Structural changes of TasA in biofilm formation of *Bacillus subtilis*

Anne Diehl[a,1], Yvette Roske[b,1], Linda Ball[a], Anup Chowdhury[a], Matthias Hiller[a], Noel Molière[c], Regina Kramer[c], Daniel Stöppler[a,d], Catherine L. Worth[a], Brigitte Schlegel[a], Martina Leidert[a], Nils Cremer[a], Natalja Erdmann[a], Daniel Lopez[e], Heike Stephanowitz[a], Eberhard Krause[a], Barth-Jan van Rossum[a], Peter Schmieder[a], Udo Heinemann[b,d,2], Kürşad Turgay[c,2], Ümit Akbey[a,f], and Hartmut Oschkinat[a,d,2]

[a]Leibniz-Forschungsinstitut für Molekulare Pharmakologie, 13125 Berlin, Germany; [b]Max-Delbrück-Center for Molecular Medicine, 13125 Berlin, Germany; [c]Institut für Mikrobiologie, Leibniz Universität Hannover, 30419 Hannover, Germany; [d]Institut für Chemie und Biochemie, Freie Universität Berlin, 14195 Berlin, Germany; [e]Centro Nacional de Biotecnologia, Universidad Autonoma de Madrid, Campus de Cantoblanco, 28049 Madrid, Spain; and [f]Aarhus Institute of Advanced Studies, Aarhus University, 8000 Aarhus C, Denmark

Edited by Scott J. Hultgren, Washington University School of Medicine, St. Louis, MO, and approved February 23, 2018 (received for review October 17, 2017)

Microorganisms form surface-attached communities, termed biofilms, which can serve as protection against host immune reactions or antibiotics. *Bacillus subtilis* biofilms contain TasA as major proteinaceous component in addition to exopolysaccharides. In stark contrast to the initially unfolded biofilm proteins of other bacteria, TasA is a soluble, stably folded monomer, whose structure we have determined by X-ray crystallography. Subsequently, we characterized in vitro different oligomeric forms of TasA by NMR, EM, X-ray diffraction, and analytical ultracentrifugation (AUC) experiments. However, by magic-angle spinning (MAS) NMR on live biofilms, a swift structural change toward only one of these forms, consisting of homogeneous and protease-resistant, β-sheet–rich fibrils, was observed in vivo. Thereby, we characterize a structural change from a globular state to a fibrillar form in a functional prokaryotic system on the molecular level.

biofilm | TasA | *Bacillus subtilis* | structure | NMR

Bacteria-forming biofilms are commonly embedded in an extracellular matrix that consists of secreted proteins that form fibrils, exopolysaccharides (EPS), and sometimes extracellular DNA (1). Additional protein components with specific functions such as hydrophobins have recently been identified and characterized (2–6). Bacteria in biofilms are generally more resistant to environmental stress and less susceptible to antibiotics; hence infections associated with biofilm formation are more difficult to treat (7).

In combination with EPS, TasA is a major component of *Bacillus subtilis* biofilms (8, 9). It was first described as a spore-associated protein with antibacterial activity (10, 11) and differs from other biofilm-forming proteins by the apparent lack of repeats in its sequence. TasA preparations from the supernatant of *B. subtilis* mutant cells lacking EPS but highly expressing TasA yield oligomeric forms with an average molecular mass of 600 kDa (8, 9). These TasA preparations fibrillize on hydrophobic surfaces and at low pH, suggesting that TasA can form fibrils that stabilize the *B. subtilis* biofilm (8, 9, 12, 13). There is also a minor biofilm protein component, TapA, which may be responsible for anchoring the fibers onto the bacterial cell wall and participating in TasA fiber formation, albeit to a lesser extent (at a 1:100 ratio relative to TasA) (14, 15).

In bacilli, biofilm formation is genetically controlled by the regulatory repressor and anti-repressor proteins, SinI and SinR (8, 16–21). In *B. subtilis*, the SinR controlled operon contains the gene triad *tapA–sipW–tasA* (13, 16, 17). The membrane-bound peptidase SipW cleaves the signal peptides of TapA and TasA before secretion (22, 23). Intriguingly, in pathogenic *Bacillus cereus* or *Bacillus anthracis* strains, the *tasA* operon contains a SipW homolog together with *tasA*-like genes named *calY1* and *calY2*. The respective proteins are members of the zinc-dependent M73 metalloproteinase family, usually termed camelysins (24, 25), and are considered important for the pathogenicity of *B. cereus* and *B. anthracis* (26–29). Double deletion of both *calY* loci in *B. cereus* leads to defects in biofilm formation (24, 25).

In this work, we present a high-resolution crystal structure of soluble, monomeric TasA in its mature secreted form. Despite its apparent homology to camelysins, biochemical experiments suggest that TasA is not an active protease. As a basis for understanding the structural changes occurring during fiber and biofilm formation, the monomer and multimeric forms were investigated in vitro and in vivo by NMR, analytical ultracentrifugation (AUC), and other biophysical techniques. In particular, we analyzed in vitro the transformation of soluble monomeric TasA into two different states, a gel-like form and fibrils. Magic-angle spinning (MAS) NMR applied to biofilms that were generated by adding soluble, monomeric and isotope-labeled TasA to the medium of *B. subtilis* Δ*tasA* cultures allowed probing of the in vivo situation, revealing the formation of homogeneous TasA fibers as the major proteinaceous extracellular matrix component. Thereby, we characterized the transition of folded

Significance

Understanding the formation and structure of protective bacterial biofilms will help to design and identify antimicrobial strategies. Our experiments with the secreted major biofilm protein TasA characterize on a molecular level in vivo the transition of a folded protein into protease-resistant biofilm-stabilizing fibrils. Such conformational changes from a globular state into fibrillar structures are so far not seen for other biofilm-forming proteins. In this context, TasA can serve as a model system to study functional fibril formation from a globular state.

Author contributions: A.D., Y.R., D.L., U.H., K.T., Ü.A., and H.O. designed research; A.D., Y.R., L.B., A.C., M.H., N.M., R.K., D.S., C.L.W., B.S., M.L., N.C., N.E., D.L., H.S., P.S., and Ü.A. performed research; H.S. and E.K. contributed new reagents/analytic tools; A.D., Y.R., L.B., A.C., M.H., N.M., R.K., D.S., C.L.W., B.S., M.L., N.C., N.E., D.L., H.S., E.K., B.-J.v.R., P.S., U.H., K.T., Ü.A., and H.O. analyzed data; and A.D., Y.R., B.-J.v.R., U.H., K.T., Ü.A., and H.O. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The atomic coordinates of $TasA_{239}$ in the presence and absence of salicylate have been deposited in the Protein Data Bank, www.wwpdb.org (PDB ID codes 5OF1 and 5OF2, respectively).

[1]A.D. and Y.R. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: heinemann@mdc-berlin.de, turgay@ifmb.uni-hannover.de, or oschkinat@fmp-berlin.de.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1718102115/-/DCSupplemental.

Published online March 12, 2018.

TasA into fibrils, both in vitro and in its natural biofilm environment on a molecular level.

Results

The Structure of Monomeric TasA Investigated by X-Ray Crystallography and NMR Reveals a Jellyroll Fold and Flexible PP-Helices. Two different TasA constructs were prepared for structural and biophysical studies (Fig. 1*A*). Mature $TasA_{261}$ (amino acids 28–261) is found outside the cell after cleavage of the N-terminal signal sequence (amino acids 1–27). $TasA_{239}$ comprises the core domain and ranges from amino acids 28 to 239 as suggested by standard bioinformatic analysis (*SI Appendix*, Fig. S1). According to gel filtration chromatography (Fig. 1*B*), both proteins are predominantly monomeric when freshly prepared at pH 7, with $TasA_{261}$ showing a higher tendency to form oligomers than $TasA_{239}$.

Furthermore, the solution NMR ^{15}N–^{1}H heteronuclear single-quantum coherence (HSQC) spectrum (*SI Appendix*, Fig. S2*A*) of $TasA_{261}$ revealed a globular fold, and the cross peaks of 220 residues were assigned. The C terminus in $TasA_{261}$ is unstructured according to the random coil chemical shifts associated with the cross peaks of the respective residues (Fig. 1*C* and *SI Appendix*, Fig. S2 *B* and *C*) and according to bioinformatics (see above). For residues 115 and 116, and in the region 174–177, peak doubling is observed in the ^{15}N–^{1}H HSQC spectra of both forms (*SI Appendix*, Fig. S2 *B* and *C*). Both $TasA_{261}$ and $TasA_{239}$ were prone to slow degradation (*SI Appendix*, Fig. S3*A*) requiring the use of protease inhibitors. The protein unfolds in the presence of 1% SDS (*SI Appendix*, Fig. S3*B*). Selenomethionine (SeMet) incorporated $TasA_{239}$ was crystallized at pH 4.6 and a structure obtained at 1.56-Å resolution by anomalous-diffraction phasing (Fig. 2 and *SI Appendix*, Table S1).

The structure of TasA in its globular monomeric form consists of a jellyroll fold composed of two antiparallel β-sheets flanked by six, in part very short helices, and longer loop regions (Fig. 2 *A* and *B*). Intriguingly, the structure shows areas of varying rigidity. In general, *B* values (Fig. 2*C*) are much higher in the lower area (referring to the orientation shown in Fig. 2 *A* and *C*; residues 30–50, 73–78, and 186–189) than in other parts of the structure, and no density was observed for residues 117–126; henceforth, this area is called the "dynamic section." The region spatially close to residues 117–126 is very well ordered and shows low *B* factors. On the other hand, multiple NMR signals were observed for residues T115, V116, and the stretch 174–177 of both $TasA_{239}$ and $TasA_{261}$.

Notably, the neighboring loop 117–126 contains a proline (P125); hence, a potential *cis-trans* isomerization in this section could cause the doubling of signals by producing two chemically different environments. The dynamic section also comprises segments with polyproline II (PPII) helical structure. One PPII helix (amino acids TPTDFD, residues 187–192) is located directly before β8; another short PPII helix (amino acids ASG, 40–42) is observed at the N terminus close to β1 (Fig. 2*A*, frames 1 and 2). These PPII helices display marked flexibility, although they are stabilized by water molecules forming a regular network of hydrogen bonds with their backbone carbonyl oxygens (Fig. 2*A*, frames 1 and 2).

The electrostatic surface potential of TasA (Fig. 2*D*) reveals two large, negatively charged patches, one located in the area with the flexible turns (at *Bottom* in Fig. 2*A*) and the second located on the surface of sheet 2. Residues anchored in the strands forming sheet 1 create a positively charged surface (Fig. 2*D*). A salicylate molecule is bound in a hydrophobic pocket between the two β-sheets (Fig. 2*A* and *SI Appendix*, Fig. S4 *A* and *B*). Wild-type $TasA_{239}$, crystallized in the absence of salicylate (*SI Appendix*, Table S1), showed an ethylene glycol molecule bound in the hydrophobic pocket, and three others in loop regions (*SI Appendix*, Fig. S4 *A* and *C*), with small structural overall differences (backbone r.m.s.d. of 0.124 Å; *SI Appendix*, Fig. S4*A*).

In summary, the presented structure of monomeric soluble TasA shows sections with an intriguing difference in dynamics and will guide experiments geared at the understanding of conformational changes before fibril formation.

Structural Homology to Camelysin Metalloproteases. A comparison of our structure with structures in the Protein Data Bank (PDB) using DALI (30) identified distant structural homologs that share the jellyroll topology (e.g., PDB ID codes 3WIN, 4DY5, 4AQB, 4DXZ, and 1SPP) and that are involved in a large variety of biological functions. The proposed structural similarity with camelysins was demonstrated by modeling (*SI Appendix*, Fig. S5*A*) after producing a structure-based sequence alignment (*SI Appendix*, Fig. S5*B*).

However, protease activity was not found for TasA and the binding of divalent cations (17, 31) was negligible according to isothermal titration calorimetry, NMR, and crystallization experiments (Fig. 2*E* and *SI Appendix*, Fig. S4 *D*–*F*), in contrast to reports for camelysin (27, 28). *B. subtilis* cells are known to

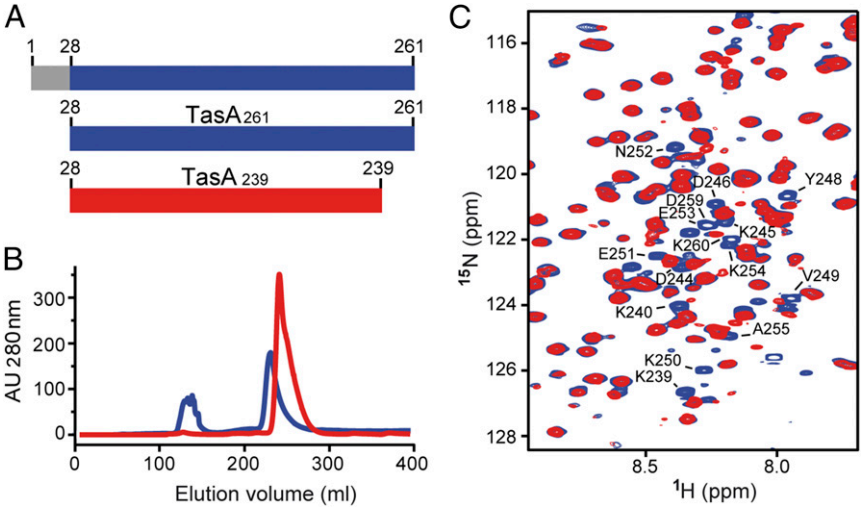

Fig. 1. (*A*) Representation of TasA (gray, signal sequence, residues 1–27; blue, secreted TasA, residues 28–261) and the two recombinant proteins investigated in this study, $TasA_{239}$ (red) and $TasA_{261}$ (blue). (*B*) Gel filtration profiles of $TasA_{239}$ (red) and $TasA_{261}$ (blue). (*C*) Zoom into a superposition of the solution ^{15}N–^{1}H correlation NMR spectra of $TasA_{239}$ (red) and $TasA_{261}$ (blue).

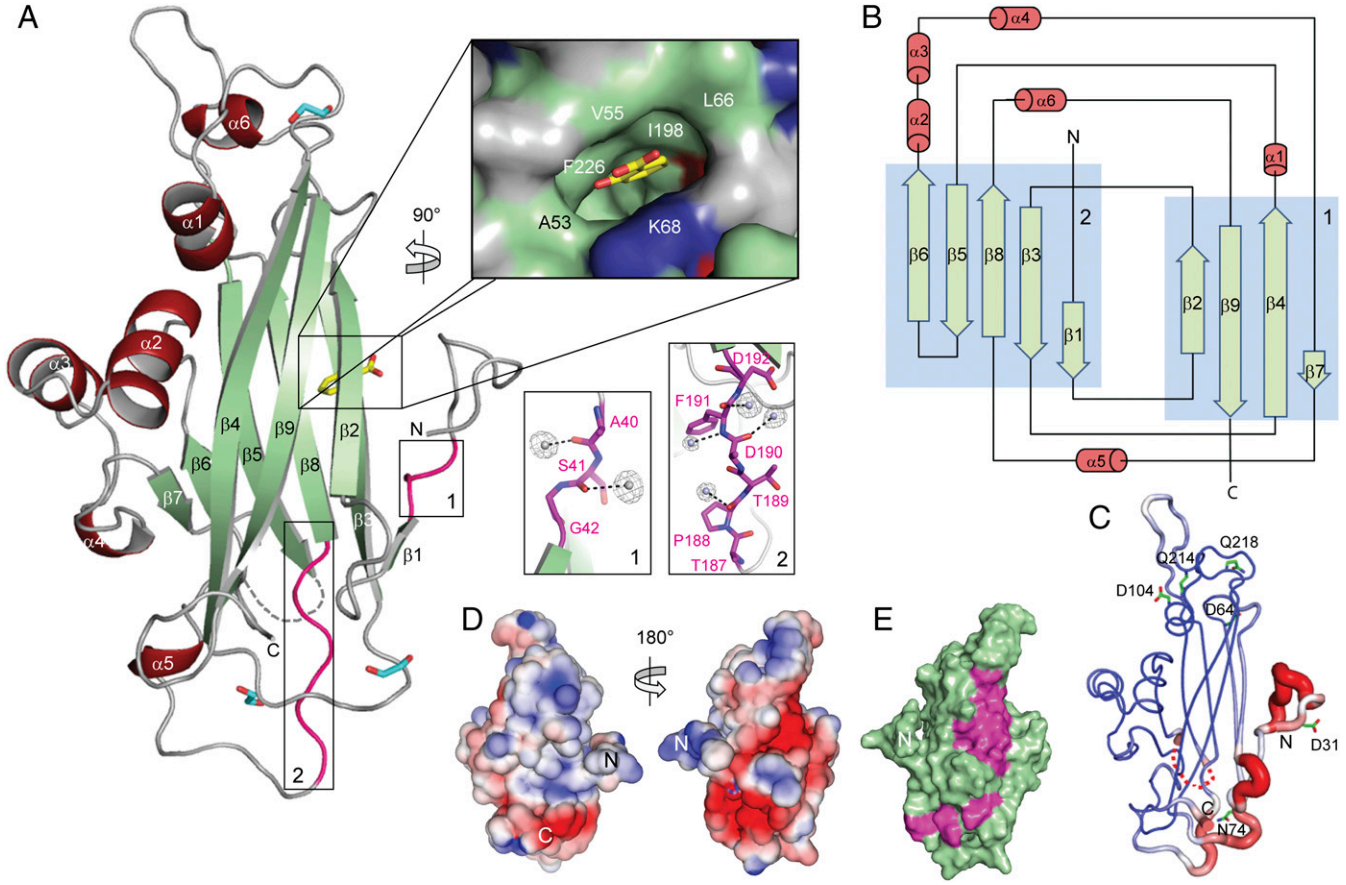

Fig. 2. (*A*) Overall crystal structure of SeMet-TasA$_{239}$ in cartoon representation. An undefined region (amino acids 117–125) is indicated by a dotted line. Bound salicylate (yellow) and ethylene glycol (cyan) molecules are depicted as sticks and with oxygen atoms in red. The magnification at *Top Right* shows the hydrophobic pocket in surface representation with the bound salicylate. Aromatic residues are colored in green, polar residues in gray, and positively charged residues in blue. The two regions forming polyproline helices (PPII) are highlighted in magenta and magnified to the *Right* in frames 1 and 2. (*B*) Secondary-structure topology. β-Strands forming the β-sheets 1 and 2 (light blue boxes) are shown as arrows and helices as red cylinders. (*C*) TasA orientated as in *A* with *B* factors indicated by ribbon thickness and color (blue indicates low and red indicates high *B* values). The amino acid regions 30–50, 73–78, and 186–189 show higher *B*-factor values. (*D*) Electrostatic surface potential of TasA in two orientations. The color scale is set from −4 kT/e (red) to +4 kT/e (blue), as calculated by Pymol. The pictures to the *Left* and *Right* show the back and front view of TasA, respectively, with regard to the orientation in *A*. (*E*) Surface representation of TasA. Residues showing strong NMR shifts upon titration of a TasA solution with $MnCl_2$ are highlighted in magenta.

secrete various extracellular proteases, also under biofilm conditions; therefore, these results suggest that such secreted proteases change to biofilm-stabilizing fibrillar structures. During such an evolutionary process, the intrinsic protease activity might have been lost in individual cases such as TasA.

TasA Assumes Monomeric, Oligomeric, and Various Polymeric Forms in Vitro, but *B. subtilis* Biofilms Contain Predominantly Homogeneous TasA Fibrils. In line with investigations by Chai et al. (12) on mature TasA isolated from the supernatant of an exopolysaccharide-negative strain, our recombinant TasA formed monomers, oligomers with a defined number of subunits, as well as polymers, often of fibrillar nature. In the following sections, we characterize the various forms of TasA occurring at different pH, temperature, and time, and establish which of those forms are taking part in biofilm formation in vivo. To this end, we applied AUC, electron microscopy (EM), mass spectrometry, and H/D exchange followed by NMR in vitro, complemented by in vivo NMR on *B. subtilis* biofilms.

pH dependence of structural transitions. When freshly prepared at pH 7.0, according to AUC, TasA$_{261}$ consists of ∼80% monomers and around 20% of presumed dodecamers with smaller oligomers appearing at a low level (Fig. 1*B* and *SI Appendix*, Fig. S6*A*). The frictional coefficient (f/fo) of 1.4 indicates a globularly folded protein (32). When adjusting pH to a value of 3.0 by overnight dialysis, the 3D structure of TasA remains intact as detected by NMR (*SI Appendix*, Fig. S7 *A* and *B*), and AUC analysis shows several oligomeric forms next to a fraction of 66% monomer (*SI Appendix*, Fig. S6*B*). However, polymeric forms appear under both conditions (pH 7.0 and pH 3.0) after incubating the samples for 2 wk at 40 °C (*SI Appendix*, Fig. S6 *A–D*). As is common for polymers, larger species distributed over a wider molecular weight range and few oligomeric forms, presumably dodecamers, are observed in AUC experiments (*SI Appendix*, Fig. S6 *A* and *B*). Interestingly, the frictional coefficients of around 2.6 for these larger species are indicative of fibrils (*SI Appendix*, Table S2) (32). In line with this observation, Thioflavin T (ThT) can stain these fibrils (*SI Appendix*, Fig. S6*C*), and they appear insoluble when analyzed by SDS/PAGE (*SI Appendix*, Fig. S6*D*). These results were corroborated by EM investigations and fiber diffraction (Fig. 3). The fiber diffraction results (Fig. 3*C*) suggest the presence of at least a subpopulation of cross-β-sheet–forming TasA species, as indicative of amyloid-like structures.

Notably, a downshift to pH 3 by dialysis leads to a vanishing of the oligomer peak(s) and polymeric forms appear (Fig. 4*A*). ThT assays (Fig. 4*B*) show a strong increase in fluorescence without lag phase within the first 100 min for both TasA$_{239}$ and TasA$_{261}$ when pH is adjusted by adding HCl. However, at pH 7, this

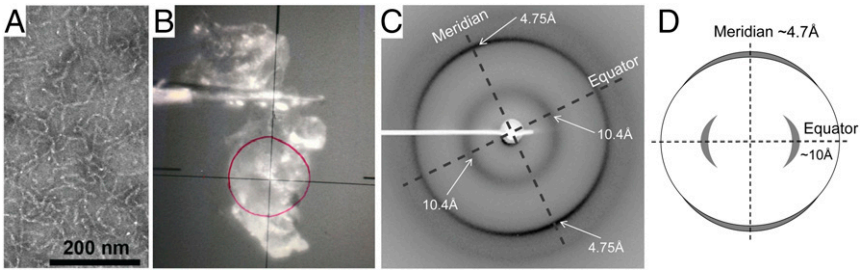

Fig. 3. (*A*) Electron micrograph of $TasA_{261}$ fibers grown at pH 3 and 40 °C over 2 wk. (*B*) The sample shown in *A* mounted onto a MiTeGen micromesh loop used for fiber diffraction experiments. The red circles indicate the X-ray beam size and position for two different beam diameters of 50 μm (solid line) and 75 μm (dashed line) as appearing on the monitor. (*C*) Fiber diffraction pattern of $TasA_{261}$ showing reflection arcs that are oriented on the meridian at 4.75 Å and on the equatorial axis at 10.4 Å, which indicate the presence of cross β-structures. (*D*) Schematic drawing of a typical cross-β diffraction pattern with a characteristic 4.7- to 4.8-Å diffraction signal on the meridian corresponding to the distance between hydrogen-bonded β-strands that run perpendicular to the fiber axis. The more diffuse signal on the equator at a distance of 10–12 Å results from the association of the β-sheets.

increase in fluorescence is not observed in the same time frame at room temperature. We assume that unfolding occurs upon rapid transition to low pH, as demonstrated by NMR spectroscopy on dialyzed or rapidly pH-adjusted TasA solutions, showing ^{15}N–^{1}H correlations of folded monomer in the former case (*SI Appendix*, Fig. S7 *A* and *B*) and of unfolded protein in the latter (*SI Appendix*, Fig. S7*C*).

We noted that solution NMR samples of TasA (pH 7.0) become gel-like over time. Such viscous protein preparation that was stored for 16 wk at −20 °C showed a much larger fraction of oligomeric TasA in AUC experiments (Fig. 4*A*) compared with freshly prepared TasA.

Initial structural characterization of high–molecular-weight forms by NMR. The structural differences between monomers and various types of high–molecular-weight aggregates characterized above are apparent from a comparison of ^{15}N–^{1}H correlation spectra (Fig. 4 *E*–*H*). The ^{15}N–^{1}H HSQC spectrum of a $TasA_{261}$ solution shows the fingerprint of a folded, soluble domain with exceptionally characteristic signals (arrows in Fig. 4*E* and *SI Appendix*, Fig. S2*B*). Those peaks are assigned to residues T115 (arrow 2) and G175 (arrow 1) and show unusual chemical shift values as well as peak doubling presumably due to their proximity to the disordered region 117–125 and a possible *cis*/*trans* isomerization of P125.

Precipitates formed by $TasA_{261}$ during concentration of a sample are of fibrillar nature, as demonstrated by EM (Fig. 4*D*). These fibrils yielded a surprisingly well-resolved MAS NMR cross-polarization (CP)–based ^{15}N–^{1}H correlation spectrum (Fig. 4*F*) that indicates a high β-structure content by the chemical shift distribution and high homogeneity due to the absence of signal splitting or doubling. Twenty-one cross peaks with chemical shifts >9.5 ppm are observed, whereas the solution spectrum exposes only 8 in this region. There are several signals with unusual ^{15}N chemical shifts $\delta < 100$ ppm and $\delta > 135$ ppm. In Fig. 4*F*, nine signals with unusual chemical shifts occur that are indicated by asterisks; these signals are not present in the solution spectrum. Vice versa, the signals indicated by arrows in the solution spectrum (Fig. 4*E*) are no longer present in the spectrum of the precipitate (Fig. 4*F*). The superposition of these two spectra shows only few accidental matches of cross-peak positions (*SI Appendix*, Fig. S8) but otherwise large differences, indicating substantial structural changes. Methods for enforcing

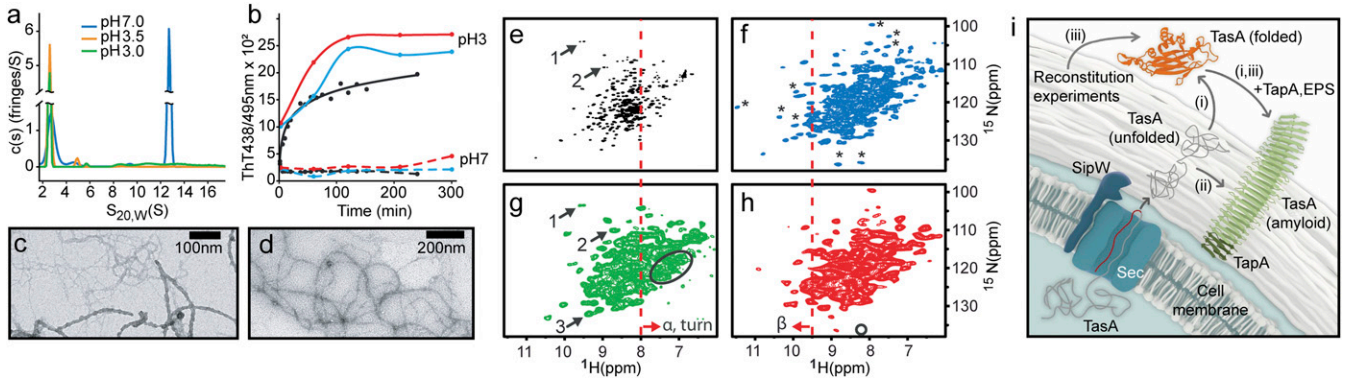

Fig. 4. (*A*) AUC of $TasA_{261}$ stored for 16 wk at −20 °C, pH 7, and dialyzed toward pH 3.5 and 3.0. (*B*) ThT assay of $TasA_{261}$ (black) incubated at pH 7 and 3, and at a temperature of 20 °C. Additionally, direct comparisons of the ThT response of $TasA_{261}$ (blue) and $TasA_{239}$ (red) at the same pH settings, concentrations, and time points are shown. (*C*) Fibrils detected by EM on $TasA_{261}$ incubated for 2 wk at pH 7.0 and 40 °C. (*D*) EM of $TasA_{261}$ precipitated during sample concentration, same sample as in *F*. (*E*) ^{15}N–^{1}H-HSQC of a $^{2}H,^{15}N,^{13}C$-$TasA_{261}$ solution. Arrows 1 and 2 indicate the signals of G175 and T115, respectively. (*F*) ^{15}N–^{1}H-MAS NMR correlation of the $^{2}H,^{15}N,^{13}C$-labeled precipitate generated by concentrating a solution; amide protons are reexchanged in 30% D_2O. The signals labeled with an asterisk show special chemical shifts and are unique to this spectrum. (*G*) ^{15}N–^{1}H-MAS NMR correlation of high–molecular-weight aggregates isolated after 5 mo from an otherwise intact, $^{2}H,^{15}N,^{13}C$-labeled $TasA_{239}$ solution NMR sample. Arrows 1 and 2 denote the same signals as in *E*; arrow 3 denotes a very strong signal unique to this spectrum. The ellipsoid indicates an area with signals whose chemical shifts are typical for residues in α-helices and turns. (*H*) ^{15}N–^{1}H-MAS NMR correlation of a biofilm formed by a *B. subtilis* Δ*tasA* strain after supply of recombinant, $^{2}H,^{15}N,^{13}C$-labeled $TasA_{261}$. The circle indicates the position of a signal appearing at lower contour levels. (*I*) Processes leading to fiber formation after TasA secretion (*i* and *ii*) and from externally provided protein in reconstitution experiments (*iii*). In all cases, but in particular for pathway *iii*, a supporting role of TapA or EPS is expected. EPS are not shown.

fibril formation such as extruding into 70% methanol lead to similar NMR spectra as shown in Fig. 4*F* but indicate the presence of heterogeneous fibrils.

Since we noted that solution NMR samples of TasA (pH 7.0) become gel-like over time, larger protein assemblies of $TasA_{239}$ were spun down from a sample stored at pH 7 and ambient temperatures over 16 wk to obtain a fingerprint by MAS NMR. The CP-based ^{15}N–^{1}H correlation (Fig. 4*G*) shows characteristic features of the folded monomer. Among those are the signals of T115 and G175 (arrows in Fig. 4 *E* and *G*), including the respective peak doubling. It is unlikely that these signals occur due to coprecipitated monomers since the cross-polarization unit of the used pulse sequence would not excite them. They most likely result from a conservation of monomer structure in those aggregates. Overall, the signals of this type of high–molecular-weight species are distributed in the same spectral range as the signals in the solution spectrum (*SI Appendix*, Fig. S9), with one important difference: a larger number of cross peaks is observed in the area between 6.5- and 8-ppm proton chemical shifts and 115- and 123-ppm nitrogen chemical shifts (ellipsoid, Fig. 4*G*), indicating a higher content of helical or loop structure. Comparing that spectrum to the spectrum of fibrils clearly shows that it is substantially different (*SI Appendix*, Fig. S10).

Taken together, the above data indicate that TasA may occur in three structurally different forms; monomers and two distinctly different but largely homogeneous types of high–molecular-weight aggregates, with characteristic NMR spectra. One of these forms (termed "precipitate" below) shows spectra with high β-content as expected for fibrils (Fig. 4*F*), in agreement with prior biophysical characterization (Fig. 4*D*), while the other is characterized by a higher content of helical or loop structure (Fig. 4*G*). Structural differences between the fibrillar and monomeric form were also detected by H/D exchange experiments (*SI Appendix*, section A). The enhancement of fibril formation in vitro by, for example, rapid pH shift strongly suggests that the concomitant forced formation of unfolded protomers might accelerate the transition into TasA fibrils that are characterized by a higher β-sheet content. Worth of note, such "enforced" fibrils were found to be heterogeneous.

Structural characterization and protease resistance of TasA in native biofilm. For obtaining a fingerprint of TasA present in biofilms, a *B. subtilis tasA* deletion strain that still secretes TapA and EPS was employed and biofilm formation successfully rescued (9) via the addition of recombinantly produced ^{2}H,^{13}C,^{15}N-$TasA_{261}$ (*SI Appendix*, Fig. S11). More than 90% of the supplied protein to the medium could be detected in the subsequently formed biofilm (*SI Appendix*, Fig. S11).

Extracellular proteases are secreted by *B. subtilis* cells also during biofilm formation (33, 34). We could detect these proteases by a high protease activity of the cell-free *B. subtilis* biofilm supernatant against Azocasein. Nevertheless, a very high stability of supplied $TasA_{261}$ toward such proteolytic attack can be observed (*SI Appendix*, Fig. S11*B*), strongly suggesting that the TasA fibrils formed in the biofilm are resistant to own proteases.

In situ solid-state MAS NMR of the harvested biofilm yielded a CP-based ^{15}N–^{1}H HSQC spectrum with decent resolution (Fig. 4*H*). This spectrum shows all special features of that obtained on fibrils (Fig. 4*F*). A superposition of both spectra (*SI Appendix*, Fig. S12) shows a surprising match, with all distinctive signals that are labeled with an asterisk in Fig. 4*F* being present in the spectrum of the biofilm, in one case at lower plot levels (indicated by a circle in Fig. 4*H*). On the contrary, a superposition with the spectrum of the gel-like form (*SI Appendix*, Fig. S13) reveals strong differences in key peak positions (e.g., see arrow 3 in Fig. 4*G* and peaks in this area, also the signals indicated by the ellipsoid). The characteristic signals of the gel-like sample also do not appear at low contour levels in the biofilm spectrum (*SI Appendix*, Fig. S13). There is only a small subset of signals that show chemical shift differences, tentatively connected by curved lines in *SI Appendix*, Fig. S12. For these reasons, and based on the in vitro characterization of this fibrillar sample by EM (Fig. 4*D*) and other techniques, we conclude that native *B. subtilis* biofilms predominantly contain homogeneous protease-resistant TasA fibrils.

Discussion

In this work, we present a high-resolution structure of TasA that provides the starting point for investigations of structural changes preceding *Bacillus* biofilm formation, representing at the same time a 3D structure of a M73 metalloprotease family member. In this way, we provide a structural basis for understanding camelysins (*SI Appendix*, Fig. S5), which play important roles in the pathogenicity of, for example, *B. anthracis*. Most importantly, our results uncover the high conformational variability of TasA, facilitating structural changes as a prerequisite for fibril formation.

The structure of the TasA monomer consists of a rigid frame and a set of loops with high flexibility in the "dynamic section" (lower section in Fig. 2*C*) where two PPII helices are also located that may function as conformational switches. PPII helices can change into right-handed α-helices or β-strands by shifting only one dihedral angle per residue (35). Intriguingly, PPII helices were shown to play an important role in conformational changes resulting in amyloid fibrils formed by the prion protein or amyloidogenic lysozyme (36–39). We assume that the structural plasticity in the dynamic section may potentially facilitate interactions that promote the formation of fibrillar structures.

The fact that TasA folds into a globular structure as a monomer represents a fundamental difference to the well-known major biofilm-forming proteins CsgA and FapC, which exist as unstructured proteins in their monomeric forms (2, 3). In case of TasA, two different scenarios (Fig. 4*I*, pathways *i* and *ii*) concerning the fate of the secreted protein may be relevant in vivo: (*i*) secreted soluble monomeric TasA could fold right after secretion and transform from there into the fibrillar structure or, alternatively, (*ii*) the still unfolded secreted monomer could also switch directly to the fibrillar form. Furthermore, the application of purified soluble monomeric recombinant TasA or secreted endogenous TasA from natural source (12) can complement the defect in biofilm formation of TasA-negative mutants (9) (Fig. 4*I*, pathway *iii*). The TasA protein supplied in our biofilm reconstitution experiment remains folded when slowly dialyzed to a pH of 3 (pathway *iii*), which suggests that the transition of TasA from a monomeric fold into a fibrillar structure could be promoted by other partners. Given the long timescale at which fibrils are formed at neutral pH and room temperature (several days to weeks), we conclude from the large amount of fibrous protein appearing more rapid in the in vivo experiment (hours) that fibril formation is promoted in a "catalytic" fashion (Fig. 4*I*, pathway *iii*) by TapA, EPS, or the characteristics of the *B. subtilis* surface (40), which facilitate the conformational transition between both structural arrangements. A possible low local pH, close to the bacterial surface and in the EPS environment, may additionally destabilize TasA, despite its overall pH stability.

The view into the in vivo situation reveals a small subset of signals in the biofilm MAS NMR spectrum that show different chemical shifts than the respective signals of fibrils obtained in vitro (*SI Appendix*, Fig. S12, curved lines). Given the nearly perfect fit of most other exposed signals, this suggests specific interactions with, for example, TapA or EPS in the biofilm, which cannot occur in the in vitro experiment.

Interestingly, the TasA structure determined here by X-ray crystallography reveals different proportions of secondary-structure elements than originally concluded from CD spectroscopy by Chai et al. (12), who suggested high α-helical content. However, we observe high β-sheet content in the monomer structure (*SI Appendix*, Table S3) and that a considerable fraction of the helical part (~50%) is formed by 3_{10} or

PPII helices (41). Our CD spectra recorded on recombinant TasA (*SI Appendix*, Fig. S14) showed a comparable shape to those published (12), displaying a minimum at 208 nm (*SI Appendix*, Fig. S14*A*), the occurrence of which is also due to contributions by the PPII helices according to Sreerama and Woody (37, 42). To our surprise, unfolding of TasA by heat denaturation did not change the spectra significantly, although a temperature of 90 °C was reached and a clear transition point (T_m = 50 °C) could be determined (*SI Appendix*, Fig. S14*B*). This was substantiated by a ^{15}N–^{1}H NMR spectrum of a sample incubated before at 60 °C, displaying the characteristic signal pattern of an unfolded protein (*SI Appendix*, Fig. S14*C*).

We considered a possible specific protease activity of the globular form of TasA due to similarities to camelysins that belong to the Peptidase_M73 superfamily (26). In the different *B. cereus* strains, two camelysins are encoded in the same operon together with the *sipW* gene. They are reported to be involved in surface biofilm formation, and they are also able to form fibrillar structures (24, 25). However, our results suggest that TasA from the nonpathogenic *B. subtilis* strains lost the ability to bind divalent cations like Zn^{2+} that is required for protease activity.

In summary, we have demonstrated that TasA undergoes substantial conformational changes in the course of *Bacillus subtilis* biofilm maturation. The X-ray structure of monomeric TasA provides a structural basis for understanding this transition, revealing flexible segments and PPII helices together in the dynamic section of the structure. Interestingly, different in vitro experiments, including X-ray fiber diffraction, suggest the presence of a population of cross-β-sheet–containing TasA species (Fig. 3) and hence amyloid-like structures in the formed, protease-resistant fibers, in accordance with Romero et al. (9). For studies of functional fibril formation from a globular state, TasA can thus serve as a model system.

Materials and Methods

EM. For EM, formvar-carbon film-coated grids were glow-discharged, and a drop of protein sample was applied on the grids, blotted via filter paper, and negatively stained by 3% aqueous uranyl acetate. Dry grids were imaged at Tecnai G2 200-kV or Zeiss 900 80-kV transmission electron microscopes. Images were taken at 50,000× magnification.

Data Availability. The atomic coordinates of $TasA_{239}$ in presence and absence of salicylate have been deposited in the Protein Data Bank (PDB ID codes 5OF1 and 5OF2, respectively).

Additional information on materials and methods is provided in *SI Appendix*.

ACKNOWLEDGMENTS. We thank the team of D. Puchkov [Leibniz-Forschungsinstitut für Molekulare Pharmakologie (FMP)] for electron microscopic investigations. We thank S. Runde and P. Ehrentraut (Freie Universität Berlin) for initial experiments and A. Wallmann (FMP) for making figures. This work was supported from a Joint Research Activity in the European Union Project iNext (Infrastructure for NMR, EM and X-rays for Translational Research; GA 653706) and Deutsche Forschungsgemeinschaft (Sonderforschungsbereich 740, OS106/9, Tu106/6, and Tu106/7).

1. Flemming H-C, et al. (2016) Biofilms: An emergent form of bacterial life. *Nat Rev Microbiol* 14:563–575.
2. Van Gerven N, Klein RD, Hultgren SJ, Remaut H (2015) Bacterial amyloid formation: Structural insights into curli biogensis. *Trends Microbiol* 23:693–706.
3. Taglialegna A, Lasa I, Valle J (2016) Amyloid structures as biofilm matrix scaffolds. *J Bacteriol* 198:2579–2588.
4. Hobley L, Harkins C, MacPhee CE, Stanley-Wall NR (2015) Giving structure to the biofilm matrix: An overview of individual strategies and emerging common themes. *FEMS Microbiol Rev* 39:649–669.
5. Schor M, Reid JL, MacPhee CE, Stanley-Wall NR (2016) The diverse structures and functions of surfactant proteins. *Trends Biochem Sci* 41:610–620.
6. Vlamakis H, Chai Y, Beauregard P, Losick R, Kolter R (2013) Sticking together: Building a biofilm the *Bacillus subtilis* way. *Nat Rev Microbiol* 11:157–168.
7. Costerton JW, Stewart PS, Greenberg EP (1999) Bacterial biofilms: A common cause of persistent infections. *Science* 284:1318–1322.
8. Branda SS, González-Pastor JE, Ben-Yehuda S, Losick R, Kolter R (2001) Fruiting body formation by *Bacillus subtilis*. *Proc Natl Acad Sci USA* 98:11621–11626.
9. Romero D, Aguilar C, Losick R, Kolter R (2010) Amyloid fibers provide structural integrity to *Bacillus subtilis* biofilms. *Proc Natl Acad Sci USA* 107:2230–2234.
10. Stöver AG, Driks A (1999) Secretion, localization, and antibacterial activity of TasA, a *Bacillus subtilis* spore-associated protein. *J Bacteriol* 181:1664–1672.
11. Serrano M, et al. (1999) A *Bacillus subtilis* secreted protein with a role in endospore coat assembly and function. *J Bacteriol* 181:3632–3643.
12. Chai L, et al. (2013) Isolation, characterization, and aggregation of a structured bacterial matrix precursor. *J Biol Chem* 288:17559–17568.
13. Branda SS, Chu F, Kearns DB, Losick R, Kolter R (2006) A major protein component of the *Bacillus subtilis* biofilm matrix. *Mol Microbiol* 59:1229–1238.
14. Romero D, Vlamakis H, Losick R, Kolter R (2011) An accessory protein required for anchoring and assembly of amyloid fibres in *B. subtilis* biofilms. *Mol Microbiol* 80:1155–1168.
15. Romero D, Vlamakis H, Losick R, Kolter R (2014) Functional analysis of the accessory protein TapA in *Bacillus subtilis* amyloid fiber assembly. *J Bacteriol* 196:1505–1513.
16. Kearns DB, Chu F, Branda SS, Kolter R, Losick R (2005) A master regulator for biofilm formation by *Bacillus subtilis*. *Mol Microbiol* 55:739–749.
17. Chu F, Kearns DB, Branda SS, Kolter R, Losick R (2006) Targets of the master regulator of biofilm formation in *Bacillus subtilis*. *Mol Microbiol* 59:1216–1228.
18. Stanley NR, Britton RA, Grossman AD, Lazazzera BA (2003) Identification of catabolite repression as a physiological regulator of biofilm formation by *Bacillus subtilis* by use of DNA microarrays. *J Bacteriol* 185:1951–1957.
19. Mandic-Mulec I, Gaur N, Bai U, Smith I (1992) Sin, a stage-specific repressor of cellular differentiation. *J Bacteriol* 174:3561–3569.
20. Bai U, Mandic-Mulec I, Smith I (1993) SinI modulates the activity of SinR, a developmental switch protein of *Bacillus subtilis*, by protein-protein interaction. *Genes Dev* 7:139–148.
21. Mielich-Süss B, Lopez D (2015) Molecular mechanisms involved in *Bacillus subtilis* biofilm formation. *Environ Microbiol* 17:555–565.
22. Terra R, Stanley-Wall NR, Cao G, Lazazzera BA (2012) Identification of *Bacillus subtilis* SipW as a bifunctional signal peptidase that controls surface-adhered biofilm formation. *J Bacteriol* 194:2781–2790.
23. Tjalsma H, et al. (2000) Conserved serine and histidine residues are critical for activity of the ER-type signal peptidase SipW of *Bacillus subtilis*. *J Biol Chem* 275:25102–25108.
24. Gao T, Foulston L, Chai Y, Wang Q, Losick R (2015) Alternative modes of biofilm formation by plant-associated *Bacillus cereus*. *MicrobiologyOpen* 4:452–464.
25. Caro-Astorga J, Pérez-García A, de Vicente A, Romero D (2015) A genomic region involved in the formation of adhesin fibers in *Bacillus cereus* biofilms. *Front Microbiol* 5:745.
26. Rawlings ND, Barrett AJ, Finn R (2016) Twenty years of the MEROPS database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res* 44:D343–D350.
27. Grass G, et al. (2004) Camelysin is a novel surface metalloproteinase from *Bacillus cereus*. *Infect Immun* 72:219–228.
28. Fricke B, et al. (2001) The cell envelope-bound metalloprotease (camelysin) from *Bacillus cereus* is a possible pathogenic factor. *Biochim Biophys Acta* 1537:132–146.
29. Pflughoeft KJ, Sumby P, Koehler TM (2011) *Bacillus anthracis* sin locus and regulation of secreted proteases. *J Bacteriol* 193:631–639.
30. Holm L, Rosenström P (2010) Dali server: Conservation mapping in 3D. *Nucleic Acids Res* 38:W545–W549.
31. Dogsa I, Brloznik M, Stopar D, Mandic-Mulec I (2013) Exopolymer diversity and the role of levan in *Bacillus subtilis* biofilms. *PLoS One* 8:e62044.
32. Erickson HP (2009) Size and shape of protein molecules at the nanometer level determined by sedimentation, gel filtration, and electron microscopy. *Biol Proced Online* 11:32–51.
33. Cairns LS, Hobley L, Stanley-Wall NR (2014) Biofilm formation by *Bacillus subtilis*: New insights into regulatory strategies and assembly mechanisms. *Mol Microbiol* 93:587–598.
34. Marlow VL, et al. (2014) The prevalence and origin of exoprotease-producing cells in the *Bacillus subtilis* biofilm. *Microbiology* 160:56–66.
35. Bochicchio B, Tamburro AM (2002) Polyproline II structure in proteins: Identification by chiroptical spectroscopies, stability, and functions. *Chirality* 14:782–792.
36. Camilloni C, De Simone A, Vranken WF, Vendruscolo M (2012) Determination of secondary structure populations in disordered states of proteins using nuclear magnetic resonance chemical shifts. *Biochemistry* 51:2224–2231.
37. Adzhubei AA, Sternberg MJE, Makarov AA (2013) Polyproline-II helix in proteins: Structure and function. *J Mol Biol* 425:2100–2132.
38. Blanch EW, et al. (2000) Is polyproline II helix the killer conformation? A Raman optical activity study of the amyloidogenic prefibrillar intermediate of human lysozyme. *J Mol Biol* 301:553–563.
39. Gill AC, et al. (2000) Post-translational hydroxylation at the N-terminus of the prion protein reveals presence of PPII structure in vivo. *EMBO J* 19:5324–5331.
40. Weiss L (1963) The pH value at the surface of *Bacillus subtilis*. *J Gen Microbiol* 32:331–340.
41. Kumar P, Bansal M (2015) Identification of local variations within secondary structures of proteins. *Acta Crystallogr D Biol Crystallogr* 71:1077–1086.
42. Sreerama N, Woody RW (2003) Structural composition of betaI- and betaII-proteins. *Protein Sci* 12:384–388.