The Functional Curli Amyloid Is Not Based on In-register Parallel β-Sheet Structure*§

From the Laboratories of *Biochemistry and Genetics, ‡Chemical Physics, and §Molecular Biology, NIDDK, National Institutes of Health, Bethesda, Maryland 20892-0830

The extracellular curli proteins of Enterobacteriaceae form fibrous structures that are involved in biofilm formation and adhesion to host cells. These curli fibrils are considered a functional amyloid because they are not a consequence of misfolding, but they have many of the properties of protein amyloid. We confirm that fibrils formed by CsgA and CsgB, the primary curli proteins of Escherichia coli, possess many of the hallmarks typical of amyloid. Moreover we demonstrate that curli fibrils possess the cross-β structure that distinguishes protein amyloid. However, solid state NMR experiments indicate that curli structure is not based on an in-register parallel β-sheet architecture, which is common to many human disease-associated amyloids and the yeast prion amyloids. Solid state NMR and electron microscopy data are consistent with a β-helix-like structure but are not sufficient to establish such a structure definitively.

Interest in amyloid is largely because of its association with many late onset human diseases, including Alzheimer disease (Aβ), Parkinson disease (α-synuclein), type II diabetes (amylin), and the transmissible spongiform encephalopathies (PrP). In each case a particular endogenous protein becomes incorporated into large aggregates known as amyloid, which was originally defined by pathologists as a tissue deposit staining like starch (1). However, the term amyloid is not based on an in-register parallel β-sheet structure (12–19).

Proteins from a variety of organisms can form amyloid both in vitro and in vivo, and the propensity to form amyloid may be a common property of many proteins (2). In addition to disease-associated amyloids, there are several confirmed cases of functional amyloid (for a review, see Ref. 3). For example, hydrophobins are amyloid-like proteins that coat the surface of fungal cells, and amyloid fibrils coating fish eggs protect them from dehydration (4, 5). The [Het-s] prion of Podospora anserina is involved in heterokaryon incompatibility, a recognition of non-self reaction believed to be important as a defense against fungal virus infection (6).

Curli are extracellular filamentous structures of Enterobacteriaceae (7) that are integral to biofilm formation and are the major protein component of the extracellular matrix of these organisms (8). Curli of Escherichia coli are composed of the secreted proteins CsgA and CsgB. The latter is believed to prime the polymerization of the former and anchor the fibrils to the outer membrane (9). Both CsgA and CsgB fibrils are β-sheet-rich and, like amyloids, stain with the dye Congo red (10, 11).

Because amyloid fibrils are non-crystalline and insoluble, solution NMR and x-ray crystallography are not directly applicable in structural studies. Solid state NMR and electron spin resonance have both been useful in obtaining constraints on amyloid structures and, in some cases, determining detailed structural information. The disease-associated amyloids formed by Aβ, amylin, α-synuclein, and tau along with the infectious amyloids of several yeast prions each have in-register parallel β-sheet structure (12–19).

Here we confirm that the fibrils formed in vitro by CsgA and CsgB proteins are amyloids and explore their structure using solid state NMR and electron microscopy. Our results indicate that, unlike the pathogenic amyloids of humans and yeast, CsgA and CsgB amyloids are not in-register parallel β-sheet structures. Solid state NMR and electron microscopy data are consistent with a β-helix-like structure but do not establish such a structure definitively.

EXPERIMENTAL PROCEDURES

CsgA and CsgB Purification—The csgA and csgB genes, with carboxyl-terminally encoded His tags, were cloned into PET101/D (Invitrogen) bacterial expression plasmids (supplemental Table 1). Protein overexpression was performed in E. coli strain BL21(DE3) grown in Luria broth at 37 °C. For single isotope labeling, cells were grown in defined amino acid medium as described previously (17, 18). For uniform 13C,15N
Biophysical Characterization of Curli Fibrils

labeling, cells were grown in a defined medium containing the following: 4 g/liter glucose, 0.2 g/liter adenine sulfate, 0.2 g/liter uracil, 0.03 g/liter thiamine, 1.2 g/liter MgSO$_4$, 13 g/liter KH$_2$PO$_4$, 10 g/liter K$_2$HPO$_4$, 9 g/liter Na$_2$HPO$_4$, 2 g/liter NH$_4$Cl, 0.05 g/liter EDTA, 0.06 g/liter FeSO$_4$, 0.06 g/liter CaCl$_2$, 12 mg/liter MnCl$_2$, 8 mg/liter CoCl$_2$, 7 mg/liter ZnSO$_4$, 3 mg/liter CuCl$_2$, 2 mg/liter H$_3$BO$_3$, and 2.5 mg/liter (NH$_4$)$_6$Mo$_7$O$_{24}$.

When the cells reached an $A_{600}$ of $0.8$, they were harvested and resuspended to twice their original volume in the same medium but with $^{13}$C-glucose and $^{15}$N$_2$H$_4$. Protein expression was induced with 1 mM isopropyl $\beta$-D-thiogalactoside when cell densities reached an $A_{600}$ of 0.5–0.7. Cells were harvested after 3–4 h of protein induction.

CsgA and CsgB were purified similarly. Following protein expression, cell pellets were resuspended in a phosphate buffer (10 mM K$_2$HPO$_4$, pH 7.4, 150 mM NaCl, and 1 mM EDTA) supplemented with Roche Applied Science Complete protease inhibitor mixture. The cells were lysed by sonication and then spun at 20,000 rpm for 20 min. The supernatant was discarded, and the pellet was resuspended in Denaturing Buffer (8 M urea, 100 mM NaCl, and 100 mM K$_2$HPO$_4$, pH 8). After 2–3 days of incubation at room temperature with occasional agitation, the lysate was spun at 30,000 rpm for 20 min. The pellet was discarded, and the supernatant was mixed with nickel-nitrilotriacetic acid-agarose (Qiagen; 1 ml/liter of culture) and was converted to intensity versus 20 plots by circular summation around the center of the two-dimensional images using the R-axis Display Program (Rigaku Americas). The backgrounds were subtracted, and the resulting plots were displayed by Origin 7 (OriginLab).

Electron Microscopy—Aqueous suspensions of curli fibrils were applied to carbon-coated copper grids, incubated for 2–3 min, briefly washed with water, and stained with 1–2% uranyl acetate for 2–3 min. Grids were then blotted and air-dried. Electron micrographs were produced with an FEI Co. Morgagni transmission electron microscope operating at 80 kV.

For electron diffraction, aqueous suspensions of CsgB fibrils were applied to carbon-coated copper grids and washed several times with water. The diffraction from the protein fibrils is sensitive to damage caused by the electron beam, so regions on the grid with CsgB fibrils were found under low magnification with low electron beam intensity prior to diffraction imaging. Immediately after diffraction imaging, regular images were taken to confirm the region used for diffraction. All electron diffraction images were taken with an 80-kV electron beam and 350-mm camera length. The diffraction distances were calibrated with the diffraction rings of thallous chloride (Electron Microscopy Sciences) under the same microscope conditions. Single crystals of molybdenum oxide (MoO$_3$; Electron Microscopy Sciences) were used to measure the rotation between regular images and diffraction images. Also MoO$_3$ crystals were used to confirm that the electron beam does not shift between regular and diffraction imaging by comparison of regular and defocused diffraction images under the measurement conditions.

Solid State NMR Spectroscopy—Solid state NMR experiments with carbonyl-labeled samples were performed at room temperature at 9.39 teslas (100.4-MHz $^{13}$C NMR frequency) with an InfinityPlus spectrometer (Varian, Palo Alto, Ca). Lyophilized CsgA fibrils were packed into 3.2-mm-diameter magic angle spinning (MAS) NMR rotors. Measurements were

Proteinase K Digestion—Comparative proteolysis was performed by incubating 20 ng of Proteinase K with 10 μg of soluble or fibrous CsgA for 0, 5, 10, and 20 min. Reactions were performed in 125 mM Tris, pH 7.5, and 0.8 M urea at room temperature in a total volume of 20 μl. The reactions were terminated by the addition of 100 μl of 88% formic acid, which also serves to depolymerize CsgA fibrils for subsequent gel electrophoresis (22). The samples were lyophilized, then resuspended in 8 M urea and 2% SDS, and heated for 3 min at 72 °C. The samples were evaluated by SDS-PAGE using a 10–20 Tricine gel.

X-ray Fiber Diffraction—CsgA and CsgB fibrils were washed extensively with water, lyophilized, and packed into a 0.7-mm-diameter, thin walled, boron-rich glass capillary by centrifugation. The samples were placed in an x-ray beam (0.15418 nm) generated by a rotating anode source operated at 50 kV and 100 mA coupled with a multilayer focusing optic. Diffraction data were collected in 20-min exposures during which the samples were continuously rotated around an axis perpendicular to the beam. Diffraction data were recorded on an XaxisIV two-dimensional imaging plate detector placed 15 cm from the capillary. Background data were collected similarly but exposing a part of the same capillary that was free of fibers. The data were converted to intensity versus 20 plots by circular summation around the center of the two-dimensional images using the R-axis Display Program (Rigaku Americas). The backgrounds were subtracted, and the resulting plots were displayed by Origin 7 (OriginLab).
performed with Varian MAS NMR probes. For certain measurements, the fibrils were rehydrated by addition of water in the rotor. Previous solid state NMR measurements on amyloid fibrils have shown that lyophilization does not disrupt amyloid structures and that rehydration of lyophilized samples leads to solid state NMR spectra with sharp lines that are indistinguishable from spectra of fully hydrated samples that had not been lyophilized (17, 23, 24).

The PITHIRDS-CT technique (25) was used to measure intermolecular $^{13}$C-$^{13}$C magnetic dipole-dipole couplings in carbonyl-labeled samples. These couplings depend on intermolecular distances as $1/R^3$. Experimental conditions were the same as in earlier PITHIRDS-CT measurements on various amyloid fibrils (19, 26, 27). Raw PITHIRDS-CT data were corrected for contributions from the 1.1% natural abundance of $^{13}$C at all carbonyl positions, assuming the 17% carbonyl signal decay observed in PITHIRDS-CT measurements on an unlabeled CsgA sample. Differences in $T_2$ decays between naturally occurring $^{13}$C and $^{13}$C labels were small (supplemental Fig. 3B) but were included in the correction for relative contribution by using the relative signals remaining following a period equal to the constant time period of the PITHIRDS-CT experiment.

Two-dimensional NMR spectra of uniformly $^{15}$N-$^{13}$C-labeled CsgA fibrils were obtained at 14.1 teslas (150.7-MHz $^{13}$C NMR frequency) with MAS at 13.50 kHz. Approximately 2 mg of lyophilized fibrils were loaded into a 3.2-mm MAS rotor and rehydrated with Teflon spacers to keep the sample centered in the radiofrequency (rf) coil of the NMR probe. Two-dimensional $^{13}$C-$^{13}$C spectra were obtained with either finite pulse radiofrequency-driven recoupling (fpRFDR) (28, 29) or rf-assisted diffusion (RAD) (30, 31) in the mixing period between $t_1$ and $t_2$ periods. In fpRFDR experiments, the mixing period was 2.37 ms, $^{13}$C π pulses were 20.0 $\mu$s, and proton decoupling fields were 105 kHz. In RAD experiments, the mixing period was 200 ms with a 6.75-kHz proton rf field. Total experiment times were 44 h for the two-dimensional fpRFDR spectrum and 112 h for the two-dimensional RAD spectrum. Two-dimensional $^{15}$N-$^{13}$C spectra were obtained with either finite pulse radiofrequency-driven recoupling (fpRFDR) (28, 29) or rf-assisted diffusion (RAD) (30, 31) in the mixing period between $t_1$ and $t_2$ periods. In fpRFDR experiments, the mixing period was 2.37 ms, $^{13}$C π pulses were 20.0 $\mu$s, and proton decoupling fields were 105 kHz. In RAD experiments, the mixing period was 200 ms with a 6.75-kHz proton rf field. Total experiment times were 44 h for the two-dimensional fpRFDR spectrum and 112 h for the two-dimensional RAD spectrum. Two-dimensional $^{15}$N-$^{13}$C spectra were obtained with either finite pulse radiofrequency-driven recoupling (fpRFDR) (28, 29) or rf-assisted diffusion (RAD) (30, 31) in the mixing period between $t_1$ and $t_2$ periods. In fpRFDR experiments, the mixing period was 2.37 ms, $^{13}$C π pulses were 20.0 $\mu$s, and proton decoupling fields were 105 kHz. In RAD experiments, the mixing period was 200 ms with a 6.75-kHz proton rf field. Total experiment times were 44 h for the two-dimensional fpRFDR spectrum and 112 h for the two-dimensional RAD spectrum. Two-dimensional $^{15}$N-$^{13}$C spectra were obtained with either finite pulse radiofrequency-driven recoupling (fpRFDR) (28, 29) or rf-assisted diffusion (RAD) (30, 31) in the mixing period between $t_1$ and $t_2$ periods. In fpRFDR experiments, the mixing period was 2.37 ms, $^{13}$C π pulses were 20.0 $\mu$s, and proton decoupling fields were 105 kHz. In RAD experiments, the mixing period was 200 ms with a 6.75-kHz proton rf field. Total experiment times were 44 h for the two-dimensional fpRFDR spectrum and 112 h for the two-dimensional RAD spectrum. Two-dimensional $^{15}$N-$^{13}$C spectra were obtained with either finite pulse radiofrequency-driven recoupling (fpRFDR) (28, 29) or rf-assisted diffusion (RAD) (30, 31) in the mixing period between $t_1$ and $t_2$ periods. In fpRFDR experiments, the mixing period was 2.37 ms, $^{13}$C π pulses were 20.0 $\mu$s, and proton decoupling fields were 105 kHz. In RAD experiments, the mixing period was 200 ms with a 6.75-kHz proton rf field. Total experiment times were 44 h for the two-dimensional fpRFDR spectrum and 112 h for the two-dimensional RAD spectrum.
CsgA fibrils are rich in β-sheet and have increased protease resistance. A, x-ray diffraction of dried CsgA fibrils indicates atomic spacing of ~0.48 (upper arrow) and ~0.9 nm (lower arrow). B, dried CsgA fibrils were compared against dried BSA by circular dichroism spectroscopy. Based on curve fitting (20), the relative secondary structure contributions of CsgA fibrils are as follows: 16 ± 2% α-helix, 40 ± 2% β-sheet, 13 ± 2% β-turn, and 31 ± 2% remainder. C, equal concentrations of soluble and fibrous CsgA were exposed to Proteinase K for 0–20 min (mdeg), millidegrees.

β-sheet structure. Also relative to soluble CsgA, fibrillar CsgA exhibited markedly higher resistance to limited Proteinase K exposure as is typical of amyloid fibrils (Fig. 2C).

CsgB Forms Amyloid Fibrils—Previously a recombinant and truncated CsgB protein was shown to form amyloid-like fibrils in vitro (9), but attempts to purify the full-length protein were reported to be unsuccessful. We attempted the same strategies reported above for CsgA to produce full-length CsgB. Two csgB constructs were made (Fig. 3A): one coding the entire open reading frame and another beginning at the second methionine residue (CsgB(1–8)) with the presumption that the export signal would be disrupted and the protein could be more easily purified from a cell lysate. Each construct yielded proteins with the correct approximate mass as determined by electrophoresis (Fig. 3B), and the truncated signal sequence did disrupt secretion as judged by plating cells on Congo red indicator plates (supplemental Fig. 2). The full-length csgB construct yielded two species of slight mass difference, which we assume to be the result of imperfect post-translational processing using this expression system. Like CsgA, both CsgB proteins formed fibrils very quickly after removal from denaturing conditions (Fig. 3, C and D). Moreover the fibrils displayed detergent and urea insensitivity but could be depolymerized with formic acid (Fig. 3B).

Fibrils formed by CsgB also have the properties of amyloid. The fluorescent dye thioflavin-T is commonly used to report amyloid formation. When CsgB fibrils were mixed with thioflavin-T, the characteristic increase in fluorescence intensity was observed (Fig. 4A). Likewise solid state circular dichroism spectroscopy yielded a spectrum consistent with high β-sheet content (Fig. 4B), and x-ray diffraction of CsgB fibrils also yielded the distinctive spacings (~4.7 and ~9 Å) of amyloid (Fig. 4C). However, high β-sheet content alone does not mean a protein aggregate is amyloid. Rather the cross-β structure, in which β-sheets are oriented such that the β-strands run perpendicular to the long axis of the fibril, is the structural hallmark of amyloid. Strong x-ray diffraction of non-aligned samples may indicate cross-β structure, but diffraction from aligned fibrils is more definitive. To this end, electron diffraction can be used to obtain such directional information about amyloid secondary structure (33, 34). The electron diffraction pattern of CsgB fibrillar aggregates on electron microscope grids shows the distinctive 4.7-Å atomic spacing of β-sheet (Fig. 5A). Moreover regions in which the CsgB fibrils are preferentially aligned in
The same direction yield asymmetrical electron diffraction patterns indicative of repeated spacing of about 4.7 Å along the long axis of the fibrils (Fig. 5B) that is the signature of cross-β structure. A single slice extending from the center of the electron diffractogram indicates an increase in intensity at about 4.7 Å (Fig. 5C). The angular dependence of the diffraction intensity around the diffraction ring shows local maxima and minima corresponding, respectively, to the diffraction parallel and perpendicular to the direction of fibril alignment (Fig. 5D), indicating that the 4.7-Å spacing is along the axis of the fibrils.
Amyloid fibrils formed by the yeast prion proteins Sup35, Ure2p, and Rnq1 are based on in-register parallel \( \beta \)-sheets (17–19, 27). In such structures, each residue is adjacent to the same residue in the neighboring polypeptide chains within the \( \beta \)-sheets, leading to intermolecular distances of roughly 5 Å. Fig. 7A shows a schematic of the Sup35 structure along with the signal decay of a \([1-^{13}C]\)tyrosine-labeled Sup35 sample in a PITHIRDS-CT experiment. Fig. 7B shows the signal decay for amyloid of \([1-^{13}C]\)-tyrosine-labeled HET-s(218–289), the fibril-forming domain of a fungal prion that forms a \( \beta \)-helix-like structure in which intermolecular carbonyl-carbonyl distances among like residues are generally greater than 9 Å (38). Accordingly there is a much slower decay of the \( ^{13}C \) signal for HET-s(218–289). The PITHIRDS-CT data for CsgA and CsgB fibril samples (corrected for signal contributions from natural abundance \( ^{13}C \) as described above) are shown in Fig. 7, C and D, respectively. None of the samples shows the rapid \( ^{13}C \) carbonyl signal decay indicative of an in-register parallel \( \beta \)-sheet. Instead the PITHIRDS-CT data indicate average carbonyl-carbonyl distances of \( \sim 7 \) Å or greater. Note that valine-labeled CsgA contains one sequential carbonyl \( ^{13}C \) pair (expected 3.4-Å intramolecular distance in a \( \beta \)-strand; see Fig. 7C legend) and one pair separated by a single unlabeled residue (expected 7-Å intramolecular distance in a \( \beta \)-strand). Leucine-labeled CsgA also contains one carbonyl \( ^{13}C \) pair separated by a single unlabeled residue. The hypothetical alignment of homologous segments in Fig. 1A would produce one pair of leucines and one pair of valines with approximate \( 5 \) Å intersegmental, intramolecular distances. These intramolecular proximities may explain the apparent average \( ^{13}C-^{13}C \) distances of \( \sim 7 \) Å for valine-labeled and leucine-labeled samples. The hypothetical alignment in Fig. 1A also leads to \( \sim 8–9 \) Å intramolecular \( ^{13}C-^{13}C \) distances for phenylalanine-labeled CsgA. We do not expect the PITHIRDS-CT technique to yield accurate interatomic distances when these distances exceed 7 Å because pulse sequence imperfections and spin relaxation processes may become dominant sources of signal decay in this regime.

**Two-dimensional Solid State NMR Spectroscopy of \( ^{13}C,^{15}N \)-Labeled CsgA**—Fibrils were prepared from uniformly \( ^{15}N,^{13}C \)-labeled CsgA and examined by two-dimensional solid state NMR. The two-dimensional \( ^{13}C,^{13}C \) NMR spectrum acquired with fpRFDR mixing shows strong cross-peaks that connect \( ^{13}C \) chemical shifts of directly bonded carbon sites, allowing partial assignments to residue types based on their known chemical shift ranges in peptides and proteins (Fig. 8, A–C). Signals from threonine, serine, and glutamine/glutamate (unresolved) residues have chemical shifts characteristic of \( \beta \)-strands. The major signals from alanine and valine residues also have \( \beta \)-strand chemical shifts, but minor signal com-
component for these residues have non-β-strand chemical shifts. Signals from asparagine/aspartate (unresolved) residues have chemical shifts that are not clearly indicative of β-strands, suggesting that the majority of these residues are in non-β-strand segments (Table 1).

In solid state NMR experiments on HET-s(218–289) fibrils, Meier and co-workers (38, 39) observed that roughly half of the amino acid sequence does not contribute to solid state NMR spectra. The invisible segments correspond to highly flexible loops (40). Similar behavior has been reported by Helmus et al. (41) in solid state NMR studies of Y145Stop PrP fibrils. To search for such highly flexible segments in CsgA fibrils, we performed two-dimensional 13C-13C NMR measurements with RAD mixing (Fig. 8D) and two-dimensional 15N-13C NMR measurements (Fig. 9) from which cross-peak signals arising from pairs of residues that occur consecutively in the CsgA sequence can be identified. The two-dimensional 13C-13C RAD spectrum shows cross-peaks attributable to glycine/asparagine, asparagine/serine, threonine/alanine, alanine/serine, and threonine/glutamine pairs. The order of residues within these pairs cannot be determined from the two-dimensional 13C-13C RAD spectrum. The two-dimensional 13N-13C spectra allow the identification of backbone 15N chemical shifts for certain residue types (Fig. 9A) and cross-peaks attributable to glycine/glycine, asparagine/glycine, asparagine/serine, glycine/asparagine, threonine/glutamine, and threonine/alanine residue pairs in the specified order (Fig. 9B). These observations argue against the possibility that the glycine-rich segments of CsgA might be highly flexible. We do not have evidence that large segments of CsgA are invisible in solid state NMR measurements.

Electron Microscopy Reveals Narrow CsgA Fibrils—Electron microscopic comparisons of CsgA fibrils with tobacco mosaic virus, HET-s prion domain fibrils, and Ure2p fibrils indicate that CsgA forms a very narrow fibril similar in size to that of HET-s(218–289) (Fig. 10). The individual CsgA fibrils are ~3 nm in diameter, are of various lengths, and frequently associate laterally to form larger fibril bundles (Fig. 10A). It is primarily these fibril bundles that are seen in Fig. 1. The single fibrils are similar in diameter to the fibrils formed by the HET-s prion domain that have diameters of 4 or 5 nm according to the dimensions of the NMR structure (38) or previous electron microscopy measurements (42). By comparison, a fiber composed of multiple CsgA fibrils is dwarfed by a single fibril of the full-length yeast prion protein Ure2p (Fig. 10C), which has a single polypeptide per 4.7 Å along the fibril axis (43) and thus has much more mass per unit length than fibers formed by HET-s(218–289).

Mass per Length Measurements—An in-register parallel β-sheet structure has one molecular mass per 4.7 Å, whereas a single β-helix should have half (or less) this value. We used dark field transmission electron microscopy to measure the mass per unit length of unstained CsgA filaments by a method developed by Chen et al.3 Measurements on the thinnest visible filaments gave an average of 15.3 kDa/nm (Fig. 11). Given that the molecular mass of CsgA is 13.916 kDa, a mass per unit length value of 29.6 kDa/nm is expected for an in-register parallel β-sheet structure. The value of 15.3 kDa/nm is a maximum value because it is uncertain that single unbundled fibers are detected under the dark field conditions.

DISCUSSION

Proteins from a variety of organisms can form amyloid both in vitro and in vivo, and the propensity to form amyloid may be a common property of many proteins (2); there is no single amino acid sequence or defining sequence motif that determines amyloid-forming proteins except a relatively low amount of charged residues (35, 44). Because of its association with many diseases, amyloid is typically viewed as an aberrant conformation that is a consequence of protein misfolding. However, there is a precedent for some proteins to form functional amyloid in which the large filamentous aggregates represent the native and functional state; this is a departure from the traditional model that amyloid is solely a result of misfolding (for a review, see Ref. 3). Such proteins form biologically functional fibrils that have the biophysical hallmarks of other amyloids but are structurally distinct from other biological filaments formed by proteins such as actin or tubulin. Implicitly native state amyloids are assumed to be based upon structural elements similarly to disease-associated amyloids.

Curli fibrils represent a functional state and are not products of misfolding, and because they have many of the characteris-

3 B. Chen, K. R. Thurber, F. Shewmaker, R. B. Wickner, and R. Tycko, manuscript in preparation.
tics of amyloid, they have been proposed to be a functional amyloid (10). Given the high sequence identity between the curli proteins CsgA and CsgB and their ability to cross-polymerize (9, 11, 45), it is likely that they adopt similar protein folds in curli fibrils. Using homology modeling, individual CsgA and CsgB subunits have previously been predicted to form \( \beta \)-helices (7, 46, 47), which are common folds for proteins secreted by bacteria (48, 49). A \( \beta \)-helix (also known as \( \beta \)-solenoid or \( \beta \)-roll) is a rodlike structural motif formed by repetitive alternation of short \( \beta \)-strands and bends with \( \beta \)-strands running perpendicular to the rod and intramolecular hydrogen bonds between \( \beta \)-strands running parallel to the rod as in a cross-\( \beta \) structure. If curli fibrils are based on linearly arranged (stacked) \( \beta \)-helices, they are structurally distinct from most of the amyloids studied to date with the exception of fibrils formed by residues 218–289 of the HET-s fungal prion protein (38).

In this work, we attempt to further characterize curli structure using several biophysical methods, including solid state NMR. Full-length curli proteins CsgA and CsgB were expressed and purified using an E. coli expression system with no specific csg mutations. Both recombinant proteins form fibrils in vitro that, like in vivo curli, have many of the characteristics of amyloid. Moreover the x-ray and electron diffraction data presented here indicate that curli fibrils have a cross-\( \beta \) structure, which is often considered the definitive characteristic of amyloid. To date, most amyloid fibrils formed by peptides or proteins longer than 15 residues have been found to be based on in-register parallel \( \beta \)-sheet structures (for a review, see Ref. 50). Some important examples of in-register parallel \( \beta \)-sheet structures are the disease-associated amyloids formed by the A\( \beta \) peptide (Alzheimer disease) (26, 51), \( \alpha \)-synuclein (Parkinson disease) (15), amylin (type II diabetes) (14, 23), and tau (Alzheimer disease) (16). Likewise three different yeast prions share a similar in-register parallel \( \beta \)-sheet structure (17–19).

The PITHIRDS-CT data in Fig. 7 imply that CsgA and CsgB fibrils do not contain in-register parallel \( \beta \)-sheet structures as we found no evidence for the \( \sim 5 \) Å intermolecular distances between \( 13C \)-labeled backbone carbonyl sites that would exist in such structures. Whether CsgA and CsgB fibrils contain \( \beta \)-helical structures or some other variant of a cross-\( \beta \) structure cannot be determined definitively from our current data. Nonetheless several features of the data merit discussion. First, \( 13C \) chemical shifts at certain residues do not follow the trends expected in \( \beta \)-strand segments, i.e. reduced chemical shift values for carbonyl and \( \alpha \)-carbon sites and increased values for \( \beta \)-carbon

**FIGURE 7.** Dipolar recoupling NMR experiments indicate curli amyloid is not based on parallel in-register \( \beta \)-sheet structure. The graphs show \( 13C \) signal decays of specifically labeled carbonyl positions using the PITHIRDS-CT method. All graphs include simulated PITHIRDS-CT curves for ideal chains of \( 13C \) with spacing of 4.7 Å. A, \( 13C \)-tyrosine-labeled Sup35NM fibrils adjacent to a representative schematic of their in-register parallel \( \beta \)-sheet amyloid structure. B, \( 13C \)-tyrosine-labeled HET-s(218–289) fibrils adjacent to a representative schematic of their \( \beta \)-helix amyloid structure. C, CsgA fibrils separately labeled at the carbonyl positions of valine, leucine, or phenylalanine. The contributions of the adjacent valines at the amino terminus of the CsgA sequence were subtracted from the valine signal values with the assumption that they have signal decays similar to the simulated 3-Å dipole-dipole distance. D, \( 13C \)-tyrosine-labeled CsgB fibrils and unlabeled CsgA fibrils.
sites (relative to random coil chemical shifts). The carbonyl line shapes for leucine and phenylalanine sites in Fig. 6 indicate that a fraction of these residues may be in non-β-strand segments. In Fig. 8C, a minor fraction of the valine and alanine residues exhibits non-β-strand chemical shifts for α- and β-carbons. Also in Fig. 8C, the strong Cα/Cβ cross-peak assigned to asparagine (16 copies in the CsgA sequence) and aspartate (eight copies) residues occurs at a non-β-strand position (Table 1). In contrast, we did not detect serine, threonine, or glutamine/glutamate NMR signals indicative of non-β-strand conformations. Thus, we have preliminary evidence for alternating β-strand and bend segments in CsgA fibrils. Second, the two-dimensional 13C-13C spectrum in Fig. 8D indicates that glycine/asparagine, asparagine-serine, alanine-serine, threonine/alanine, and threonine/glutamine pairs that occur consecutively in the CsgA sequence (in either order) are sufficiently rigid that they contribute to solid state NMR signals. The two-dimensional 13N-13C spectra in Fig. 9 indicate that glycine/glycine, asparagine/glycine, glycine/asparagine, asparagine/serine, alanine/serine, threonine/alanine, and threonine/alanine pairs (in this order) are rigid. The significance of these observations is that we did not find evidence for the existence of flexible loops that separate structurally ordered segments as has been reported for HET-s(218–289) fibrils (38, 39). Third, CsgA fibrils have an apparent diameter of 3–4 nm in negatively stained transmission electron microscopy images (Fig. 10). If ~100 or more residues participate in the fibril core structure, this diameter seems inconsistent with any structure in which each 4.7-Å β-sheet spacing along the fibril axis contains an entire CsgA molecule. For example, 2-fold symmetric and 3-fold symmetric structures for 40-residue Aβ peptide fibrils (containing ~60 and 90 residues per β-sheet spacing in the fibril core) have significantly larger diameters in negatively stained transmission electron microscopy images (26). The small diameter of CsgA fibrils would be consistent with a structure in which each CsgA molecule extends over several β-sheet spacings so that the structural repeat distance along the fibril axis is a multiple of 4.7 Å. Consistent with these estimates, our mass per unit length measurements place an upper limit of about one-half a molecular mass

### TABLE 1

13C NMR chemical shifts in CsgA fibrils determined from the two-dimensional spectrum in Fig. 8C

| Residue type       | Cα chemical shift | Cα random coil value (36) | Cβ chemical shift | Cβ random coil value (36) | Predicted secondary structure |
|--------------------|-------------------|----------------------------|-------------------|---------------------------|-------------------------------|
| Alanine (major)    | 48.9 ppm          | 50.8 ppm                   | 20.8 ppm          | 17.4 ppm                  | β                             |
| Alanine (minor)    | 53.2 ppm          | 50.8 ppm                   | 16.9 ppm          | 17.4 ppm                  | Non-β                         |
| Valine (major)     | 58.4 ppm          | 60.5 ppm                   | 33.1 ppm          | 31.2 ppm                  | β                             |
| Valine (minor)     | 59.7 ppm          | 60.5 ppm                   | 30.7 ppm          | 31.2 ppm                  | Non-β                         |
| Asparagine/aspartate | 52.1 ppm        | 51.4/52.5 ppm              | 37.7 ppm          | 37.2/39.4 ppm             | Non-β                         |
| Glutamine/glutamate | 52.1 ppm      | 54.0 ppm                   | 31.2 ppm          | 27.7 ppm                  | β                             |
| Serine             | 55.4 ppm          | 56.6 ppm                   | 64.2 ppm          | 62.1 ppm                  | β                             |
| Threonine*         | 59.9 ppm          | 60.1 ppm                   | 68.8 ppm          | 68.1 ppm                  | β                             |

* The threonine CO chemical shift is 171.2 ppm, whereas the random coil value is 173.0 ppm.
per 4.7 Å. A β-helix-like structure is one possibility but not the only one.

One impetus for studying curli fibrils was the possibility that uniformly 15N,13C-labeled CsgA would exhibit exceptionally sharp, well resolved solid state NMR lines, which would facilitate full structure determination as demonstrated previously for HET-s(218–289) fibrils (38, 39). The functional nature of curli fibrils suggests that the CsgA and CsgB sequences are evolved to adopt specific molecular structures without the structural variations or polymorphism observed in fibrils formed by A/H9252, amylin, and yeast prion proteins (for reviews, see Refs. 35, 50, and 52). A single molecular structure without internal disorder might then be expected to result in sharp solid state NMR lines. The two-dimensional solid state NMR spectra in Figs. 8 and 9 do not support this expectation. Thus, the observation of exceptionally sharp solid state NMR lines in HET-s(218–289) fibrils (and in fibrils formed by a truncated version of mammalian prion protein (41)) may not be solely a consequence of an evolved biological function. Instead the solid state NMR line widths may also be affected by lateral association of fibrils (or the absence thereof), by the amplitudes and time scales of internal molecular motions, or by other factors not yet identified.

The definition of amyloid includes several architectures with cross-β-strands as a common feature. Functional amyloid may represent the cell’s adoption of a pathology to do useful work,
but there may prove to be a general structural distinction between pathologic and functional amyloids. The in-register parallel structure of yeast prions and amyloids is compatible with the existence of prion variants (for a review, see Ref. 53) and amyloid variants (26), but a functional amyloid probably should have only one structure.

Acknowledgments—We thank Ulrich Baxa for helpful suggestions, Eric Anderson for mass spectrometry, and Bo Chen for help with mass per length measurements.

REFERENCES

1. Virchow, R. (1853) C. R. Acad. Sci. 37, 860–861
2. Stefani, M., and Dobson, C. M. (2003) J. Mol. Med. 81, 678–699
3. Fowler, D. M., Koulov, A. V., Balch, W. E., and Kelly, J. W. (2007) Trends Biochem. Sci. 32, 217–224
4. Podrabsky, J. E., Carpenter, J. F., and Hand, S. C. (2001) Am. J. Physiol. Regul. Integr. Comp. Physiol. 280, R123–R131
5. Mackay, J. P., Matthews, J. M., Winefield, R. D., Mackay, L. G., Haverkamp, R. G., and Templeton, M. D. (2001) Structure 9, 83–91
6. Baxa, U. (2007) Biochemistry 46, 13149–13162
7. Paravastu, A. K., Leapman, R. D., Yau, W. M., and Tycko, R. (2008) Proc. Natl. Acad. Sci. U.S.A. 105, 18349–18354
8. McPhie, P. (2001) Anal. Biochem. 293, 109–119
9. Kryndushkin, D. S., Shevchenko, F., and Wickner, R. B. (2008) EMBO J. 27, 2725–2735
10. Collins, S. K., Emödy, L., Müller, K. H., Trust, T. J., and Kay, W. W. (1991) J. Bacteriol. 173, 4773–4781
11. Fearon, T. D., Yau, W. M., and Tycko, R. (2000) J. Mol. Biol. 302, 793–804
12. Baxa, U., Wickner, R. B., Steven, A. C., Anderson, D. E., Marekov, L. N., Yau, W. M., and Tycko, R. (2007) Biochemistry 46, 13149–13162
13. Shewmaker, F., Wickner, R. B., and Tycko, R. (2006) Proc. Natl. Acad. Sci. U.S.A. 103, 19754–19759
14. Wickner, R. B., Dyda, F., and Tycko, R. (2008) Proc. Natl. Acad. Sci. U.S.A. 105, 2403–2408
15. McPhie, P. (2001) Anal. Biochem. 293, 109–119
16. Kryndushkin, D. S., Shevchenko, F., and Wickner, R. B. (2008) EMBO J. 27, 2725–2735
17. Collins, S. K., Emödy, L., Müller, K. H., Trust, T. J., and Kay, W. W. (1991) J. Bacteriol. 173, 4773–4781
18. Fearon, T. D., Yau, W. M., and Tycko, R. (2000) J. Mol. Biol. 302, 793–804
19. Baxa, U., Wickner, R. B., Steven, A. C., Anderson, D. E., Marekov, L. N., Yau, W. M., and Tycko, R. (2007) Biochemistry 46, 13149–13162
20. Shewmaker, F., Wickner, R. B., and Tycko, R. (2006) Proc. Natl. Acad. Sci. U.S.A. 103, 19754–19759
21. Wickner, R. B., Dyda, F., and Tycko, R. (2008) Proc. Natl. Acad. Sci. U.S.A. 105, 2403–2408
22. McPhie, P. (2001) Anal. Biochem. 293, 109–119
23. Kryndushkin, D. S., Shevchenko, F., and Wickner, R. B. (2008) EMBO J. 27, 2725–2735
24. Collins, S. K., Emödy, L., Müller, K. H., Trust, T. J., and Kay, W. W. (1991) J. Bacteriol. 173, 4773–4781
25. Fearon, T. D., Yau, W. M., and Tycko, R. (2000) J. Mol. Biol. 302, 793–804
26. Paravastu, A. K., Leapman, R. D., Yau, W. M., and Tycko, R. (2008) Proc. Natl. Acad. Sci. U.S.A. 105, 18349–18354
27. Shewmaker, F., Ross, E. D., Tycko, R., and Wickner, R. B. (2008) Biochemistry 47, 4000–4007
28. Balbach, J. J., Ishii, Y., Antzutkin, O. N., Leapman, R. D., Rizzo, N. W., Dyda, F., Reed, J., and Tycko, R. (2000) Biochemistry 39, 13748–13759
29. Ishii, Y., Balbach, J. J., and Tycko, R. (2001) Chem. Phys. 266, 231–236
30. Morcombe, C. R., Gaponenko, V., Byrd, R. A., and Zilm, K. W. (2004)
Biophysical Characterization of Curli Fibrils

J. Am. Chem. Soc. 126, 7196–7197
31. Takegoshi, K., Nakamura, S., and Terao, T. (2001) Chem. Phys. Lett. 344, 631–637
32. Wang, X., Smith, D. R., Jones, J. W., and Chapman, M. R. (2007) J. Biol. Chem. 282, 3713–3719
33. Baxa, U., Cheng, N., Winkler, D. C., Chiu, T. K., Davies, D. R., Sharma, D., Inouye, H., Kirschner, D. A., Wickner, R. B., and Steven, A. C. (2005) J. Struct. Biol. 150, 170–179
34. Serpell, L. C., Berriman, J., Jakes, R., Goedert, M., and Crowther, R. A. (2000) Proc. Natl. Acad. Sci. U.S.A. 97, 4897–4902
35. Tycko, R. (2006) Q. Rev. Biophys. 39, 1–55
36. Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S., and Sykes, B. D. (1995) J. Biomol. NMR 5, 67–81
37. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1991) J. Mol. Biol. 222, 311–333
38. Wasmer, C., Lange, A., Van Melckebeke, H., Siemer, A. B., Riek, R., and Meier, B. H. (2008) Science 319, 1523–1526
39. Ritter, C., Maddelein, M. L., Siemer, A. B., Lührs, T., Ernst, M., Meier, B. H., Saupe, S. I., and Riek, R. (2005) Nature 435, 844–848
40. Siemer, A. B., Arnold, A. A., Ritter, C., Westfeld, T., Ernst, M., Riek, R., and Meier, B. H. (2006) J. Am. Chem. Soc. 128, 13224–13228
41. Helmus, J. J., Surewicz, K., Nadaud, P. S., Surewicz, W. K., and Jaroniec, C. P. (2008) Proc. Natl. Acad. Sci. U.S.A. 105, 6284–6289
42. Sabaté, R., Baxa, U., Benkemoun, L., Sánchez de Groot, N., Coulal-Salin, B., Maddelein, M. L., Malato, L., Ventura, S., Steven, A. C., and Saupe, S. J. (2007) J. Mol. Biol. 370, 768–783
43. Baxa, U., Taylor, K. L., Wall, J. S., Simon, M. N., Cheng, N., Wickner, R. B., and Steven, A. C. (2003) J. Biol. Chem. 278, 43717–43727
44. Chiti, F., and Dobson, C. M. (2006) Annu. Rev. Biochem. 75, 333–366
45. Hammar, M., Bian, Z., and Normark, S. (1996) Proc. Natl. Acad. Sci. U.S.A. 93, 6562–6566
46. Collinson, S. K., Parker, J. M., Hodges, R. S., and Kay, W. W. (1999) J. Mol. Biol. 290, 741–756
47. White, A. P., Collinson, S. K., Banser, P. A., Gibson, D. L., Paetzel, M., Strynadka, N. C., and Kay, W. W. (2001) J. Mol. Biol. 311, 735–749
48. Kajava, A. V., and Steven, A. C. (2006) Adv. Protein Chem. 73, 55–96
49. Jenkins, J., and Pickersgill, R. (2001) Prog. Biophys. Mol. Biol. 77, 111–175
50. Margittai, M., and Langen, R. (2008) Q. Rev. Biophys. 41, 265–297
51. Petkova, A. T., Yau, W. M., and Tycko, R. (2006) Biochemistry 45, 498–512
52. Baxa, U., Cassese, T., Kajava, A. V., and Steven, A. C. (2006) Adv. Protein Chem. 73, 125–180
53. Wickner, R. B., Shewmaker, F., Kryndushkin, D., and Edskes, H. K. (2008) BioEssays 30, 955–964