Single molecule studies of the bacterial curli protein CsgA reveal a structurally dynamic monomeric structure

Daniel A.N. Foster, Derek R. Dee.

Food, Nutrition and Health, Faculty of Land and Food Systems, The University of British Columbia, 2205 East Mall, Vancouver, BC, V6T 1Z4, Canada

For correspondence: Derek.Dee@ubc.ca

Abstract

Bacterial biofilm constructed with functional amyloid are a substantial health concern. The major subunit of curli fibrils, CsgA, has recently had its amyloid structure explored by cryoEM, but the folding of the subunit had yet to be carefully examined at the level of the single molecule. We describe the first *in vitro* single molecule folding trajectories of CsgA, demonstrating a dynamic mixture of metastable states or a molten globule, but also unfolding of the R domains. In order to content with the aggregation propensity of this protein, we developed an assay to characterise the observation of individual folding trajectories of a single molecule of CsgA using force spectroscopy with optical tweezers. This work was made possible by an engineered disulfide trap and the controlled reduction of its cysteines. When folding, CsgA does not get locked into a β-solenoid, but may be found in a molten globule or metastable state in about a quarter of the times it is unfolded.

Introduction

A biofilm serves as a bacterial community's home, protecting against environmental stressors such as antibotics. The functional amyloid known as curli, present in *E. coli* biofilms, is one of the best studied amyloidal systems. CsgA is the subunit of the curli fibers, though there are many folding chaperones and inhibitors involved in its synthesis pathway. However, atomistic details of fibril assembly and the CsgA amyloid structure remain unresolved. Progress to understanding CsgA's structure have come from modelling and cryoEM, among other techniques. Monomeric CsgA is intrinsically disordered (*1*) and highly prone to aggregation, confounding high-resolution studies at the ensemble level. Single molecule force spectroscopy (SMFS) is well placed to elucidate the folding pathway of CsgA, and resolve the minutia of CsgA's assembly into amyloid by resolving the protein's behaviour that is hidden from ensemble averaging.

Curli accounts for 85% of the matrix materials in curli-related biofilms (*2*). It organises into aggregates with their beta-sheet structures aligned perpendicular to the long fibril-axis (cross-β architecture) (*3*), forming amyloid. These fibrils are thermodynamically stable, capable of self-propagation, have a high aspect ratio (*e.g.* 10-20 nm wide and 10^2-10^3 nm long), provide high mechanical strength (*4*) and are protease resistant (*5–8*). Given amyloid's key role in the scaffolding of biofilm, it is not surprising that such amyloidal proteins are highly conserved across genera (*9*, *10*). Thus, to further development of novel anti-biofilm/anti-bacterial treatments (*11*) a detailed understanding of the molecular basis the amyloid structure is needed.

The curli proteins from *E. coli* are the best characterised bacterial amyloid system. Six proteins encoded by curli-specific genes (csg) play a unique role in the tightly controlled process of CsgA fibril synthesis (Figure 1). Aggregation of the intrinsically disordered protein CsgA is nucleated at the outer cellular membrane by CsgB (*12*), but cytotoxic aggregates are inhibited in the periplasm by chaperone proteins CsgC and CsgH (*13*, *14*). The secretion of CsgA to the outer membrane is controlled by CsgG, CsgE and CsgF (*15*), where CsgB nucleates CsgA fibril formation (*16*). CsgF is anchored to the outer membrane and may anchor CsgB to the membrane (*17*). Currently, most of the mechanistic details are unknown (*2*, *16*, *18*), such as the dynamic structure of CsgA or CsgB monomers, how CsgB nucleates CsgA amyloid formation, or how CsgC, CsgE and CsgH inhibit fibril formation (*2*, *14*).

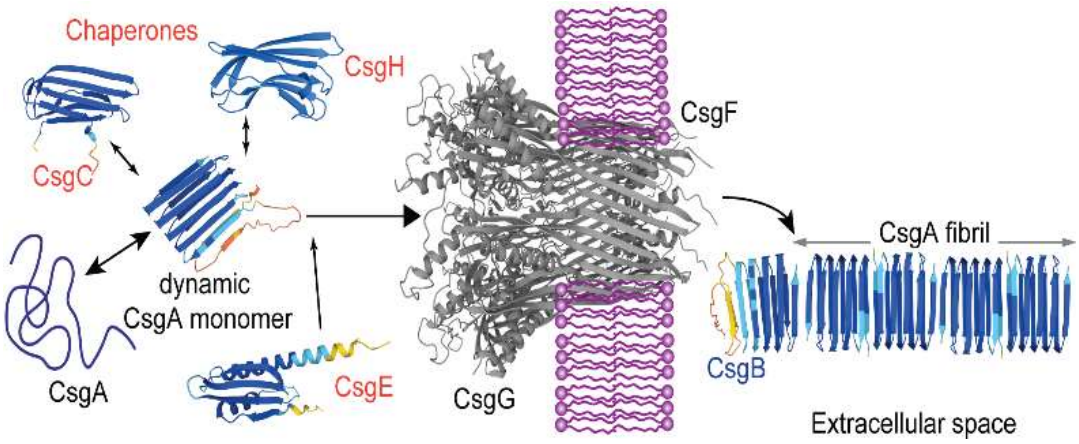

Figure 1 **Overview of regulated CsgA amyloid formation**. Solubility of CsgA in the periplasm is maintained by CsgC and CsgH amyloid inhibitors while it is exported out of the cell through the CsgG-CsgF β-barrel pore complex in the outer membrane, for incorporation into fibrils with CsgB acting as a nucleator for biofilm supporting scaffolds.

Progress towards the amyloid structure has been made recently through cryoEM imaging of a CsgA R-domain analogue (*19*) and with an engineered CsgA with residue substitutions to control aggregation onset (*20*). Ensemble studies of CsgA have not concluded on set structures, rather indications from circular dichroism spectrum pointed towards CsgA as unstructured (*21*). Moreover, cryoEM structures raise questions about monomer orientation within the amyloid fibril (*22*). Some protein-fusion CsgA magnetic tweezers measurements have been published (*23*), but a clear understanding of behaviour observed in that assay is not yet ready.

Why single molecule experiments? Rare or transient subpopulations are averaged out in ensemble measurements. To understand the folding pathways comprehensively, we leverage optical tweezers (OT) for *in vitro* measurements, while controlling the onset of the CsgA folding with an engineered disulfide reduction trigger. OT measurements of intrinsically disordered proteins linked to disease amyloid (*such as the prion and α-synuclein proteins*) have established the type of results that can be expected from this mode of investigation (*24–26*).

The structure of CsgA amyloid has been studied in at least two cryoEM reconstructions. Sleutel *et al.* took a wide-ranging view of CsgA analogues in bacteria and opted for a 15.5 R repeat structure (*19*). Our colleagues made use of the internal cysteine amino acid change CsgA protein (A63C/V140C) developed by the Chapman lab (*27*), allowing fibrillation to be turned on by reducing agent [PDB: 8ENR]. Preceding these structural determinations, foundational modelling work was done by DeBenedictis *et al.*, examining the results of structural predictions and their stability under molecular dynamics simulations out to 150 ns (*28*). Since that time, publicly facing webservers are available for structure prediction, such as Robetta (*29*) (using *de novo* structure prediction methods) and AlphaFold 3 (AF3; using machine learning based on the records of the PDB), which predict similar CsgA structure. However, bulk studies show that these proteins are ostensibly disordered, and importantly, the deep learning methods are biased from sampling amyloid fragments that crystallise well. The β-hairpin solenoid shown by DeBenedictis *et al.* also appears in machine learning results.

The N22 region of the protein sequence is generally predicted as unstructured or containing β-sheet regions, while the modelling used for the cryoEM structure (*20*) expects this to be unstructured. For the structure reconstruction during cryoEM, which is an endpoint ensemble average, the N22 was assumed to point out into the solvent like a noodle, such that it was not included in the predicted structure docking.

Aggregation kinetics of CsgA using thioflavin T (ThT), indicate a concerted sigmoid behaviour with an exponential rise to plateau within 10 hours. We have replicated these ThT kinetics with our CsgA constructs as well (see Figure 2B). However, before this fibrillation process, the CsgA is not locked-in to this β-solenoid form, as shown by circular dichroism (*21*). Thus, the exploration of CsgA's folding process for intrinsically disordered protein (IDP) behaviour is warranted. Single molecule methods with optical tweezers (OT) offer *in vitro* measurements in desired buffer and temperature conditions for down to angstrom resolution with picoNewtons (pN) of force (*30*). OT and other single molecule methods can also be leveraged to probe the dynamics and interactions of proteins within complexes (*31*). Our work herein focuses on the folding pathways and intermediate states revealed by OT, which were previously hidden at the ensemble level.

Results

Controlling the onset of fibrillation was important for producing CsgA linked with dsDNA for single molecule measurements. We used CsgA with internal cysteines and with ybbR linkers on its flanks (see table 1) in which fibrillation is prevented when the disulfide is formed. The ybbR tags allowed for enzymatic attachment to DNA handles, which were linked to functionalised beads in a dumbbell arrangement.

Both the wild type CsgA and that with residue substitutions A63C & V140C were purified using FPLC then tested for fibrillation propensity with ThT fluorescence and amyloid fibrils compared under transmission electron microscopy to verify its similarity (see Figure 2). The introduction of a reducing agent quickly led to the exponential growth curve of amyloid fibrils (Figure 2B).

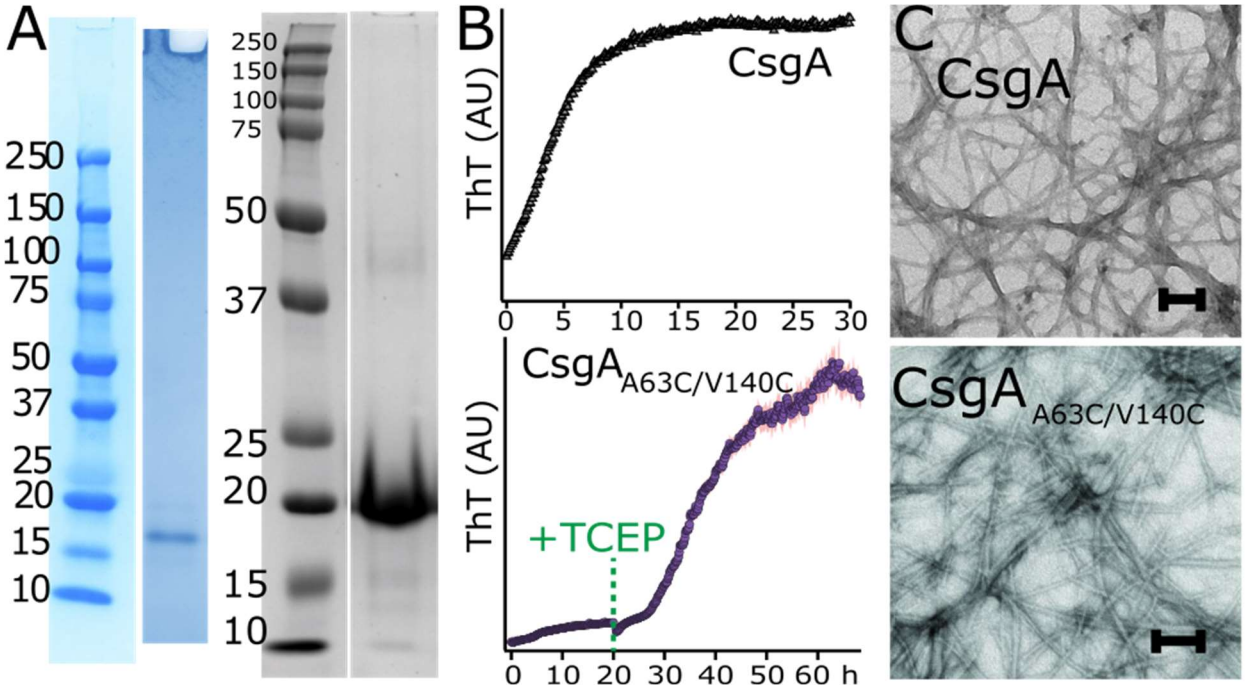

Figure 2 **Purification and qualification of CsgA**. **A**: SDS-PAGE of wild type (left) and mutant (right) CsgA. **B**: Aggregation kinetics from thioflavin T (ThT) binding fluorescence (AU) of CsgA wt (top) and mutant CsgA (bottom). **C**: TEM microscopy of wt (top) and mutant (bottom) CsgA (scale bar 100 nm).

CsgA constructs were initially screened for a single tether in running buffer without reducing agent. The data was fit offline to a worm-like chain (WLC) (*32*) to confirm that it was a single tether (Figure 3A). Before reduction, with the internal disulfide intact, force extension curves (FECs) were smooth with no apparent transitions. After the addition of a reducing agent (Figure 3B inset), the FECs showed transitions and a total contour length consistent with the length of the dsDNA plus reduced protein construct. The unfolding of CsgA showed two different qualitative behaviours: large and discrete unfolding events consistent with the β-solenoid made up of β-sheets arranged into hairpins for each R domain, as shown in modelling in the literature as well as RoseTTAFold and by AF3. The other pattern was of a lower force unfolding, often without discrete steps and with greatly reduced hysteresis. Examples are shown in Figure 3B, 4B with WLC fit lines. In Figure 3, the process of disulfide bond reduction is illustrated: when a reducing agent is introduced, by moving the dumbbells into a flow channel where mercaptoethylamine (MEA) is in solution, at which point when the cysteines are reduced, the contour length of flexible polypeptide increases significantly, and the FEC has a notable elongation, and subsequent unfolding FECs may show discrete evens characteristic of R domain unfolding. However, while not in a cysteine reducing buffer, CsgA would occasionally show a small unfolding event; see Figure 3C. In Figure 4A, is present a general unfolding example for CsgA, from all R domains folded together to the molecule fully unfolded. In this example, divergence from the R1-R5 fit line can be seen, as well as with the subsequent FEC that happens to finally unfold from approximately

3.5 R domains instead. A rough shoulder feature appears in the FECs before large unfolding events, where the contour length change from WLC fittings indicates the involvement of the N22 domain.

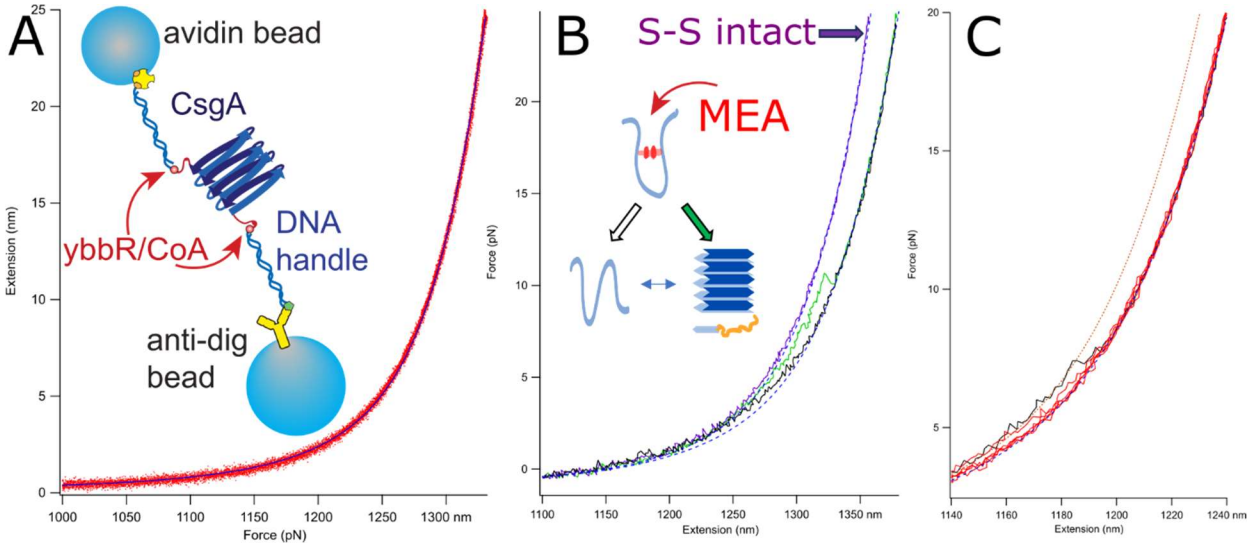

Figure 3 **Example single molecule FEC of CsgA**. **A**: initial overlaid force extension curves (FECs) without reducing agent and an intact disulfide bond, engineered to trap CsgA in a disordered or non-amyloid fibril competent state. A WLC fit used to screen for single tethers is overlaid. Inset: diagram of CsgA with linkers to dsDNA handles (not to scale). **B**: Inset: addition of reducing agent MEA to break the disulfide bond and thus allow wild type CsgA folding behaviour. In violet, the FEC with the bridge intact, fit with 35.1 nm of free polypeptide, then black fully unfolded refolding FEC with 62.5 nm of free polypeptide. In green, unfolding FEC with cysteines reduced. **C**: Structure of intact disulfide bridge CsgA: occasionally a 10-13 nm unzip is seen before the introduction of a reducing agent, which is consistent with the free regions of R1 and R5 outside of the disulfide forming structure.

We analysed eight CsgA molecules with over 200 pulls overall, examining the effects of pulling velocities (50, 100 and 200 nm/s) and refolding times at low force (200 ms to several seconds) on unfolding of CsgA. Results of FEC behaviour are summarised in the scatterplot in Figure 4B. Unfolding data from less discrete FECs is not included, but it falls generally in the 4-8 pN range, which is empty in the plot. All discrete unfolding events are reduced to histogram form to look for clustering, in bins of 1 nm for contour length change and 1 pN for unfolding force in Figure 5. The clustering of events between 5 and 10 pN was also the non-discrete un- and re-folding range seen in our FECs.

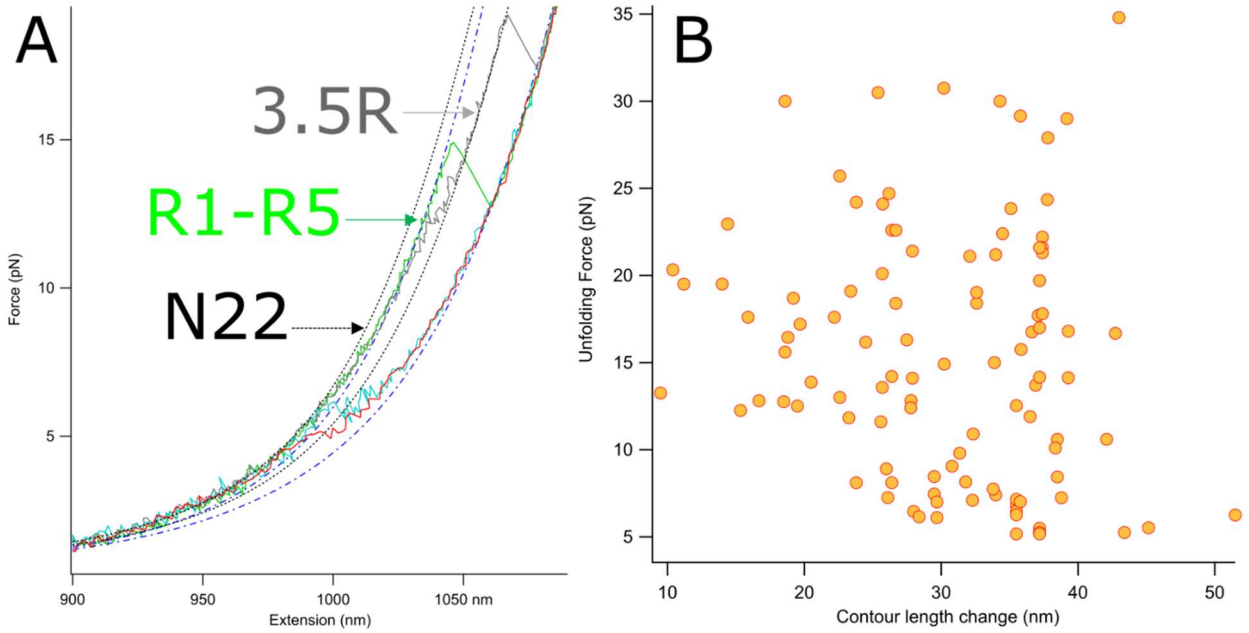

Figure 4 **Characteristic CsgA FEC behaviour and unfolding force vs contour length change of the last unfolding events**. **A**: Example FECs showing discrete unfolding of states: R1 through R5 in green, approximately 3.5 R domains in grey. While another unfolding FEC in cyan went through a rough unfolding process, very close to the refolding FEC in red. WLC fits for the fully unfolded state and for R1-R5 are shown as blue dashed lines. The range to the N22 region is in dashed black, as is the upper ~3.5R fit. The beads were moved apart at a constant velocity of 200 nm/s, with a 200 ms refolding time between curves. **B**: The last discrete unfolding event data from all the FECs, as with the example in panel 4A is shown in the scatterplot. (Much of the contour length change from the fully folded state may not occur discretely.)

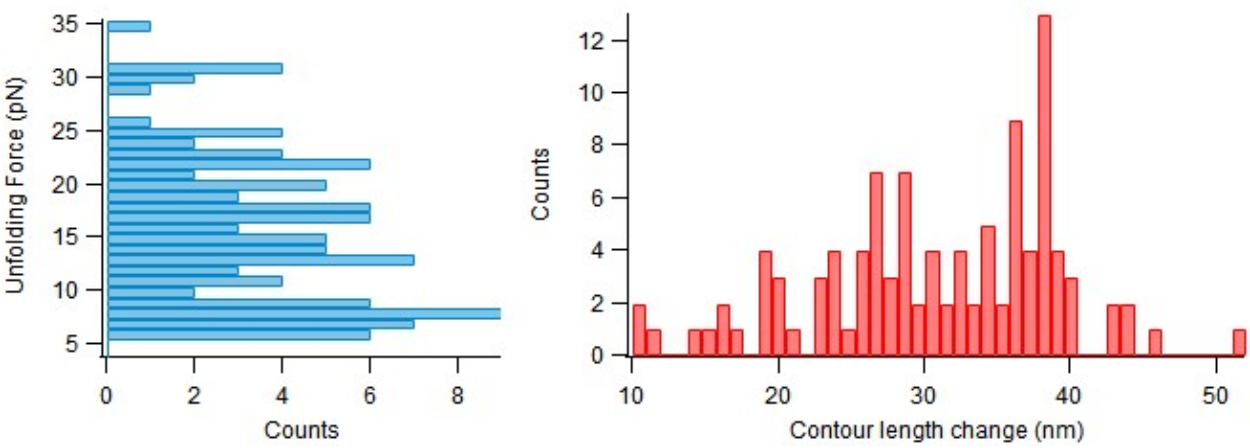

Figure 5 **Clustering of unfolding force vs contour length change of the last unfolding events**. Left: histogram of the unfolding forces of CsgA, with accompanying contour length changes on the right.

A limited number of FECs also showed two or three *strongly discrete* unfolding steps on their path to becoming fully unfolded, as shown in Figure 6. The unfolding length change goes from an unstructured 15.3 nm (for ybbR and GSSS spacer sequences) to a fully unfolded length of 62.5 nm. Most transitions are fairly flat in term of unfolding force, meaning that they typically occur in a very short period after the initial contour length change, though there are significant increases in force that occur at a much smaller rate.

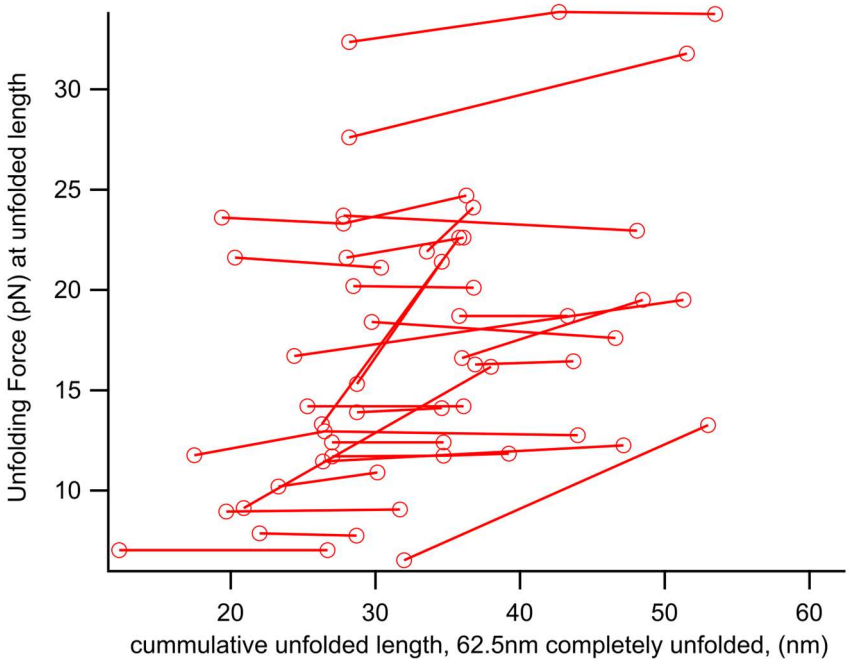

Figure 6 **Discrete unfolding paths of CsgA.** In FECs tabulated from four molecules, multiple discrete unfolding steps can be observed, up to the point where they unfold to 62.5 nm, which is the fully unfolded length of CsgA. The protein starts out with 15.3 nm of unstructured polypeptide.

In force extension curves samples from a refolding time limit of 200 ms, only six of 46 curves (13%), had final unfolding events classifiable as discrete. With a refolding time-limit of 500 ms, this rate rises to 33 of 42 (79%), though nine of these are within 10 pN that overlaps with the non-discrete folders, and with a 1 s refolding time the rate stays similar at 37 of 50 (74%). We have fewer curves restricted to a 2 s refolding time, showing a 16 of 22 (73%) discrete unfolding rate. In one outlier, a non-discrete state is seen after 53 s to refold, but generally, longer waiting times for refolding often show a discrete unfolding event.

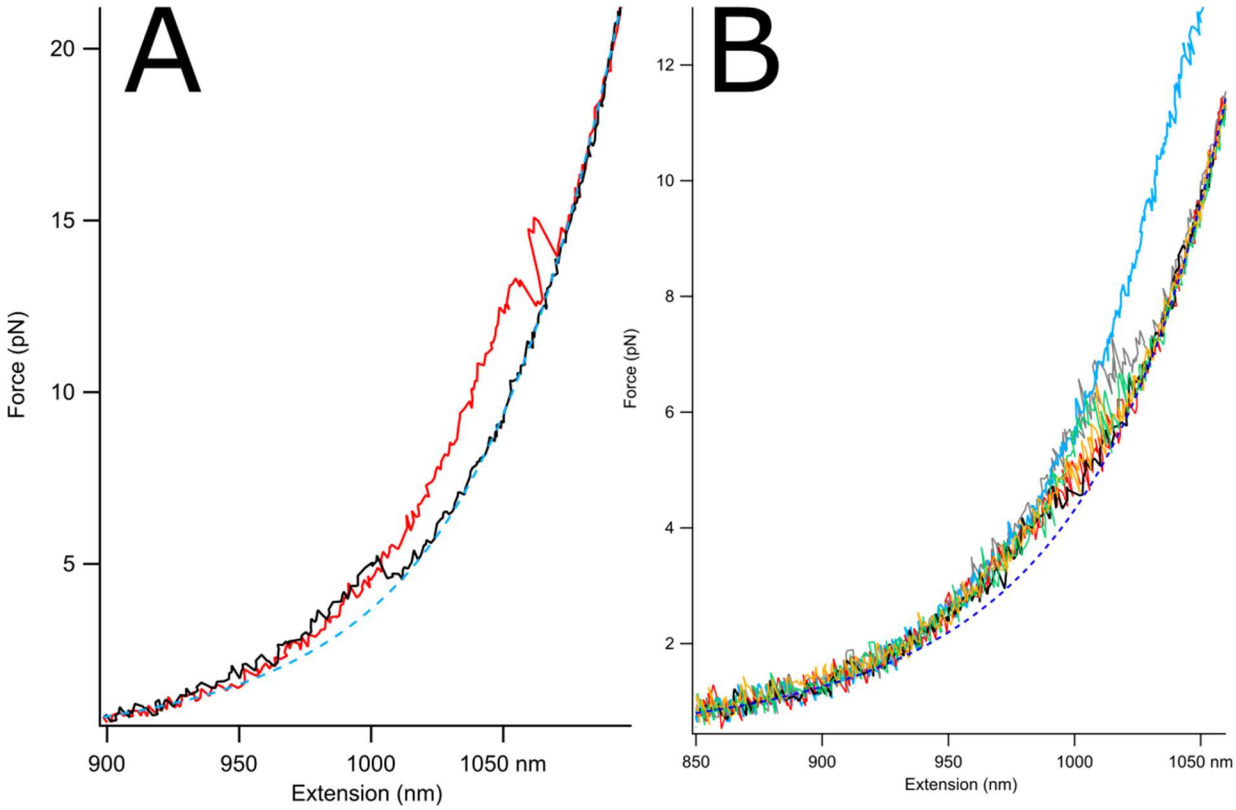

Figure 7 **State switching and metastability in FECs**. **A:** In this example, we have a molecule that unfolded at 13.9 pN yielding 20.5 nm to the fully unfolded state, but switched back into an unfolded state breaking momentarily at 15.6 pN, yielding 18.6 nm. Unfolding curve in red, refolding in black, WLC for the unfolded state in light blue. **B**: Metastability of CsgA: all our refolding FECs pass through a subtle shoulder, but more than 20% of unfolding FEC did not display a *lasting* discrete behaviour near the refolding region. The bright blue FEC here shows a discrete unfolding of ~3.5 R domains at ~18pN, to the fully unfolded WLC fit dashed dark blue line, with a refolding FEC in black. Other unfolding FECs in grey, green, orange and red oscillated between short extension with varying length. Data in panel B is from a 100 nm/s ramp with a 200 ms pause at zero force.

Refolding and Unfolding was also observed to not always be in one direction with the ramping of force. Typically, state switching was observed at forces between 4 to 10 pN, but this phenomenon has been observed as high as 15 pN, as shown in Figure 7A, with a 200 nm/s ramp. (Faster loading rates are farther from equilibrium.) Across all our refolding curves a "shoulder" was observed where the trace diverges from a WLC (Figure 7B). Importantly, in curves that cannot be assigned a clear lasting discrete unfold, a rough trace occurred, rather than the monotonic rise usually seen with dsDNA or with our CsgA pinned with a disulfide bond (Figure 3C). Currently our time resolution limits very from ~5 to 7 ms per data point.

Discussion

CsgA undergoes conformational rearrangement from an initially soluble, intrinsically disordered monomer to form a β-sheet rich amyloid through an unknown mechanism. The first steps of nuclei formation are hypothesized to include either transient folding of two CsgA monomers that then bind together to form the amyloid nucleus, or co-operative binding and folding of two CsgA monomers to form structured dimers (*18*).

In E. coli after CsgA protein synthesis, the signal peptide is cleaved, leaving the five R domains as well as 22 residues on its N terminus (N22), which is then exported out of the cell before forming fibrils nucleated by CsgB. It has been assumed that CsgA is maintained as an intrinsically disordered protein (IDP) in the cell until fibril assembly, and indeed around ~73-79% of FECs show weak, non-discrete, folding and refolding that could be described as a molten globule or a metastable fold (e.g. in Figures 3 and 7). Moreover, there is a data gap (0-10 pN, 10-20 nm) in Figure 3 due to this metastable protein state we observe, where cooperative length changes occur. Clustering of contour length changes can be seen at 20, 27, 37 nm, which can be attributed to 3, 4 and 5 R domains, as shown in Figure 4. However, many contour length changes appear to involve half R domain as well, e.g. 3.5 R domains in Figure 4A. Half of an R domain may be interpreted as the shearing of a β-hairpin. What we can see from the discrete step unfolding data of Figure 6 is that most of these folded states have a limited lifetime, and rapidly unfold at almost the same force.

The non-monotonic rise of our FECs, with a "shoulder" deviation from the WLC is an important nuance. This metastable behaviour of CsgA seen in 20 to 25% of FEC we attribute to a metastable or molten-globule quality of CsgA, and is distinct from the observations of the IDP α-synuclein reported by Solanki *et al.*(*25*, *26*) where they described a monotonic rise without fluctuation. CsgA appears to have a more intermediate behaviour between something like α-synuclein and CsgA's stable or trapped structure present in amyloid fibrils. Examination of a hydropathy plot for CsgA shows significant hydrophilicy, suggesting competition between forming hydrogen bonds with itself and with the solvent, and perhaps that entropy maximisation keeps CsgA metastable. Additionally, the N22 region participates in this shoulder and occasionally appears to be included in a discrete unfolding structure. AF3 predicts a complementary β-strand to R1 with a plDDT between 50 to 90 (depending on the prediction run) within the N22 region as well. Ongoing work in our research group is examining a truncation mutant with this region removed.

Experiments here have two compromises that limits the lifetime of our samples and the forces that can be applied. The ssDNA-CoA oligo attached to the protein of interest are merely annealed to the overhang of the dsDNA handles without ligation (*see Materials and Methods for more detail*). In practice, this means limiting forces to within 40 pN of applied tension of the handles, because high force leads to oligo shearing from the overhang. Five pulls of CsgA did not unfold within 200 ms at 25 or 30 pN during initial screening. The second challenge is the targeted reduction of the internal cysteine disulfide bridge. Aggregation kinetics of this construct were demonstrated with ThT binding (see Figure 2), triggered by the introduction of tris(2-carboxyethyl) phosphine (TCEP). However, TCEP is too strong a reducing agent to incubate with anti-dig functionalised beads, as immunoglobulin G (IgG) has structural disulfide bonds. Instead, MEA is carefully introduced, which is known selectively to reduce the hinge region of IgG, leaving antigen binding sites and Fab regions intact. However, some softening of the IgG was indicated a minor drop in the persistence length and 30% in the modulus of the WLC.

The observation of structure of R1 & R5 partial regions with A63C/V140C disulfide bridge intact: During screening, most tethers in the absence of reducing agent show no unfolding events during pulling, however, occasionally an unzip of 10 to 13 nm in the ~5 to 7 pN range was notable (Figure 3C). Outside of the disulfide bridge at 63C and 140C, R1 and R5 have 32 free amino acid residues, which normally would be part of the β-hairpin structure of their respective R domain. Residues 64 through 139 are bound into a loop with no structure prediction. Balistreri *et al.* hypothesized that the disulfide bridge keeps the CsgA in a disordered state (*14*), but the rapid onset of fibrillation of CsgA A63C/V140C after the introduction of a reducing agent (TCEP) suggests that some seeding aggregates may already be possible. OT results suggest, usually on a first pull, a partial β-hairpin or β-sheet formation sequestering the 32 free residues is possible given the change in contour length. AF3 and RoseTTAFold do not offer a structure prediction with the disulfide intact.

In order to better interpret and understand the unfolding behaviour of CsgA under tension when pulled by optical tweezers, preliminary steered molecular dynamics simulations (LAMMPS) have also been undertaken by our collaborators (K. Kudriavtsev & S. Srebik). They were guided by the work of DeBenedictis and Keten (*33*), with the AF3 predicted structure as a starting point. Contact maps show a dominance of parallel β-sheets under tension with the transient formation of antiparallel β-sheets on the ends of the central R domains, as R3 is the last to unfold. Moreover, the tracking of contact ratios shows those from R2 to R3 being the most persistent. At the high loading rate present in simulations, shearing of the protein structure is a prominent feature. In our FECs, the odd half integer R domain ratios appear to be a reflection of this shearing. By contrast, any IDP, molten-globule or metastable form of CsgA is not modelled by these simulations.

Conclusions

We have presented the first single molecule folding data of CsgA, which presents a unique blend of metastable - molten globule behaviour and the structures found within amyloid fibrils imaged by cryoEM. The R domains are a fixture of discrete unfolding of CsgA, but CsgA does not appear to be locked into an amyloid subunit form. This is in contrast to the assumption from ensemble experiments, such as circular dichroism, pointing towards intrinsically disordered states. Equilibrium measurements with optical tweezers could produce more insights into dominant folding pathways of the metastable state of CsgA. There are engineered mutations of CsgA worth investigating further, namely the 'slowgo' mutant (*34*) and clarifying if the N22 region modifies protein fold stability.

Materials and Methods

Protein expression: C-terminal His_6-tagged CsgA (A63C, V140C) is expressed and purified as previously described with modifications (*18*, *35*, *36*). NEB 3016 ΔslyD cells harboring a pET11d vector encoding C-terminal His_6-tagged CsgA are grown to OD_{600} ~1 in LB broth containing 100 µg/mL ampicillin at 37°C. CsgA expression is induced with 0.5 mM IPTG for 2-3 h and cells are harvested (8000 rpm, 10 min) and stored at -80°C. Cell pellets from 1 L of the induced cell culture are resuspended in 20 mL lysis buffer (8 M GdnHCl, 50 mM K_2HPO_4, pH 7.3) per gram of pellet, sonicated for a few min (20 s intervals), and shaken for 1 h at RT. The supernatant of the cell lysate is collected following centrifugation ($4 \cdot 10^4 \times g$ for 40 min) and mixed with an equivalent volume

of equilibration buffer (50 mM Na_2HPO_4, 300 mM NaCl, 6 M GdnHCl, 10 mM imidazole, pH 7.4) and purified using FPLC and a *HisPur* Cobalt column (Cytiva). Protein is eluted (6 M GdnHCl, 50 mM Na_2HPO_4, 300 mM NaCl, 150 mM imidazole, pH 7.4). Fractions containing CsgA are pooled together and stored at 4°C prior to analysis. All purification steps are conducted under 6 M GdnHCl to prevent premature fibrillation. See Figure 2A for PAGE gel purification results.

Bioconjugate Chemistry: SMFS requires extensive use of bioconjugate chemistry to connect the molecule(s) of interest, DNA handles and beads together to form 'dumbbell tethers' to keep the trapped beads far enough apart to avoid interference. As CsgA contains no native cysteines, thiol chemistry could be used to attach DNA handles, however, this would preclude the use of the internal cysteine amino acid change (*14*) that can be use to pin the protein in a metastable, disordered or non-fibril forming state. Therefore, we made use of ybbR attachment chemistry (*37*) with the commercial kit provided by Lumicks (Netherlands). The kit has ssDNA CoA-oligos of 36 nt, with complementary-overhang handles of 1506 bp on both flanks of the protein. The oligo labelled CsgA is annealed with the complementary biotin and digoxigenin handles at 37°C for 15 minutes, then the construct is diluted ~100 fold before being incubated with 2.08 µm anti-dig beads (Spherotech). Lumicks reports their handles can reach 40-45 pN before melting of the overhang/oligo, which has a GC content of 58%. In practice forces near 40 pN were avoided as this reduces tether lifetime.

SMFS – dumbbell fishing: with 1.14 µm streptavidin beads (Spherotech) was performed within a five-channel microfluidic flowcell of the Lumicks C-trap (dual beam/trap optical tweezers). Bead position within the traps was recorded using back focal plane detection with position sensitive detectors and sampled at 78.125 kHz with an NI-4472 (national instruments) acquisition card with 64× anti-aliasing (5 MS/s per channel). Data was imported into Igor Pro 9 (Wavemetrics), where it was downsampled to millisecond time resolution with filtering for anti-aliasing. Force extension curves (FECs) were extracted then with two worm-like chains in series as described previously (*24*). dsDNA fitting parameter were a persistence length (L_p) of ~40 nm and elastic modulus (K) of ~1000 pN, while unfolded polypeptide lengths used an L_p of 0.85 nm and a K of 2000 pN.

$$F(x) = \frac{k_B T}{L_p}\left[\frac{1}{4}\left(1 - \frac{x}{L_c} + \frac{F}{K}\right)^{-2} - \frac{1}{4} + \frac{x}{L_c} - \frac{F}{K}\right]$$

The disulfide bridge of the CsgA (A63C/V140C) was broken using ~100 mM MEA (2-MercaptoEthylAmine) in a fourth channel of the microfluidic cell. However, the compliance of the handles changes due to partial reduction of disulfides of the IgG, depressing the modulus to ~700 pN when fitting the handles with the WLC. After establishing the break of the internal disulfide bridge, the dumbbell was moved back into a running buffer of PBS with an oxygen scavenging system as previously described (*24*) with 50 U/mL glucose oxidase (GOx), 140 U/mL catalase and 0.01% D-glucose, and 10 –20 mM of MEA to keep the protein's cysteines reduced. MEA levels need to be moderated as GOx (*A. niger*) has a disulfide bond between residues 164 and 206 required for enzyme function (*38*), though our catalase (bovine) is not known to require disulfide bridges for enzymatic function.

Analysis*:* Force extension curves (FECs) were fit using the worm like chain (WLC) model from Marko and Siggia encoded in Igor Pro (Wavemetrics). The Lumicks OT instrument (C-trap) uses two traps by polarisation splitting and then back focal plane detection with position sensitive devices (PSDs). When the two trapped beads are close together deflected beams produce secondary

spots on the detectors. This crosstalk is very similar to a Bessel function of the first kind seen with an Airy disc. Given the trapping wavelength of 1064 nm (λ), this crosstalk is quite strong within one λ of the bead contact surfaces. The curvature of the FECs is very sensitive to this correction (subtraction of a large polynomial fit), but when done carefully, we maintain a persistence length (Lp) of ~40 nm and a modulus (K) of ~1000 pN. Without detection scheme modification this crosstalk can be further addressed with smaller beads (~1 μm pair) (*39*) and longer dsDNA handles.

Table 1 CsgA construct sequence.

ybbR and linker sequence	MDSLEFIASKLAGGGGSGGGGS
N22 (21-42)	GVVPQYGGGGNHGGGGNNSGPN
R1 (43-65)	SELNIYQYGGGNSALALQTD**C**RN
R2 (66-87)	SDLTITQHGGGNGADVGQGSDD
R3 (88-110)	SSIDLTQRGFGNSATLDQWNGKN
R4 (111-132)	SEMTVKQFGGGNGAAVDQTASN
R5 (133-151)	SSVNVTQ**C**GFGNNATAHQY
Linker, ybbR and His tag sequence	GGGGSGGGGSDSLEFIASKLAGGGGSGGGGSHHHHHH

Acknowledgements

We have many colleagues to thank who supported this work in some form. In particular, Prof. Matthew Chapman shared details about his research group's panel of engineered mutants for CsgA and their *E. coli* cell strain for expression. Adaptation and additional development of the expression protocol was done by Fan Bu, who also continues to provide insights into CsgA amyloid formation. Our summer undergrads Hrishika Dekate and Rachel Rosenberg worked on protein expression used for experimental assays such as ThT and TEM results.

Prof. Johannes Stigler helped us get started with his import code for Igor. Additionally, Noel Hoffer (Lumicks) provided helpful discussions about instrument operation and bead optimisation, while Krishna P. Neupane offered helpful discussions in regards to handle length and trap crosstalk.

References

1. N. Jain, M. R. Chapman, Bacterial functional amyloids: Order from disorder. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1867**, 954–960 (2019).

2. N. Van Gerven, R. Klein, S. Hultgren, H. Remaut, Bacterial amyloid formation: structural insights into curli biogensis. *Trends Microbiol* **23**, 693–706 (2015).

3. D. Romero, C. Aguilar, R. Losick, R. Kolter, Amyloid fibers provide structural integrity to Bacillus subtilis biofilms. *Proceedings of the National Academy of Sciences* **107**, 2230–2234 (2010).

4. J. Li, F. Zhang, Amyloids as Building Blocks for Macroscopic Functional Materials: Designs, Applications and Challenges. *International Journal of Molecular Sciences* **22**, 10698 (2021).

5. M. R. Chapman, L. S. Robinson, J. S. Pinkner, R. Roth, J. Heuser, M. Hammar, S. Normark, S. J. Hultgren, Role of Escherichia coli Curli Operons in Directing Amyloid Fiber Formation. *Science* **295**, 851–855 (2002).

6. S. K. Collinson, J. M. R. Parker, R. S. Hodges, W. W. Kay, Structural predictions of AgfA, the insoluble fimbrial subunit of Salmonella thin aggregative fimbriae11Edited by W. Baumeister. *Journal of Molecular Biology* **290**, 741–756 (1999).

7. S. K. Collinson, L. Emödy, K. H. Müller, T. J. Trust, W. W. Kay, Purification and characterization of thin, aggregative fimbriae from Salmonella enteritidis. *Journal of Bacteriology* **173**, 4773–4781 (1991).

8. F. Shewmaker, R. P. McGlinchey, K. R. Thurber, P. McPhie, F. Dyda, R. Tycko, R. B. Wickner, The Functional Curli Amyloid Is Not Based on In-register Parallel β-Sheet Structure *. *Journal of Biological Chemistry* **284**, 25065–25076 (2009).

9. M. S. Dueholm, D. Otzen, P. H. Nielsen, Evolutionary Insight into the Functional Amyloids of the Pseudomonads. *PLOS ONE* **8**, e76630 (2013).

10. A. Bleem, G. Christiansen, D. J. Madsen, H. Maric, K. Strømgaard, J. D. Bryers, V. Daggett, R. L. Meyer, D. E. Otzen, Protein Engineering Reveals Mechanisms of Functional Amyloid Formation in Pseudomonas aeruginosa Biofilms. *Journal of Molecular Biology* **430**, 3751–3763 (2018).

11. K. H. Hansen, C. H. Byeon, Q. Liu, T. Drace, T. Boesen, J. F. Conway, M. Andreasen, Ü. Akbey, Structure of biofilm-forming functional amyloid PSMα1 from Staphylococcus aureus. *Proceedings of the National Academy of Sciences* **121**, e2406775121 (2024).

12. N. D. Hammer, J. C. Schmidt, M. R. Chapman, The curli nucleator protein, CsgB, contains an amyloidogenic domain that directs CsgA polymerization. *Proceedings of the National Academy of Sciences* **104**, 12494–12499 (2007).

13. J. D. Taylor, W. J. Hawthorne, J. Lo, A. Dear, N. Jain, G. Meisl, M. Andreasen, C. Fletcher, M. Koch, N. Darvill, N. Scull, A. Escalera-Maurer, L. Sefer, R. Wenman, S. Lambert, J. Jean, Y. Xu, B. Turner, S. G. Kazarian, M. R. Chapman, D. Bubeck, A. de Simone, T. P. J. Knowles, S. J. Matthews, Electrostatically-guided inhibition of Curli amyloid nucleation by the CsgC-like family of chaperones. *Sci Rep* **6**, 24656 (2016).

14. A. Balistreri, D. Kolli, S. W. Jayaweera, D. Lundahl, Y. Han, L. Kalcec, E. Goetzler, R. Alessio, B. Ruotolo, A. Olofsson, M. R. Chapman, The bacterial chaperone CsgC inhibits functional amyloid CsgA formation by promoting the intrinsically disordered pre-nuclear state. *Journal of Biological Chemistry* **301** (2025).

15. L. Sewell, F. Stylianou, Y. Xu, J. Taylor, L. Sefer, S. Matthews, NMR insights into the pre-amyloid ensemble and secretion targeting of the curli subunit CsgA. *Sci Rep* **10**, 7896 (2020).

16. M. Andreasen, G. Meisl, J. D. Taylor, T. C. T. Michaels, A. Levin, D. E. Otzen, M. R. Chapman, C. M. Dobson, S. J. Matthews, T. P. J. Knowles, Physical Determinants of Amyloid Assembly in Biofilm Formation. *mBio* **10**, e02279-18 (2019).

17. W. Hawthorne, S. Rouse, L. Sewell, S. J. Matthews, Structural insights into functional amyloid inhibition in Gram −ve bacteria. *Biochem Soc Trans* **44**, 1643–1649 (2016).

18. M. Sleutel, I. Van den Broeck, N. Van Gerven, C. Feuillie, W. Jonckheere, C. Valotteau, Y. F. Dufrêne, H. Remaut, Nucleation and growth of a bacterial functional amyloid at single-fiber resolution. *Nat Chem Biol* **13**, 902–908 (2017).

19. M. Sleutel, B. Pradhan, A. N. Volkov, H. Remaut, Structural analysis and architectural principles of the bacterial amyloid curli. *Nat Commun* **14**, 2822 (2023).

20. F. Bu, D. R. Dee, B. Liu, Structural insight into Escherichia coli CsgA amyloid fibril assembly. *mBio* **15**, e00419-24 (2024).

21. X. Wang, D. R. Smith, J. W. Jones, M. R. Chapman, In Vitro Polymerization of a Functional Escherichia coli Amyloid Protein *. *Journal of Biological Chemistry* **282**, 3713–3719 (2007).

22. F. Bu, D. R. Dee, B. Liu, Reply to Sleutel and Remaut, “Structural insights into Escherichia coli CsgA amyloid fibril assembly revisited.” *mBio* **16**, e0110625 (2025).

23. K. Zhou, Y. Liu, K. Li, Y. Li, S. Wei, Y. Zhou, H. Yu, Y. Zhang, J. Lv, B. Xiao, C. Zhong, Force Spectra of a Single CsgA Molecule and Amyloid Nanofibers Assembled from Chimeric Mfp5 and CBD Proteins: Implications for a Nanomaterial Testing Machine. *ACS Appl. Nano Mater.* **5**, 1758–1766 (2022).

24. H. Yu, D. R. Dee, X. Liu, A. M. Brigley, I. Sosova, M. T. Woodside, Protein misfolding occurs by slow diffusion across multiple barriers in a rough energy landscape. *Proceedings of the National Academy of Sciences* **112**, 8308–8313 (2015).

25. A. Solanki, K. Neupane, M. T. Woodside, Single-Molecule Force Spectroscopy of Rapidly Fluctuating, Marginally Stable Structures in the Intrinsically Disordered Protein $\ensuremath{\alpha}$-Synuclein. *Phys. Rev. Lett.* **112**, 158103 (2014).

26. K. Neupane, A. Solanki, I. Sosova, M. Belov, M. T. Woodside, Diverse Metastable Structures Formed by Small Oligomers of α-Synuclein Probed by Force Spectroscopy. *PLOS ONE* **9**, e86495 (2014).

27. A. Balistreri, E. Kahana, S. Janakiraman, M. R. Chapman, Tuning Functional Amyloid Formation Through Disulfide Engineering. *Frontiers in Microbiology* **11** (2020).

28. E. P. DeBenedictis, D. Ma, S. Keten, Structural predictions for curli amyloid fibril subunits CsgA and CsgB. *RSC Adv.* **7**, 48102–48112 (2017).

29. D. E. Kim, D. Chivian, D. Baker, Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res* **32**, W526–W531 (2004).

30. C. Bustamante, L. Alexander, K. Maciuba, C. M. Kaiser, Single-Molecule Studies of Protein Folding with Optical Tweezers. *Annual Review of Biochemistry* **89**, 443–470 (2020).

31. A. Narayan, M. T. Woodside, Probing structural dynamics and interactions in macromolecular complexes with single-molecule force spectroscopy. *Current Opinion in Structural Biology* **93**, 103110 (2025).

32. M. D. Wang, H. Yin, R. Landick, J. Gelles, S. M. Block, Stretching DNA with optical tweezers. *Biophysical Journal* **72**, 1335–1346 (1997).

33. E. P. DeBenedictis, S. Keten, Mechanical unfolding of alpha- and beta-helical protein motifs. *Soft Matter* **15**, 1243–1252 (2019).

34. X. Wang, M. R. Chapman, Sequence determinants of bacterial amyloid formation. *J Mol Biol* **380**, 570–580 (2008).

35. X. Wang, Y. Zhou, J.-J. Ren, N. D. Hammer, M. R. Chapman, Gatekeeper residues in the major curlin subunit modulate bacterial amyloid fiber biogenesis. *Proceedings of the National Academy of Sciences* **107**, 163–168 (2010).

36. A. Balistreri, E. Goetzler, M. Chapman, Functional Amyloids Are the Rule Rather Than the Exception in Cellular Biology. *Microorganisms* **8**, 1951 (2020).

37. M. J. Avellaneda, E. J. Koers, D. P. Minde, V. Sunderlikova, S. J. Tans, Simultaneous sensing and imaging of individual biomolecular complexes enabled by modular DNA–protein coupling. *Commun Chem* **3**, 1–7 (2020).

38. S. Ittisoponpisan, I. Jeerapan, In Silico Analysis of Glucose Oxidase from Aspergillus niger: Potential Cysteine Mutation Sites for Enhancing Protein Stability. *Bioengineering (Basel)* **8**, 188 (2021).

39. M. Freitag, S. Jaklin, F. Padovani, E. Radzichevici, S. Zernia, K. M. Schmoller, J. Stigler, Single-molecule experiments reveal the elbow as an essential folding guide in SMC coiled-coil arms. *Biophys J* **121**, 4702–4713 (2022).