Article

Characterization of heterologously expressed Fibril, a shape and motility determining cytoskeletal protein of the helical bacterium *Spiroplasma*

Shrikant Harne,
Pananghat
Gayathri

gayathri@iiserpune.ac.in

Highlights

Codon-optimized
Spiroplasma citri Fibril
was expressed in *E. coli*

SDS treatment does not
denature Fibril filaments

Fibril possesses α helices
and β sheet and is not an
amyloid-like protein

Domain boundaries were
predicted using mass
spectrometry of
proteolytic fragments

Spiroplasma Fibril

Expression in *E. coli*

Filament bundles
in TEM

FT-IR
Secondary structure

Article

Characterization of heterologously expressed Fibril, a shape and motility determining cytoskeletal protein of the helical bacterium *Spiroplasma*

Shrikant Harne[1] and Pananghat Gayathri[1,2,*]

SUMMARY

Fibril is a constitutive filament-forming cytoskeletal protein of unidentified fold, exclusive to members of genus *Spiroplasma*. It is hypothesized to undergo conformational changes necessary to bring about *Spiroplasma* motility through changes in cell helicity. However, the mechanism driving conformational changes in Fibril remains unknown. We expressed Fibril from *S. citri* in *E. coli* for its purification and characterization. Sodium dodecyl sulfate solubilized Fibril filaments and facilitated purification by affinity chromatography. An alternative protocol for obtaining enriched insoluble Fibril filaments was standardized using density gradient centrifugation. Electron microscopy of Fibril purified by these protocols revealed filament bundles. Probable domain boundaries of Fibril protein were identified based on mass spectrometric analysis of proteolytic fragments. Presence of α-helical and β-sheet signatures in FT-IR measurements suggests that Fibril filaments consist of an assembly of folded globular domains, and not a β-strand-based aggregation like amyloid fibrils.

INTRODUCTION

Spiroplasma are the group of cell-wall-less, helical bacteria that exhibit a characteristic kinking motility. In the absence of external locomotory appendages, *Spiroplasma* motility is thought to be driven by the cytoskeletal ribbon composed of Fibril and the bacterial actin homolog MreB (Kürner et al., 2005; Williamson et al., 1991). Fibril (59,000 Da protein called Fib or Fibril) is a *Spiroplasma*-specific nucleotide-independent filament-forming protein (Townsend et al., 1980; Townsend and Archer, 1983). It is hypothesized to be a bi-domained protein consisting of an N-terminal domain, which shares 21.6% identity with 5′-methylthioadenosine nucleosidase (MTAN; PDB: 1JYS) from *Escherichia coli* (*E. coli*) and a C-terminal domain of unidentified fold (Cohen-Krausz et al., 2011; Sasajima et al., 2021). Fibril forms a homopolymeric assembly that is a component of the cytoskeletal ribbon in *Spiroplasma* (Trachtenberg and Gilad, 2001).

The cytoskeletal ribbon of *Spiroplasma* is constituted of a set of Fibril and MreB filaments that are present on the cytoplasmic side of the cell membrane (Figure 1A i) (Kürner et al., 2005; Trachtenberg et al., 2008). However, the organization of MreBs and Fibril filaments within the cytoskeletal ribbon inside the cell is unclear (Harne et al., 2020b). According to Kürner et al. (2005), Fibril filaments flank the MreB filaments (Figure 1A ii) while Trachtenberg et al. (2008) propose that the MreB filaments are sandwiched between Fibril filaments and membrane (Figure 1A iii). Thus, the physiologically relevant arrangement of cytoskeletal filaments within the ribbon is currently unknown. The ribbon has been proposed to confer helical shape and to function as a linear motor contributing to motility of spiroplasmas (Harne et al., 2020a; Kürner et al., 2005; Razin, 1978; Shaevitz et al., 2005; Trachtenberg et al., 2008; Trachtenberg and Gilad, 2001). However, the detailed molecular architecture of Fibril and the cytoskeletal ribbon remains to be determined.

According to the current model for *Spiroplasma* motility, Fibril undergoes conformational changes to bring about length variations in the Fibril filaments (Cohen-Krausz et al., 2011; Razin, 1978; Trachtenberg and Gilad, 2001). Co-ordinated cycles of alternate contraction and extension of the Fibril filaments can cause change in handedness of the cell body that are reflected as kinks. The switching of helicity pushes surrounding liquid and propels the cell forward (Kürner et al., 2005; Shaevitz et al., 2005; Trachtenberg and Gilad, 2001). Indeed, low-resolution electron microscopy data suggest the possibility of different conformations of Fibril (Cohen-Krausz et al., 2011; Liu et al., 2017) but the molecular mechanism of kink propagation based

[1]Indian Institute of Science Education and Research (IISER) Pune, Dr. Homi Bhabha Road, Pashan, Pune 411008, India

[2]Lead contact

*Correspondence: gayathri@iiserpune.ac.in

https://doi.org/10.1016/j.isci.2022.105055
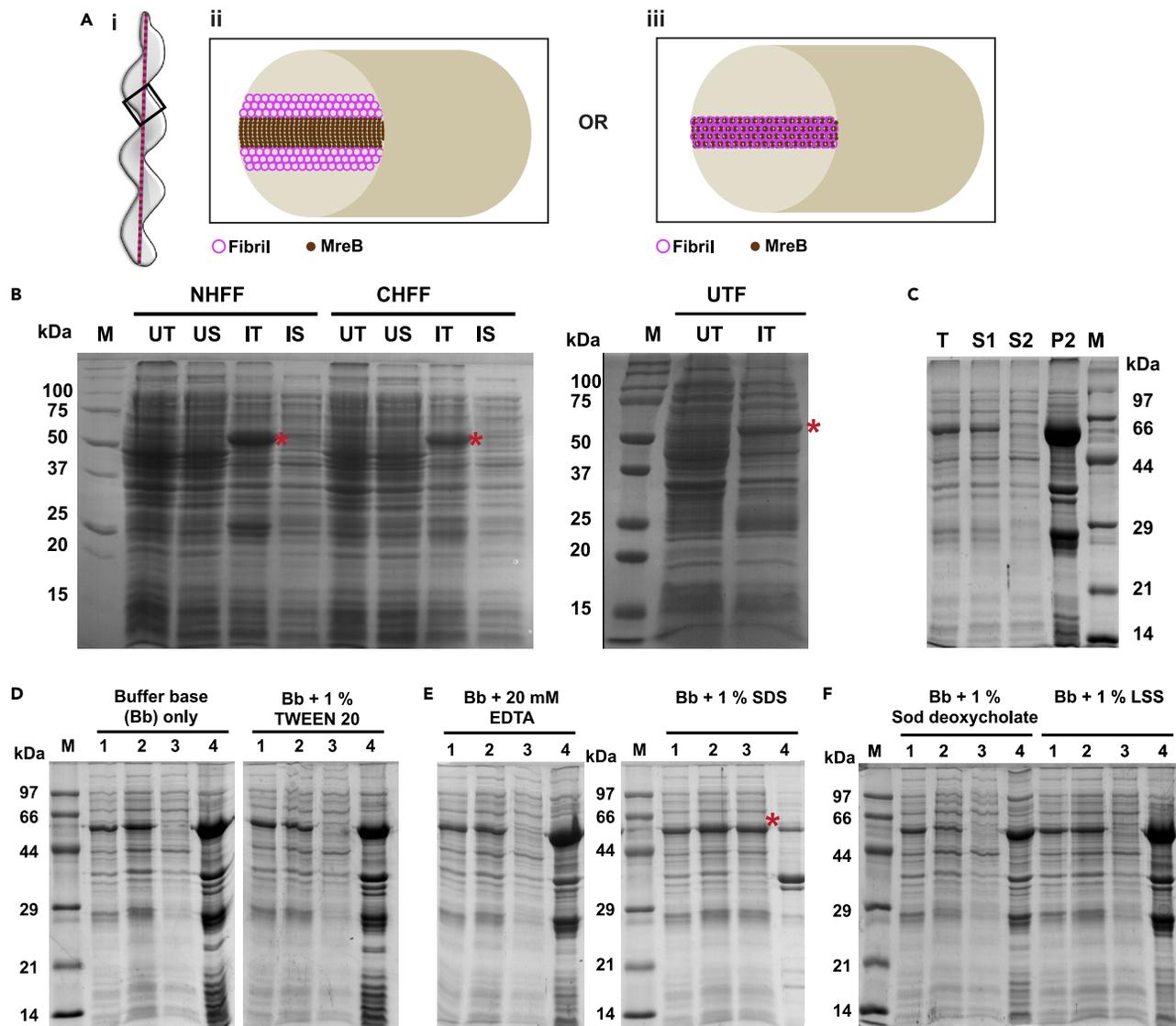
Figure 1. Fibril: cytoskeletal organization, expression, and solubility check

(A) Pictorial representation of intracellular localization of cytoskeletal ribbon of *Spiroplasma* (i). Longitudinal section (black box in i) of *Spiroplasma* cell showing models for Fibril and MreB organization as proposed by Kürner et al. (2005) (ii) and Trachtenberg et al. (2008) (iii).

(B) Expression of desired recombinant Fibril construct in *E. coli* BL21 (AI) cells transformed with the construct, induced with L-arabinose and further incubation revealed the appearance of an approximately 59,000 Da band (see lanes labeled `IT´; band shown with a red asterisk) on a 12% SDS-PAGE gel. Corresponding band was not observed in *E. coli* BL21 (AI) cells transformed with the construct but were not induced (lanes labeled `UT´ and `US´). M— known molecular weight protein markers, NHFF [N - (His$_6$)-Fibril; N-His Full-length Fibril], CHFF [C - Fibril-(His$_6$); C-His Full-length Fibril], UTF – untagged Fibril, UT and US respectively represent total lysate and supernatant of uninduced *E. coli* BL21 (AI) cells transformed with the desired construct obtained after spinning the lysate at 21,000 xg, IT (induced total lysate), IS (induced supernatant; soluble proteins obtained by spinning `IT´ at 21,000 xg).

(C) Initial solubility check revealed that the Fibril from lysate (lane T) remains in supernatant (lane S1) upon spinning at 4,629 xg. Further spinning the supernatant (lane S1) at 159,000 xg results in appearance of most of the Fibril in the pellet (lane P2) as compared to supernatant (lane S2). Attempts to re-suspend the pellet (lane P2) resulted in turbid solution, an indication of the insoluble nature of Fibril. Lanes represent T – total lysate, S1 – supernatant upon spinning the lysate at 4,629 xg, S2 and P2 – supernatant and pellet, respectively, obtained by spinning S1 at 159,000 xg, M— protein marker with known molecular weights.

(D–F) Solubility check of Fibril using different additives, namely, Triton X-100 only (panel D left side) or Triton X-100 in combination with TWEEN 20 (panel D right side); ethylenediaminetetraacetic acid (EDTA; panel E left side) and sodium dodecyl sulfate (SDS; panel E right side), sodium deoxycholate (Sod deoxycholate; panel F left side) and N-lauroylsarcosine sodium salt (LSS; panel F right side). The additives were added to lysate (Buffer base, Bb; containing Triton X-100) and allowed to act for 30 minutes on ice. The lysate was then clarified by spinning at 4,629 xg and supernatant from this stage was further spun at

Figure 1. *Continued*
159,000 xg and supernatant and pellet were visualized on SDS-PAGE gels to check for Fibril solubilization. Lanes represent protein marker with known molecular weights (M), total lysate (1), supernatant after spinning the lysate at 4,629 xg (2), supernatant (3) and pellet (4) after 159,000 xg spin of the clarified lysate. Addition of SDS (1% w/v) helped obtain soluble Fibril upon 159,000 xg spin (Fibril band obtained in supernatant after treatment with SDS followed by 159,000 xg spin is highlighted with a red asterisk).

on conformational changes of a constitutive filament such as the Fibril filament is unknown. Recent studies involving expression of *Spiroplasma* MreBs and Fibril in *Mycoplasma capricolum* or a synthetic bacterium revealed that Fibril is not essential for conferring motility or helicity (Kiyama et al., 2021; Lartigue et al., 2021). It is hypothesized that Fibril supports propagation of kinks necessary for *Spiroplasma* movement by potentially transmitting the force generated by itself and MreBs to the membrane (Kiyama et al., 2021; Lartigue et al., 2021). Force generation could arise due to the ATP hydrolysis activity of MreBs (Pande et al., 2022; Sasajima and Miyata, 2021). Further studies are required to understand whether the conformational changes in Fibril, if any, and its interaction with MreBs (Harne et al., 2020a) are necessary for generating or relaying the force.

Fibril has been proposed to be a key player in the *Spiroplasma* shape determination and motility since a very long time (Razin, 1978). However, the dearth of efficient tools for facile genetic modification of *Spiroplasma* and the lack of advanced imaging techniques have limited mechanistic insights into Fibril function (Harne et al., 2020b). Recent advances in electron cryomicroscopy allow us to obtain three-dimensional reconstruction of proteins and their assemblies at atomic resolution (Fitzpatrick et al., 2017; Merino et al., 2018). High-resolution structures of Fibril may facilitate identification of conformational changes proposed to be necessary for bringing about kinking motility of *Spiroplasma* (Sasajima et al., 2021; Shaevitz et al., 2005). Structural information may also provide insights about evolution of Fibril and the uncharacterized mode of motility driven by cytoskeletal filaments.

To overcome the challenges of genetic manipulation in *Spiroplasma* and gain insights into structure of Fibril, we heterologously expressed Fibril in *E. coli*, and standardized protocols for its purification. We demonstrate that Fibril expressed in *E. coli* forms filamentous assemblies that are very stable and can withstand sodium dodecyl sulfate (SDS) treatment. Solubilization of expressed Fibril using SDS enabled purification of a hexahistidine-tagged construct using affinity chromatography. Secondary structure content remained unaffected by SDS treatment, based on characteristic signatures observed using FT-IR (Fourier Transform Infrared) spectra. Alternatively, Fibril filaments were also purified using density gradient centrifugation without the use of SDS. Visualization of Fibril obtained by these protocols using electron microscope indeed confirmed the presence of filaments, despite treatment by SDS. Mass spectrometric characterization of proteolytic fragments confirmed probable domain boundaries for the folded Fibril monomers.

RESULTS

Fibril expressed in *E. coli* forms insoluble filamentous assemblies

For heterologous expression of *Spiroplasma* Fibril, the TGA codons in *fibril* gene were modified to TGG to enable the expression of full-length Fibril in *E. coli*. This modification was necessary since TGA is a codon for tryptophan in *Spiroplasma* while it functions as a stop codon in *E. coli* (Stamburski et al., 1992). The expression of Fibril in *E. coli* was confirmed by the appearance of an additional protein band with a molecular weight of about 60,000 kilodalton (kDa) on an SDS-PAGE gel (Figure 1B) in the cells transformed with vector and induced with arabinose. In contrast, protein band of corresponding molecular weight was not seen in the cells transformed with the vector but were not induced (Figure 1B). This confirmed the heterologous expression of recombinant Fibril in *E. coli*.

Having overexpressed Fibril constructs in *E. coli*, we checked the Fibril solubility. Fibril was found in the pellet fraction upon 159,000 xg spin of the cell lysate, suggestive of the presence of filament assembly of Fibril (Figure 1C). Re-suspension of Fibril from the 159,000 xg pellet in lysis buffer resulted in a turbid solution, indicating that it was insoluble. Thus, to obtain Fibril in soluble form for its further purification, re-suspension and solubilization was attempted using different detergents (Tween 20, LSS, sodium deoxycholate, SDS, and Triton X-100) and EDTA. EDTA was selected to depolymerize any nucleotide-dependent filament-forming proteins such as MreB, if relevant, and prevent them from pelleting down along with Fibril

A Lysate [1]

↓ Spin @ 4,629 xg
/4 °C/ 15 mins

Supernatant [2]

↓ Spin @ 159,000 xg
/4 °C/ 30 mins (Supernatant [3])

Pellet [4]

↓ Resuspend in
buffer and add SDS

Clear solution

↓ Spin @ 21,000 xg
/25 °C/ 10 mins

Supernatant [5]

↓ Spin @ 159,000 xg
/25 °C/ 30 mins

Soluble Fibril [6]

↓ Affinity
chromatography
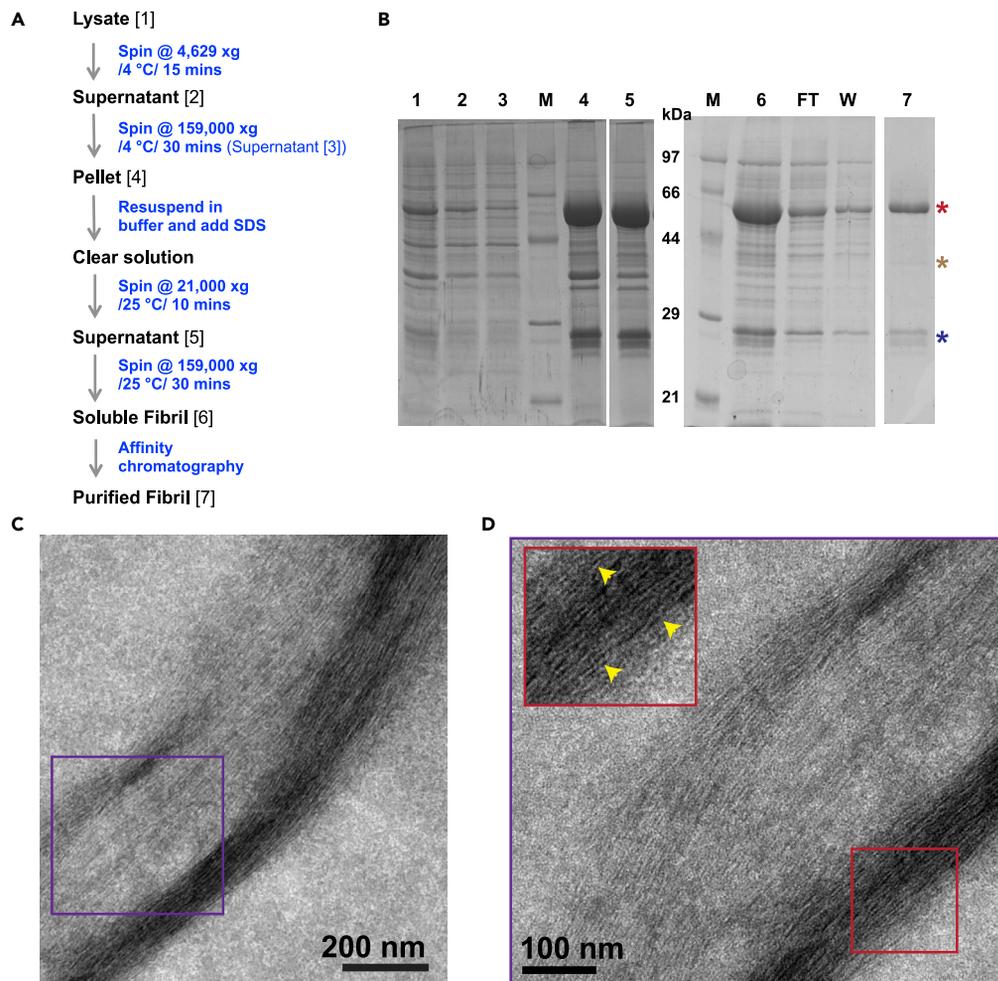
Purified Fibril [7]

Figure 2. Purification of hexahistidine-tagged Fibril by SDS treatment and affinity chromatography

(A) Schematic showing protocol for purification of hexahistidine-tagged Fibril by solubilization with SDS followed by affinity chromatography. Numbers in square bracket correspond to the lanes in SDS-PAGE shown in panel B.

(B) 12 % SDS-PAGE gel showing purity of Fibril at different stages of purification. Lanes 1–7 represent the samples at various stages as marked within square brackets in panel A. During purification steps, Fibril is obtained in pellet/insoluble form (lane 4) until it is treated with SDS. Upon SDS treatment, Fibril is solubilized and does not pellet even upon spinning at 159,000 xg (lane 6). The asterisks point to proteins of interest; red – Fibril (59 kDa), tan and blue – unknown proteins (∼36 kDa and ∼ 26 kDa) associated with Fibril. Lane representation is as follows- M— Marker, total lysate (1), clarified lysate/supernatant after spinning the lysate at 4,629 xg (2), supernatant (3) and pellet (4) after 159,000 xg spin of the clarified lysate, supernatant obtained by resuspension of pellet 159,000 xg pellet using SDS and followed by spin at 21,000 xg (5), supernatant obtained by spinning soluble protein (supernatant) from 21,000 xg spin further at 159,000 xg (6), flowthrough/ unbound protein (FT), wash/loosely bound protein eluted by passing wash buffer (W), purified protein from Ni-NTA matrix using buffer containing imidazole (7).

(C and D) Visualization of purified, SDS-treated Fibril using negative staining transmission electron microscopy revealed filament bundles. Region in the purple box in panel C is shown in panel D. Area in the red box is magnified and shown in inset. Yellow arrows point to the "donut-shaped" structures commonly observed in *Spiroplasma* Fibril.

when spun at 159,000 xg. Out of the various additives attempted, only the addition of SDS at a final concentration of 1% (w/v) was found to help in obtaining Fibril in solution (Figures 1D–1F).

Fibril filaments withstand 1% SDS treatment

The small amount of Fibril in the lysate (without addition of SDS) did not bind to the Ni-NTA matrix, probably because of occlusion of the filament bundles from the matrix or non-exposure of the hexa-histidine tag. However, the SDS-solubilized $(His)_6$-Fibril bound to Ni-NTA affinity matrix and could be further purified

by affinity chromatography (Figures 2A and 2B). Despite the purification step, additional bands of molecular weight of about 26 kDa were consistently observed along with Fibril protein band at ∼ 59 kDa (Figure 2B, lane 7 blue asterisk).

Observation of Fibril assemblies using transmission electron microscopy (TEM) revealed that these bundles (Figures 2C and 2D) were indeed filament assemblies. These filaments formed twisted bundles of filaments (Figures 2C and 2D) and appeared to be wider than the (∼55 nm wide) filament assemblies observed *in vivo* in *Spiroplasma* (Kürner et al., 2005). It is interesting to note that Fibril retained the filamentous structure despite treatment with SDS. The beady appearance of the filaments in the micrographs resembled the characteristic features observed for the higher resolution TEM characterization of Fibrils isolated from *Spiroplasma* (Cohen-Krausz et al., 2011).

Fibril filament purified using density gradient ultracentrifugation was present as bundles

To confirm that filament integrity was not affected by SDS treatment, an alternative protocol was standardized for purification of Fibril, using density gradient ultracentrifugation. We used urografin density gradients (with a density range between 1.2 gm/cm^3 and 1.35 gm/cm^3) to purify Fibril since the density of the former (1.52 g/cm^3) is higher than Fibril (1.22 g/cm^3; Townsend et al., 1980). Consequently, we were able to separate Fibril from other proteins based on their density (Figures 3A and 3B). Resuspension of the Fibril pellet obtained after 159,000 xg ultracentrifugation followed by separation on the urografin density gradient resulted in two bands (Figure 3B). Visualization of the protein in the two bands on SDS-PAGE gel after removal of urografin by dilution and pelleting revealed that both the bands contained enriched Fibril (marked with red asterisk in Figure 3C, lanes 5 and 6). Careful observation of both the bands on the SDS-PAGE gel revealed that additional proteins (marked with tan and blue asterisks in Figure 3C, lanes 5 and 6) accompanied Fibril. The most prominent of these accompanying proteins were those corresponding to molecular weights of ∼36 and ∼26 kDa (Figure 3C lanes 5, 6). The 26 kDa protein migrated similar to that observed during purification of His$_6$-tagged Fibril using SDS and affinity chromatography (Figure 2B).

The lower band obtained on the urografin density gradient appeared to be purer than the upper band (Figure 3C, lanes 5 and 6, respectively). Thus, the lower band was used for visualization using electron microscopy. Electron micrographs (Figures 3D–3F) of the protein from lower band (Figure 3C, lane 6) revealed filament bundles similar to those obtained with SDS treatment. Fibril purified using the density gradient also showed twisted filament assemblies of varying widths. Regions of Fibril filaments with good staining showed the "donut shaped" structures (highlighted with yellow arrowheads) and occasionally, thin filaments (Figure 3F) were also seen on the grid.

Fibril filaments exhibited secondary structure features of α-helices and β-sheets

Next, we investigated if the SDS treatment affected Fibril filaments by comparing secondary structures of Fibril purified with (Figure 2B, lane 7) and without use of SDS (Figure 3C, lane 6) using FT-IR spectroscopy. The FT-IR spectra for Fibril purified with or without SDS (using urografin density gradient) showed similar profiles (Figure 4). All the three protein constructs showed the presence of peaks corresponding to α-helices (1655 cm^{-1}), β-turns (1685 cm^{-1}), and β-sheets (1636 cm^{-1}) (Adochitei and Drochioiu, 2011; Kim et al., 1994). This observation suggested the presence of these secondary structures in Fibril filaments purified using the different protocols. Based on FT-IR data, we concluded that Fibril is not an amyloid-like aggregate of protein with only β-strands, but has specific secondary structures, such as α-helices and β-sheets.

Proteolytic fragments of Fibril identify its approximate domain boundaries

We hypothesized that ∼ 36 and 26 kDa proteins were either breakdown products of Fibril or contaminant proteins having affinity to the Ni-NTA matrix or Fibril. Hence, we performed mass spectrometry study to find the identity of these proteins. Peptide fingerprinting analysis of the ∼36 and ∼26 kDa proteins accompanying enriched Fibril purified by density gradient and ∼26 kDa proteins associated with purified His-tagged Fibril by in-gel trypsin digestion revealed that these are indeed breakdown products of full-length Fibril (Figure 5A and Table S1). Analysis of Fiibril protein sequence revealed that its proteolysis in the residue stretch extending from 220 to 312 can result in complementary fragments with molecular weights of ∼26 and 36 kDa (Figure 5B). A proteolytic cleavage at either of the sites gave rise to bands of sizes 26 or 36 kDa from either ends of the protein. The proximal residues at 220 and 312 could be loop regions exposed to the surface of the filament, while the middle portion consisting of residues around 220 to
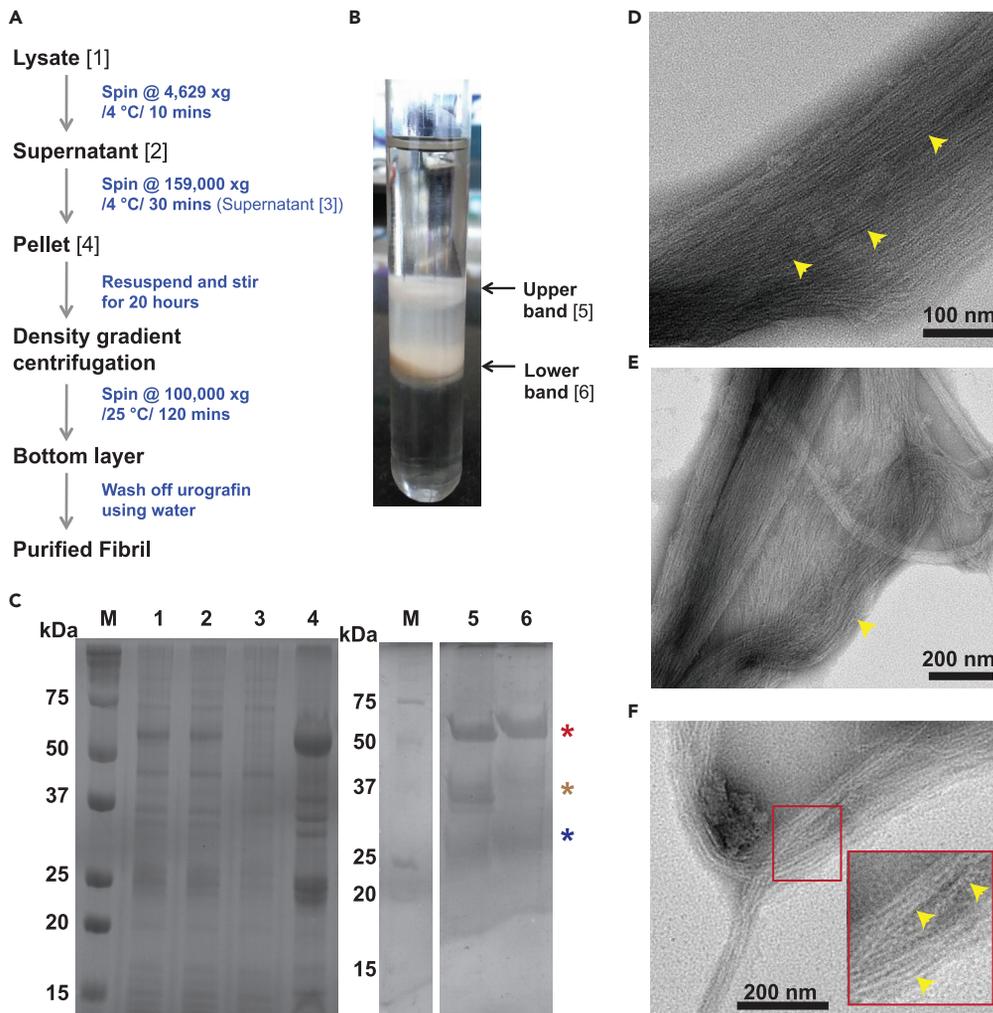
Figure 3. Purification of untagged Fibril using urografin density gradient

(A) Schematic of the protocol for purification of Fibril using urografin density gradient. The lane numbers in the gels for other Figure 3 panels are shown within square brackets.

(B) Image of Fibril filaments separated into 2 layers on a urografin gradient. The two layers {upper [5] and lower [6]} are labeled.

(C) 12 % SDS-PAGE gel showing purity of Fibril at different stages of purification. Visualization of fractions from urografin gradient revealed presence of Fibril in two bands [lanes 5 (upper band), 6 (lower band)] with the latter being purer than former. Fibril and associated prominent proteins are marked with asterisks; red – Fibril (59 kDa), tan and blue – Fibril-associated unknown proteins (∼36 and ∼26 kDa). The lanes of the gel are smeary due to the contamination by urografin despite multiple washes. Lanes represent samples at the various stages of the protocol, as labeled within square brackets in panels A and B. M— Marker, total lysate (1), clarified lysate/supernatant after spinning the lysate at 4,629 xg (2), supernatant (3) and pellet (4) after 159,000 xg spin of the clarified lysate, upper (5) and lower (6) bands formed by separation of 159,000 xg pellet on urografin density gradient.

(D–F) The lower band containing Fibril was visualized using negative staining transmission electron microscopy (TEM) and Fibril bundles were observed. Individual thin filaments are clearly visible in the images. Region within the red box in panel F is zoomed and shown in inset. Yellow arrows point to "donut-shaped" structures observed in *Spiroplasma* Fibril.

312 could be a linker region connecting the N-terminal and C-terminal domains of Fibril (Figure 5B). This suggested that the peptides consisting of residues 1–220 could constitute the amino-terminal domain while residues 312–512 made up the carboxy-terminal domain (Figure 5B). The existence of a smeared band (for example, Figure 2B, lane 7 at approximately 26 kDa), rather than a single band, suggested the possibility of multiple exposed protease cleavage sites at this stretch of residues, as represented schematically in Figure 5B.
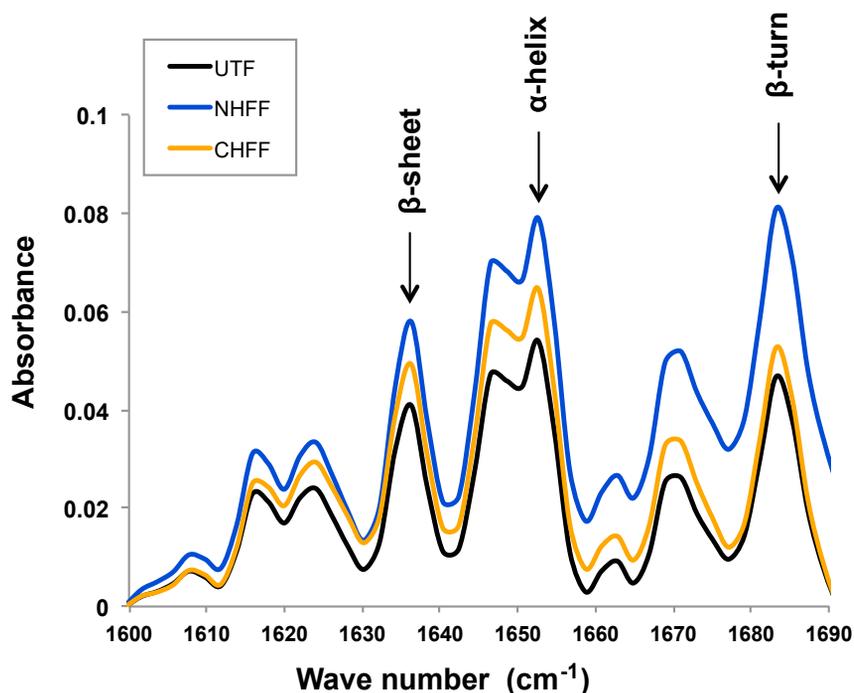
Figure 4. Secondary structure of Fibril remains unaffected by SDS treatment

FT-IR analysis of untagged Fibril (UTF) purified using urografin density gradient, His_6-Fibril (N-His_6 Fibril; NHFF) and Fibril-His_6 (C-His_6 Fibril; CHFF) purified by SDS treatment followed by affinity chromatography show similar peak pattern. All the three protein constructs display peaks around 1655, 1636, and 1685 cm^{-1} corresponding to α-helices, β-sheets and β-turn secondary structures, respectively.

DISCUSSION

The challenges associated with purification of constitutive filament-forming proteins have limited our understanding about their structure and function. Fibril is one such nucleotide-independent polymerizing protein classified as a cytoskeletal element (Townsend et al., 1980; Williamson, 1974). We successfully solubilized Fibril filaments using SDS and purified using affinity chromatography. We could also purify Fibril filaments using density gradient ultracentrifugation-based methods, a technique widely used for the separation of macromolecular assemblies such as protein polymers, viruses, and lipoproteins from other biomolecules (Chung et al., 1980; Mamada et al., 2017; Redgrave et al., 1975; Reimer et al., 1967; Samuel et al., 2017; Soellner et al., 1985; Yin et al., 2018).

Our attempts to purify Fibril by incubating the lysate with Ni-NTA resins revealed that Fibril did not bind to the Ni-NTA matrix. This is an indication that the hexahistidine tag is probably not exposed because the amino and carboxy termini of Fibril are buried either within the Fibril molecule or between the adjacent Fibril molecules in the filament bundles. Another possibility of Fibril not binding to the Ni-NTA matrix is due to the occlusion of filament bundles of Fibril from entering the affinity matrix. Addition of SDS solubilized the filaments as observed from the loss of turbidity, probably by opening up the bundles of filaments. This enabled the tag to bind the column, despite the filament form being retained.

Thick bundles of Fibril were obtained in the heterologous expression system. This difference in the width is potentially because of the surface properties of the Fibril that causes the filaments to associate with each other *in vitro* and form wider bundles. These possibilities are supported by previous observations of wider filament assemblies for Fibril purified from *Spiroplasma citri* or *Spiroplasma melliferum* (Trachtenberg et al., 2003). It is probable that the width of Fibril *in vitro* in a heterologous overexpression system or upon purification from the native organism cannot be controlled, in the absence of membrane/lipids and other Fibril assembly regulators that may be present in an *in vivo* scenario within *Spiroplasma*. Nevertheless, the purified Fibril filaments showed characteristic "donut-shaped" structures (Figures 2D and 3F
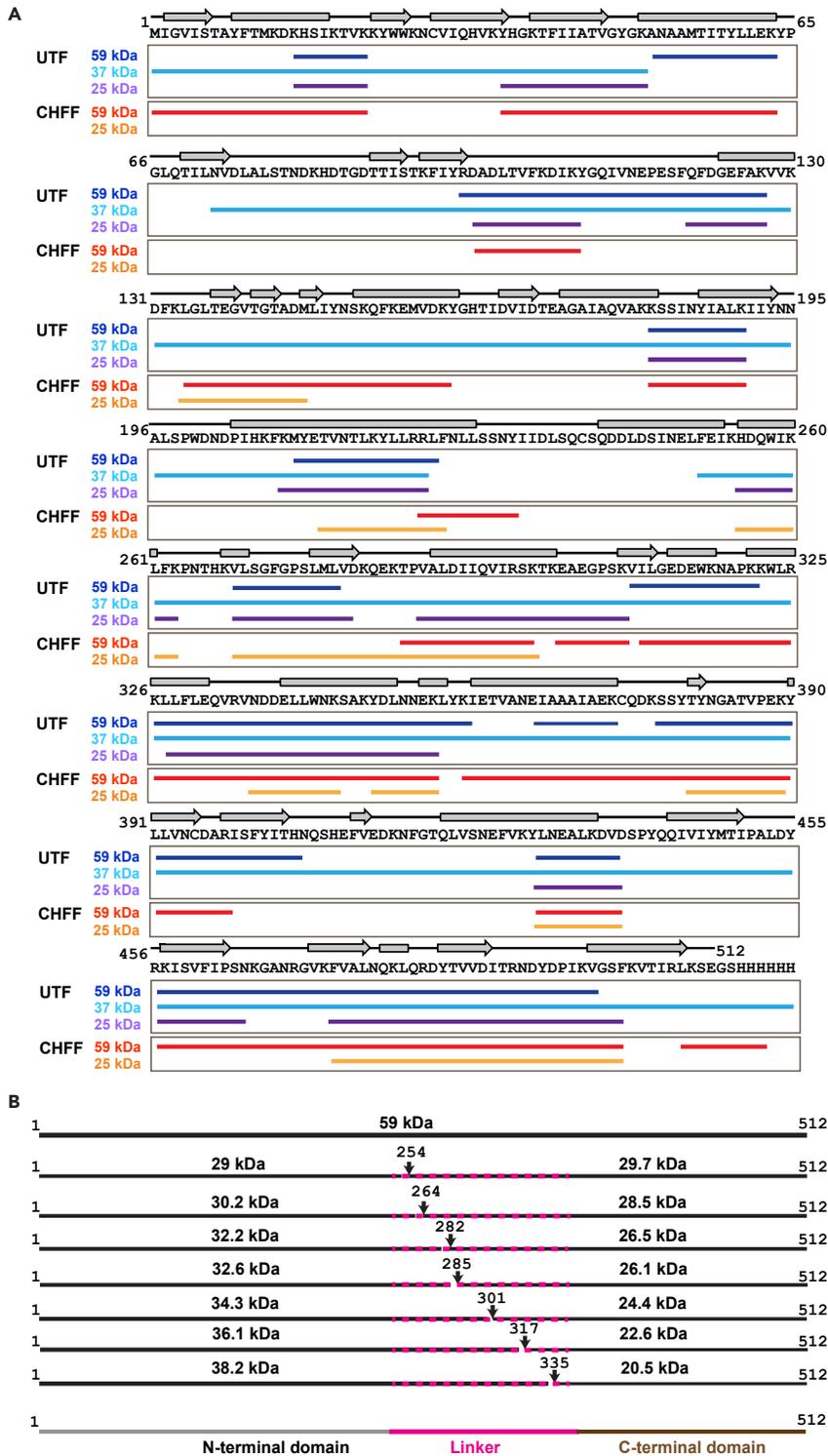
Figure 5. Fibril is proteolyzed in the cell

(A) Mass spectrometry analysis of trypsin-digested peptides of about 59, 36, and 26 kDa bands observed during the purification of untagged Fibril (UTF) and Fibril-His$_6$ (C-His$_6$ Fibril; CHFF). The analysis revealed that the 36,000 Da and 26,000 Da bands are indeed breakdown products of Fibril. Amino acid sequences of individual peptides detected by mass spectrometer mapped to the Fibril sequence are represented by lines below the corresponding sequence while secondary structures (gray arrows - strands, gray sheets - helices) as predicted using PSIPRED

Figure 5. *Continued*

(http://bioinf.cs.ucl.ac.uk/psipred/) are shown above the sequence and residue numbers are shown in superscript. See also Table S1 for the full list of annotated peptides.

(B) Schematic representation of the proposed domain boundaries and linker connecting the two domains in Fibril. The pink stripes represent the potentially exposed trypsin protease target sites and black arrows indicate the amino acid numbers at the indicated site. Masses in kilodalton (kDa) of the amino and carboxy-terminal fragment if the trypsin cleaved Fibril at the site indicated by black arrow position is indicated on the corresponding schematic. Values at the ends of lines indicate amino acid numbers.

insets) as have been reported for Fibril filaments purified from *Spiroplasma* (Cohen-Krausz et al., 2011; Trachtenberg and Gilad, 2001).

The observation of lower molecular weight bands, with peptide composition corresponding to that of Fibril, indicated that Fibril got proteolyzed during heterologous expression in *E. coli* or the purification process. However, the fragments were retained with the Fibril polymers most likely because they form an integral core of the Fibril filament, and proteolysis potentially happens at the exposed loops in the filament. Residues between the N and C-terminal domains must be forming a linker domain, the boundaries of which are exposed toward the surface of the filament and hence accessible to protease digestion. This information is useful to design short constructs of Fibril to identify its soluble domains, perform structural characterization, and study their polymerization properties.

We have visualized the Fibril purified from *E. coli* using TEM and compared its secondary structure composition, confirming the presence of folded protein. The lack of high-resolution TEM images because of challenges associated with bundling has prevented us from obtaining repeat distances and other helical parameters of the filament and compare with the values reported from *Spiroplasma*. Using the heterologous protein expression system *E. coli*, further attempts for separating out the filament bundles will enable us to obtain high-resolution information using TEM.

The genetically refractive nature of *Spiroplasma* has prevented researchers from dissecting Fibril to identify its domain boundaries and polymerization interface. In such a situation, use of a genetically facile organism becomes obligatory. Recently, the interactions between *Spiroplasma* MreB isoforms were investigated by expressing *Spiroplasma poulsonii* MreBs with a fluorescent tag in *E. coli* (Masson et al., 2021). The successful expression and characterization of Fibril filaments based on expression from a heterologous system opens up avenues for various functional assays, which include testing interaction of Fibril with other *Spiroplasma* proteins such as MreB.

Limitations of the study

Our study reports the protocols for purification of Fibril using a heterologous expression system, *E. coli*. Although we have visualized the Fibril purified from *E. coli* using TEM, the lack of its high-resolution structure has prevented us from comparing structures of Fibril purified with and without use of SDS as well as to the one from *Spiroplasma*. In absence of the structural information, the Fibril domains predicted by us remain to be verified.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Cultures
- METHOD DETAILS
  - Cloning
  - Expression and cell lysis
  - Standardization for solubility of fibril
  - Purification of fibril expressed in *E. coli* by Ni-NTA affinity chromatography
  - Purification of insoluble fibril using density gradient

- Purification of fibril from *E. coli* by separation on urografin density gradient
- Visualization of Fibril using transmission electron microscopy (TEM)
- Fourier-transform infrared (FT-IR) spectroscopy
- In-gel trypsin digestion and mass spectroscopy of the protein

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2022.105055.

AUTHOR CONTRIBUTIONS

S.H. designed, conceptualized, performed all experiments, carried out data and sequence analyses, and wrote the manuscript; P.G. performed experiments for cloning of the *fibril* gene, designed, conceptualized, and supervised and wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

Adochitei, A., and Drochioiu, G. (2011). Rapid characterization of peptide secondary structure by FT-IR spectroscopy. Rev. Roum. Chim. 56, 783–791.

Chung, B.H., Wilkinson, T., Geer, J.C., and Segrest, J.P. (1980). Preparative and quantitative isolation of plasma lipoproteins : rapid , single discontinuous density gradient ultracentrifugation in a vertical rotor. J. Lipid Res. 21, 284–291.

Cohen-Krausz, S., Cabahug, P.C., and Trachtenberg, S. (2011). The monomeric, tetrameric, and fibrillar organization of Fib: the dynamic building block of the bacterial linear motor of *Spiroplasma melliferum* BC3. J. Mol. Biol. 410, 194–213. https://doi.org/10.1016/j.jmb.2011.04.067.

Fitzpatrick, A.W.P., Falcon, B., He, S., Murzin, A.G., Murshudov, G., Garringer, H.J., Crowther, R.A., Ghetti, B., Goedert, M., and Scheres, S.H.W. (2017). Cryo-EM structures of tau filaments from Alzheimer's disease. Nature 547, 185–190. https://doi.org/10.1038/nature23002.

Harne, S., Duret, S., Pande, V., Bapat, M., Béven, L., and Gayathri, P. (2020a). MreB5 is a

determinant of rod-to-helical transition in the cell-wall-less bacterium *Spiroplasma*. Curr. Biol. 30, 4753–4762.e7. https://doi.org/10.1016/j.cub.2020.08.093.

Harne, S., Gayathri, P., and Béven, L. (2020b). Exploring *Spiroplasma* biology: opportunities and challenges. Front. Microbiol. 11, 589279–589288. https://doi.org/10.3389/fmicb.2020.589279.

Jagtap, P., Bandhakavi, S., Higgins, L., Mcgowan, T., Sa, R., Stone, M.D., Chilton, J., Arriaga, E.A., Seymour, S.L., and Griffin, T.J. (2012). Workflow for analysis of high mass accuracy salivary data set using MaxQuant and ProteinPilot search algorithm. Proteomics 12, 1726–1730. https://doi.org/10.1002/pmic.201100097.

Kim, Y., Rose, C.A., Liu, Y., Ozaki, Y., Datta, G., and Tu, A.T. (1994). FT-IR and near-infared FT-Raman studies of the secondary structure of insulinotropin in the solid state: α-helix to β-sheet conversion induced by phenol and/or by high shear force. J. Pharm. Sci. 83, 1175–1180. https://doi.org/10.1002/jps.2600830819.

Kiyama, H., Kakizawa, S., Sasajima, Y., Tahara, Y.O., and Miyata, M. (2021). Reconstitution of

Spiroplasma swimming by expressing two bacterial actins in synthetic minimal bacterium. Preprint at bioRxiv 1–13. https://doi.org/10.1101/2021.11.16.468548.

Kunzelmann, S., and Webb, M.R. (2009). A biosensor for fluorescent determination of ADP with high time resolution. J. Biol. Chem. 284, 33130–33138. https://doi.org/10.1074/jbc.M109.047118.

Kürner, J., Frangakis, A.S., and Baumeister, W. (2005). Cryo – electron tomography reveals the cytoskeletal structure of *Spiroplasma melliferum*. Science 307, 436–438. https://doi.org/10.1126/science.1104031.

Lartigue, C., Lambert, B., Rideau, F., Decossas, M., Hillion, M., Douliez, J.-P., Hardouin, J., Lambert, O., Blanchard, A., and Béven, L. (2021). Tuning spherical cells into kinking helices in wall-less bacteria. Preprint at bioRxiv 16, 467908. https://doi.org/10.1101/2021.11.16.467908.

Liu, P., Zheng, H., Meng, Q., Terahara, N., Gu, W., Wang, S., Zhao, G., Nakane, D., Wang, W., and Miyata, M. (2017). Chemotaxis without conventional two-component system, based on cell polarity and aerobic conditions in

helicity-switching swimming of *Spiroplasma eriocheiris*. Front. Microbiol. *8*, 58. https://doi.org/10.3389/fmicb.2017.00058.

Mamada, N., Tanokashira, D., Ishii, K., Tamaoka, A., and Araki, W. (2017). Mitochondria are devoid of amyloid β-protein (Aβ)-producing secretases: evidence for unlikely occurrence within mitochondria of Aβ generation from amyloid precursor protein. Biochem. Biophys. Res. Commun. *486*, 321–328. https://doi.org/10.1016/j.bbrc.2017.03.035.

Masson, F., Pierrat, X., Lemaitre, B., and Persat, A. (2021). The wall-less bacterium *Spiroplasma poulsonii* builds a polymeric cytoskeleton composed of interacting MreB isoforms. iScience *24*, 103458. https://doi.org/10.1016/j.isci.2021.103458.

McGuffin, L.J., Bryson, K., and Jones, D.T. (2000). The PSIPRED protein structure prediction server. Bioinformatics *16*, 404–405. https://doi.org/10.1093/bioinformatics/16.4.404.

Merino, F., Pospich, S., Funk, J., Wagner, T., Küllmer, F., Arndt, H.D., Bieling, P., and Raunser, S. (2018). Structural transitions of F-actin upon ATP hydrolysis at near-atomic resolution revealed by cryo-EM. Nat. Struct. Mol. Biol. *25*, 528–537. https://doi.org/10.1038/s41594-018-0074-0.

Pande, V., Mitra, N., Bagde, S.R., Srinivasan, R., and Gayathri, P. (2022). Filament organization of the bacterial actin MreB is dependent on the nucleotide state. J. Cell Biol. *221*, e202106092.

Rappsilber, J., Mann, M., and Ishihama, Y. (2007). Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nat. Protoc. *2*, 1896–1906. https://doi.org/10.1038/nprot.2007.261.

Razin, S. (1978). The mycoplasmas. Microbiol. Rev. *42*, 414–470.

Redgrave, T.G., Roberts, D.C., and West, C.E. (1975). Separation of plasma lipoproteins by density-gradient untracentrifugation. Anal. Biochem. *65*, 42–49.

Reimer, C.B., Baker, R.S., VanFrank, R.M., Newlin, T.E., Cline, G.B., and Anderson, N.G. (1967). Purification of large quantities of influenza virus by density gradient centrifugation. J. Virol. *1*, 1207–1216.

Samuel, M., Chisanga, D., Liem, M., Keerthikumar, S., Anand, S., Ang, C.S., Adda, C.G., Versteegen, E., Jois, M., and Mathivanan, S. (2017). Bovine milk-derived exosomes from colostrum are enriched with proteins implicated in immune response and growth. Sci. Rep. *7*, 5933–6010. https://doi.org/10.1038/s41598-017-06288-8.

Sasajima, Y., Kato, T., Miyata, T., Namba, K., and Miyata, M. (2021). Elucidation of fibril structure responsible for swimming in Spiroplasma using electron microscopy. Preprint at bioRxiv. https://doi.org/10.1101/2021.02.24.432793.

Sasajima, Y., and Miyata, M. (2021). Prospects for the mechanism of *Spiroplasma* swimming. Front. Microbiol. *12*, 706426. https://doi.org/10.3389/fmicb.2021.706426.

Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). Image to ImageJ: 25 years of image analysis. Nat. Methods *9*, 671–675. https://doi.org/10.1038/nmeth.2089.

Shaevitz, J.W., Lee, J.Y., and Fletcher, D.A. (2005). *Spiroplasma* swim by a processive change in body helicity. Cell *122*, 941–945. https://doi.org/10.1016/j.cell.2005.07.004.

Shilov, I.V., Seymour, S.L., Patel, A.A., Loboda, A., Tang, W.H., Keating, S.P., Hunter, C.L., Nuwaysir, L.M., and Schaeffer, D.A. (2007). The paragon algorithm, a next generation search engine that uses sequence temperature values sequence temperature values and feature probabilities to identify peptides from tandem mass spectra. Mol. Cell. Proteomics *6*, 1638–1655. https://doi.org/10.1074/mcp.T600050-MCP200.

Soellner, P., Quinlan, R.A., and Franke, W.W. (1985). Identification of a distinct soluble subunit of an intermediate filament protein : tetrameric vimentin from living cells. Proc. Natl. Acad. Sci. USA *82*, 7929–7933.

Stamburski, C., Renaudin, J., and Bové, J.M. (1992). Mutagenesis of a tryptophan codon from TGG to TGA in the cat gene does not prevent its expression in the helical mollicute *Spiroplasma citri*. Gene *110*, 133–134. https://doi.org/10.1016/0378-1119(92)90458-2.

Tang, W.H., Shilov, I.V., and Seymour, S.L. (2008). Nonlinear fitting method for determining local false discovery rates from decoy database searches. J. Proteome Res. *7*, 3661–3667. https://doi.org/10.1021/pr070492f.

Townsend, R., and Archer, D.B. (1983). A Fibril protein antigen specific to *Spiroplasma*. Microbiology *129*, 199–206.

Townsend, R., Archer, D.B., and Plaskitt, K.A. (1980). Purification and preliminary characterization of *Spiroplasma* fibrils. J. Bacteriol. *142*, 694–700.

Trachtenberg, S., Andrews, S.B., and Leapman, R.D. (2003). Mass distribution and spatial organization of the linear bacterial motor of *Spiroplasma citri* R8A2. J. Bacteriol. *185*, 1987–1994. https://doi.org/10.1128/JB.185.6.1987-1994.2003.

Trachtenberg, S., Dorward, L.M., Speransky, V.V., Jaffe, H., Andrews, S.B., and Leapman, R.D. (2008). Structure of the cytoskeleton of *Spiroplasma melliferum* BC3 and its interactions with the cell membrane. J. Mol. Biol. *378*, 778–789. https://doi.org/10.1016/j.jmb.2008.02.020.

Trachtenberg, S., and Gilad, R. (2001). A bacterial linear motor: cellular and molecular organization of the contractile cytoskeleton of the helical bacterium *Spiroplasma melliferum* BC3. Mol. Microbiol. *41*, 827–848. https://doi.org/10.1046/j.1365-2958.2001.02527.x.

Wheeler, D.L., Barrett, T., Benson, D.A., Bryant, S.H., Canese, K., Chetvernin, V., Church, D.M., Dicuccio, M., Edgar, R., Federhen, S., et al. (2008). Database resources of the national center for Biotechnology information. Nucleic Acids Res. *36*, D13–D21. https://doi.org/10.1093/nar/gkm1000.

Williamson, D.L. (1974). Unusual fibrils from the spirochete-like sex ratio organism. J. Bacteriol. *117*, 904–906.

Williamson, D.L., Renaudin, J., and Bové, J.M. (1991). Nucleotide sequence of the *Spiroplasma citri* fibril protein gene. J. Bacteriol. *173*, 4353–4362.

Yin, C., Goonawardane, N., Stewart, H., and Harris, M. (2018). A role for domain I of the hepatitis C virus NS5A protein in virus assembly. PLoS Pathog. *14*, e1006834. https://doi.org/10.1371/journal.ppat.1006834.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial strains		
Invitrogen™ *E. coli* BL21-AI™	Fisher Scientific	Cat# C607003
E. coli NEB® TURBO	New England Biolabs Inc.	C2984H
Chemicals, peptides, and recombinant proteins		
Urografin	Cadila healthcare Ltd, Kundaim, India	N/A
L-arabinose	SRL Pvt. Ltd.	52392
Tris-base	Sigma-Aldrich	10708976001
Triton X-100	Sigma-Aldrich	T8787
Glycerol	Sigma-Aldrich	G6279
N-lauroyl sarcosine sodium salt	Sigma-Aldrich	61743
Sodium deoxycholate	Sigma-Aldrich	D6750
Tween-20	Sigma-Aldrich	P9416
Sodium dodecyl sulphate	Sigma-Aldrich	L4390
EDTA	Sigma-Aldrich	E6758
Imidazole	Sigma-Aldrich	12399
Uranyl acetate	SRL	Cat#81405
Trypsin	Promega	Cat#V5280
Oligonucleotides		
Oligonucleotides used in this work	This work	See Table S2
Recombinant DNA		
Plasmid: pHIS17	Kunzelmann and Webb (2009)	Addgene Cat#78201
Software and algorithms		
PSIPRED	McGuffin et al. (2000)	http://bioinf.cs.ucl.ac.uk/psipred/
ImageJ 1.52p	Schneider et al. (2012)	https://imagej.net/Fiji
NCBI (RefSeq protein database)	Wheeler et al. (2008)	https://www.ncbi.nlm.nih.gov/
Protein pilot 2.0.1	SciEx	https://sciex.com/products/software/proteinpilot-software
Other		
Carbon-Formvar coated copper grids (300 mesh)	Electron Microscopy Sciences	Cat#01843
Luria Bertani broth	HiMedia	M1245
Spiroplasma citri genomic DNA	Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ)	accession number 21833
Ampicillin	Sigma-Aldrich	A9518
Ni-NTA agarose	Qiagen	Cat#30210

RESOURCE AVAILABILITY

Lead contact

Information and requests for resources and reagents should be directed to and fulfilled by the lead contact, Pananghat Gayathri (gayathri@iiserpune.ac.in).

Materials availability

E. coli BL21 (AI) strain and the plasmids generated in the study are available with Pananghat Gayathri (gayathri@iiserpune.ac.in).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cultures

E. coli NEB ®TURBO and *E. coli* BL21 (AI) cells used for cloning and protein expression in this study were grown in LB (Luria Bertani) broth supplemented with ampicillin at final concentration of 100 μg/mL at 37°C.

METHOD DETAILS

Cloning

The *fibril* gene was amplified from genomic DNA of *Spiroplasma citri* (*S. citri*; DSMZ accession number 21833) using appropriate primers (Table S2). Six TGA codons (corresponding to amino acid number 26, 200, 258, 316, 323 and 342 from amino terminus) in *fibril* gene were mutated to TGG and full-length *fibril* gene with modified tryptophan codons was cloned into pHIS17 vector [Addgene plasmid #78201; (Kunzel-mann and Webb, 2009)] with or without a terminal hexa-histidine (His₆) tag either at the N-terminus or at the C-terminus. *E. coli* NEB ®TURBO cells were used for cloning. All the clones were confirmed by sequencing using appropriate primers (Table S2).

Expression and cell lysis

All Fibril constructs were expressed in *E. coli* BL21 (AI) cells (Invitrogen) using the protocol described earlier (Harne et al., 2020a). Expression check of desired protein construct in cells with or without induction was confirmed by resuspending cells from 10 mL culture in 1 mL lysis buffer [50 mM Tris (pH 8.0), 200 mM NaCl and 10% glycerol] and lysing cells using probe sonicator followed by spinning the lysate at 21,000 xg for 10 min at 4°C. Protein overexpression was confirmed by the presence of high intensity band with molecular weight of ~59,000 Da on the SDS-PAGE gel. Cells overexpressing desired *fibril* construct were pelleted, pellet flash-frozen in liquid nitrogen and stored at −80°C until further processing.

Frozen pellet of *E. coli* cells expressing desired *fibril* construct was thawed on ice and resuspended in appropriate lysis buffer [50 mM Tris (pH 8.0), 200 mM NaCl and 10% glycerol]. The cells were lysed on ice by sonication using a probe sonicator (VibraCell™). The lysate was spun at 4629 xg for obtaining the clarified lysate. During initial attempts to purify Fibril, the clarified lysate was spun at 159,000 xg to separate filamentous proteins from monomeric or oligomeric components. Accordingly, Fibril was obtained in the pellet fraction. However, Fibril from the pellet could not be solubilized and hence required screening of different compounds for their ability to solubilize Fibril.

Standardization for solubility of fibril

Pellet obtained from 1200 mL of *E. coli* culture expressing His₆-Fibril was re-suspended in 270 mL of lysis buffer base [10 mM Tris (pH 7.6), 1% (v/v) Triton X-100 and 2 M glycerol]. Cells were lysed by sonication (described above) and the lysate was equally divided into 6 tubes. Multiple solubilization conditions for the lysate were tested with one of the following detergents added to buffer base: N-lauroyl sarcosine sodium salt (LSS; 0.5% w/v), sodium deoxycholate (1% w/v), Tween-20 (1 % v/v), sodium dodecyl sulphate (SDS; 1 % w/v), EDTA (ethylenediaminetetraacetic acid; 20 mM w/v). All these conditions show the effect of respective compounds in addition to Triton X-100, a constituent of lysis buffer base. The effect of only Triton X-100 (1% v/v) was checked by addition of lysis buffer base to the sixth tube containing lysate. Values in bracket indicate final concentration of respective compounds in the 50 mL lysate. The tubes containing lysate were incubated on ice for 30 min and then spun at 4,629 xg to remove cell debris as pellet. The supernatant was further spun at 159,000 xg at 4°C for 30 min. Supernatants obtained after centrifugation at 159,000 xg were transferred to fresh tubes and each pellet was re-suspended in 2 mL T₁₀E₁₀ buffer [10 mM Tris (pH 7.6), 10 mM EDTA]. 10 μL sample from each condition at all the stages was mixed with 10 μL SDS-PAGE loading dye (2×), heated at 99°C for 10 min and visualized on SDS-PAGE gel. 7 μL was loaded on the gel for total lysate and 159,000 xg pellet while 15 μL was loaded for 4,629 xg supernatant and 159,000 xg supernatant.

Purification of fibril expressed in *E. coli* by Ni-NTA affinity chromatography

A 200 mL culture pellet of *E. coli* BL21 (AI) cells expressing Fibril with a terminal His$_6$ tag (either N-terminal or C-terminal end) was re-suspended in 50 mL lysis buffer [50 mM Tris (pH 7.6), 200 mM NaCl and glycerol 10% (v/v)]. Cells were lysed by sonication (described above) and lysate spun at 4,629 xg at 4°C for 15 min to remove cell debris as pellet. Supernatant was further spun at 159,000 xg at 4°C for 30 min to pellet down Fibril. The pellet was re-suspended using 1.8 mL buffer A$_{200}$ [50 mM Tris (pH 7.6), 200 mM NaCl] at room temperature. 200 μL of 10% (w/v) SDS solution was added and mixed with the re-suspended pellet. The protein suspension was spun at 21,000 xg at 25°C for 10 min to remove any large insoluble aggregates. The soluble fraction from this step was spun at 159,000 xg at 25°C for 30 min and Fibril was obtained in the supernatant. All the steps after addition of SDS were performed at room temperature (∼25°C) since SDS precipitates in cold condition (4°C).

The supernatant was mixed with 2 mL Ni-NTA resin (Ni-NTA agarose; Qiagen) pre-equilibrated with buffer A$_{200}$ (50 mM Tris pH 7.6, 200 mM NaCl) and left for mixing with gentle agitation at room temperature for 30 min. The mixture was put into an empty column with filter. Additional (∼20 mL) buffer A$_{200}$ was passes through the resins to remove loosely bound proteins and collected as wash. The bound protein was eluted using elution buffer containing increasing concentrations (25 mM, 50 mM, 100 mM, 250 mM and 500 mM) of imidazole in binding buffer. Fractions containing protein of interest were identified on a 12% SDS-PAGE gel and pooled (about 8 mL). The pooled protein was concentrated to about 500 μL using concentrators. To remove the salts and imidazole, cycles of dilution of the concentrated protein with sterile distilled water, followed by concentration were performed as follows. Sterile distilled water was added to concentrated protein and volume made to 15 mL. It was again concentrated to about 500 μL using concentrators. During the process of concentrating the protein, the protein solution was mixed by pipetting several times to ensure uniform solution. This cycle of protein dilution followed by concentration was repeated 4 times to remove maximum salts and SDS. Finally, the protein was concentrated to minimum possible volume (about 300 μL). It was aliquoted as 20 μL fractions in thin walled (polymerase chain reaction; PCR) tubes, flash frozen in liquid nitrogen and stored in - 80°C freezer until use. The salt-free, concentrated protein obtained by this method was used for visualization using Transmission Electron Microscope (TEM), FT-IR and mass spectroscopy studies.

Purification of insoluble fibril using density gradient

Preparation of urografin density gradient

Sterile urografin solution (76% w/v; density 1.52 gm/cm^3; Cadila healthcare Ltd, Kundaim, India) was obtained for preparation of density gradients. Urografin solutions of 30.4% and 53.2% (w/v) were prepared in buffer containing 10 mM Tris (2 M; pH 7.6) and 10 mM EDTA and vortexed to ensure uniform solution in each tube. The density gradient was prepared by layering 4.5 mL of 30.4% urografin solution on top of 4.5 mL of 53.2% urografin solution in an ultra-clear centrifuge tube (Beckman Coulter; catalogue number 344059). The tubes were left at room temperature (25°C) for about 20 h to allow the formation of a linear gradient.

Purification of fibril from *E. coli* by separation on urografin density gradient

A 500 mL culture pellet of *E. coli* BL21 (AI) cells expressing full length untagged *fibril* (UTF) gene was re-suspended in 60 mL of lysis buffer [10 mM Tris pH 7.6, 1% (v/v) Triton X-100] and lysed using a probe sonicator (described above). The lysate was spun at 4,629 xg at 4°C for 10 min and cell debris in the form of pellet was discarded. The supernatant so obtained was further spun at 159,000 xg to separate Fibril as pellet. The pellet was re-suspended in 1 mL T$_{10}$E$_{10}$ buffer [10 mM Tris (pH 7.6), 10 mM EDTA] and left for stirring at 4°C for about 20 h to break clumps into finer particles.

1 mL of stirred enriched Fibril was loaded on top of 9 mL urografin density gradient and spun in a swinging bucket rotor at 100,000 xg at 25°C for 2 h. The two protein layers from gradient were collected as separate fractions. Urografin was removed from the fractions by two rounds of washes by diluting it with 65 mL distilled water and spinning at 159,000 xg at 25°C for 30 min. The pellet obtained after second round of washing was re-suspended in sterile distilled water (500 μL) and used for visualization on 12% SDS-PAGE gel along with marker proteins to identify fraction containing Fibril.

Visualization of Fibril using transmission electron microscopy (TEM)

Carbon-Formvar coated copper grids (300 mesh; Ted-Pella, Inc) were used without or with glow-discharging using plasma cleaner (Quorum technologies) just before use. 2–5 μL of purified protein (Figure 2B lane 7 or Figure 3C lane 6) was applied to grids and allowed to stand at room temperature for 30 s. Excess buffer was blotted and 2 - 4 μL stain [0.5% w/v uranyl acetate in water] was applied to the grid and immediately absorbed from bottom. Grids were scanned and imaged using transmission electron microscope (JEM-2200FS, Jeol Ltd.). Images were visualized using ImageJ software (Schneider et al., 2012).

Fourier-transform infrared (FT-IR) spectroscopy

Purified His_6-tagged and untagged Fibril [obtained using SDS treatment protocol (Figure 2B lane 7) and urogafin gradients (Figure 3C lane 6) respectively] was flash frozen in liquid nitrogen and lyophilized. The dried protein samples were used for collecting FT-IR spectra. Sample transmittance was recorded in the IR spectra (1600–1690 cm^{-1}) with a resolution of 2 cm^{-1} and 24 scans at each point were acquired in FT-IR Spectrophotometer (NICOLET 6700) using KBr pellet. The transmittance of a sample was normalized by considering its transmittance at 1600 cm^{-1} as 100% and the transmittance at other wavelengths as the percentage of transmittance at 1600 cm^{-1}. Values obtained after normalization were subtracted from 1 to obtain absorbance (as percentage) and it is plotted against wave number (cm^{-1}) to identify peaks corresponding to secondary structures.

In-gel trypsin digestion and mass spectroscopy of the protein

The most prominent bands from purified protein (corresponding to masses ~60,000 Da, ~ 36,000 Da and ~26,000 Da) [corresponding to red and blue asterisk in Figure 2B lane 7 and red, tan and blue asterisks in Figure 3C lane 5] were excised from SDS-PAGE gel stained with Coomassie and transferred to fresh tubes. Gel pieces were de-stained by incubating them with mixture of 40% acetonitrile and 50 mM TEABC [triethylammonium bicarbonate]. Protein reduction was performed by incubating gel pieces in 10 mM DTT (dithiothreitol) at 60°C for 30 min. Excess DTT was then removed, and alkylation carried out by addition of 20 mM iodoacetamide and incubation in dark for 15 min. Iodoacetamide was then removed, and gel pieces were dehydrated by 4 cycles of washing using acetonitrile (100%). The completely dried gel pieces were rehydrated by re-suspension in sequencing grade trypsin (10 ng/μL in 50 mM TEABC) on ice for 60 min. Excess trypsin was removed without centrifugation and then gel pieces were re-suspended in 20 mM TEABC buffer overnight at 37°C. Next, trypsin-digested peptides were obtained by spinning the samples (21000 xg at 25°C for 5 min) and collection of supernatants. Remnants of the peptides were extracted from the gel pieces by re-suspending them in varying concentrations of acetonitrile: formic acid (3%: 0.4%, 40%: 0.4% and 100% acetonitrile) followed by incubation for 10 min at room temperature, spinning at 10,000 xg for 10 min and harvesting the supernatant. The aliquots of liquid containing peptides were pooled and evaporated using CentriVap DNA vacuum concentrator until dry. The desalting and cleaning of tryptic peptides was performed using StageTip protocol (Rappsilber et al., 2007).

For LC-MS/MS analysis, a Sciex TripleTOF6600 mass spectrometer interfaced with an Eksigent nano-LC 425 instrument was used. Trypsin-digested peptides (~1 μg) were loaded onto an Eksigent C18 trap column (5 μg capacity) followed by elution on an Eksigent C18 analytical column using a linear gradient of acetonitrile. A typical liquid chromatography (LC) run consisted of post-loading for 60 min onto the trap with solvent A (water and 0.1% formic acid) and solvent B (acetonitrile). The gradient schedule for the LC run was 5% to 10% (v/v) B for 2 min, a linear gradient of B from 10% to 30% (v/v) over 50 min, 30% to 90% (v/v) over 4 min, 90% (v/v) B for 5 min and equilibration with 5% (v/v) B for 2 min. For all the samples data was acquired in information-dependent acquisition (IDA) mode over a mass range of m/z 300 − 1600. Each full MS survey scan was followed by MS/MS of the 13 most intense peptides. Dynamic exclusion was enabled for all experiments (repeat count 2; exclusion duration 5 s).

The peptides were identified out using Protein Pilot (version 2.0.1, Sciex) using Pro-Group and Paragon algorithms (Jagtap et al., 2012; Shilov et al., 2007; Tang et al., 2008) against the RefSeq protein database (Wheeler et al., 2008) of *E. coli* (last downloaded 2021-03-01) with manual addition of Fibril sequence (*S. citri*; NCBI RefSeq id WP_071937222.1) without a tag or with a C-terminal 6xHis tag separated by amino acids GS (GSHHHHHH). The peptides detected from 2 different batches of each sample were pooled and non-redundant peptides marked on the Fibril sequence. For analysis the peptides arising out of non-trypsin cleavage were excluded. Secondary structure prediction of Fibril was performed using PSIPRED (McGuffin et al., 2000) and used for analysis of the peptides.