

# $\beta$ -Helical architecture of cytoskeletal bactofilin filaments revealed by solid-state NMR

Suresh Vasa<sup>a,1</sup>, Lin Lin<sup>b,c,1</sup>, Chaowei Shi<sup>a,d,1</sup>, Birgit Habenstein<sup>a,2</sup>, Dietmar Riedel<sup>e</sup>, Juliane Kühn<sup>b,c</sup>, Martin Thanbichler<sup>b,c,f,3</sup>, and Adam Lange<sup>a,d,g,3</sup>

<sup>a</sup>Department of NMR-Based Structural Biology and <sup>e</sup>Electron Microscopy Group, Max Planck Institute for Biophysical Chemistry, 37077 Göttingen, Germany; <sup>b</sup>Prokaryotic Cell Biology Group, Max Planck Institute for Terrestrial Microbiology, 35043 Marburg, Germany; <sup>f</sup>Faculty of Biology and <sup>1</sup>LOEWE Center for Synthetic Microbiology, Philipps-Universität, 35043 Marburg, Germany; <sup>d</sup>Department of Molecular Biophysics, Leibniz-Institut für Molekulare Pharmakologie, 13125 Berlin, Germany; and <sup>3</sup>Institut für Biologie, Humboldt-Universität zu Berlin, 10115 Berlin, Germany

Edited by Robert Tycko, National Institutes of Health, Bethesda, MD, and accepted by the Editorial Board December 9, 2014 (received for review September 24, 2014)

**Bactofilins are a widespread class of bacterial filament-forming proteins, which serve as cytoskeletal scaffolds in various cellular pathways. They are characterized by a conserved architecture, featuring a central conserved domain (DUF583) that is flanked by variable terminal regions. Here, we present a detailed investigation of bactofilin filaments from *Caulobacter crescentus* by high-resolution solid-state NMR spectroscopy. De novo sequential resonance assignments were obtained for residues Ala39 to Phe137, spanning the conserved DUF583 domain. Analysis of the secondary chemical shifts shows that this core region adopts predominantly  $\beta$ -sheet secondary structure. Mutational studies of conserved hydrophobic residues located in the identified  $\beta$ -strand segments suggest that bactofilin folding and polymerization is mediated by an extensive and redundant network of hydrophobic interactions, consistent with the high intrinsic stability of bactofilin polymers. Transmission electron microscopy revealed a propensity of bactofilin to form filament bundles as well as sheet-like, 2D crystalline assemblies, which may represent the supramolecular arrangement of bactofilin in the native context. Based on the diffraction pattern of these 2D crystalline assemblies, scanning transmission electron microscopy measurements of the mass per length of BacA filaments, and the distribution of  $\beta$ -strand segments identified by solid-state NMR, we propose that the DUF583 domain adopts a  $\beta$ -helical architecture, in which 18  $\beta$ -strand segments are arranged in six consecutive windings of a  $\beta$ -helix.**

assembles into short filamentous structures that emanate from the cell poles, recruiting and thus controlling a small GTPase involved in type IV pili-dependent motility (12). As another well-characterized example, the  $\epsilon$ -proteobacterium *Helicobacter pylori*, a human pathogen notorious for causing peptic ulcers, was shown to depend on a bactofilin homolog (CcmA) for maintaining its characteristic helical cell shape, a feature required for cells to efficiently colonize the gastric mucus (13). Moreover, in the  $\gamma$ -proteobacterium *Proteus mirabilis*, a homolog of the bactofilin CcmA has been implicated in cell shape and swarming motility (14).

Bactofilins are usually small proteins (~20 kDa) that are composed of a central conserved domain of unknown function (DUF583) and flanking N- and C-terminal regions of variable length and sequence. A characteristic of bactofilins is their ability to polymerize spontaneously in the absence of nucleotides or other cofactors (8). Native BacM protofilaments have been isolated from *M. xanthus* whole-cell lysates by sucrose density centrifugation (11). Moreover, polymers of *C. crescentus* BacA and BacB were obtained after heterologous expression in *Escherichia coli*, a species that lacks chromosomally encoded bactofilin homologs (8). Similarly, polymerization was observed for heterologously produced *M. xanthus* BacN, BacO, and BacP (8, 12). In all cases, the filamentous structures formed were biochemically inert and

solid-state NMR | cytoskeleton | protein structure | filaments

**S**imilar to eukaryotes, bacteria use a number of different cytoskeletal elements to ensure the proper temporal and spatial organization of their cellular machinery. Various studies proved the existence of bacterial homologs of eukaryotic cytoskeleton proteins including tubulin homologs such as FtsZ (1), actin homologs such as MreB (2), and intermediate filament (IF)-like proteins (3), which together have important roles in cell division, morphogenesis, polarity determination, and DNA segregation (4–7). In addition, several groups of polymer-forming proteins that are limited to the bacterial domain have been described (6).

A recent addition to these bacteria-specific cytoskeletal proteins are the so-called bactofilins (8), a class of proteins that is widespread among most bacterial lineages and involved in a variety of different cellular processes. In the prosthecate  $\alpha$ -proteobacterium *Caulobacter crescentus*, for instance, the two bactofilin paralogues BacA and BacB (Fig. 1A) assemble into membrane-associated polymeric sheets that are specifically localized to the cell pole carrying the stalk (8), a thin protrusion of the cell body involved in cell attachment and nutrient acquisition (9). These assemblies serve as spatial landmarks mediating the polar localization of a cell wall biosynthetic enzyme, PbpC, involved in stalk biogenesis (8) and organization (10). The  $\delta$ -proteobacterial species *Myxococcus xanthus*, by contrast, possesses four bactofilin paralogues (BacMNOP) with, at least partly, distinct functions. BacM was shown to form cable- or rod-like structures that are critical for proper cell shape (8, 11). BacP, on the other hand,

## Significance

**Bactofilins are a new class of cytoskeletal proteins that are involved in key cellular processes. For instance, in the human pathogen *Helicobacter pylori*, they are responsible for maintaining its characteristic helical cell shape, a feature required for cells to efficiently colonize the gastric mucus. So far the atomic structure of bactofilin filaments has remained elusive, as the large bactofilin assemblies are not amenable to standard methods for 3D structure determination. Here, we have applied a combination of solid-state NMR and electron microscopy and discovered that bactofilins adopt a  $\beta$ -helical architecture, which has not been observed before for other cytoskeletal filaments. Interestingly, however, the structure bears similarities to that of the fungal prion protein HET-s.**

Author contributions: M.T. and A.L. designed research; S.V., L.L., C.S., and D.R. performed research; S.V., L.L., C.S., B.H., D.R., J.K., M.T., and A.L. analyzed data; and S.V., L.L., C.S., M.T., and A.L. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. R.T. is a guest editor invited by the Editorial Board.

<sup>1</sup>S.V., L.L., and C.S. contributed equally to this work.

<sup>2</sup>Present address: Université de Bordeaux, Chimie et Biologie des Membranes et des Nano-objets (CBMN) (UMR 548), Institut Européen de Chimie et de Biologie, 33607 Pessac, France.

<sup>3</sup>To whom correspondence may be addressed. Email: alange@mp-biochemie.de or thanbichler@uni-marburg.de.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1418450112/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1418450112/-DCSupplemental).

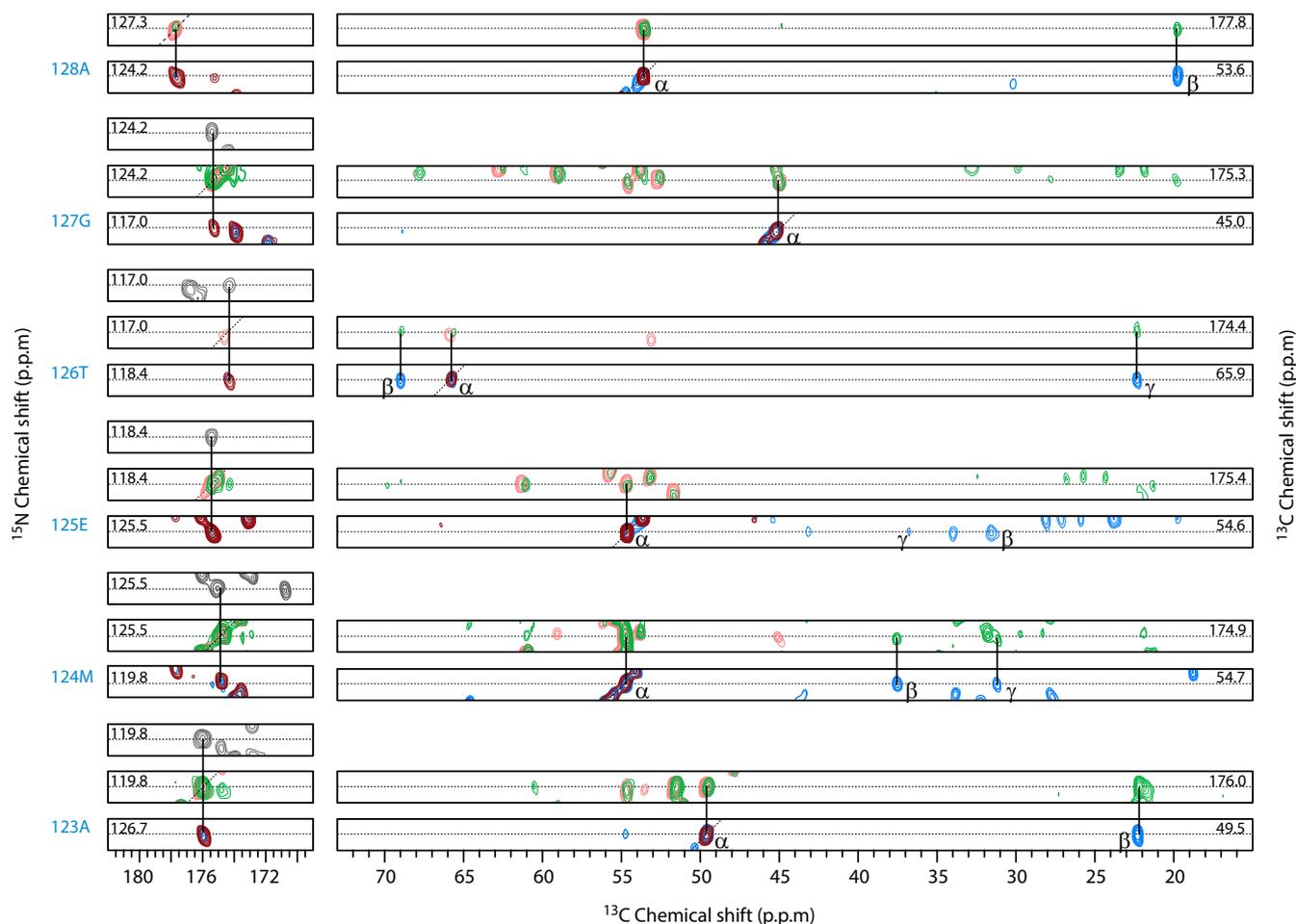


spectra of two independently produced samples (Fig. 1C and *SI Appendix*, Fig. S1). The two samples yielded essentially identical results, with a very high spectral resolution for carbons of around 0.25–0.6 ppm (full-width at half-height, FWHH) and for nitrogens of 0.5–0.8 ppm (see  $^{13}\text{C}$  trace in Fig. 1C and  $^{13}\text{C}$  and  $^{15}\text{N}$  traces in *SI Appendix*, Fig. S1). The quality of the data was comparable to that obtained for benchmarks in ssNMR spectral quality, such as well-structured microcrystalline proteins (27, 28),  $\beta$ -helical HET-s(218–289) prion fibrils (29), or the type III secretion system needles (30, 31). Because the CC and NCA experiments were based on through-space dipolar couplings, which can be averaged out by dynamics, the observed peaks had to represent the rigid core of the protein. We found that the number of residue type-specific spin systems closely matched the number of residues constituting the conserved DUF583 domain. This result indicates that the whole spectrum of the rigid core is fully explained by the DUF583 domain (*SI Appendix*, Table S3).

To obtain sequential  $^{13}\text{C}$  and  $^{15}\text{N}$  backbone and side-chain assignments of the rigid core of BacA filaments, a set of dipolar-based 2D and 3D correlation experiments was performed. A sequential walk along the backbone resonances can be easily achieved by a combination of 3D experiments (32). All 3D experiments conducted and the optimized experimental parameters are listed in *SI Appendix*, Table S1. Intraresidue spin system correlations were obtained from 3D NCACX and NCACO spectra, whereas interresidue connections were provided by 3D NCOCX and

NCOCA spectra. The interresidue connections were further confirmed by 3D CANCO and 2D CANCECO spectra (*SI Appendix*, Fig. S3). The latter experiment and also the 3D NCACO and NCOCA experiments benefitted from efficient CO–CA transfer, which was achieved by using band-selective homonuclear cross-polarization (CP) (33–35). For illustration, a sequential walk from Ala128 to Ala123 is shown in Fig. 2. Collectively, this approach resulted in unambiguous and reliable de novo assignments for residues Ala39 to Phe137. An almost complete assignment of the rigid core (95% of backbone and 90% of side chain nuclei) was obtained in this region (*SI Appendix*, Table S2). No unambiguous assignments could be obtained for residues Leu46 and Val81. In summary, the analysis clearly shows that the rigid core of BacA filaments comprises residues Ala39 to Phe137, which corresponds to the complete DUF583 domain.

Importantly, the chemical shifts of the backbone nuclei are sensitive to the backbone dihedral angles. This fact allowed us to use the obtained  $^{13}\text{C}$  and  $^{15}\text{N}$  chemical shifts to determine the secondary structure of the BacA filament core. First, structural predictions were derived from secondary chemical shift values ( $\Delta\delta_{\text{CA}} - \Delta\delta_{\text{CB}}$ ) (Fig. 3). This approach is based on the observation that for all residues (apart from glycine), three or more consecutive negative secondary chemical shifts are indicative of  $\beta$ -strand conformation, whereas positive values point to an  $\alpha$ -helix, and values close to zero indicate random coil conformation (36). Second, a statistical databank approach as implemented in the



**Fig. 2.** Sequential walk using 3D correlation spectra. Shown are strip plots of 3D NCACX (sky blue), NCOCX (light green), CANCO (black), NCACO (maroon), and NCOCA (red) spectra analyzed to perform a sequential walk along amino acids Ala128–Ala123. The connections used in the sequential walk are indicated by solid lines. The corresponding nitrogen and carbon chemical shifts are listed.

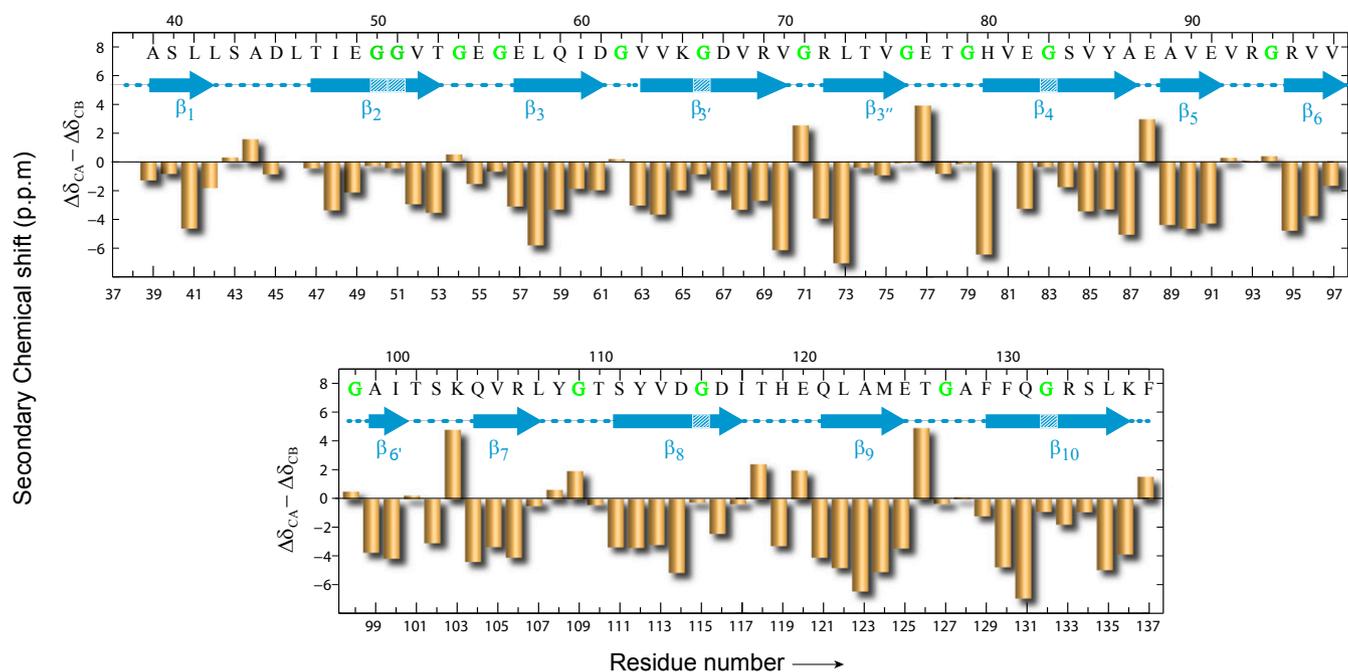
TALOS+ program (37) was used to corroborate the results from the secondary chemical shift analysis (*SI Appendix*, Fig. S4). This combined analysis clearly demonstrated that most of the residues in the core region (i.e., the DUF583 domain) adopt a  $\beta$ -sheet conformation. More specifically, we identified at least 10  $\beta$ -strands of different lengths, where strands  $\beta_3$  and  $\beta_6$  are likely interrupted by short kinks or turns at residues Gly62/Gly71 ( $\beta_3$ ) and Gly98 ( $\beta_6$ ). Although TALOS+ predicts  $\beta$ -sheet conformation for the glycines in strands  $\beta_2$ ,  $\beta_3$ ,  $\beta_4$ ,  $\beta_8$ , and  $\beta_{10}$ , these residues may give rise to kinks in the corresponding strands as well, resulting in a total of 18 potential  $\beta$ -strand segments. Given that these glycines are highly conserved (see Fig. S4), the formation of kinked  $\beta$ -strands may be an important structural feature of bactofilins.

**The N- and C-Terminal Regions of Polymerized BacA Remain Highly Flexible.** The rigid core region exhibits differences in mobility, which can be observed at first glance in the NCA correlation spectrum (*SI Appendix*, Fig. S1). Because this correlation spectrum is based on dipolar NCA transfer, the peak intensities are a direct measure of the backbone dynamics (38). Compared with other alanine residues, Ala39 and Ala99 exhibit considerably weaker peak signal intensities. Similarly, Ser40 displays less intensity compared with other serine residues. In general, the flanking residues of the DUF583 domain show less intensity compared with the middle part, suggesting an increase in dynamics from the center of the domain to the ends (*SI Appendix*, Fig. S2). In addition to variations in the backbone dynamics, also side chain dynamics appear to vary between different residues. For instance, the side-chain cross-peak intensities of Ile48 are reduced compared with the corresponding correlations of Ile60, Ile100, and Ile117 (Fig. 1C).

To investigate the highly dynamic regions that were not detected in dipolar-mediated experiments (based on CP), we additionally performed through-bond correlation experiments [based on insensitive nuclei enhanced by polarization transfer (INEPT); 39]. This type of experiment yields spectra in which dynamic residues give rise to cross-peaks, whereas rigid core residues do not

contribute. An initial comparison of 1D- $^{13}\text{C}$  or  $^{15}\text{N}$  CP and INEPT spectra (*SI Appendix*, Fig. S5) revealed the presence of different observable spin systems. More detailed information was then obtained from 2D  $^1\text{H}$ - $^{13}\text{C}$  and  $(^1\text{H})$ - $^{13}\text{C}$ - $^{13}\text{C}$  through-bond correlation experiments (40) (*SI Appendix*, Fig. S6). As bactofilin polymerizes in a nucleotide-independent manner, most of the BacA protein in the ssNMR sample should exist in polymerized form. Consistently, only a low number of threonine, isoleucine, leucine, glycine, and valine spin systems could be counted in the INEPT-type spectra. This result clearly excluded the presence of monomers in the sample, because otherwise signals from all residues of the sequence would be expected. At the same time, the number of observed residues was smaller than the number of residues constituting the proline-rich N- and C-terminal regions of BacA (*SI Appendix*, Table S3), indicating that only parts of these terminal segments are highly flexible. All of the observed residues appeared with random-coil chemical shifts, confirming the unstructured and highly dynamic nature of the N and C termini. The situation observed here resembles the case of fibrils formed by amyloids such as tau (41),  $\alpha$ -synuclein (42), and the human prion protein (43), which were also reported to feature a rigid core that is surrounded by flexible termini. One could therefore speculate that, because of their  $\beta$ -sheet-rich nature, BacA filaments are amyloid fibrils. However, as shown below, on the basis of MPL measurements, we can unambiguously exclude the possibility that bactofilin filaments are classical amyloid fibrils with a single layer of (parallel, in-register)  $\beta$ -sheet architecture.

**The  $\beta$ -Sheet-Rich Core of Polymerized BacA Adopts a  $\beta$ -Helical Architecture.** In addition to the spectroscopic studies, the ssNMR samples were also investigated by TEM (Fig. 4). Various morphologies were observed, including single filaments, filament bundles, and even 2D crystalline arrangements. Single filaments were approx. 3 nm wide and showed a clear tendency to assemble laterally into bundles or sheets. Bundles were the most abundant higher order structures in the sample (Fig. 4B,



**Fig. 3.** Secondary structure analysis of BacA filaments by means of secondary chemical shifts. The secondary chemical shifts ( $\Delta\delta_{\text{CA}} - \Delta\delta_{\text{CB}}$ , in ppm units) are shown as a function of the residue number for the rigid DUF583 domain. The blue arrows represent  $\beta$ -strands, numbered such that two continuous  $\beta$ -strands separated by a single glycine residue are considered as one strand. Potential kinks due to glycine residues are indicated by slashed lines. All glycine residues are highlighted in green color in the sequence.

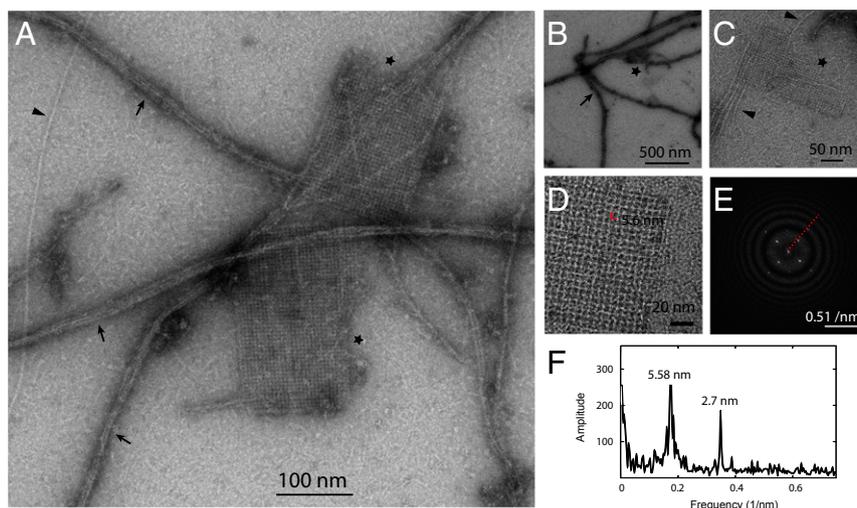
arrows), with bundle widths ranging from 15 to 30 nm. Occasionally, clear gaps were observed between adjacent filaments, possibly resulting from repulsive forces between the flexible N- and C-terminal regions. Interestingly, we additionally detected 2D crystalline arrays of BacA in all samples investigated (Fig. 4 C and D and *SI Appendix*, Fig. S7). Their architecture resembled a woven carpet-like pattern with an average distance between two adjacent “knots” of around 5.6 nm. Diffraction data of this crystal arrangement showed two prominent bright spots at 5.6 nm and 2.7 nm (Fig. 4 E and F), where the latter may correspond to the periodic arrangement in the filament structure. Notably, in *C. crescentus*, BacA was shown to form sheets that line the inner cell membrane (8). It is thus tempting to speculate that the 2D crystalline arrangement observed in this study for the first time, to our knowledge, reveals the configuration of BacA adopted in the biological context.

Next, we estimated the MPL of BacA filaments by means of STEM. Similar to the TEM results, dark-field STEM images revealed a tendency of BacA filaments to form bundles with widths ranging from 50 to 70 nm (Fig. 5). A typical bundle with a width of (approximately) 63 nm consisted of six well-separated filaments, with an average width of about 3 nm and a spacing of around 4–7 nm. The interfilament region is likely occupied by the mobile N and C termini. To obtain the MPL of a single filament, we quantified the electron density of different segments (each 25 nm long) in well-separated filaments and determined the corresponding masses, using tobacco mosaic virus (TMV) particles with a known MPL as a standard. Although the images were of low contrast, the data were generally in good agreement with a “3-nm solid rod” model. Other models with larger or smaller widths led to significantly poorer fits. Fitting of the MPL histogram with a single Gaussian distribution function yielded an MPL for BacA of  $4.33 \pm 0.09$  (FWHM,  $0.58 \pm 0.07$ ) kDa/nm (Fig. 5D). This value corresponds to a length of 4.46 nm per BacA subunit, with a molecular mass of 19.3 kDa. However, the disordered N- and C-terminal regions of BacA are likely to contribute only little to the observed MPL. Assuming that only the compact DUF583 domain (9.63 kDa) is detectable in the STEM measurements, the length of one subunit would thus be 2.2 nm.

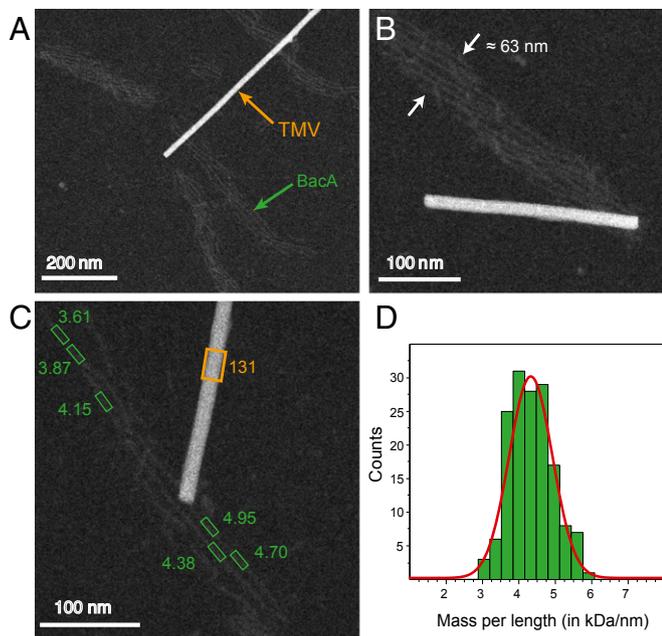
The MPL is a characteristic and important constraint in defining the molecular structure of polymer-forming proteins. As  $\beta$ -strands constitute the major secondary structural element of

BacA, we can use the MPL value to test the hypothesis that bactofilin polymers are functional amyloids. In a cross- $\beta$  structure, the MPL is approximately  $\eta \times MW / 0.47$  kDa/nm, where MW is the molecular mass and  $\eta$  is the number of molecules in a single  $\beta$ -sheet layer (44). The corresponding MPL values for BacA would be 41.1 and 82.1 kDa/nm for  $\eta = 1$  and 2, respectively (or 20.5 and 41.0 kDa/nm when considering only the DUF583 domain without the flexible N- and C-terminal regions). These values are much higher than the observed MPL of 4.33 kDa/nm. Hence, we can conclude that bactofilins do not form classical single-layer amyloid fibrils. Assuming multiple layers of  $\beta$ -structure per molecule, one would expect  $\eta$  values of 0.11 and 0.21 for full-length BacA or the isolated DUF583 core domain, respectively, which correspond to approx. nine and five layers per single BacA molecule. Considering the number of turns and  $\beta$ -strands observed by ssNMR spectroscopy, we propose a  $\beta$ -helical arrangement for the structure of bactofilins. Such a model was also suggested by Koch et al. (11), based on the frequent repetition of glycine residues. Coming back to the reflection data (Fig. 4F; see *Results and Discussion*), we interpret the 2.7 nm spot by assuming a  $\beta$ -helical fold with six windings ( $6 \times 0.47$  nm  $\sim$  2.8 nm). This interpretation is speculative, but the proposed architecture is consistent with the STEM data, considering that the flexible termini contribute only little to the observed mass, and additionally supported by homology modeling (as detailed below). In such a  $\beta$ -helical arrangement, the observed maximum number of 18  $\beta$ -strand segments would be arranged in three segments per winding, resulting in a triangular hydrophobic core structure. Fig. 6 gives a schematic summary of the restraints obtained for the structure of BacA filaments by ssNMR, TEM, and STEM analysis.

**Homology Model of BacA Based on Crystal Structure of the Acetyltransferase GlmU.** To evaluate the conclusions drawn from the previously described analyses, we generated a homology model of BacA using the Protein Homology/analogy Recognition Engine V2.0 (Phyre2) server (45), with only the primary sequence of BacA as input information. Interestingly and in line with our ssNMR- and EM-based analysis, the resulting structural model exhibits a  $\beta$ -helical fold with six windings and a triangular hydrophobic core. The homology model (Fig. 7A and *SI Appendix*, Fig. S8) is based on the known crystal structure of the acetyltransferase GlmU from *Mycobacterium tuberculosis*



**Fig. 4.** TEM images of the uniformly [ $^{13}\text{C}$ ,  $^{15}\text{N}$ ]-labeled BacA sample. (A) Image showing an overview of different morphologies detected by TEM. (B) BacA filament bundles (arrows). (C) 2D crystalline sheets (asterisks) and single filaments (arrowheads). (D) High-resolution image of the 2D crystalline assemblies. (E) Diffraction analysis of the 2D arrangement as obtained by Fourier transform of the image shown in D. A slice through the pattern (red dotted line) is shown in F. Individual scale bars are included in the images. For clarity, C is shown in larger format in *SI Appendix*, Fig. S7.



**Fig. 5.** MPL measurements of BacA filaments by STEM. (A) Annular dark-field STEM image of BacA filament bundles. TMV particles are included for independent mass calibration. (B) A selected filament bundle with a width of  $\sim 63$  nm, comprising six individual filaments. (C) Filament segments used for the MPL analysis (boxes). The corresponding MPL values (in kDa/nm) are indicated. (D) MPL histogram fitted with a Gaussian function centered at  $4.33 \pm 0.09$  kDa/nm and with an FWHM equal to  $0.58 \pm 0.07$  kDa/nm.

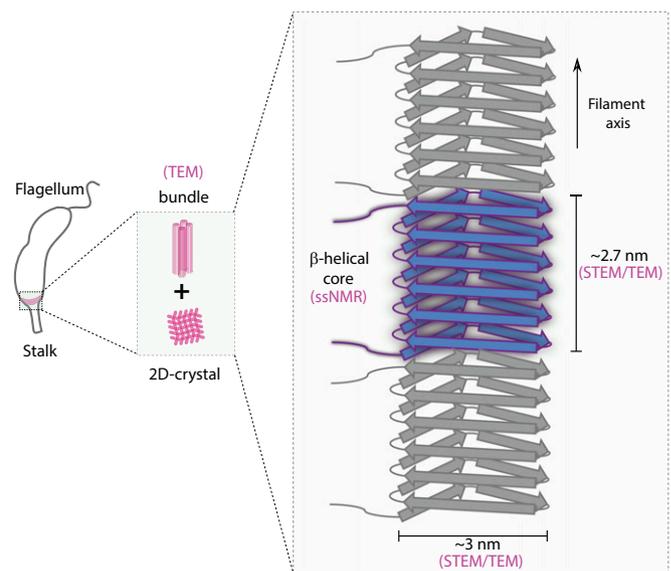
(Protein Data Bank ID code 2qkx), a bifunctional enzyme catalyzing the two last consecutive steps in the synthesis of UDP-*N*-acetylglucosamine (46). Although the template structure shows only  $\sim 13\%$  sequence identity to BacA, 86 out of 100 residues of the DUF583 domain were modeled at greater than 90% confidence. Importantly, the predicted model is fully consistent with the secondary structure elements identified by ssNMR. It is furthermore corroborated by some reliably assigned long-range and medium-range correlation peaks in an nitrogen-hydrogen-hydrogen-nitrogen (NHHC) (47, 48) spectrum (Fig. 7B) and a proton-driven spin diffusion (PDSF) spectrum recorded with long mixing time (Fig. 7C), both of which report on long-range structural information (SI Appendix, Table S4). The modeled  $\beta$ -helix measures  $\sim 2.7$  nm in length and width (including the amino acid side chains), which closely corresponds to the values obtained by TEM and STEM analysis. Notably, the highly conserved glycine residues in the DUF583 domain are all found at kink or loop positions (highlighted in blue in Fig. 7A). Collectively, these results provide strong evidence for the proposed  $\beta$ -helical conformation. Further evidence for the presence of a hydrophobic core structure comes from a systematic mutagenesis study of conserved hydrophobic residues in the DUF583 domain (see following section).

**Conserved Hydrophobic Residues in the  $\beta$ -Strands Are Critical for BacA Polymerization.** The DUF583 domain is characterized by a series of conserved bulky hydrophobic amino acids (49), all of which are located in the  $\beta$ -strand segments of the BacA filament core (Fig. 8A and SI Appendix, Fig. S9). To examine the role of these residues *in vivo*, we performed site-directed mutagenesis of the core region. Initially, we chose 15 hydrophobic residues (V52, L58, I60, V64, V68, L73, V75, V81, V85, V96, V105, V113, L122, M124, and F130) (Fig. 8A) and substituted them with alanine. As a negative control, four nonconserved charged residues located in the loop regions (E77, E88, K103, and D116)

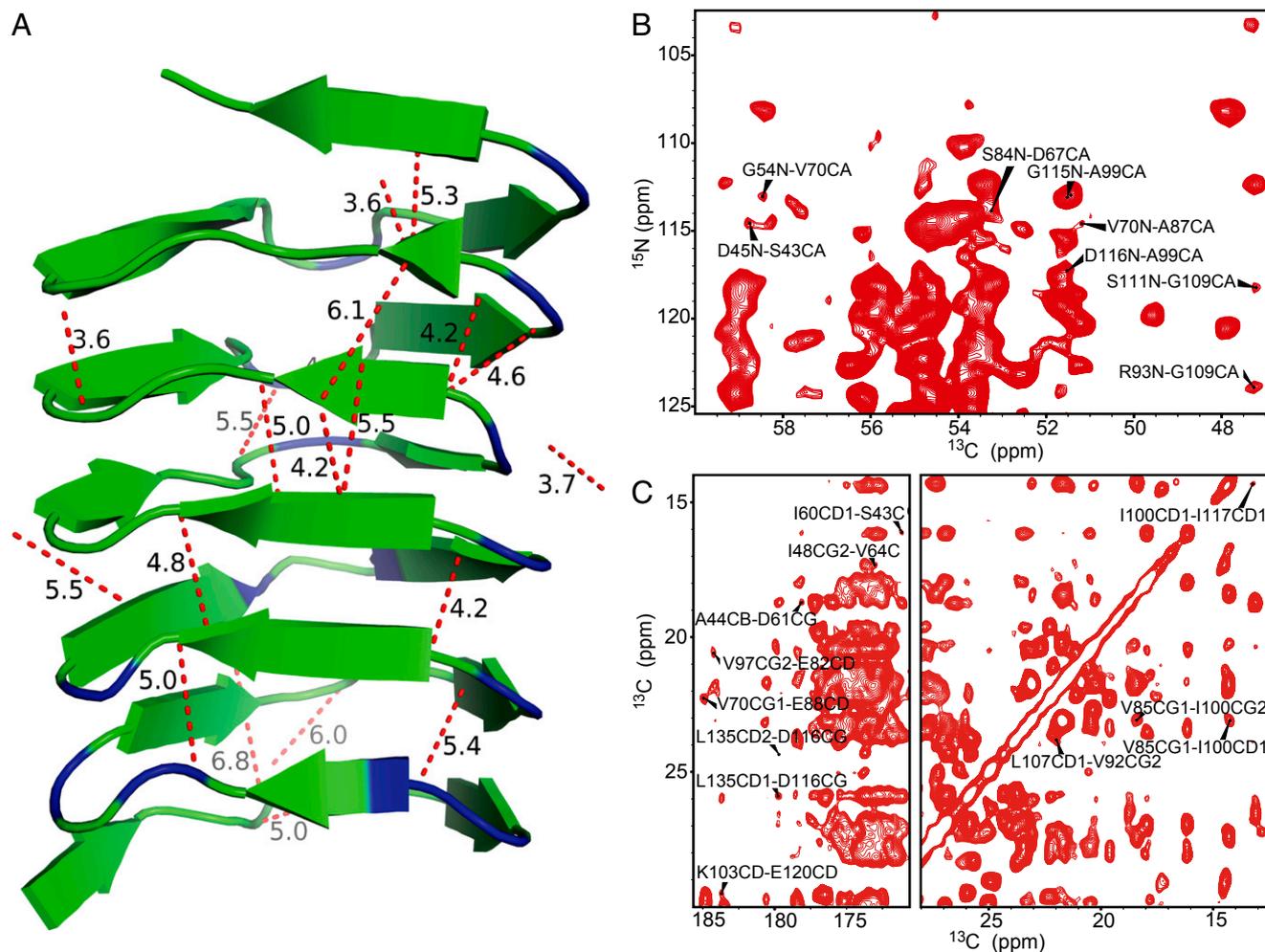
were mutated to alanine or to amino acids carrying the respective opposite charge. To visualize the effect of these amino acid substitutions on BacA assembly, we took advantage of the fact that BacA is still able to polymerize when fused to the yellow fluorescent protein Venus, forming bright foci (mostly) at the stalked pole of the *C. crescentus* cell (8). We thus produced the wild-type BacA-Venus fusion or its different mutant derivatives under the control of the xylose-inducible  $P_{xyI}$  promoter and determined their localization patterns. The majority of mutations, including all negative controls, did not affect BacA assembly (SI Appendix, Fig. S10A). However, three mutations (V75A, M124A, and F130A) in  $\beta$ -strands  $\beta_3$ ,  $\beta_9$ , and  $\beta_{10}$ , respectively, led to significantly altered localization patterns (Fig. 8B). The fusions carrying the V75A or F130A exchange were evenly distributed within the cytoplasm, indicating the absence of polymers (Fig. 8B). The M124A variant, by contrast, had a patchy subcellular distribution and still showed occasional polar complexes, which may suggest some residual polymerization activity. It was conceivable that substitution of bulky hydrophobic residues with a small hydrophobic amino acid such as alanine was in many cases not sufficient to cause a destabilizing effect. To test this possibility, a subset of residues was additionally mutated to the polar amino acid serine, which would completely disrupt potential hydrophobic interactions. For four (L73S, V75S, L122S, and M124S) out of the six mutations analyzed (L41SL42S, L73S, V75S, L122S, A123S, and M124S), the incorporation of serine led to a strong defect in BacA assembly (Fig. 8B and SI Appendix, Fig. S10B and C). Immunoblot analysis showed that all fusion proteins were stable (SI Appendix, Fig. S11), excluding the possibility that the changes in protein localization observed were due to cleavage of the fluorescent protein tag. Notably, in the homology model, the side chains of most of the conserved hydrophobic residues point into the hydrophobic core of BacA. Together, these results suggest that BacA folding and polymerization is mediated through a redundant network of hydrophobic interactions involving the conserved hydrophobic residues that characterize the DUF583 domain.

## Conclusions

We have determined the secondary structure of the filaments formed by the *C. crescentus* bactofilin homolog BacA using



**Fig. 6.** Cartoon representation of the BacA filament structure. The structural parameters that were derived from the ssNMR, TEM, and STEM data collected in this work are indicated.

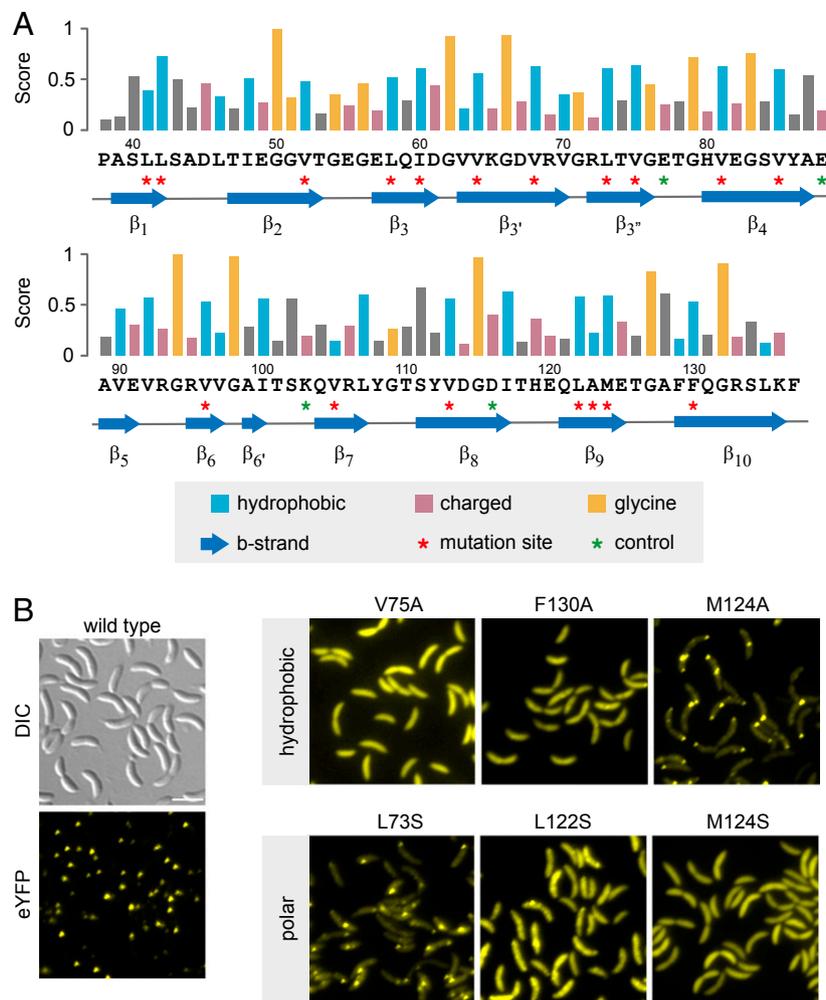


**Fig. 7.** (A) Homology model of the DUF583 domain of BacA generated using the program Phyre2. Assigned long-range and medium-range correlations are illustrated with red dashed lines (distances in Å). For better visibility, side chains are not shown explicitly, despite the fact that some of the assigned restraints also involve side chains. Glycines are shown in blue. The cartoon representation of the  $\beta$ -strands follows the ssNMR secondary structure analysis (Fig. 3). (B) 2D NHC spectrum of BacA with a 200  $\mu\text{s}$  proton spin diffusion mixing time. (C) PDS spectrum of BacA with a mixing time of 200 ms, suitable for the detection of long-range restraints. Both spectra were recorded on a 20 T (850 MHz proton Larmor frequency) wide-bore spectrometer at a magic-angle spinning rate of 11 kHz and at a temperature of 4  $^{\circ}\text{C}$ .

ssNMR. The filament core is formed by residues Ala39 to Phe137 and thus largely corresponds to the conserved DUF583 domain. It comprises  $\sim 10$   $\beta$ -strands, which are likely further interrupted by kinks at conserved glycine residues, resulting in a maximum number of 18  $\beta$ -strand segments. The proline-rich N- and C-terminal regions were found to be, to a large extent, unstructured and flexible, in agreement with the idea that they are involved in protein-protein or protein-membrane interactions. Additional restraints for the previously unknown structure come from STEM and TEM measurements. STEM yields a value for the MPL of single filaments, providing basic boundaries for structural models of bactofilin filaments. The obtained value is incompatible with a single-layer amyloid arrangement as suggested, for example, for A $\beta$ , Tau, and  $\alpha$ -synuclein fibrils, but rather suggestive of a multilayer  $\beta$ -helix as found for the prion HET-s(218–289) (two layers). The TEM data confirmed the previous findings that BacA assembles into protofilaments of about 3 nm in width, which show a high propensity to interact laterally, thereby forming large filament bundles in vitro. Intriguingly, under our experimental conditions, we also observed the formation of sheet-like 2D crystalline arrays. Previous electron cryotomographic studies of *C. crescentus* cells have suggested that BacA polymerizes into

sheet-like structures that line the cytoplasmic membrane. Thus, these 2D crystalline arrays may represent the actual arrangement of BacA polymers in the native context. Their formation may be favored in vivo through interaction of the BacA subunits with the 2D membrane surface. Diffraction data of these assemblies show a bright spot at 2.7 nm that can be interpreted in the context of a  $\beta$ -helical arrangement in which six layers of  $\beta$ -strands are stacked at a distance of 0.47 nm. In such a six-winding  $\beta$ -helix, the up to 18  $\beta$ -strands observed by ssNMR would arrange in three segments per layer, similar to the triangular hydrophobic core arrangement found for the prion HET-s(218–289). Support for this proposal comes from a homology model of BacA based on the crystal structure of the acetyltransferase GlmU, which also exhibits six windings of a  $\beta$ -helix with a triangular hydrophobic core structure. This architecture is consistent with the ssNMR-based secondary structure analysis and is further corroborated by a number of long-range correlations between residues of consecutive windings.

Bactofilins share a common domain (DUF583) that is characterized by arrays of conserved bulky hydrophobic residues. Alanine scanning mutagenesis showed that some of these residues are essential for proper BacA assembly, although the majority



**Fig. 8.** Mutational analysis of the BacA filament core. (A) Sequence of the BacA DUF583 domain. Numbers give the position of residues in the primary sequence of *C. crescentus* BacA. The bar graph indicates the conservation scores of individual residues, derived from an alignment of 150 representative bactofilin homologs, ranging from 0 (no conservation) to 1 (complete conservation). The  $\beta$ -strands identified in this study are shown as blue arrows. Asterisks denote residues that were targeted by site-directed mutagenesis. (B) Localization of different BacA–Venus variants in *C. crescentus*. Cells of strain JK5 ( $\Delta bacAB$ ) carrying the indicated alleles of *bacA-venus* under the control of the xylose-inducible *P<sub>xyI</sub>* promoter were grown to late exponential phase and diluted to an OD<sub>600</sub> of ~0.1. After incubation for another hour, the cells were induced with 0.005% xylose for 1 h before visualization by differential interference contrast (DIC) and fluorescence microscopy. Note that wild-type BacA occasionally forms additional, nonpolar foci due to slight overexpression of the fusion protein. (Scale bar, 3  $\mu$ m.)

of amino acid substitutions did not significantly change the localization behavior of the protein. A more pronounced effect was, however, observed after mutation of conserved residues to the polar amino acid serine, which fully disrupts potential hydrophobic interactions. This finding supports the idea that contacts between conserved hydrophobic residues are critical for polymer formation. These interactions may either stabilize the tertiary structure of the bactofilin DUF583 domain or mediate interaction between adjacent subunits in the polymer. In addition, they could potentially contribute to the membrane attachment of BacA, which may facilitate polymerization by increasing the local protein concentration at the membrane surface. Previous work has shown that BacA levels do not vary over the course of the *C. crescentus* cell cycle. Nevertheless, polar complexes are only observed after the start of S-phase, suggesting the existence of mechanisms that tightly control the polymerization process. It will be interesting to determine the molecular principles underlying bactofilin folding and polymerization and investigate the determinants that control its positioning of bactofilin polymers within the cell.

## Materials and Methods

**Preparation of Uniformly [<sup>13</sup>C, <sup>15</sup>N]-Labeled BacA Filaments.** Uniformly [<sup>13</sup>C, <sup>15</sup>N]-labeled BacA–His<sub>6</sub> was purified essentially as described (8). Briefly, plasmid pMT879 was transferred into *E. coli* Rosetta (DE3)/pLysS (Invitrogen), and a single transformant was grown at 37 °C in 3 L of [<sup>13</sup>C, <sup>15</sup>N]-labeled *E. coli* OD 2 complete medium (Silantes) supplemented with appropriate antibiotics. Overexpression was induced by the addition of 0.5 mM IPTG. After 3 more hours of cultivation, the culture was harvested by centrifugation at 4,500  $\times$  *g* and 4 °C for 10 min. The cells were washed with buffer B1 (50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, 10 mM imidazole, pH 8.0, adjusted with NaOH), resuspended in 60 mL buffer B2 (50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, 10 mM imidazole, 1 mM  $\beta$ -mercaptoethanol, pH 8.0, adjusted with NaOH) supplemented with 10  $\mu$ g/mL DNase I and 100  $\mu$ g/mL phenylmethylsulfonyl fluoride, and lysed by three passages through a French press (16,000 psi). After the removal of cell debris by centrifugation at 20,000  $\times$  *g* for 30 min, the cleared lysates were mixed with 8 mL Ni-NTA agarose beads (Qiagen) equilibrated with buffer B2 and incubated with gentle agitation for 2 h at 4 °C. The beads were then collected by centrifugation at 3,000  $\times$  *g* for 5 min and washed four times with buffer B3 (50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, 20 mM imidazole, 1 mM  $\beta$ -mercaptoethanol, pH 8.0, adjusted with NaOH). Proteins were finally eluted with 48 mL buffer B4 (50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, 250 mM imidazole, 1 mM  $\beta$ -mercaptoethanol, pH 8.0, adjusted with NaOH). After centrifugation at 3,000  $\times$  *g* for 10 min, the elution fractions

were dialyzed against 4 L of buffer B5 (50 mM sodium phosphate, 5 mM MgCl<sub>2</sub>, 0.1 mM EDTA, 1 mM β-mercaptoethanol, pH 7) containing 50 mM NaCl. The dialyzed protein was then applied to a HiPrep 16/10 SP FF cation exchange column (GE Healthcare) that had been preequilibrated with buffer B5 containing 50 mM NaCl. The column was washed with 75 mL buffer B5 containing 50 mM NaCl, and protein was eluted with 250 mL of a linear NaCl gradient (50–500 mM NaCl in buffer B5). Fractions containing pure BacA-His<sub>6</sub> were combined and dialyzed against two changes of 3 L buffer B6 [50 mM Hepes/NaOH pH 7.2, 50 mM NaCl, 5 mM MgCl<sub>2</sub>, 0.1 mM EDTA, 10% (vol/vol) glycerol, 1 mM β-mercaptoethanol]. The solution was snap-frozen in liquid N<sub>2</sub> and stored at –80 °C until further use.

**ssNMR Spectroscopy and Analysis.** Purified BacA-His<sub>6</sub> was collected by ultracentrifugation at 540,000 × *g* and 4 °C for 6–8 h. The pelleted protein was dissolved in 20 mM Tris-HCl (pH 7.5) and collected by ultracentrifugation again. The pellets were then transferred into magic-angle spinning (MAS) rotors and subjected to ssNMR analysis. All spectra were recorded on either a Bruker Avance II 850 MHz wide-bore or an 800 MHz narrow-bore spectrometer equipped with 3.2 and 4 mm triple-resonance (<sup>1</sup>H, <sup>13</sup>C, <sup>15</sup>N) MAS probe heads. The sample temperature was kept at around 4 °C, as measured by the temperature-dependent position of the water resonance peak with respect to internal DSS. The chemical shift referencing of all spectra was done with respect to the DSS proton chemical shift as an internal reference. All of the spectra were processed with TOPSPIN and analyzed using CcpNMR (50). A complete list of ssNMR experiments and the corresponding experimental parameters is provided in *SI Appendix, Table S1*.

**TEM.** TEM images of negatively stained BacA polymers were acquired using a commercial Philips 100 kV BioTwin microscope equipped with a 1024 × 1024 pixel GATAN CCD camera. Imaging was performed on the same samples that were used for the ssNMR experiments.

**STEM Measurements.** STEM analyses for accurate determination of the MPL of BacA filaments were carried out at Brookhaven National Laboratory. The measurements were performed on unstained freeze-dried protein at a temperature of –160 °C. Mass calibration was achieved by means of TMV particles that were added to the same sample. A correction was made for mass loss, resulting mainly from the high dose of electron exposure, by adjusting the MPL value of TMV to 131 kDa/nm. Image analysis was performed using the PCmass software (51).

**Bacterial Growth.** *E. coli* was grown at 37 °C in LB medium (Carl Roth), whereas *C. crescentus* was cultivated at 28 °C in peptone-yeast extract (PYE) medium (52). When required, antibiotics were added at the following concentrations

(in μg/mL in liquid/solid medium): ampicillin (100/200), chloramphenicol (20/30), kanamycin (30/50) for *E. coli*, and kanamycin (5/25) for *C. crescentus*. To activate the *C. crescentus* *P*<sub>xyI</sub> promoter, media were supplemented with 0.005% or 0.03% xylose.

**Plasmid and Strain Construction.** The plasmids and strains used in this study are listed in *SI Appendix, Tables S5 and S6*, respectively. Mutant *bacA* alleles were generated by PCR-based site-directed mutagenesis, using the Quik-Change protocol (Stratagene) and plasmid pMT879 (8) as a template. The mutated genes were isolated by restriction of the resulting plasmids with NdeI and EcoRI and cloned into equally treated pXVENC-2 (53). All constructs were verified by sequence analysis and integrated via single homologous recombination at the *xyI/X* locus of strains CB15N (wild type) or JK5 ( $\Delta$ *bacA*B) to allow expression of the different fluorescent protein gene fusions under the control of the xylose-inducible *xyI/X* promoter (54).

**Microscopic Analysis.** Cells were immobilized on pads made of 1.5% agarose in H<sub>2</sub>O and imaged using a Zeiss Axio Imager.Z1 microscope equipped with a Zeiss 100×/1.46 Oil DIC Plan-Apochromat objective and a pco.edge sCMOS camera (PCO) or a CoolSNAP-ES2 CCD camera (Photometrics). An X-Cite120PC metal halide light source (EXFO) and an ET-YFP filter cube (Chroma) were used for fluorescence detection. Images were processed using Metamorph 7.7 (Molecular Devices).

**Immunoblot Analysis.** Immunoblot analysis was performed as described (55), using a monoclonal anti-BacA antibody (8) at a dilution of 1:10,000.

**Bioinformatic Analysis.** The boundaries of the BacA bactofilin/DUF583 domain were obtained from the PFAM (protein families) database (49). To calculate conservation scores, the seed alignment forming the basis of the profile hidden Markov model of the DUF583 domain was retrieved from the PFAM database and analyzed using the Scorecons server (scoring method, entropic, 21 types) (56).

**ACKNOWLEDGMENTS.** We thank Julia Rosum and Nicole Schnaß for help with the mutational analysis of BacA and Joe Wall for collecting the scanning transmission electron microscopy (STEM) data. This work was supported by the Max Planck Society, the Leibniz-Institut für Molekulare Pharmakologie, the German Research Foundation (DFG) Emmy Noether Fellowship (to A.L.), Human Frontier Science Program Young Investigator Grant RGY0076/2013 (to M.T.), and European Union Seventh Framework Program Grant Agreement 261863 (Bio-NMR). C.S. acknowledges funding from the Max Planck Society and Chinese Academy of Sciences Joint Doctoral Promotion Program. L.L. acknowledges funding from DFG Research Training Group GRK 1216.

- Bi EF, Lutkenhaus J (1991) FtsZ ring structure associated with division in *Escherichia coli*. *Nature* 354(6349):161–164.
- van den Ent F, Amos LA, Löwe J (2001) Prokaryotic origin of the actin cytoskeleton. *Nature* 413(6851):39–44.
- Ausmees N, Kuhn JR, Jacobs-Wagner C (2003) The bacterial cytoskeleton: An intermediate filament-like function in cell shape. *Cell* 115(6):705–713.
- Cabeen MT, Jacobs-Wagner C (2010) The bacterial cytoskeleton. *Annu Rev Genet* 44:365–392.
- Gerdes K, Howard M, Szardenings F (2010) Pushing and pulling in prokaryotic DNA segregation. *Cell* 141(6):927–942.
- Lin L, Thanbichler M (2013) Nucleotide-independent cytoskeletal scaffolds in bacteria. *Cytoskeleton (Hoboken)* 70(8):409–423.
- Ozaymak E, Kollman JM, Komeili A (2013) Bacterial actins and their diversity. *Biochemistry* 52(40):6928–6939.
- Kühn J, et al. (2010) Bactofilins, a ubiquitous class of cytoskeletal proteins mediating polar localization of a cell wall synthase in *Caulobacter crescentus*. *EMBO J* 29(2):327–339.
- Klein EA, et al. (2013) Physiological role of stalk lengthening in *Caulobacter crescentus*. *Commun Integr Biol* 6(4):e24561.
- Hughes HV, et al. (2013) Co-ordinate synthesis and protein localization in a bacterial organelle by the action of a penicillin-binding-protein. *Mol Microbiol* 90(6):1162–1177.
- Koch MK, McHugh CA, Hoiczky E (2011) BacM, an N-terminally processed bactofilin of *Mycococcus xanthus*, is crucial for proper cell shape. *Mol Microbiol* 80(4):1031–1051.
- Bulyha I, et al. (2013) Two small GTPases act in concert with the bactofilin cytoskeleton to regulate dynamic bacterial cell polarity. *Dev Cell* 25(2):119–131.
- Sycuro LK, et al. (2010) Peptidoglycan crosslinking relaxation promotes *Helicobacter pylori*'s helical shape and stomach colonization. *Cell* 141(5):822–833.
- Hay NA, Tipper DJ, Gygi D, Hughes C (1999) A novel membrane protein influencing cell shape and multicellular swarming of *Proteus mirabilis*. *J Bacteriol* 181(7):2008–2016.
- Goldbourt A, Gross BJ, Day LA, McDermott AE (2007) Filamentous phage studied by magic-angle spinning NMR: Resonance assignment and secondary structure of the coat protein in Pf1. *J Am Chem Soc* 129(8):2338–2344.
- Jehle S, et al. (2010) Solid-state NMR and SAXS studies provide a structural basis for the activation of alphaB-crystallin oligomers. *Nat Struct Mol Biol* 17(9):1037–1042.
- Loquet A, et al. (2012) Atomic model of the type III secretion system needle. *Nature* 486(7402):276–279.
- Lu JX, et al. (2013) Molecular structure of β-amyloid fibrils in Alzheimer's disease brain tissue. *Cell* 154(6):1257–1268.
- Lv G, et al. (2012) Structural comparison of mouse and human α-synuclein amyloid fibrils by solid-state NMR. *J Mol Biol* 420(1–2):99–111.
- Wasmer C, et al. (2008) Amyloid fibrils of the HET-s(218–289) prion form a beta-solenoid with a triangular hydrophobic core. *Science* 319(5869):1523–1526.
- Bertini I, Gonnelli L, Luchinat C, Mao J, Nesi A (2011) A new structural model of Aβ40 fibrils. *J Am Chem Soc* 133(40):16013–16022.
- Andronesi OC, et al. (2005) Determination of membrane protein structure and dynamics by magic-angle-spinning solid-state NMR spectroscopy. *J Am Chem Soc* 127(37):12965–12974.
- Cady SD, et al. (2010) Structure of the amantadine binding site of influenza M2 proton channels in lipid bilayers. *Nature* 463(7281):689–692.
- Wang S, et al. (2013) Solid-state NMR spectroscopy structure determination of a lipid-embedded heptahelical membrane protein. *Nat Methods* 10(10):1007–1012.
- McDermott A (2009) Structure and dynamics of membrane proteins by magic angle spinning solid-state NMR. *Annu Rev Biophys* 38:385–403.
- Hu F, Luo W, Hong M (2010) Mechanisms of proton conduction and gating in influenza M2 proton channels from solid-state NMR. *Science* 330(6003):505–508.
- Sengupta I, Nadaud PS, Helmus JJ, Schwieters CD, Jaroniec CP (2012) Protein fold determined by paramagnetic magic-angle spinning solid-state NMR spectroscopy. *Nat Chem* 4(5):410–417.
- Franks WT, et al. (2008) Dipole tensor-based atomic-resolution structure determination of a nanocrystalline protein by solid-state NMR. *Proc Natl Acad Sci USA* 105(12):4621–4626.
- Ritter C, et al. (2005) Correlation of structural elements and infectivity of the HET-s prion. *Nature* 435(7043):844–848.
- Demers JP, et al. (2013) The common structural architecture of *Shigella flexneri* and *Salmonella typhimurium* type three secretion needles. *PLoS Pathog* 9(3):e1003245.

31. Loquet A, Lv G, Giller K, Becker S, Lange A (2011)  $^{13}\text{C}$  spin dilution for simplified and complete solid-state NMR resonance assignment of insoluble biological assemblies. *J Am Chem Soc* 133(13):4722–4725.
32. Schuetz A, et al. (2010) Protocols for the sequential solid-state NMR spectroscopic assignment of a uniformly labeled 25 kDa protein: HET-s(1-227). *ChemBioChem* 11(11):1543–1551.
33. Chevelkov V, Giller K, Becker S, Lange A (2013) Efficient CO-CA transfer in highly deuterated proteins by band-selective homonuclear cross-polarization. *J Magn Reson* 230:205–211.
34. Chevelkov V, Shi C, Fasshuber HK, Becker S, Lange A (2013) Efficient band-selective homonuclear CO-CA cross-polarization in protonated proteins. *J Biomol NMR* 56(4):303–311.
35. Shi C, et al. (2014) BSH-CP based 3D solid-state NMR experiments for protein resonance assignment. *J Biomol NMR* 59(1):15–22.
36. Luca S, et al. (2001) Secondary chemical shifts in immobilized peptides and proteins: A qualitative basis for structure refinement under magic angle spinning. *J Biomol NMR* 20(4):325–331.
37. Shen Y, Delaglio F, Cornilescu G, Bax A (2009) TALOS+: A hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR* 44(4):213–223.
38. Lange A, et al. (2009) A combined solid-state NMR and MD characterization of the stability and dynamics of the HET-s(218-289) prion in its amyloid conformation. *ChemBioChem* 10(10):1657–1665.
39. Morris GA, Freeman R (1979) Enhancement of nuclear magnetic-resonance signals by polarization transfer. *J Am Chem Soc* 101(3):760–762.
40. Baldus M, Meier BH (1996) Total correlation spectroscopy in the solid state. The use of scalar couplings to determine the through-bond connectivity. *J Magn Reson A* 121(1):65–69.
41. Daebel V, et al. (2012)  $\beta$ -sheet core of tau paired helical filaments revealed by solid-state NMR. *J Am Chem Soc* 134(34):13982–13989.
42. Heise H, et al. (2005) Molecular-level secondary structure, polymorphism, and dynamics of full-length alpha-synuclein fibrils studied by solid-state NMR. *Proc Natl Acad Sci USA* 102(44):15871–15876.
43. Helmus JJ, Surewicz K, Surewicz WK, Jaroniec CP (2010) Conformational flexibility of Y145Stop human prion protein amyloid fibrils probed by solid-state nuclear magnetic resonance spectroscopy. *J Am Chem Soc* 132(7):2393–2403.
44. Chen B, Thurber KR, Shewmaker F, Wickner RB, Tycko R (2009) Measurement of amyloid fibril mass-per-length by tilted-beam transmission electron microscopy. *Proc Natl Acad Sci USA* 106(34):14339–14344.
45. Kelley LA, Sternberg MJE (2009) Protein structure prediction on the Web: A case study using the Phyre server. *Nat Protoc* 4(3):363–371.
46. Zhang Z, Bulloch EMM, Bunker RD, Baker EN, Squire CJ (2009) Structure and function of GlmU from *Mycobacterium tuberculosis*. *Acta Crystallogr D Biol Crystallogr* 65(Pt 3):275–283.
47. Lange A, Luca S, Baldus M (2002) Structural constraints from proton-mediated rare-spin correlation spectroscopy in rotating solids. *J Am Chem Soc* 124(33):9704–9705.
48. Lange A, Seidel K, Verdier L, Luca S, Baldus M (2003) Analysis of proton-proton transfer dynamics in rotating solids and their use for 3D structure determination. *J Am Chem Soc* 125(41):12640–12648.
49. Finn RD, et al. (2014) Pfam: The protein families database. *Nucleic Acids Res* 42(Database issue):D222–D230.
50. Stevens TJ, et al. (2011) A software framework for analysing solid-state MAS NMR data. *J Biomol NMR* 51(4):437–447.
51. Wall JS, Hainfeld JF (1986) Mass mapping with the scanning transmission electron microscope. *Annu Rev Biophys Biophys Chem* 15:355–376.
52. Poindexter JS (1964) Biological properties and classification of the *Caulobacter* group. *Bacteriol Rev* 28(3):231–295.
53. Thanbichler M, Iniesta AA, Shapiro L (2007) A comprehensive set of plasmids for vanillate- and xylose-inducible gene expression in *Caulobacter crescentus*. *Nucleic Acids Res* 35(20):e137.
54. Meisenzahl AC, Shapiro L, Jenal U (1997) Isolation and characterization of a xylose-dependent promoter from *Caulobacter crescentus*. *J Bacteriol* 179(3):592–600.
55. Thanbichler M, Shapiro L (2006) MipZ, a spatial regulator coordinating chromosome segregation with cell division in *Caulobacter*. *Cell* 126(1):147–162.
56. Valdar WS (2002) Scoring residue conservation. *Proteins* 48(2):227–241.