *Nature* **441**, 300-302 (2006).
- K. K. Jefferson, What drives bacteria to produce a biofilm? *FEMS Microbiol. Lett.* **236**, 163-173 (2004).
- A. Dragos *et al.*, Division of labor during biofilm matrix production. *Curr. Biol.* **28**, 1903-1913.e1905 (2018).
- C. Beloin, J. M. Ghigo, Finding gene-expression patterns in bacterial biofilms. *Trends Microbiol.* **13**, 16-19 (2005).
- D. An, M. R. Parsek, The promise and peril of transcriptional profiling in biofilm communities. *Curr. Opin. Microbiol.* **10**, 292-296 (2007).
- L. Vidakovic, P. K. Singh, R. Hartmann, C. D. Nadell, K. Drescher, Dynamic biofilm architecture confers individual and collective mechanisms of viral protection. *Nat. Microbiol.* **3**, 26-31 (2018).
- P. S. Stewart, J. W. Costerton, Antibiotic resistance of bacteria in biofilms. *Lancet* **358**, 135-138 (2001).
- N. Hoiby, T. Bjarnsholt, M. Givskov, S. Molin, O. Ciofu, Antibiotic resistance of bacterial biofilms. *Int. J. Antimicrob. Agents* **35**, 322-332 (2010).
- D. Sharma, L. Misba, A. U. Khan, Antibiotics versus biofilm: An emerging battleground in microbial communities. *Antimicrob. Resist. Infect. Control* **8**, 76 (2019).
- L. Hall-Stoodley, J. W. Costerton, P. Stoodley, Bacterial biofilms: From the natural environment to infectious diseases. *Nat. Rev. Microbiol.* **2**, 95-108 (2004).
- H. Vlamakis, Y. Chai, P. Beauregard, R. Losick, R. Kolter, Sticking together: Building a biofilm the *Bacillus subtilis* way. *Nat. Rev. Microbiol.* **11**, 157-168 (2013).
- S. Amaouteli, N. C. Bamford, N. R. Stanley-Wall, A. T. Kovács, *Bacillus subtilis* biofilm formation and social interactions. *Nat. Rev. Microbiol.* **19**, 600-614 (2021).
- S. B. Otto *et al.*, Privatization of biofilm matrix in structurally heterogeneous biofilms. *mSystems* **5**, e00425-20 (2020).
- D. Romero, C. Aguilar, R. Losick, R. Kolter, Amyloid fibers provide structural integrity to *Bacillus subtilis* biofilms. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 2230-2234 (2010).
- A. Diehl *et al.*, Structural changes of TasA in biofilm formation of *Bacillus subtilis*. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 3237-3242 (2018).
- D. N. Azulay *et al.*, Colloidal-like aggregation of a functional amyloid protein. *Phys. Chem. Chem. Phys.* **22**, 23286-23294 (2020).
- D. Romero, H. Vlamakis, R. Losick, R. Kolter, An accessory protein required for anchoring and assembly of amyloid fibres in *B. subtilis* biofilms. *Mol. Microbiol.* **80**, 1155-1168 (2011).
- C. Earl *et al.*, The majority of the matrix protein TapA is dispensable for *Bacillus subtilis* colony biofilm architecture. *Mol. Microbiol.* **114**, 920-933 (2020).
- A. G. Stöver, A. Driks, Control of synthesis and secretion of the *Bacillus subtilis* protein YqxM. *J. Bacteriol.* **181**, 7065-7069 (1999).
- A. G. Stöver, A. Driks, Secretion, localization, and antibacterial activity of TasA, a *Bacillus subtilis* spore-associated protein. *J. Bacteriol.* **181**, 1664-1672 (1999).
- M. van Wolferen, A. Orell, S. V. Albers, Archaeal biofilm formation. *Nat. Rev. Microbiol.* **16**, 699-713 (2018).
- H. C. Flemming, S. Wuertz, Bacteria and archaea on Earth and their abundance in biofilms. *Nat. Rev. Microbiol.* **17**, 247-260 (2019).
- A. Koerd, J. Gödecke, J. Berger, K. M. Thormann, S. V. Albers, Crenarchaeal biofilm formation under extreme conditions. *PLoS One* **5**, e14104 (2010).
- J. Megaw, B. F. Gilmore, Archaeal persisters: Persister cell formation as a stress response in *Haloflex volcanii*. *Front. Microbiol.* **8**, 1589 (2017).
- C. Lapaglia, P. L. Hartzell, Stress-induced production of biofilm in the hyperthermophile *Archaeoglobus fulgidus*. *Appl. Environ. Microbiol.* **63**, 3158-3163 (1997).
- M. Z. DeMaere *et al.*, High level of intergenera gene exchange shapes the evolution of haloarchaea in an isolated Antarctic lake. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 16939-16944 (2013).
- S. Chimileski, M. J. Franklin, R. T. Papke, Biofilms formed by the archaeon *Haloflex volcanii* exhibit cellular differentiation and social motility, and facilitate horizontal gene transfer. *BMC Biol.* **12**, 65 (2014).
- S. Krause *et al.*, The importance of biofilm formation for cultivation of a Micrarchaeon and its interactions with its Thermoplasmatales host. *Nat. Commun.* **13**, 1735 (2022).
- G. Wegener, V. Krukenberg, D. Riedel, H. E. Tegetmeyer, A. Boetius, Intercellular wiring enables electron transfer between methanotrophic archaea and bacteria. *Nature* **526**, 587-590 (2015).
- A. J. Probst, C. Moissl-Eichinger, "Altiarchaeales": Uncultivated archaea from the subsurface. *Life (Basel)* **5**, 1381-1395 (2015).
- C. Moissl-Eichinger *et al.*, Archaea are interactive components of complex microbiomes. *Trends Microbiol.* **26**, 70-85 (2018).
- A. L. Henche, A. Koerd, A. Ghosh, S. V. Albers, Influence of cell surface structures on crenarchaeal biofilm formation using a thermostable green fluorescent protein. *Environ. Microbiol.* **14**, 779-793 (2012).
- R. N. Esquivel, R. Xu, M. Pohlschroder, Novel archaeal adhesion pilins with a conserved N terminus. *J. Bacteriol.* **195**, 3808-3818 (2013).
- R. N. Esquivel, S. Schulze, R. Xu, M. Hippler, M. Pohlschroder, Identification of *Haloflex volcanii* Pilin N-Glycans with diverse roles in pilus biosynthesis, adhesion, and microcolony formation. *J. Biol. Chem.* **291**, 10602-10614 (2016).
- F. Wang *et al.*, The structures of two archaeal type IV pili illuminate evolutionary relationships. *Nat. Commun.* **11**, 3424 (2020).

39. F. Wang *et al.*, An extensively glycosylated archaeal pilus survives extreme conditions. *Nat. Microbiol.* **4**, 1401–1410 (2019).
40. S. Fröls, M. Dyall-Smith, F. Pfeifer, Biofilm formation by haloarchaea. *Environ. Microbiol.* **14**, 3159–3174 (2012).
41. S. Jachlewski *et al.*, Isolation of extracellular polymeric substances from biofilms of the Thermoacidophilic Archaeon *Sulfolobus acidocaldarius*. *Front. Bioeng. Biotechnol.* **3**, 123 (2015).
42. A. Koerdet *et al.*, Complementation of *Sulfolobus solfataricus* PBL2025 with an  $\alpha$ -mannosidase: Effects on surface attachment and biofilm formation. *Extremophiles* **16**, 115–125 (2012).
43. T. Amo *et al.*, *Pyrobaculum calidifontis* sp. nov., a novel hyperthermophilic archaeon that grows in atmospheric air. *Archaea* **1**, 113–121 (2002).
44. J. Jumper *et al.*, Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
45. J. J. Almagro Armenteros *et al.*, SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* **37**, 420–423 (2019).
46. C. N. Spaulding *et al.*, Functional role of the type 1 pilus rod structure in mediating host-pathogen interactions. *eLife* **7**, e31662 (2018).
47. M. K. Hospenthal *et al.*, The cryoelectron microscopy structure of the type 1 chaperone-usher pilus rod. *Structure* **25**, 1829–1838.e1824 (2017).
48. M. K. Hospenthal *et al.*, Structure of a chaperone-usher pilus reveals the molecular basis of rod uncoiling. *Cell* **164**, 269–278 (2016).
49. S. Shibata *et al.*, Structure of polymerized type V pilin reveals assembly mechanism involving protease-mediated strand exchange. *Nat. Microbiol.* **5**, 830–837 (2020).
50. D. Kohda, Structural basis of protein Asn-Glycosylation by Oligosaccharyltransferases. *Adv. Exp. Med. Biol.* **1104**, 171–199 (2018).
51. F. Gabler *et al.*, Protein sequence analysis using the MPI bioinformatics toolkit. *Curr. Protoc. Bioinformatics* **72**, e108 (2020).
52. G. Grass *et al.*, Camelysin is a novel surface metalloproteinase from *Bacillus cereus*. *Infect. Immun.* **72**, 219–228 (2004).
53. D. Romero, H. Vlamakis, R. Losick, R. Kolter, Functional analysis of the accessory protein TapA in *Bacillus subtilis* amyloid fiber assembly. *J. Bacteriol.* **196**, 1505–1513 (2014).
54. F. Chu *et al.*, A novel regulatory protein governing biofilm formation in *Bacillus subtilis*. *Mol. Microbiol.* **68**, 1117–1127 (2008).
55. Y. Chai, F. Chu, R. Kolter, R. Losick, Bistability and biofilm formation in *Bacillus subtilis*. *Mol. Microbiol.* **67**, 254–263 (2008).
56. S. S. Branda, F. Chu, D. B. Kearns, R. Losick, R. Kolter, A major protein component of the *Bacillus subtilis* biofilm matrix. *Mol. Microbiol.* **59**, 1229–1238 (2006).
57. L. Friedman, R. Kolter, Genes involved in matrix formation in *Pseudomonas aeruginosa* PA14 biofilms. *Mol. Microbiol.* **51**, 675–690 (2004).
58. A. Driks, Tapping into the biofilm: Insights into assembly and disassembly of a novel amyloid fibre in *Bacillus subtilis*. *Mol. Microbiol.* **80**, 1133–1136 (2011).
59. L. Holm, P. Rosenström, Dali server: Conservation mapping in 3D. *Nucleic Acids Res.* **38**, W545–W549 (2010).
60. J. Böhning *et al.*, Molecular architecture of the TasA biofilm scaffold in *Bacillus subtilis*. *bioRxiv* 10.1101/2022.03.14.484220 (2022).
61. D. H. Parks *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** (D1), D785–D794 (2022).
62. K. Mendlar *et al.*, AnnoTree: Visualization and exploration of a functionally annotated microbial tree of life. *Nucleic Acids Res.* **47**, 4442–4448 (2019).
63. C. Rinke *et al.*, A standardized archaeal taxonomy for the Genome Taxonomy Database. *Nat. Microbiol.* **6**, 946–959 (2021).
64. G. Waksman, S. J. Hultgren, Structural biology of the chaperone-usher pathway of pilus biogenesis. *Nat. Rev. Microbiol.* **7**, 765–774 (2009).
65. R. Kreitner *et al.*, Complete sequential assignment and secondary structure prediction of the canulae forming protein CanA from the hyperthermophilic archaeon *Pyrodicticum abyssi*. *Biomol. NMR Assign.* **14**, 141–146 (2020).
66. M. M. Sauer *et al.*, Catch-bond mechanism of the bacterial adhesin FimH. *Nat. Commun.* **7**, 10738 (2016).
67. C. Berne, A. Ducret, G. G. Hardy, Y. V. Brun, Adhesins involved in attachment to abiotic surfaces by Gram-negative bacteria. *Microbiol. Spectr.* **3**, 10.1128/microbiolspec.MB-0018-2015 (2015).
68. M. Pohlschroder, A. Ghosh, M. Tripepi, S. V. Albers, Archaeal type IV pilus-like structures—Evolutionarily conserved prokaryotic surface organelles. *Curr. Opin. Microbiol.* **14**, 357–363 (2011).
69. C. L. Giltner, Y. Nguyen, L. L. Burrows, Type IV pilin proteins: Versatile molecular modules. *Microbiol. Mol. Biol. Rev.* **76**, 740–772 (2012).
70. S. Q. Zheng *et al.*, MotionCor2: Anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
71. A. Punjani, J. L. Rubinstein, D. J. Fleet, M. A. Brubaker, cryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
72. A. Rohou, N. Grigorieff, CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
73. E. H. Egelman, Reconstruction of helical filaments and tubes. *Methods Enzymol.* **482**, 167–183 (2010).
74. E. H. Egelman, A robust algorithm for the reconstruction of helical filaments using single-particle methods. *Ultramicroscopy* **85**, 225–234 (2000).
75. F. Wang, O. Gnewou, A. Solemanifar, V. P. Conticello, E. H. Egelman, Cryo-EM of helical polymers. *Chem. Rev.* 10.1021/acs.chemrev.1c00753. (2022).
76. E. F. Pettersen *et al.*, UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
77. L. Chang *et al.*, DeepTracer ID: De novo protein identification from Cryo-EM Maps. *Biophys. J.*, in press.
78. J. Pfab, N. M. Phan, D. Si, DeepTracer for fast de novo cryo-EM protein structure modeling and special studies on CoV-related complexes. *Proceedings of the National Academy of Sciences of the United States of America* **118** (2021).
79. P. Emsley, K. Cowtan, Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
80. P. V. Afonine *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).
81. C. J. Williams *et al.*, MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci.* **27**, 293–315 (2018).
82. A. Krogh, B. Larsson, G. von Heijne, E. L. Sonnhammer, Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *J. Mol. Biol.* **305**, 567–580 (2001).
83. R. Zallot, N. Oberg, J. A. Gerlt, The EFI Web resource for genomic enzymology tools: Leveraging protein, genome, and metagenome databases to discover novel enzymes and metabolic pathways. *Biochemistry* **58**, 4169–4182 (2019).
84. J. Obersto, SyntTax: A web server linking synteny to prokaryotic taxonomy. *BMC Bioinformatics* **14**, 4 (2013).
85. F. Wang, V. Cvirkaite-Krupovic, M. Krupovic, E. H. Egelman, 7UEG, Cryo-EM of bundling pili from *Pyrobaculum calidifontis*. Protein Data Bank. <https://www.rcsb.org/structure/7UEG>. Deposited 21 March 2022.
86. F. Wang, V. Cvirkaite-Krupovic, M. Krupovic, E. H. Egelman, EMD-26474. Electron Microscopy Database. <https://www.ebi.ac.uk/emdb/EMD-26474>. Deposited 21 March 2022.