

# Amyloid-like Fibrils from a Domain-swapping Protein Feature a Parallel, in-Register Conformation without Native-like Interactions<sup>\*[S]</sup>

Received for publication, May 18, 2011, and in revised form, June 21, 2011. Published, JBC Papers in Press, June 28, 2011, DOI 10.1074/jbc.M111.261750

Jun Li, Cody L. Hoop, Ravindra Kodali, V. N. Sivanandam<sup>1</sup>, and Patrick C. A. van der Wel<sup>2</sup>

From the Department of Structural Biology, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania 15260

The formation of amyloid-like fibrils is characteristic of various diseases, but the underlying mechanism and the factors that determine whether, when, and how proteins form amyloid, remain uncertain. Certain mechanisms have been proposed based on the three-dimensional or runaway domain swapping, inspired by the fact that some proteins show an apparent correlation between the ability to form domain-swapped dimers and a tendency to form fibrillar aggregates. Intramolecular  $\beta$ -sheet contacts present in the monomeric state could constitute intermolecular  $\beta$ -sheets in the dimeric and fibrillar states. One example is an amyloid-forming mutant of the immunoglobulin binding domain B1 of streptococcal protein G, which in its native conformation consists of a four-stranded  $\beta$ -sheet and one  $\alpha$ -helix. Under native conditions this mutant adopts a domain-swapped dimer, and it also forms amyloid-like fibrils, seemingly in correlation to its domain-swapping ability. We employ magic angle spinning solid-state NMR and other methods to examine key structural features of these fibrils. Our results reveal a highly rigid fibril structure that lacks mobile domains and indicate a parallel in-register  $\beta$ -sheet structure and a general loss of native conformation within the mature fibrils. This observation contrasts with predictions that native structure, and in particular intermolecular  $\beta$ -strand interactions seen in the dimeric state, may be preserved in “domain-swapping” fibrils. We discuss these observations in light of recent work on related amyloid-forming proteins that have been argued to follow similar mechanisms and how this may have implications for the role of domain-swapping propensities for amyloid formation.

Amyloid fibril formation is characteristic of a variety of human disorders, including Huntington and Alzheimer diseases (1, 2). In amyloid-related diseases, one or more proteins are found in fibrillar aggregates, in a non-native, highly  $\beta$ -sheet-rich conformation. Depending on the disease context, the propensity for amyloid formation may be traced to mutations and cleavage events, combined with poorly understood external triggers. Many proteins can also be made to form amyloid fibrils

*in vitro*. Intriguingly, there are many accounts of proteins that form amyloid-like fibrils that are not associated with pathologies (2). Whether disease-related or not, amyloid fibrils share key biochemical and biophysical characteristics, suggesting common structural features. Understanding the fibril formation pathway is of interest not only as there may be a correlation between disease onset and protein aggregation, but also because transient oligomeric precursors may act as toxic species. However, a lack of high resolution structures of most fibrils and their precursors limits our knowledge of the mechanism of formation.

One seemingly common structural motif for amyloid fibrils is an in-register parallel (IP)<sup>3</sup> assembly into pleated  $\beta$ -sheets, stabilized by backbone-to-backbone hydrogen bonding, combined with favorable side-chain interactions (*e.g.* Asn or Gln ladders or hydrophobic clustering) (3–5). Such “serpentine” models have been invoked for various proteins and explain some intriguing features of amyloid formation, such as insensitivity to sequence scrambling (6). One consequence of serpentine assemblies is the loss of the hydrogen bonding interactions characteristic of the native tertiary structure.

In seeming contrast, there are cases where a persistence of the native structure has been reported, even upon amyloid formation. Such observations may not be contradictory, if the native conformations occur in non-amyloidogenic domains that remain attached to the amyloid itself (7, 8). However, native interactions could well be preserved within the fibril core for proteins that natively contain  $\beta$ -sheets, as proposed for  $\beta$ 2m and transthyretin, for instance (9, 10). Similar ideas form the basis for proposed mechanisms of three-dimensional domain swapping or runaway domain swapping (11, 12). The basic principle here is that structural motifs that are present in domain-swapped dimers can also accommodate more extensive oligomerization leading to, *e.g.*, amyloid formation. These models have gained support from the fact that various amyloidogenic proteins are able to form domain-swapped dimers. The amyloid-forming protein cystatin C, which is the cause of a cerebral amyloidosis, is a key example. Certain mutants form amyloid fibrils and also undergo domain swapping, and inhibition of domain swapping also affects fibril formation (13–17). Domain swapping has also been proposed as an oligomeriza-

<sup>\*</sup> This work was supported by the National Center for Research Resources (Grant UL1 RR024153) and startup funds from the University of Pittsburgh.

[S] The on-line version of this article (available at <http://www.jbc.org>) contains supplemental Figs. S1–S3 and Tables S1 and S2.

<sup>1</sup> Present address: Dept. of Chemistry & Biochemistry, 251 Nieuwland Science Hall, University of Notre Dame du Lac, Notre Dame, IN 46556.

<sup>2</sup> To whom correspondence should be addressed: 3501 Fifth Ave., BST3 2044, Pittsburgh, PA 15260. Tel.: 412-383-9896; Fax: 412-648-9008; E-mail: pdvwel@pitt.edu.

<sup>3</sup> The abbreviations used are: IP, in-register, parallel; MAS, magic angle spinning, ssNMR solid-state NMR; GB1, immunoglobulin binding domain B1 of streptococcal protein G; ds, domain-swapping;  $\beta$ 2m,  $\beta$ 2-microglobulin; PARIS, phase-alternated recoupling irradiation scheme; ThT, thioflavin T; H/D, hydrogen-deuterium; T, tesla(s).

tion mechanism for transthyretin (18),  $\beta 2m$  (19–21), and the prion protein (22, 23), although arguably with less direct evidence to support these cases. More data are available on non-disease-related proteins where domain swapping and amyloid formation seem to correlate (24–29). In one case a domain-swapped oligomeric species even shows toxicity, paralleling amyloid-related oligomers (29).

In such cases,  $\beta$ -sheet elements of the native protein structure are preserved in the domain-swapped dimer: native strand-strand hydrogen bonding that is *intramolecular* in the monomer becomes *intermolecular* in the dimer. Partial unfolding and structural rearrangements are needed to attain a dimeric structure, and it is proposed that similar re-arrangements can open up pathways to more extensive oligomerization. Hinge loop regions have different conformations in monomeric and domain-swapped structures and may facilitate continued oligomerization (20). Although there are various reports of apparent correlations between domain swapping and amyloid formation, there is a lack of structural data that unequivocally demonstrate the preservation of native structure within the mature fibril, a key feature of its most basic principles.

Here, we use MAS ssNMR to examine fibrils formed by one of the proteins that displays such a correlated tendency for domain swapping and amyloid formation, a mutant of the GB1 protein. GB1 itself has served as a model for protein folding and stability studies, and a number of mutants form amyloid fibrils (31–34). One of these is a domain-swapping mutant (dsGB1)<sup>4</sup> with mutations that destabilize the hydrophobic core and the native  $\alpha$ -helix (26, 27). Relative to the highly stable and well studied T2Q version (35) of GB1 (henceforth referred to simply as GB1), the mutations are L5V, F30V, Y33F, and A34F (26). These mutations destabilize the monomeric state and instead favor a domain-swapped dimer (Fig. 1), which is less stable than monomeric T2Q-GB1. In analogy to studies on other domain-swapping amyloid-formers, a combination of solution NMR, mutational studies, and disulfide bond formation suggests a correlation between amyloid formation and domain-swapping ability (26). Accordingly, it was suggested that native-like contacts may be conserved in the fibrillar state, as predicted by certain runaway domain-swapping models.

MAS ssNMR provides direct molecular insight into the structure of amyloid fibrils (36), allowing us to obtain key structural information on the mature dsGB1 fibrils. Combined with hydrogen-deuterium exchange experiments, it reveals a highly rigid and largely  $\beta$ -sheet conformation. The ssNMR also indicates intermolecular same-to-same interactions consistent with a parallel  $\beta$ -sheet geometry, as commonly seen in other amyloid-forming proteins. Our data indicate a loss of key interactions that characterize the monomeric and domain-swapped dimer conformations, especially loss of both the native helical segment and specific strand-strand interactions of the native  $\beta$ -sheet assembly. We conclude with a discussion of implications for domain-swapping mechanisms invoked for amyloid formation for this and other amyloidogenic proteins.

<sup>4</sup> Previously also referred to as the HS#124<sup>F26A</sup> mutant.

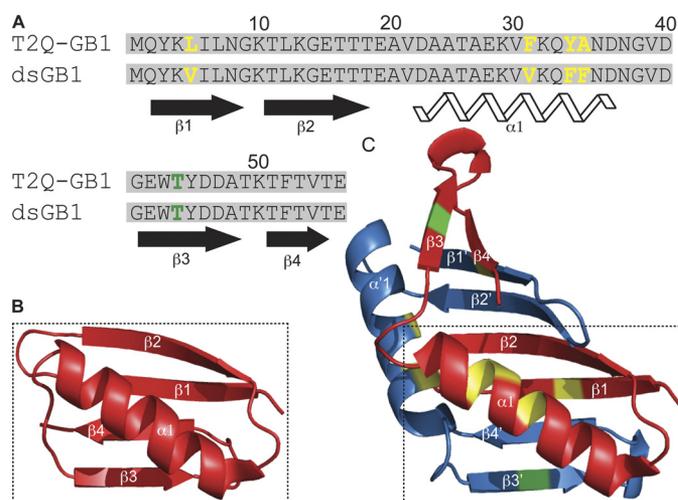


FIGURE 1. A, sequence alignment between T2Q-GB1 and dsGB1. B and C, schematic illustration of dsGB1 native and domain-swapped structures. The dashed box highlights the preservation of a monomer-like fold in domain-swapped dimeric dsGB1. Residues highlighted in yellow indicate mutation sites from T2Q-GB1. Upon fibril formation, mutation of T44 (green) to Cys permits formation of an intermolecular disulfide bond. This indicates an intermonomer proximity of residue 44 in the fibrils as also found in the domain-swapped dimer.

## EXPERIMENTAL PROCEDURES

**Protein Expression and Purification**—The dsGB1 protein was expressed, isotopically labeled, and purified using standard methods (26). Briefly, plasmid pET11A coding for dsGB1 was a gift from Angela Gronenborn (University of Pittsburgh). Uniformly and selectively <sup>13</sup>C- and/or <sup>15</sup>N-labeled protein was obtained by expression in *Escherichia coli* HMS174 (DE3), using M9 minimal medium and isotopically labeled materials from Cambridge Isotope Laboratories (Andover, MA) and Isotec (Sigma-Aldrich). <sup>13</sup>CO Tyr-labeled sample was prepared as previously described (37).

**Fibril Sample Preparation**—Protein in 50 mM sodium phosphate buffer (pH 5.5) was concentrated to ~0.7 mM, then centrifuged at 100,000 × g for 1 h (Optima ultracentrifuge, Beckman Coulter) to eliminate pre-existing aggregates, giving a final concentration of 0.5–0.6 mM. Fibrils are formed over several days, under agitation at 58 °C and 600 rpm (26) and then isolated by centrifugation. For MAS ssNMR experiments, 8–9 mg of pelleted fibrillized protein was packed into a 3.2-mm MAS rotor (Bruker BioSpin, Billerica, MA) using a customized sample packing tool under 130,000 × g centrifugation.

**Electron Microscopy**—An aliquot of 5  $\mu$ l from the aggregation mixture was placed on freshly glow-discharged, carbon-coated, 400-mesh-size copper grids and adsorbed for 2 min. Grids were washed with deionized water and then stained with freshly filtered 1% (w/v) uranyl acetate for 2–15 s. Grids were imaged in the Structural Biology Department's EM facility on a Tecnai T12 microscope (FEI) operating at 120 kV and 30,000× magnification and equipped with an UltraScan 1000 charge-coupled device camera (Gatan) with post-column magnification of 1.4×.

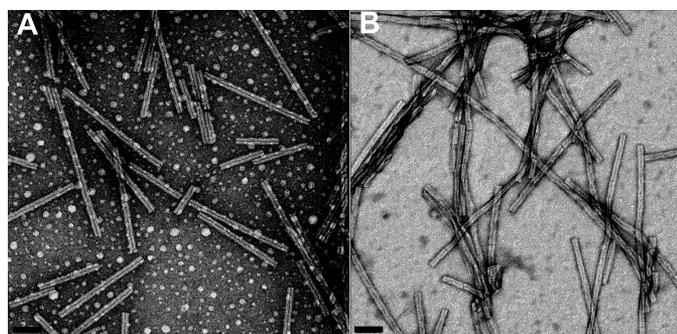
**H/D Exchange Measurements**—Sequestration of amide protons within the fibrils was examined via hydrogen-deuterium (H/D) exchange experiments monitored by solution NMR (9, 38). Amyloid fibrils of <sup>15</sup>N-labeled dsGB1 were centrifuged at

## Fibril Structure of a Domain-swapping Protein

5000  $\times$  *g* for 10 min, washed with 50 mM sodium phosphate buffer (pH 5.5), and lyophilized. 0.8 mg of the fibril was weighed and resuspended in deuterated buffer (50 mM sodium phosphate and D<sub>2</sub>O, pD 5.5). At the indicated time points, aliquots were centrifuged and frozen in liquid nitrogen to quench exchange. The frozen samples were dissolved in perdeuterated dimethyl sulfoxide (*d*<sub>6</sub>-DMSO) containing 5% D<sub>2</sub>O and 0.01% *d*<sub>2</sub>-dichloroacetic acid. After dissolution, the solution was immediately transferred into an NMR tube, and the heteronuclear single quantum coherence spectrum was recorded at 25°C. ssNMR experiments were recorded on 18.8-T (800-MHz <sup>1</sup>H frequency), 16.4-T (700 MHz), and 14.1-T (600 MHz) spectrometers (Bruker BioSpin) equipped with cryoprobes. More details can be found in the [supplemental materials](#).

**ssNMR Spectroscopy**—MAS ssNMR experiments were conducted using 3.2-mm HCN Efree MAS probes (Bruker BioSpin) at static magnetic fields of 14.1 and 18.8 T (<sup>1</sup>H frequencies of 600 and 800 MHz). Spinning rates were typically 13 kHz (at 14.1 T) or 17 kHz (at 18.8 T) unless stated otherwise, with sample cooling at 267 K. Standard two-dimensional and three-dimensional MAS ssNMR assignment experiments were employed to

assign the resonances, mostly via NCOX and NCACX experiments, along with two-dimensional <sup>13</sup>C-<sup>13</sup>C spectra obtained with dipolar-assisted rotational resonance (39) and phase-alternated recoupling irradiation scheme (PARIS) (40) <sup>13</sup>C-<sup>13</sup>C mixing. A series of two-dimensional <sup>13</sup>C-<sup>13</sup>C NMR spectra with 23-, 54-, and 100-ms PARIS mixing were used to probe longer distance <sup>13</sup>C-<sup>13</sup>C contacts (at 17-kHz MAS and 800-MHz <sup>1</sup>H frequencies). To probe the inter-strand register, PITHIRDS-CT measurements (41, 42) were carried out using fibrils prepared from Tyr-1-<sup>13</sup>CO-labeled dsGB1 at 600 MHz (<sup>1</sup>H frequency) and an MAS frequency of 20 kHz, with up to 38.4 ms of total dipolar recoupling. Two fibril samples were prepared: one entirely from Tyr-1-<sup>13</sup>CO labeled protein and the other from the same labeled protein diluted 1:2 with unlabeled protein. PITHIRDS-CT reference dephasing curves were calculated with the SPINEVOLUTION program (43), assuming a linear chain of five equally spaced <sup>13</sup>C labels consistent with an IP  $\beta$ -sheet. More details can be found in the [supplemental materials](#).

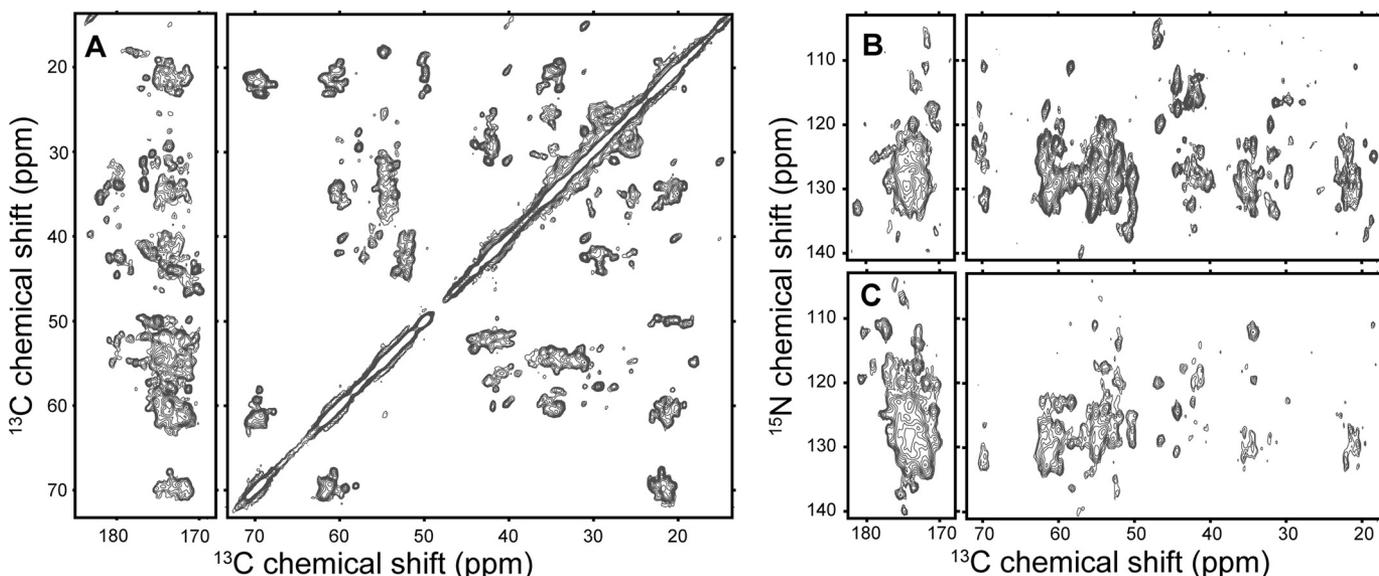


**FIGURE 2. Negative stain transmission electron microscopy images of dsGB1 during the aggregation process, showing circular assemblies as well as fibrils (A), and of mature dsGB1 fibrils having a homogeneous and periodically twisted morphology (B).** Scale bars indicate 50 nm.

## RESULTS

**Fibril Preparation and Characterization**—In line with previous reports (26), we observed fibril formation by dsGB1 under agitation at 58 °C. After an initial lag phase, an increase in thioflavin T binding was observed that completed within several days. During early stages, alongside fibrils, circular assemblies could be seen by EM (Fig. 2A) that appear similar to oligomeric species previously reported for other amyloids, such as A $\beta$  (44). Over time, these transient species disappeared yielding the mature fibrils (Fig. 2B).

**MAS ssNMR Chemical Shifts and Secondary Structure**—MAS ssNMR was used to study mature fibrils formed from uniformly and selectively labeled dsGB1. Assignments were obtained for most of the residues, with some exceptions due to signal overlap (see [supplemental Table S2](#) and Fig. 3). The <sup>13</sup>C line width in homonuclear spectra was typically  $\sim$ 1 ppm, indi-



**FIGURE 3. MAS ssNMR spectra on <sup>13</sup>C,<sup>15</sup>N-labeled dsGB1 fibrils.** A, <sup>13</sup>C-<sup>13</sup>C two-dimensional PARIS spectrum with 23-ms mixing time; B, NCACX; and C, NCOX two-dimensional spectra with 5-ms <sup>13</sup>C-<sup>13</sup>C dipolar-assisted rotational resonance mixing. These data were acquired at 800 MHz (<sup>1</sup>H frequency), 17 kHz MAS rate, and 267 K.

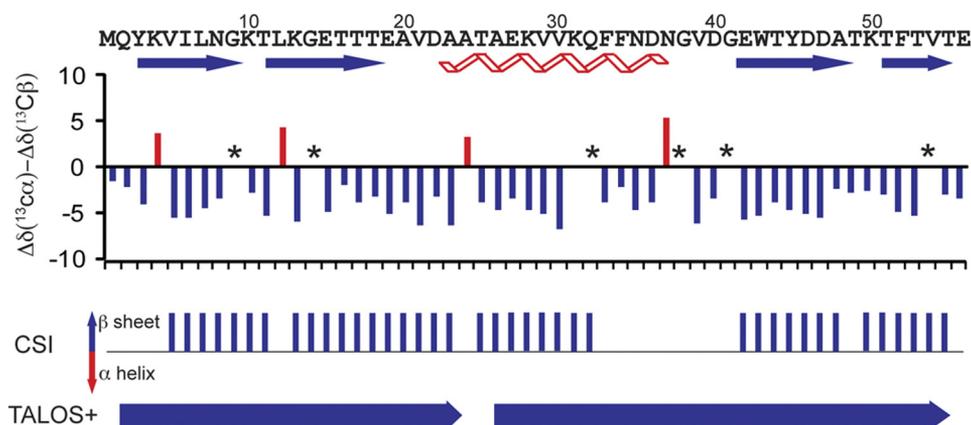


FIGURE 4. **Secondary chemical shift analysis of the fibrils.** The primary sequence and native conformation are indicated at the top. The next row shows  $\Delta\delta(^{13}\text{C}\alpha) - \Delta\delta(^{13}\text{C}\beta)$ , the difference between experimental  $^{13}\text{C}\alpha$ - $^{13}\text{C}\beta$  ssNMR chemical shifts and corresponding "random-coil" values. Residues with missing assignments are indicated with asterisks. Significantly positive values indicate  $\alpha$ -helical structure (red), and negative values indicate  $\beta$ -strands (blue). Chemical shift indexing (CSI) or the TALOS+ program also indicate high  $\beta$ -sheet content (45, 46). Although TALOS+ appears to suggest two long  $\beta$ -strands, it is possible that there are additional loop regions not identified by this algorithm.

cating a relatively homogeneous and well ordered structure. A number of sites were found to display two sets of peaks, in particular residues Asn<sup>8</sup> and Gly<sup>9</sup>, the segment Asp<sup>22</sup>-Ala<sup>26</sup>, Gly<sup>38</sup> and Val<sup>39</sup>, and Ala<sup>48</sup> (e.g. supplemental Fig. S2). Such peak doubling is indicative of variations in the local environment or structure. In some cases it appeared due to partial deamidation of specific Asn side chains (also seen in solution), but elsewhere it may be explained by increased conformational freedom for residues in less restricted loop regions (e.g. residues Asp<sup>22</sup>-Ala<sup>26</sup>).

The assigned  $^{13}\text{C}\alpha$ ,  $^{13}\text{C}\beta$ ,  $^{13}\text{CO}$ , and  $^{15}\text{N}$  shifts were used for CSI (45) and TALOS+ analysis (46), which indicate a predominance of  $\beta$ -sheet structure (Fig. 4). The lack of contiguous residues with helix-typical chemical shifts indicates that the native  $\alpha$ -helix is lost in the fibrils and instead converted into  $\beta$ -sheet structure. These observations also hold true for the secondary conformations detected within the fibrils. More generally, comparison of the fibril chemical shifts to those of the native structure, as reported for dsGB1 by solution NMR (27) or microcrystalline GB1 (47), indicates a large difference in most of the shifts (see supplemental Fig. S1). These differences exceed those expected for experimental differences between solution and ssNMR, for instance due to intermolecular interactions (48). Rather, the chemical shift differences indicate substantial structural differences that are not limited to the hinge loops or swapped domains.

**Rigidity and Solvent Exposure**—Exposure of the fibrils to  $\text{D}_2\text{O}$  allows exchange of mobile sites, which can reveal the non-exchanging fibril core, as it is expected to be highly stable under aqueous conditions. To detect exchangeability in a residue-specific fashion, we quenched the exchange, dissolved the fibrils in high concentrations of DMSO (where exchange is suppressed), and performed solution NMR to detect residual protons (9, 38, 49).  $^1\text{H}$ - $^{15}\text{N}$  heteronuclear single quantum coherence spectra of dsGB1 dissolved in 95%  $\text{DMSO}-d_6/5\%$   $\text{H}_2\text{O}$  at pH 5.5 and room temperature reveal backbone  $\text{H}_\text{N}$  chemical shifts consistent with a denatured state (supplemental Fig. S3). Assignments obtained in  $\text{U-}^{13}\text{C}$ ,  $^{15}\text{N}$ -labeled reference samples were employed to site-specifically identify the exchanged sites. Sim-

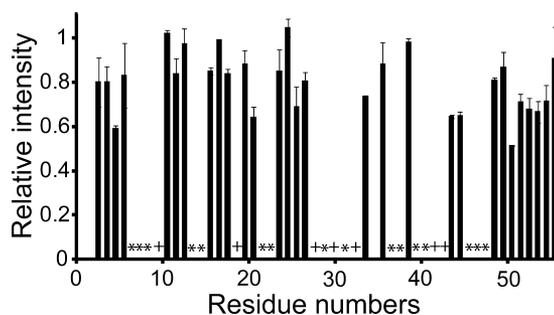


FIGURE 5. **Relative peak intensities in the 24-h H/D exchange spectrum** (see supplemental Fig. S3C) plotted against residue number. Peak intensities were normalized to the reference spectrum (supplemental Fig. S3B), but incompletely separated peaks (+) and peaks highly influenced by fast exchange in DMSO (\*) are omitted.

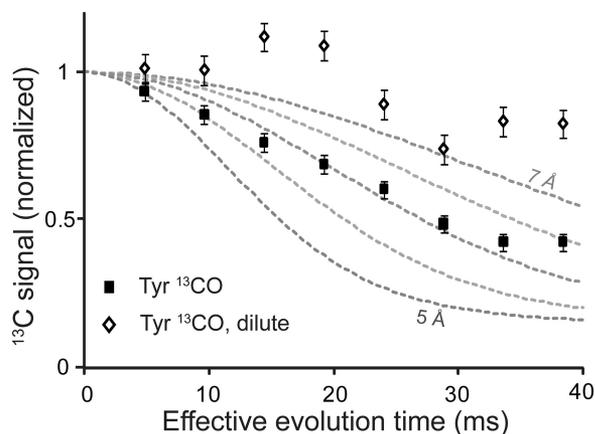
ilar to previous reports (9, 38), certain residues are found to undergo rapid H/D exchange in the DMSO solvent system, preventing a robust analysis of their exchange in the fibrils. These sites are mostly charged residues and their neighbors, consistent with previous observations and mechanistic discussions (38, 50). From the remaining sites, we generated the H/D exchange profile shown in Fig. 5, showing the exchange after 24 h. The fact that many peaks throughout the sequence had a residual intensity of  $>50\%$  implies that dsGB1 fibrils are rigid and highly protected. Even after exposure of fibrils to  $\text{D}_2\text{O}$  for 4 weeks many amides were still highly protonated (supplemental Fig. S3D). Remarkably, the N-terminal segment Tyr<sup>3</sup>-Val<sup>5</sup> and the very C-terminal segment Phe<sup>52</sup>-Thr<sup>55</sup> showed particularly high levels of protection. In general, the exchange of all observable sites was relatively low and no dynamic loops or termini are apparent.

**Tertiary Structure of the Fibrils**—A key feature of domain-swapped dimers is that some native hydrogen bonding patterns are preserved, but in an inter- rather than intramolecular fashion. We performed selected MAS ssNMR experiments to probe intermolecular interactions in the fibrils. The PITHIRDS-CT experiment is designed for the detection of intermolecular contacts in IP fibrils, prepared with site-specific carbonyl and/or methyl group  $^{13}\text{C}$ -labeling (41). Fibrils were prepared from dsGB1, which was  $^{13}\text{CO}$ -labeled in both tyrosines (Tyr<sup>3</sup> and Tyr<sup>45</sup>). After correction for natural abundance background sig-

## Fibril Structure of a Domain-swapping Protein

nals, the signal decay indicated a distance of close to  $\sim 6$  Å between neighboring labeled carbonyl sites (Fig. 6). These interactions were unequivocally intermolecular, as amyloid fibrils formed from co-fibrillization of labeled and unlabeled protein in a 1:2 ratio showed a dramatic decrease in the extent of dephasing (Fig. 6). Our data curves and fit results were highly similar to previous reports on other IP protein fibrils (5, 41, 51, 52).

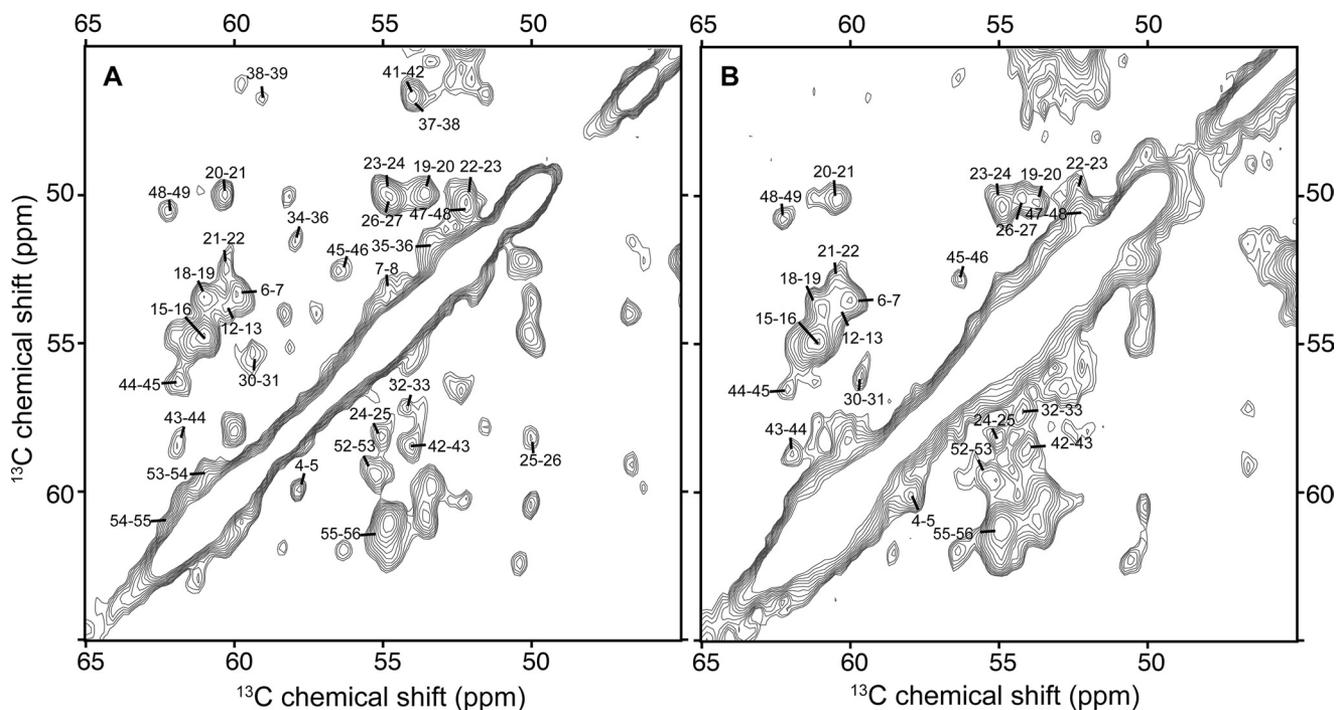
To explore the tertiary structure of the dsGB1 fibrils, we also examined a series of  $^{13}\text{C}$ - $^{13}\text{C}$  experiments applied to both fully labeled and isotopically diluted samples. These experiments



**FIGURE 6.**  $^{13}\text{C}$  PITHIRDS-CT data for Tyr- $^{13}\text{C}$ O-labeled dsGB1 fibrils, indicating a close intermolecular proximity of the Tyr carbonyls, consistent with an IP  $\beta$ -sheet structure and similar to previous studies on other amyloid fibrils. Fully labeled fibrils show significant dephasing, which is largely eliminated in fibrils containing a 1:2 mixture of labeled and unlabeled protein (dilute). Simulated dephasing curves are shown that correspond to interatomic distances from 5.0 to 7.0 Å in 0.5-Å increments. Measurements were done at 14.1 T, 20-kHz MAS, and 267 K.

employed the PARIS mixing scheme (40, 53), a  $^{13}\text{C}$ - $^{13}\text{C}$  transfer experiment that permits cross-peaks up to 4.8 Å to be easily observed, with distances up to 5.7 Å reported (53). Long distance intermolecular  $\text{C}\alpha$ - $\text{C}\alpha$  contacts can be indicative of the supramolecular alignment (54). Intramolecular sequential  $\text{C}\alpha_i$ - $\text{C}\alpha_{i+1}$  contacts are expected to be  $\sim 3.8$  Å. Peptide model structures (55) suggest that, in an IP  $\beta$ -sheet, the intermolecular  $\text{C}\alpha_i$ - $\text{C}\alpha_i$  contacts are  $\sim 4.9$  Å, whereas  $\text{C}\alpha_i$ - $\text{C}\alpha_{i+1}$  distances are  $>6$  Å. In anti-parallel  $\beta$ -sheets (55), the  $\text{C}\alpha$ - $\text{C}\alpha$  distances were  $\sim 4.7$ – $5.1$  Å, but in this case intermolecular contacts would not consistently involve the same or neighboring residues in the primary sequence, giving  $\text{C}\alpha_i$ - $\text{C}\alpha_j$  contacts with  $j \neq i (+1)$ . Thus, long mixing PARIS experiments should reveal primarily sequential  $\text{C}\alpha$ - $\text{C}\alpha$  contacts for an IP  $\beta$ -sheet, but, for an anti-parallel  $\beta$ -sheet, other  $\text{C}\alpha_i$ - $\text{C}\alpha_j$  contacts are expected (that differ from local contacts seen for shorter mixing times).

$^{13}\text{C}$ - $^{13}\text{C}$  spectra of dsGB1 fibrils with mixing times of 23 and 100 ms are shown in Fig. 3A and Fig. 7. 30 out of 55 sequential  $\text{C}\alpha$ - $\text{C}\alpha$  cross-peaks were identified at 100-ms mixing time, along with  $\sim 7$  peaks that are too close to the diagonal or overlapped with intra-residue contacts. Fibrils prepared from a 1:1 ratio of  $^{15}\text{N}$ -only- and  $^{13}\text{C}$ -only-labeled protein give very similar spectra with mostly the same cross-peaks, although they are limited by the reduced signal to noise and with some changes in the relative peak intensities. The lack of “additional” peaks in the fully labeled sample indicates that intermolecular contacts are analogous to the intramolecular  $\text{C}\alpha$ - $\text{C}\alpha$  contacts and is consistent only with a parallel  $\beta$ -sheet geometry in which each residue in the fibril is in-register and each monomer is in nearly the same local conformation (54).



**FIGURE 7.** Section of  $^{13}\text{C}$ - $^{13}\text{C}$  two-dimensional PARIS spectra with 100-ms mixing time, showing assignments of sequential  $\text{C}\alpha$ - $\text{C}\alpha$  cross-peaks for both uniformly labeled (A) and 1:1 isotopically diluted (B) dsGB1 fibrils. The observed  $\text{C}\alpha$ - $\text{C}\alpha$  peaks in the undiluted sample largely map onto those observed in the diluted sample (aside from intensity changes) and reflect local ( $i \rightarrow i + 1$ ) interactions. Such observations indicate an IP conformation (54).

## DISCUSSION

We have characterized key structural features of the fibrils formed by the domain-swapping model protein dsGB1, which has served as an example of how domain-swapping propensity may result in an ability to form amyloid-like fibrils. In models of three-dimensional or runaway domain swapping in amyloid formation, it is often proposed that the resulting fibrils would incorporate structural features that reflect strand-strand interactions that are present in the domain-swapped intermonomer interface. However, our results point to a loss of the native structure. Large differences in chemical shift are seen throughout the primary sequence. Secondary structure analysis by ssNMR indicates a loss of the native  $\alpha$ -helix. Rather, ssNMR and H/D exchange point to a highly rigid amyloid core that spans the entire protein. Furthermore, the ssNMR-based detection of intermolecular contacts shows that these fibrils adopt an IP conformation, similar to many other amyloid-like fibrils (3, 5). Such "serpentine" amyloid structures are often characterized by short  $\beta$ -strands with intervening loop regions. Several data suggest at least one loop region near residues 22–26, where we see local non- $\beta$  structure as well as a propensity for multiple conformations as reflected in doubled peak assignments. The presence or absence of other loops remains uncertain, because indications of their presence are ambiguous. Nonetheless, the possibility of more than two  $\beta$ -strand segments cannot be excluded. Remarkably, the loop regions and termini appear to be highly protected from H/D exchange, suggesting a very rigid and compact amyloid structure that lacks the flexible segments outside the amyloid core seen in many other amyloid fibrils.

Previous experiments on this particular dsGB1 mutant reported a correlation between the ability to undergo domain swapping into a dimeric form and the propensity to aggregate into amyloid-like fibrils (26). As the dimeric form was found to resemble the native monomeric conformation (Fig. 1) and various experimental data suggested similar interactions in the aggregated state, it was proposed that aspects of the native  $\beta$ -sheet structure persisted in the fibrils. These data combined solution NMR and the examination of disulfide formation by cysteine mutants, methods also used on other domain-swapping amyloidogenic proteins. T44C mutants (26) displayed intermolecular disulfide bond formation consistent with the anti-parallel interface seen in the domain-swapped dimer, supporting a structural model for the fibrils that incorporated native-like  $\beta$ -sheet interactions. Our data are consistent with these experimental data observed before, although we suggest an alternate explanation. In particular, the IP fibril assembly that we observed would define a stacked alignment of the Cys residues in neighboring strands, close enough to allow an intermolecular disulfide bond.

Serpentine fibril structures consisting of IP stacks of pleated  $\beta$ -strands appear to be increasingly common, as more detailed structural data is gained on proteins within amyloid-like fibrils (4). For instance, an earlier x-ray diffraction study on amyloid fibrils formed by different amyloid-forming mutants of GB1 indicated such a conformation for those GB1 fibrils (31). The authors described a twisted structure featuring IP-stacked

slabs, each containing four tightly packed  $\beta$ -strands that are part of a single protein. It remained unclear what the fate of the native helix was, how the native  $\beta$ -strands map onto the amyloid core strands, and to what degree there were unstructured loops or termini. Thus, the extent of correspondence between the amyloid conformations of these different mutants and the dsGB1 variant remains uncertain. Note that it is well known that polymorphic fibrils can be obtained from a single protein (56, 57), let alone from amyloidogenic proteins with different primary sequences.

*Mechanism of dsGB1 Aggregation*—The structural states of the monomer-dimer equilibrium of dsGB1 and related mutants have been well characterized by solution NMR (26, 27, 58, 59). Compared with the domain-swapped dimer, the monomer is substantially more flexible, in particular near the N terminus (59). The more stable dimer avoids structural clashes in the hydrophobic core that prevent a stable monomeric conformation. Instead, the monomeric form is more like a molten globular structure that maintains the  $\alpha$ -helix/C-terminal-hairpin core but lacks interactions that would stabilize the N-terminal  $\beta$ -hairpin (59). Thus, the spontaneous monomer-dimer refolding involves a preservation of certain secondary and tertiary structure interactions (*e.g.* conserving the highly stable  $\beta$ 3- $\beta$ 4 hairpin). Dimerization seems to be facilitated by antiparallel interactions of pre-existing  $\beta$ -hairpins, with very specific side-chain reorientations involved in the stabilization of the resulting structure (58).

In contrast, amyloid formation by dsGB1 requires both elevated temperatures and agitation, not unlike other amyloidogenic proteins. Agitation enhances amyloid formation through mechanisms that may include enhanced seeding, structural deformation, as well as modulation of local protein concentrations (60, 61). The need for elevated temperatures suggests that further destabilization of the structure is needed to initiate amyloid formation. Is this due to the need for sufficient disruption of the residual stable structure that is present in the monomers at lower temperatures? Amyloid formation may initiate with the coming together of these less well structured (although not necessarily fully unfolded) proteins, possibly mediated by hydrophobic interactions of exposed core residues. Note that the dsGB1 mutations (in particular the mutation A34F) affect the orientation and integrity of the  $\alpha$ -helix and its interactions with the  $\beta$ -strands and cause an extension of the neighboring loop region, thus opening up the hydrophobic core (58). It could then be that resulting oligomeric species permit extended structural elements (*e.g.* unstructured loop regions or isolated  $\beta$ -strands) to form a core of *parallel*  $\beta$ -sheets, which nucleate fibril growth. Once initiated, nucleation and seeded growth could further destabilize residual native secondary structure present in the oligomers to yield the completely non-native fibrils that we observed. Naturally, the detailed aggregation process, as well as the nature of on-pathway oligomers, remains ill-defined and speculative, reminiscent of the oligomeric precursors of many other amyloid-forming proteins.

*Domain Swapping in Other Amyloid-forming Proteins*—Clearly, the dsGB1 protein is just a single example where a correlation between domain-swapping and amyloid formation has been reported. So the question arises whether our observa-

## Fibril Structure of a Domain-swapping Protein

tion, that this protein's mature fibrils are not native-like, is an isolated case that does not hold true for other domain-swapping proteins. Here, it is worth noting that recent studies have suggested similar features in other proteins. For  $\beta 2m$  and prion proteins, domain-swapped dimers have been observed (19, 20), but for both proteins ssNMR and ESR studies on fibrils have indicated an IP sheet structure and a loss of the native conformation (51, 62–65). Thus, at least in a number of cases, similar observations have been made, when mature fibrils were studied directly.

**Role of Domain Swapping in Amyloid Formation**—Even if native structure may not be preserved in the fibrils, there remain seemingly compelling examples of a correlation between the ability to form domain-swapped dimers and the propensity to form amyloid. A somewhat trivial explanation would be that those proteins indeed form amyloid via a domain-swapping mechanism and that simply the “right” proteins have not been studied yet. Or, it may even be that a protein like dsGB1 forms polymorphic fibrils, some of which do indeed have a fibril structure that is native-like and was formed via domain swapping.

There may also be a true correlation between domain swapping and fibril formation that affects steps occurring along the fibril formation pathway without resulting in native-like structure in the fibrils. It is important to realize that our experiments probe the final mature fibrillar state, which will be predominantly assembled post-nucleation via seeded growth. Although it is generally thought that this templated mechanism would reflect the structure of the original seed, this may not be universally true. In other words, even if the earliest aggregates may form via domain swapping, the subsequent fibril extension may follow a distinctly different mechanism and result in the non-native structure that we observed. Similarly, it also is possible that domain swapping occurs prior to nucleation, resulting in the assembly of native-like but transient oligomeric aggregates. Note that various oligomeric assemblies of amyloidogenic proteins feature  $\alpha$ -helical or native-like structures (66, 67) that typically, but not always (68), are lost during fibril maturation.

Another manner by which the propensity for domain swapping may correlate to amyloidogenicity is in the earliest stages, prior to oligomer formation. It could simply be that domain swapping necessitates destabilization of the monomeric (or native) structure and that destabilization of the native conformation is also a prerequisite for amyloid formation. Proteins are dynamic in solution, and multiple conformations are populated at the same time, but the majority is usually in the native structure, as it has the lowest energy. Aggregation-prone conformations may be present, but only at a low population, resulting in a long lag time of aggregation. Exposure to denaturing conditions, but also mutations, can destabilize the native structure, increase the population of partially unfolded conformations, and thus encourage aggregation. For GB1 and its variants, the highly stable T<sub>2</sub>Q-GB1 variant (35) has not been reported to form fibril, but destabilization of the GB1 fold allows formation of amyloid-like fibrils under appropriate conditions, in correlation to the extent of destabilization (26, 32–34). Nonetheless, it is unclear that destabilization is the only factor, because this may suggest that all proteins would be prone and able to form

amyloid fibrils under appropriate conditions. There are examples, for instance among SH3 domains, where stability alone does not explain differences in amyloidogenicity (30). Similarly, a closely similar GB1 mutant with a melting temperature lower than dsGB1, but lacking its domain-swapping propensity, fails to form fibrils despite a reduced stability (26, 59). Rather, specific conformations may exist that are crucial for the protein to form amyloid fibrils, and destabilized proteins may not always have access to appropriate conformations. Domain swapping may thus fulfill one of the requirements for amyloid formation by destabilizing the native structure but *also* favor structural propensities that help generate on-pathway assemblies and structures that facilitate amyloid formation. Although this is thus far completely hypothetical, a possibly related observation is that loop regions appear to be important for domain swapping (11) but also for amyloid formation (*e.g.* in defining the fibril-forming ability of different SH3 domains) (30).

**Summary**—We have shown the loss of the native conformation upon amyloid formation by the domain-swapping model protein dsGB1, in contrast with certain models of domain swapping-induced amyloid formation. The protein undergoes a complete structural rearrangement to form the IP  $\beta$ -sheet, more reminiscent of the increasingly common “serpentine” amyloid conformations than structural models commonly proposed in the context of the runaway domain swapping-related amyloid formation. How applicable these observations are to other domain-swapping proteins remains to be seen, but they appear to echo a number of other studies and at the least raise questions about the precise role of domain swapping and the structural requirements for the conversion of native to amyloid-like conformations.

**Acknowledgments**—We thank Angela Gronenborn for helpful discussions and providing the plasmids. We also acknowledge Lin Liu, Elizabeth Landrum, and Tanxing Cui for their help and helpful discussions, and Mike Delk for technical assistance. We acknowledge Ron Wetzel for the use of laboratory equipment.

## REFERENCES

1. Chiti, F., and Dobson, C. M. (2006) *Annu. Rev. Biochem.* **75**, 333–366
2. Eisenberg, D., Nelson, R., Sawaya, M. R., Balbirnie, M., Sambashivan, S., Ivanova, M. I., Madsen, A. Ø., and Riek, C. (2006) *Acc. Chem. Res.* **39**, 568–575
3. Margittai, M., and Langen, R. (2008) *Q. Rev. Biophys.* **41**, 265–297
4. Kajava, A. V., Baxa, U., Wickner, R. B., and Steven, A. C. (2004) *Proc. Natl. Acad. Sci. U.S.A.* **101**, 7885–7890
5. Shewmaker, F., Ross, E. D., Tycko, R., and Wickner, R. B. (2008) *Biochemistry* **47**, 4000–4007
6. Ross, E. D., Baxa, U., and Wickner, R. B. (2004) *Mol. Cell Biol.* **24**, 7206–7213
7. Loquet, A., Bousset, L., Gardiennet, C., Sourigues, Y., Wasmer, C., Habenstein, B., Schütz, A., Meier, B. H., Melki, R., and Böckmann, A. (2009) *J. Mol. Biol.* **394**, 108–118
8. Baxa, U., Speransky, V., Steven, A. C., and Wickner, R. B. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 5253–5260
9. Hoshino, M., Katou, H., Hagihara, Y., Hasegawa, K., Naiki, H., and Goto, Y. (2002) *Nat. Struct. Biol.* **9**, 332–336
10. Olofsson, A., Ippel, J. H., Wijmenga, S. S., Lundgren, E., and Ohman, A. (2004) *J. Biol. Chem.* **279**, 5699–5707
11. Bennett, M. J., Sawaya, M. R., and Eisenberg, D. (2006) *Structure* **14**, 811–824

12. Bennett, M. J., Schlunegger, M. P., and Eisenberg, D. (1995) *Protein Sci.* **4**, 2455–2468
13. Rodziewicz-Motowidlo, S., Wahlbom, M., Wang, X., Lagiewka, J., Janowski, R., Jaskólski, M., Grubb, A., and Grzonka, Z. (2006) *J. Struct. Biol.* **154**, 68–78
14. Janowski, R., Kozak, M., Abrahamson, M., Grubb, A., and Jaskólski, M. (2005) *Proteins* **61**, 570–578
15. Janowski, R., Abrahamson, M., Grubb, A., and Jaskólski, M. (2004) *J. Mol. Biol.* **341**, 151–160
16. Nilsson, M., Wang, X., Rodziewicz-Motowidlo, S., Janowski, R., Lindström, V., Onnerfjord, P., Westermark, G., Grzonka, Z., Jaskólski, M., and Grubb, A. (2004) *J. Biol. Chem.* **279**, 24236–24245
17. Janowski, R., Kozak, M., Jankowska, E., Grzonka, Z., Grubb, A., Abrahamson, M., and Jaskólski, M. (2001) *Nat. Struct. Biol.* **8**, 316–320
18. Laidman, J., Forse, G. J., and Yeates, T. O. (2006) *Acc. Chem. Res.* **39**, 576–583
19. Domanska, K., Vanderhaegen, S., Srinivasan, V., Pardon, E., Dupeux, F., Marquez, J. A., Giorgetti, S., Stoppini, M., Wyns, L., Bellotti, V., and Steyaert, J. (2011) *Proc. Natl. Acad. Sci. U.S.A.* **108**, 1314–1319
20. Liu, C., Sawaya, M. R., and Eisenberg, D. (2011) *Nat. Struct. Mol. Biol.* **18**, 49–55
21. Eakin, C. M., and Miranker, A. D. (2005) *Biochim. Biophys. Acta* **1753**, 92–99
22. Yang, S., Levine, H., Onuchic, J. N., and Cox, D. L. (2005) *FASEB J.* **19**, 1778–1782
23. Knaus, K. J., Morillas, M., Swietnicki, W., Malone, M., Surewicz, W. K., and Yee, V. C. (2001) *Nat. Struct. Biol.* **8**, 770–774
24. Liu, Y., Gotte, G., Libonati, M., and Eisenberg, D. (2001) *Nat. Struct. Biol.* **8**, 211–214
25. Sambashivan, S., Liu, Y., Sawaya, M. R., Gingery, M., and Eisenberg, D. (2005) *Nature* **437**, 266–269
26. Louis, J. M., Byeon, I. J., Baxa, U., and Gronenborn, A. M. (2005) *J. Mol. Biol.* **348**, 687–698
27. Byeon, I. J., Louis, J. M., and Gronenborn, A. M. (2003) *J. Mol. Biol.* **333**, 141–152
28. Guo, Z., and Eisenberg, D. (2006) *Proc. Natl. Acad. Sci. U.S.A.* **103**, 8042–8047
29. Rousseau, F., Wilkinson, H., Villanueva, J., Serrano, L., Schymkowitz, J. W., and Itzhaki, L. S. (2006) *J. Mol. Biol.* **363**, 496–505
30. Ventura, S., Lacroix, E., and Serrano, L. (2002) *J. Mol. Biol.* **322**, 1147–1158
31. Wang, J., Gülich, S., Bradford, C., Ramirez-Alvarado, M., and Regan, L. (2005) *Structure* **13**, 1279–1288
32. Ramirez-Alvarado, M., Cocco, M. J., and Regan, L. (2003) *Protein Sci.* **12**, 567–576
33. Ramirez-Alvarado, M., and Regan, L. (2002) *J. Mol. Biol.* **323**, 17–22
34. Ramirez-Alvarado, M., Merkel, J. S., and Regan, L. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 8979–8984
35. Smith, C. K., Withka, J. M., and Regan, L. (1994) *Biochemistry* **33**, 5510–5517
36. Tycko, R. (2011) *Annu. Rev. Phys. Chem.* **62**, 279–299
37. Shewmaker, F., McGlinchey, R. P., Thurber, K. R., McPhie, P., Dyda, F., Tycko, R., and Wickner, R. B. (2009) *J. Biol. Chem.* **284**, 25065–25076
38. Hoshino, M., Katou, H., Yamaguchi, K., and Goto, Y. (2007) *Biochim. Biophys. Acta* **1768**, 1886–1899
39. Takegoshi, K., Nakamura, S., and Terao, T. (2001) *Chem. Phys. Lett.* **344**, 631–637
40. Weingarth, M., Demco, D. E., Bodenhausen, G., and Tekely, P. (2009) *Chem. Phys. Lett.* **469**, 342–348
41. Tycko, R. (2007) *J. Chem. Phys.* **126**, 064506
42. Wickner, R. B., Dyda, F., and Tycko, R. (2008) *Proc. Natl. Acad. Sci. U.S.A.* **105**, 2403–2408
43. Veshkort, M., and Griffin, R. G. (2006) *J. Magn. Reson.* **178**, 248–282
44. Chimon, S., Shaibat, M. A., Jones, C. R., Calero, D. C., Aizezi, B., and Ishii, Y. (2007) *Nat. Struct. Mol. Biol.* **14**, 1157–1164
45. Wishart, D. S., and Sykes, B. D. (1994) *J. Biomol. NMR* **4**, 171–180
46. Shen, Y., Delaglio, F., Cornilescu, G., and Bax, A. (2009) *J. Biomol. NMR* **44**, 213–223
47. Franks, W. T., Zhou, D. H., Wylie, B. J., Money, B. G., Graesser, D. T., Frericks, H. L., Sahota, G., and Rienstra, C. M. (2005) *J. Am. Chem. Soc.* **127**, 12291–12305
48. Robustelli, P., Cavalli, A., and Vendruscolo, M. (2008) *Structure* **16**, 1764–1769
49. Alexandrescu, A. T. (2001) *Pac. Symp. Biocomput.* **6**, 67–78
50. Zhang, Y. Z., Paterson, Y., and Roder, H. (1995) *Protein Sci.* **4**, 804–814
51. Walsh, P., Simonetti, K., and Sharpe, S. (2009) *Structure* **17**, 417–426
52. Shewmaker, F., Kryndushkin, D., Chen, B., Tycko, R., and Wickner, R. B. (2009) *Biochemistry* **48**, 5074–5082
53. Weingarth, M., Bodenhausen, G., and Tekely, P. (2009) *J. Am. Chem. Soc.* **131**, 13937–13939
54. Iwata, K., Fujiwara, T., Matsuki, Y., Akutsu, H., Takahashi, S., Naiki, H., and Goto, Y. (2006) *Proc. Natl. Acad. Sci. U.S.A.* **103**, 18119–18124
55. Sawaya, M. R., Sambashivan, S., Nelson, R., Ivanova, M. I., Sievers, S. A., Apostol, M. I., Thompson, M. J., Balbirnie, M., Wiltzius, J. J., McFarlane, H. T., Madsen, A. Ø., Riek, C., and Eisenberg, D. (2007) *Nature* **447**, 453–457
56. Fändrich, M., Meinhardt, J., and Grigorieff, N. (2009) *Prion* **3**, 89–93
57. Kodali, R., and Wetzel, R. (2007) *Curr. Opin. Struct. Biol.* **17**, 48–57
58. Jee, J., Byeon, I. J., Louis, J. M., and Gronenborn, A. M. (2008) *Proteins* **71**, 1420–1431
59. Byeon, I. J., Louis, J. M., and Gronenborn, A. M. (2004) *J. Mol. Biol.* **340**, 615–625
60. Sasahara, K., Yagi, H., Sakai, M., Naiki, H., and Goto, Y. (2008) *Biochemistry* **47**, 2650–2660
61. Sicorello, A., Torrasa, S., Soldi, G., Gianni, S., Travaglini-Allocatelli, C., Taddei, N., Relini, A., and Chiti, F. (2009) *Biophys. J.* **96**, 2289–2298
62. Ladner, C. L., Chen, M., Smith, D. P., Platt, G. W., Radford, S. E., and Langen, R. (2010) *J. Biol. Chem.* **285**, 17137–17147
63. Debelouchina, G. T., Platt, G. W., Bayro, M. J., Radford, S. E., and Griffin, R. G. (2010) *J. Am. Chem. Soc.* **132**, 10414–10423
64. Debelouchina, G. T., Platt, G. W., Bayro, M. J., Radford, S. E., and Griffin, R. G. (2010) *J. Am. Chem. Soc.* **132**, 17077–17079
65. Cobb, N. J., Sönnichsen, F. D., McHaourab, H., and Surewicz, W. K. (2007) *Proc. Natl. Acad. Sci. U.S.A.* **104**, 18946–18951
66. Abedini, A., and Raleigh, D. P. (2009) *Protein Eng. Des. Sel.* **22**, 453–459
67. Liu, G., Prabhakar, A., Aucoin, D., Simon, M., Sparks, S., Robbins, K. J., Sheen, A., Petty, S. A., and Lazo, N. D. (2010) *J. Am. Chem. Soc.* **132**, 18223–18232
68. Sivanandam, V. N., Jayaraman, M., Hoop, C. L., Kodali, R., Wetzel, R., and Van der Wel, P. C. (2011) *J. Am. Chem. Soc.* **133**, 4558–4566