BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

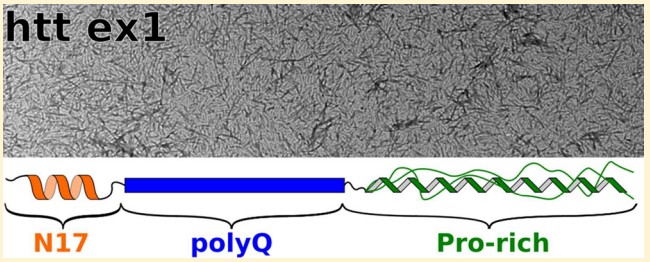
Solid-State Nuclear Magnetic Resonance on the Static and Dynamic Domains of Huntingtin Exon-1 Fibrils

J. Mario Isas, Ralf Langen, and Ansgar B. Siemer*

Department of Biochemistry and Molecular Biology, Zilkha Neurogenetic Institute, Keck School of Medicine, University of Southern California, 1501 San Pablo Street, Los Angeles, California 90033, United States

Supporting Information

ABSTRACT: Amyloid-like fibrils formed by huntingtin exon-1 (htt_ex1) are a hallmark of Huntington's disease (HD). The structure of these fibrils is unknown, and determining their structure is an important step toward understanding the misfolding processes that cause HD. In HD, a polyglutamine (polyQ) domain in htt_ex1 is expanded to a degree that it gains the ability to form aggregates comprising the core of the resulting fibrils. Despite the simplicity of this polyQ sequence, the structure of htt_ex1 fibrils has been difficult to determine. This study provides a detailed structural investigation of fibrils formed by htt_ex1 using solid-state nuclear magnetic resonance (NMR) spectroscopy. We show that the polyQ domain of htt_ex1 forms the static amyloid core similar to polyQ model peptides. The Gln residues of this domain exist in two distinct conformations that are found in separate domains or monomers but are relatively close in space. The rest of htt_ex1 is relatively dynamic on an NMR time scale, especially the proline-rich C-terminus, which we found to be in a polyproline II helical and random coil conformation. We observed a similar dynamic C-terminus in a soluble form of htt_ex1, indicating that the conformation of this part of htt_ex1 is not changed upon its aggregation into an amyloid fibril. From these data, we propose a bottlebrush model for the fibrils formed by htt_ex1. In this model, the polyQ domains form the center and the proline-rich domains the bristles of the bottlebrush.



Huntington's disease (HD) is a heritable, fatal neurodegenerative disease with symptoms of motor dysfunctions, cognitive impairments, and psychiatric disorders.[1] HD is the most common of a class of diseases in which a polyglutamine (polyQ) domain is pathologically extended above a certain threshold (36 repeats in the case of HD).[2] Besides changing the flexibility of the monomeric state,[3] pathologically expanded polyQ domains have the tendency to form fibrillar, amyloid-like aggregates in vivo and in vitro. Fibril forming kinetics and the onset of HD are faster the longer the polyQ domain.[4,5] In HD, the polyQ domain is part of the protein huntingtin (htt) and is located on htt exon-1 (htt_ex1).[6] Furthermore, htt_ex1 has been shown to be significant for HD because it is prominently found in the amyloid deposits of post-mortem brains[7] and can be produced by an aberrant splice variant.[8] Htt_ex1 has an N-terminal amphiphilic domain often termed N17, followed by the polyQ domain, whose aggregation is aided by the presence of the N17 domain.[9−11] The C-terminus of htt_ex1 has two pure polyproline stretches interrupted by a proline-rich sequence (see Figure 1). Such polyproline flanking sequences were shown to have an inhibitory effect on polyQ aggregation.[12] How the polyQ expansion results in HD is unknown. The mechanism of htt toxicity is an active field of research, and there are nontoxic and toxic fibril species. Furthermore, there are toxic protofibrils and oligomeric forms of htt.[13−15] To understand the molecular origins of toxicity and protein misfolding in HD, it is important to know the molecular structure and dynamic properties of the fibrils that are the end product of this misfolding process.

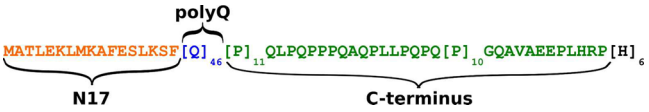


Figure 1. Sequence of htt_ex1 Q46. N17, polyQ, and the C-terminal domain are colored orange, blue, and green, respectively.

Until recently, structural studies of htt fibrils have focused on simple polyQ model peptides and htt_ex1 mimics with polyQ domains shorter than those found in HD.[16−21] A recent EPR study conducted with fibrils formed by htt_ex1 Q46 showed that N17 and the polyQ domain are relatively static, whereas the Pro-rich domain becomes increasingly dynamic toward the C-terminus. Interestingly, EPR also showed that, contrary to many other amyloid fibrils, the polyQ domain is not in an in-register β-sheet conformation.[22] However, the precise structural organization of htt_ex1 fibrils remains unknown.

To provide detailed structural information, this study uses solid-state NMR data on htt_ex1 fibrils grown at 4 °C, the

Received: March 15, 2015
Revised: May 27, 2015
Published: May 28, 2015

DOI: 10.1021/acs.biochem.5b00281
Biochemistry 2015, 54, 3942−3949

same type of fibrils employed in the previous EPR study. Temperature was shown to modulate the mechanism of misfolding, the saturation concentrations, and the fibril forming kinetics of htt.[11] Moreover, fibrils grown at 4 °C were previously shown to be more toxic and less rigid than fibrils grown at 37 °C.[14] Our data for the polyQ domain of htt_ex1 allow the comparison with the polyQ domain of htt model peptides, and we show that the proline-rich domain of htt_ex1 is dynamically and structurally more complex than previously thought.

MATERIALS AND METHODS

Protein Expression, Purification, and Fibril Formation. Uniformly ^{13}C- and ^{15}N-labeled wild-type htt_ex1 Q46 fibrils were expressed, and they were purified as described by Fodale et al.[23] with modification following a protocol by Marley et al.[24] that allows efficient isotope labeling. Overnight cultures of BL21(DE3) cells transformed with the pET32a-HDx46Q plasmid were diluted 50-fold into LB medium and grown at 37 °C to an A_{600} of 0.6. Pellets were collected by centrifugation at 3500*g*, resuspended in M9 wash buffer, pelleted again, and resuspended in a quarter of the original volume using M9 medium containing 4 g/L [U-^{13}C]glucose and 0.5 g/L [^{15}N]ammonium chloride. After incubation for 1 h at 30 °C and 200 rpm, protein expression was initiated by adding 1 mM isopropyl 1-thio-D-galactopyranoside. Cell pellets were collected by centrifugation after incubation for 4 h, resuspended in 20 mM Tris-HCl (pH 8.0), 300 mM NaCl, and 10 mM imidazole containing 1× CelLytic B Cell Lysis reagent (Sigma-Aldrich), and incubated for 20 min at room temperature on a shaker. Lysates were clarified by centrifugation at 21000*g* for 10 min and incubated with nickel-nitrilotriacetic acid-agarose beads (Qiagen) for 1 h at 4 °C on a shaker. Beads were decanted into an Econo-Pac chromatography column (Bio-Rad) and washed with several column volumes of 20 mM Tris-HCl (pH 8.0), 300 mM NaCl, and 20 mM imidazole. The pure protein was eluted with 20 mM Tris-HCl (pH 8.0), 300 mM NaCl, and 250 mM imidazole.

Following concentration of the proteins via Amicon Ultra-15 10000 MWCO centrifugal filters (Millipore), proteins were rediluted in 20 mM Tris (pH 7.4) and then purified on a HiTrap Q XL column (GE Healthcare) with an AKTA FPLC system (Amersham Pharmacia Biotech) using a NaCl gradient. The protein eluted at ∼150 mM NaCl and was consequently diluted to 20 μM (560 μg/mL). To cleave the N-terminal thioredoxin tag and initiate fibril formation, EKMax (Invitrogen) was added to a concentration of 1 unit/280 μg of protein. The reaction mixture was incubated without agitation at 4 °C for 3 days. Finally, fibril samples were washed with deionized water, sedimented by ultracentrifugation (150000*g* for 20 min), and packed into 1.6 mm magic angle spinning (MAS) rotors.

Uniformly ^{13}C- and ^{15}N-labeled htt_ex1 Q7 was expressed and purified as described above but without cleaving the thioredoxin tag. The protein was concentrated with an Amicon Ultra-15 10000 MWCO centrifugal filter, and the concentration was determined via absorbance at 280 nm; 25% glycerol was added to the final buffer of 20 mM Tris-HCl (pH 7.4) and 150 mM NaCl, and the protein solution was pipetted in a 4 mm NMR rotor.

Solid-State NMR Spectroscopy. All spectra were recorded on an Agilent DD2 600 MHz solid-state NMR spectrometer. Fibril spectra were recorded using a T3 1.6 mm probe operating at 25 kHz MAS; 200 and 100 kHz hard pulses were applied on ^{1}H and ^{13}C, respectively. The recycle delay was 3 s for all spectra. $^{1}H-^{13}C$ cross-polarizations (CPs) were conducted using a Hartman−Hahn match of 60 kHz on ^{13}C and 85 kHz on ^{1}H with a 10% amplitude ramp; 120 kHz two-pulse phase modulation (TPPM) ^{1}H decoupling was applied during direct and indirect detection. $^{13}C-^{13}C$ spectra were recorded using a 25 kHz ^{1}H recoupling field during the 50 ms mixing time (DARR[25]) and no recoupling field for the spectrum with a 500 ms mixing time (PDSD). The ^{13}C spectral width was 50 kHz in both dimensions, and 400 complex t_1 increments and 16 and 32 acquisitions were co-added per increment for the spectra with mixing times of 50 and 500 ms, respectively. $^{15}N-^{13}C$ correlation spectra were recorded using a SPECIFIC double CP (DCP) pulse sequence[26] with rf-fields of 40 kHz on ^{15}N and 65 kHz on ^{1}H for the first and 40 kHz on ^{15}N and 15 kHz on ^{13}C for the second CP. The NCA spectrum was recorded with a transmitter at 52 ppm and the NCOCX spectrum with a ^{13}C transmitter at 191 ppm and an additional $^{13}C-^{13}C$ DARR mixing of 50 ms. Both $^{15}N-^{13}C$ correlation spectra were recorded with a 50 kHz spectral width in the direct ^{13}C dimension and 10 kHz spectral width in the indirect ^{15}N dimension, and 32 and 64 acquisitions were co-added for each of the 96 complex t_1 increments for the NCA and NCOCX spectra, respectively. Direct polarization (DP) constant-time uniform-sign cross-peak (CTUC) correlation spectroscopy (COSY) spectra were recorded using hard pulses for the CA−CO correlations and using 420 and 180 μs r-SNOB pulses on carbonyl and aliphatic regions, respectively, for the aliphatic correlations.[27,28] The spectral width of these DP CTUC COSY spectra was 50 kHz in both dimensions, and 32 acquisitions were co-added for each of the 340 and 400 complex t_1 increments for the aliphatic and CA−CO spectra, respectively. The $^{1}H-^{13}C$ heteronuclear correlation (HETCOR) spectrum of the fibrils was recorded using a refocused INEPT transfer. The indirect ^{1}H dimension had a 180° ^{13}C refocusing pulse and a spectral width of 10 kHz; the direct ^{13}C dimension had a spectral width of 50 kHz, and 16 acquisitions were co-added for each of the 128 complex t_1 increments.

The refocused INEPT HETCOR spectrum of soluble htt_ex1 Q7 with an N-terminal thioredoxin tag was recorded in a 4 mm T3 probe at a MAS frequency of 3 kHz; 75 and 50 kHz hard pulses were applied on ^{1}H and ^{13}C, respectively. Twenty kilohertz TPPM decoupling was applied in the direct ^{13}C dimension. The spectral width was 50 kHz for the direct ^{13}C dimension and 10 kHz for the indirect ^{1}H dimension, and 80 acquisitions were co-added for each of the 128 complex t_1 increments. All chemical shifts were referenced externally to DSS using adamantane.[29]

RESULTS

For this NMR study, we expressed uniformly ^{13}C- and ^{15}N-labeled htt_ex1 with a 46-amino acid polyQ domain in *E. coli*. The N-terminal thioredoxin tag was cleaved off to induce fibril formation, and fibrils were formed at 4 °C. As one can see from the electron microscopy (EM) image in Figure S1 of the Supporting Information, these conditions lead to relatively homogeneous fibril preparations with occasional bundling. Frequently, fibrils appeared to be hindered from crossing over in the EM images and were spaced approximately 28−33 nm (center-to-center) apart. Mature fibrils were packed in a 1.6

DOI: 10.1021/acs.biochem.5b00281
Biochemistry 2015, 54, 3942−3949

mm MAS rotor. To study the homogeneity and structure of the htt_ex1 fibril core, we recorded dipolar CP-based solid-state NMR spectra that are sensitive to the static domains of the htt_ex1 fibrils.[30] As one can seen from the two-dimensional (2D) DARR spectrum shown in Figure 2, the resolution of the

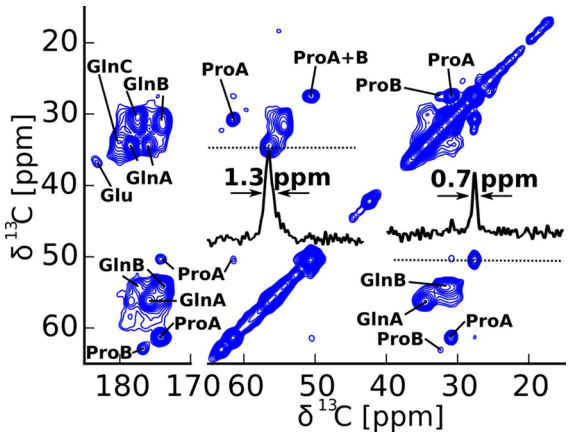


Figure 2. The fibril core of htt_ex1 Q46 is very homogeneous. 2D DARR spectrum of htt_ex1 fibrils recorded at 14.1 T, 25 kHz MAS, 0 °C, and a mixing time of 50 ms. The amino acid assignments of the two types of Gln and Pro are indicated, and the line width of Gln A C^{α} and Pro A C^{γ} is illustrated via one-dimensional slices in the direct dimension. The thus determined line widths of 0.8 ppm for Pro A C^{γ} and 1.3 ppm for Gln A C^{α} are remarkably narrow considering that these peaks result from the superposition of several resonances.

spectrum was very good. For example, the Pro C^{γ} line width highlighted in Figure 2 is 0.7 ppm, and one of the Gln C^{α} lines has a width of 1.3 ppm, which is comparable to line widths observed for other homogeneous amyloid preparations,[31] even though both of these lines arise from several different sites. These narrow lines indicate a high degree of structural homogeneity of the static parts of our htt_ex1 fibrils. The amino acid type assignment of this spectrum identifies three different forms of Gln, two different forms of Pro, and a Glu. Two forms of Gln, termed Gln A and Gln B, have roughly the same intensity, and their assignments match the previously observed chemical shift assignments,[17,18] where Gln B is compatible with an extended β-strand sheet conformation and Gln A has chemical shifts different from average C^{α}, C^{β}, and C′ chemical shifts reported for Gln in an α-helical, β-sheet, or random coil conformation.[32] From Gln C, only the $C^{\gamma}-C^{\delta}$ peak can be identified in the 2D DARR spectrum. The other peaks of Gln C partially overlap with Gln B but are the dominant form of Gln and more visible in the spectra focusing on the dynamic domains of the htt_ex1 fibrils discussed below. The two forms of Pro observed in this spectrum have very different intensities, where Pro A gives intense cross-peaks and Pro B is barely detectable. The chemical shift of Pro A, which has been previously reported on spectra of $N_{17}Q_{30}P_{10}K_2$,[19] is in good agreement with that of a polyproline II helix bound to an SH3 domain[33] but fits less well to the Pro shift measured in polyproline powder samples.[34,35] Pro B, which has not been reported previously, has a chemical shift compatible with a random coil conformation.[32]

An important question regarding the two forms of Gln that dominate the 2D DARR of htt_ex1 fibrils is whether they either are located in separate domains or fibrils or alternate in sequence (i.e., type A is followed by type B, which is followed by type A, etc.) or are randomly distributed in sequences (see Figure 3A). Distinguishing these three possibilities will allow us

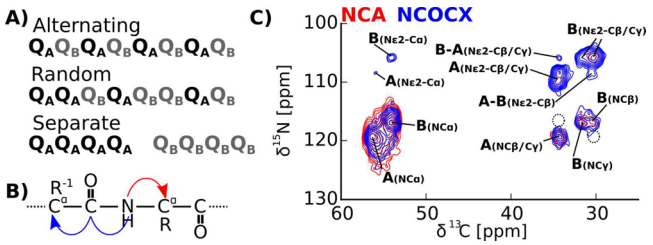


Figure 3. Gln A and Gln B are sequentially clustered. (A) Illustration of three possible arrangements of Gln A and B in the primary sequence of htt_ex1. (B) Schematic showing the magnetization pathway of the NCA (red) and NCOCX (blue) experiment. (C) 2D NCA (red) spectra correlate the backbone nitrogen to C^{α} of the same residue, and NCOCA spectra (blue) correlate the backbone nitrogen with C^{α} of the preceding residue. The almost perfect overlap of the $N-C^{\alpha}$ peaks in these spectra and also the overlap of the weaker $N-C^{\beta/\gamma}$ peaks indicate that Gln A has only Gln A as sequence neighbor and Gln B only Gln B. Cross-peaks that would be expected in different arrangements are indicated with dashed circles. Despite the clear separation of Gln A and B in sequence, weak side-chain cross-peaks between Gln A and Gln B were detected in the NCOCX spectrum. Spectra were recorded under conditions similar to those used for the spectrum in Figure 2 using a SPECIFIC DCP pulse sequence for the NCA spectrum and a SPECIFIC DCP followed by a 50 ms $^{13}C-^{13}C$ DARR mixing period for the NCOCX experiment.

to exclude and confirm different structural models. For example, an alternate arrangement of the two Gln types could point toward an α-sheet structure that has been proposed as a possible conformation of polyQ fibrils.[36] In α-sheets, the backbone dihedral angles alternate between two different regions of the Ramachandran plot, which is an interesting model because it would explain that we observe two different types of Gln in a 1:1 ratio. To distinguish between the three arrangements of Gln A and B, we recorded the NCA and NCOCX spectra. The NCA spectrum shows cross-peaks between the backbone nitrogen (N) and the C^{α} atom and other aliphatic carbons of the same residue. The NCOCX spectrum shows cross-peaks between N and the C′ and aliphatic carbons of the preceding residue in the sequence (see Figure 3B). If the two Gln forms alternate in sequence, the NCOCX spectrum would show cross-peaks of the N of Gln A with the aliphatic carbons of Gln B and of the N of Gln B with the aliphatic carbons of Gln A. In the case of random distribution, we would expect cross-peaks from N of Gln A and Gln B with the aliphatic carbons of both Gln A and Gln B. If Gln A and Gln B are separated in sequence, we expect to see only cross-peaks from Gln A to Gln A and from Gln B to Gln B. As one can see in Figure 3C, the latter is correct because the NCA and the NCOCX spectra overlap very well and no backbone Gln A to Gln B cross-peaks were observed.

The experiments described above show that Gln A and Gln B are separated in sequence or even found in different monomers. However, are Gln A and Gln B also separated in space, e.g., found in separate fibrils? To answer this question, we recorded 2D $^{13}C-^{13}C$ DARR spectra with a long mixing time of 500 ms. DARR spectra of this kind can have cross-peaks between amino acid residues that are up to 7 Å apart in space. However, nucleus-specific distance information is often obscured by relayed magnetization transfer that occurs during the long

DOI: 10.1021/acs.biochem.5b00281
Biochemistry 2015, 54, 3942−3949

mixing time.[37] As one can see in Figure 4, we observed clear cross-peaks between the Gln A and Gln B carbons, for example,

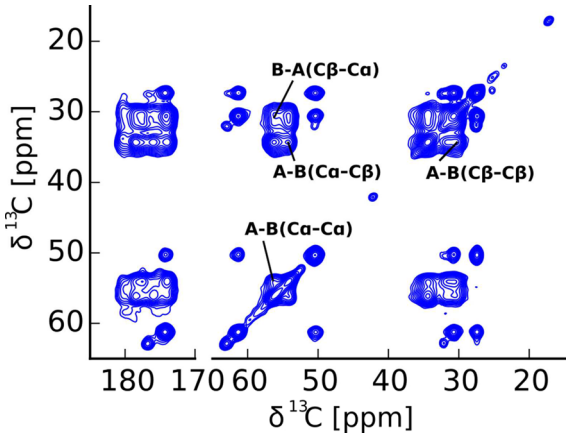


Figure 4. Although not neighboring in sequence, Gln A and Gln B are close in space. $^{13}C-^{13}C$ DARR spectrum similar to the one shown in Figure 2 but with a mixing time of 500 ms. For the sake of clarity, only additional cross-peaks belonging to Gln A−Gln B correlations are indicated. These cross-peaks show that Gln A and Gln B are within 7 Å of each other in space. Interestingly, no additional cross-peaks between the two types of Pro or between Gln and Pro are detected in these spectra.

the strong $C^{\alpha}-C^{\alpha}$ cross-peak close to the diagonal. If Gln A and Gln B were found in completely separate fibrils, no such cross-peaks would be detectable. In addition to these $C^{\alpha}-C^{\alpha}$ cross-peaks, our NCOCX spectrum in Figure 3C shows weak $N^{\varepsilon 2}$−carbon cross-peaks compatible with Gln A−Gln B side-chain−side-chain contacts besides the stronger $N^{\varepsilon 2}-C^{\gamma}$ and $N^{\varepsilon 2}-C^{\beta}$ cross-peaks within the same amino acid type. Together, these data show that although the two forms of Gln are in separate domains or even monomers, they are relatively close in space.

Besides the static domains of htt_ex1, we investigated the dynamic domains of htt_ex1 using *J* coupling-based spectroscopy that works best in the case of narrow, dynamically averaged ^{1}H and ^{13}C lines.[30] As one can see from the direct excitation (DE) $^{13}C-^{13}C$ CTUC−COSY spectrum[27,28] shown in Figure 5, we were able to detect an additional $C^{\beta}-C^{\gamma}$ cross-peak for Glu and the full spin system of Gln C, as well as Pro A and Pro B, which were also detected using CP-based spectroscopy. In addition, we were able to detect cross-peaks that we assigned to His, Ala, Leu, Val, and Lys that could not be observed in the 2D DARR spectrum. From these residues, His (from the His tag) and the only Val (V105) are exclusive to the C-terminus, Glu, Ala, and Leu can be found in both the C-terminus and the N17 domain, and Lys is exclusive to the N17 domain. Interestingly, Pro A and Pro B are observed with comparable intensity in these spectra, indicating that Pro A might be found in the static as well as dynamic domains of htt_ex1 whereas Pro B is mostly located in dynamic domains (see Figure S2 of the Supporting Information).

These dynamic residues are also observed in the $^{1}H-^{13}C$ INEPT-HETCOR spectrum shown in Figure 6. In addition, the INEPT-HETCOR spectrum also has resonances we assigned to Arg and Met. Met and Lys are the only dynamic residues that are exclusive to the N-terminus of htt_ex1, and only signals from atoms located toward the ends of their side chains (i.e., Lys $C^{\varepsilon}-H^{\varepsilon}$ and Met $C^{\varepsilon}-H^{\varepsilon}$) could be detected. The rest of the

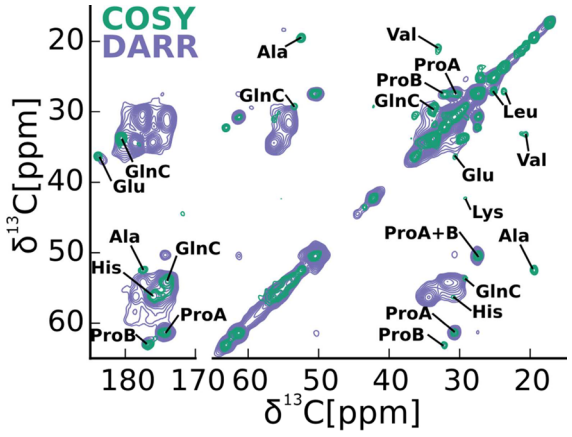


Figure 5. Dynamic parts of htt_ex1 include both types of Pro, the third form of Gln (Gln C), and a Glu residue as well as resonances compatible with Leu, His, Ala, Val, and Lys that were not previously detected using dipolar-based spectroscopy. Overlay of $^{13}C-^{13}C$ DP-CTUC-COSY (green) and $^{13}C-^{13}C$ DARR (purple) spectra of htt_ex1 fibrils recorded at 25 kHz MAS. The residues identified in the $^{13}C-^{13}C$ DP-CTUC-COSY spectrum, which is sensitive to the more dynamic parts of the fibril, are indicated.

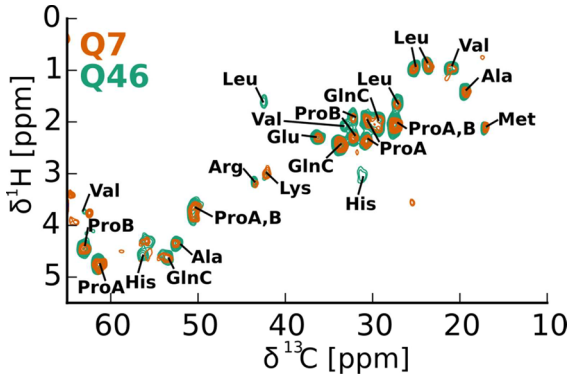


Figure 6. Dynamic parts of htt_ex1 fibrils are compatible with the C-terminal proline-rich domain. Overlay of $^{1}H-^{13}C$ INEPT-HETCOR spectra of htt_ex1 Q46 fibrils (green) and soluble htt_ex1 Q7 with an N-terminal thioredoxin tag (orange). The Q46 fibril spectrum was recorded at 14.1 T, 25 °C, and 25 kHz MAS in a 1.6 mm rotor. The soluble htt_ex1 Q7 spectrum (300 μM protein) was recorded at 14.1 T and 5 °C in a 4 mm rotor spinning at 3 kHz MAS. The tentative amino acid type assignments of the htt_ex1 fibril spectrum are indicated. Except for Met and Lys side-chain resonances, all residues identified are found in the proline-rich C-terminus of htt_ex1. The good overlap with the spectrum of soluble htt_ex1 Q7 indicates that in the most dynamic domains the fibrils and the soluble form are the same.

dynamic residues are either found throughout htt_ex1 or specific to the C-terminus where in this case also backbone resonances were observable (i.e., His $H^{\alpha}-C^{\alpha}$, Pro A and Pro B $H^{\alpha}-C^{\alpha}$). Of the amino acid types that could not be identified in any of our spectra, three are exclusive to the N-terminus (Thr, Ser, and Phe) and only one is found in the C-terminus (Gly). Taken together, these data indicate that most of the dynamic residues are located at the C-terminus of htt_ex1 and only a few side chains of the N17 peptide are dynamic enough or static enough to be detected by *J* coupling-based or dipolar coupling-based spectroscopy, respectively.

The chemical shifts of all the dynamic residues detected using *J*-based spectroscopy are compatible with a random coil

DOI: 10.1021/acs.biochem.5b00281
Biochemistry 2015, 54, 3942−3949

conformation except for Gln C, which has a particular set of chemical shifts different from those of Gln A and Gln B also not matching any of the shift patterns observed for α-helical, β-sheet, or random coil conformation.[32] (See Tables S2 and S3 of the Supporting Information for all chemical shifts reported in this paper.)

Is this dynamic C-terminus a particular feature of the fibril formed by htt_ex1 or similar in the soluble state of htt_ex1?[10,13] To answer this question, we recorded a $^{1}H-^{13}C$ INEPT-HETCOR spectrum of soluble htt_ex1. We used htt_ex1 that still had its N-terminal thioredoxin tag, contained a polyQ domain only seven residues in length, and was dissolved in 25% (v/v) glycerol to prevent aggregation. This spectrum is colored orange in Figure 6 and overlaps well with the equivalent spectrum recorded for htt_ex1 Q46 fibrils. No signals specific to the N-terminal thioredoxin tag (e.g., Asp or Asn) could be detected in these spectra. The notable differences between these two spectra are limited to a few but not all of the Leu, His, and Val resonances as well as the glycerol peaks that come from the solvent of htt_ex1 Q7. The striking overlap of the two spectra indicates that the C-termini of the soluble htt_ex1 Q7 and the htt_ex1 Q46 fibril are conformationally related.

DISCUSSION

Our spectra of the static domains of 4 °C htt_ex1 fibrils show that these are dominated by two different conformations of Gln (Gln A and Gln B) in a 1:1 ratio and a Pro that is compatible with a polyproline II conformation. Furthermore, signals of a third Gln (Gln C), a second Pro (Pro B), and a Glu could be detected in these spectra. Except for Gln A and Gln B, all resonances detected in the static domains of our htt_ex1 fibrils were also detected in the dynamic domains (i.e., in the CTUC-COSY and INEPT-HETCOR spectra) with the same chemical shifts. This suggests that the residues flanking the polyQ domain of htt_ex1 experience some dynamical heterogeneity that was also observed via EPR, especially at the interface of the polyQ and Pro-rich domain.[22]

Gln A and B were also observed in fibrils formed by selectively labeled $N_{17}Q_{30}P_{10}K_2$ and on polyQ peptides of different lengths.[17−19] Latter fibrils also showed a third Gln conformation. However, the Gln C observed in our spectra had different backbone chemical shifts and was the dominant Gln observed in the dynamic domains of the fibril and less dominant in the 2D DARR spectra. NMR data on selectively labeled $N_{17}Q_{30}P_{10}K_2$ samples showed that all Gln residues of the polyQ domain except for the last show the splitting into Gln A and Gln B.[18,19] These data, however, did not give any information about whether this splitting occurs randomly, alternatingly, or in separate domains. Our NCA and NCOCX spectra of htt_ex1 fibrils refine this assessment by showing no evidence of Gln A−Gln B contacts in these sequential backbone walks, which indicates that the two types of Gln either are in large separated domains of one monomer or are found in different monomers. These data already exclude the possibility that htt_ex1 is forming an α-sheet structure that was suggested as a possible structure of polyQ fibril structures.[36] Nevertheless, by using $^{13}C-^{13}C$ 2D DARR spectra, we observed through-space couplings between Gln A and Gln B in our htt_ex1 fibrils. These data show that, although separated in primary structure, both forms of Gln must be located within ~7 Å of each other. The weak side-chain cross-peaks between Gln A and Gln B we detected in the NCOCX spectrum indicate that these contacts come from side-chain rather than backbone interactions. Combined with the previous observation that all Gln residues in the polyQ domain can adopt both conformations, our data show that Gln A and Gln B are located either in separate monomers or inside large separate domains of the same monomer that have no fixed location in the primary structure (see Figure 7A). In the latter case, the one possible

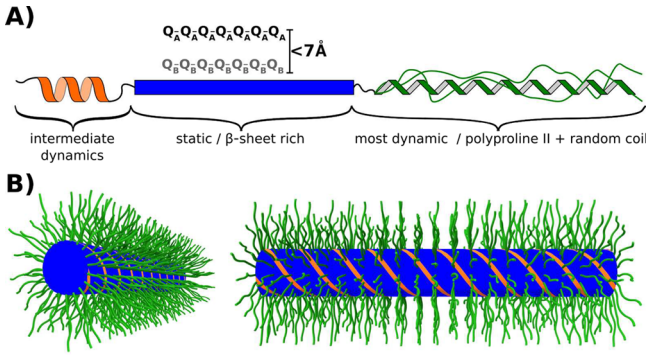


Figure 7. (A) Model illustrating the dynamic and structural heterogeneity of the htt_ex1 monomers found in 4 °C fibrils. The N17 fragment, which was shown to be in an α-helical conformation,[18] was difficult to detect in our spectra probably because of intermediate dynamics. The polyQ domain is static and features two types of Gln, A and B, which are separate in sequence but relatively close in space. The C-terminus has Pro residues in a polyproline II conformation and Pro and other non-Pro residues in a random coil conformation. The C-terminus is the most dynamic domain in htt_ex1 fibrils, albeit with considerable dynamic heterogeneity because of relatively static Pro residues as detected in our spectra. (B) Bottlebrush model of htt_ex1 4 °C fibrils. The N17 (orange) and polyQ (blue) domains form the center of the fibril. The proline-rich C-termini (green) are flexible, partially disordered, and pointing away from the center forming the bristles in this bottlebrush model.

sequential Gln A and Gln B contact would give signals that would be too weak to be detected in our spectra. One possibility is that Gln A and Gln B are clustered in different β-sheets and the through-space couplings observed between Gln A and Gln B are intersheet contacts. Such sheet−sheet interactions could occur (1) within a monomer, (2) because of lateral association of different filaments that constitute a fibril, or (3) because of lateral association of fibrils in a fibril bundle. However, option 3 is unlikely because most of the fibrils shown in Figure S1 of the Supporting Information do not bundle and the 1:1 ratio observed for Gln A and B would require that there be independently formed type A and B fibrils that have exactly the same conformational free energy.

What is the origin of the different chemical shifts observed for Gln A and Gln B? One possibility could be a difference in side-chain dihedral angles as suggested by Schneider et al.[17] Another possibility could be that the chemical shift difference is the result of a difference in β-sheet twist, e.g., found in different filaments or on different sides of a β-sandwich.

It is interesting to note that Gln alone has a relatively low propensity to form amyloid especially when arranged in an in-register parallel β-sheet conformation.[38,39] The different conformations of Gln A and Gln B might improve the complementarity of Gln in a β-sheet environment, thereby compensating for its low aggregation propensity. However, further experiments are needed to determine the origins of the two Gln forms.

DOI: 10.1021/acs.biochem.5b00281
Biochemistry 2015, 54, 3942−3949

Gln C gave weak signals in the CP-based spectra but was the dominant signal in the *J* coupling-based spectroscopy, whereas Gln A and Gln B were not detected in the latter spectra. The chemical shift pattern of Gln C is, similar to that of Gln A, not well compatible with the average chemical shift patterns observed for either α-helical, β-sheet, or random coil conformation. Because of their dynamics, we think that these Gln residues are found either at the edge of the polyQ domain or amid the seven Gln residues that are not part of the polyQ domain but are found in the proline-rich C-terminal domain between the first and second polyproline stretches and briefly after the second polyproline stretch. The origin of this particular chemical shift pattern is not clear. One possibility could be that these Gln residues are part of the C-terminal polyproline II helical structure. Another possibility could be that these Gln residues are transiently forming part of the structure formed by the polyQ domain. Both possibilities would explain the partial immobilization of these Gln residues.

The other amino acid types we detected in the dynamic domains of our htt_ex1 fibrils using *J* coupling-based spectroscopy match the C-terminus more than the N17 domain of htt_ex1. Because of the absence of Thr, Ser, and Phe resonance in any of our spectra, we conclude that the dynamic residues we detect in our *J*-based spectra are located at the C-terminus of htt_ex1. Except for two side-chain resonances that are often relatively flexible, the N17 domain is not visible in our spectra, possibly because of their intermediate dynamics described by Hoop et al. that might make their detection with both our CP- and *J* coupling-based experiments difficult.[19]

Our CP- and *J* coupling-based experiments indicate that Pro is found in both the static and dynamic parts of our htt_ex1 fibrils. While these residues are certainly not present in fibrils formed by polyQ peptides, Hoop et al. detected relatively immobile Pro in fibrils formed by $N_{17}Q_{30}P_{10}K_2$ peptides and assigned it to a polyproline II helical structure, correlating with previous circular dichroism (CD) spectroscopy data of htt_ex1.[19,22,33] Our NMR spectra, in contrast, show two different confirmations of Pro, Pro A and Pro B. Pro A, which gives intense signals in both CP- and *J* coupling-based spectra, matches the Pro described by Hoop et al. Pro B, which has not been described previously, has a chemical shift compatible with a random coil conformation. Pro B gives only faint cross-peaks in the CP-based spectra that highlight the static domains of the fibril but gives strong cross-peaks comparable in intensity to those of Pro B in the *J* coupling-based spectra, which are dominated by the dynamic domains of the fibril. This can also be seen from one-dimensional spectra (see Figure S2 of the Supporting Information), where the Pro B:Pro A ratio is 1:2.2 using dipolar CP and 1:1.1 in the refocused INEPT spectrum. In the direct pulse spectrum, which excites all spins independent of dynamics, the Pro B:Pro A ratio is 1:1.7, which is comparable to the ratio of Pro that has non-Pro residues as neighbors (12) to Pro that is surrounded by other Pro residues (18). Taken together with previous EPR and NMR observations,[19,22] our data suggest that Pro A is found closer to the polyQ domain whereas Pro B is found in the more dynamic parts toward the C-terminus. Alternatively, Pro B is the preferred conformation of Pro that has non-Pro residues as neighbors. Both of these models would also explain why previous studies of $N_{17}Q_{30}P_{10}K_2$ revealed only Pro A because this construct contains only the first ten Pro residues following the polyQ domain.

Interestingly, the highly dynamic C-terminal residues are found not only in fibrils of htt_ex1 but also in our nonfibrillar soluble preparations of htt_ex1 as shown by the almost identical INEPT-HETCOR spectra we recorded on fibrillar htt_ex1 Q46 and soluble htt_ex1 Q7, which served as a reference state of nonfibrillized htt_ex1. These data show that the dynamic, mixed polyproline II helical and random coil C-terminus is not specific to the fibrillar form of htt_ex1 and that the most dynamic domains of htt_ex1 are structurally not affected by fibril formation under these conditions.

We consequently propose a bottlebrush model for the structure of htt_ex1 fibrils that is shown in Figure 7B. In this model, the static, β-sheet-rich polyQ domains form the center of the brush. The N17 peptides with intermediate dynamics are appended to the polyQ domain, similar to a previous model of Williamson and co-workers.[40] The proline-rich C-terminal domains form the bristles of the bottlebrush pointing to the outside of the fibril. This model is supported not only by the solid-state NMR data presented in this paper but also by previous EPR data[22] as described in the following. (1) The EPR line widths of labeled sites in the C-terminus were much narrower than those observed in the N17 and polyQ domain of htt_ex1, indicating that the C-terminus is the most dynamic domain of the fibril. (2) Our solid-state NMR spectra confirm that the proline-rich domain is much more dynamic than the relatively static polyQ domain. (3) The EPR line widths of the proline-rich domain decrease toward the C-terminus until they are below 2 G and comparable to EPR line widths found in soluble protein domains lacking tertiary or quaternary contacts.[41,42] (4) The chemical shifts measured with solid-state NMR show that the proline-rich domain is in a polyproline II helical and random coil conformation. (5) Previous CD data[22] confirm the existence of a polyproline II helix in htt_ex1. (6) Polyproline II helices are known to be extended and act as bristles.[43] For that reason, they are often used as molecular rulers. (7) These bristles could principally be facing into the fibril, or they could be on the outside. Were they to face inward, the interior of the fibrils would have to have a large aqueous cavity long enough to accommodate bristles >12 nm in length assuming fully extended polyproline II helices (~40 amino acids at 3.1 Å per amino acid in polyproline II helix). Our EM data show no evidence of such a cavity, whose diameter would exceed that size of most fibrils [approximately 10−13 nm (Figure S1 of the Supporting Information)]. Moreover, the crowding in such a cavity would also likely lead to tertiary or quaternary contacts, which would be inconsistent with the EPR and NMR data. Thus, the polyproline region is likely to face outward. Disordered protein domains outside the fibril core are usually not detected using negatively stained EM, and thus, we cannot directly detect the presence of polyproline bristles via EM.[44] We did note, however, that fibrils often appear to have a regular spacing between them. This center-to-center spacing is approximately 28−33 nm, a range that is consistent with fibrils being held apart by polyproline bristles.

Polyproline or proline-rich domains have been previously suggested to act as entropic spacers or bristles, and a recent molecular dynamics simulation of polyproline showed that these polymers can form heterogeneous conformational ensembles that feature relatively rigid extended segments that are interrupted by flexible kinks.[45] This heterogeneity could explain our findings that the C-terminal Pro residues of htt_ex1 can be found in both the static and dynamic parts of the fibril.

DOI: 10.1021/acs.biochem.5b00281
Biochemistry 2015, 54, 3942−3949

In summary, we described the structural and dynamic properties of fibrils formed by htt_ex1. The polyQ domain is dominated by two different forms of Gln that are sequentially separate, but still relatively close in space. The proline-rich domain has Pro in two different conformations, one being compatible with a polyproline II helix and the other with a random coil conformation. Parts of polyproline II Pro A are relatively static, while other parts of polyproline II Pro A are found in the dynamic domains of htt_ex1 together with Pro B, which is compatible with a random coil conformation. Most of the other amino acids identified in the dynamic domains are compatible with the C-terminus, confirming that the proline-rich domain of htt_ex1 fibrils becomes increasingly dynamic toward the C-terminus.

■ ASSOCIATED CONTENT

Supporting Information

Figures S1 and S2 and Tables S3 and S4. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.biochem.5b00281.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: asiemer@usc.edu. Telephone: (323) 442-2720.

Funding

This work was supported by the Hereditary Disease Foundation (R.L.), startup funds from the University of Southern California (A.B.S.), and the National Institute of Neurological Disorders and Stroke of the National Institutes of Health via Grant R01NS084345.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Franziska Meier for initial work on the expression of labeled htt_ex1, Natalie C. Kegulian for carefully proofreading the manuscript, Alexander Falk for fruitful discussions, and Erin N. Johnson for making the bottlebrush model depicted in Figure 7.

■ ABBREVIATIONS

htt, huntingtin; htt_ex1, huntingtin exon-1; HD, Huntington's disease; polyQ, polyglutamine; NMR, nuclear magnetic resonance; EPR, electron paramagnetic resonance; MAS, magic angle spinning; MWCO, molecular weight cutoff; TPPM, two-pulse phase modulation; CP, cross-polarization; DARR, dipolar assisted rotational resonance; PDSD, proton-driven spin diffusion; DCP, double cross-polarization; HETCOR, heteronuclear correlation; DP, direct polarization; CTUC, constant-time uniform-sign cross-peak; COSY, correlation spectroscopy; INEPT, insensitive nuclei enhanced by polarization transfer; EM, electron microscopy; CD, circular dichroism.

■ REFERENCES

(1) Vonsattel, J. P., and DiFiglia, M. (1998) Huntington disease. *J. Neuropathol. Exp. Neurol. 57*, 369−384.

(2) Gatchel, J. R., and Zoghbi, H. Y. (2005) Diseases of unstable repeat expansion: Mechanisms and common principles. *Nat. Rev. Genet. 6*, 743−755.

(3) Caron, N. S., Desmond, C. R., Xia, J., and Truant, R. (2013) Polyglutamine domain flexibility mediates the proximity between flanking sequences in huntingtin. *Proc. Natl. Acad. Sci. U.S.A. 110*, 14610−14615.

(4) Brinkman, R. R., Mezei, M. M., Theilmann, J., Almqvist, E., and Hayden, M. R. (1997) The likelihood of being affected with Huntington disease by a particular age, for a specific CAG size. *Am. J. Hum. Genet. 60*, 1202−1210.

(5) Chen, S., Ferrone, F. A., and Wetzel, R. (2002) Huntington's disease age-of-onset linked to polyglutamine aggregation nucleation. *Proc. Natl. Acad. Sci. U.S.A. 99*, 11884−11889.

(6) The Huntington's Disease Collaborative Research Group (1993) A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell 72*, 971−983.

(7) DiFiglia, M., Sapp, E., Chase, K. O., Davies, S. W., Bates, G. P., Vonsattel, J. P., and Aronin, N. (1997) Aggregation of huntingtin in neuronal intranuclear inclusions and dystrophic neurites in brain. *Science 277*, 1990−1993.

(8) Sathasivam, K., Neueder, A., Gipson, T. A., Landles, C., Benjamin, A. C., Bondulich, M. K., Smith, D. L., Faull, R. L. M., Roos, R. A. C., Howland, D., Detloff, P. J., Housman, D. E., and Bates, G. P. (2013) Aberrant splicing of HTT generates the pathogenic exon 1 protein in Huntington disease. *Proc. Natl. Acad. Sci. U.S.A. 110*, 2366−2370.

(9) Kokona, B., Rosenthal, Z. P., and Fairman, R. (2014) Role of the coiled-coil structural motif in polyglutamine aggregation. *Biochemistry 53*, 6738−6746.

(10) Jayaraman, M., Kodali, R., Sahoo, B., Thakur, A. K., Mayasundari, A., Mishra, R., Peterson, C. B., and Wetzel, R. (2011) Slow Amyloid Nucleation via α-Helix-Rich Oligomeric Intermediates in Short Polyglutamine-Containing Huntingtin Fragments. *J. Mol. Biol. 415*, 881−899.

(11) Crick, S. L., Ruff, K. M., Garai, K., Frieden, C., and Pappu, R. V. (2013) Unmasking the roles of N- and C-terminal flanking sequences from exon 1 of huntingtin as modulators of polyglutamine aggregation. *Proc. Natl. Acad. Sci. U.S.A. 110*, 20075−20080.

(12) Darnell, G. D., Derryberry, J., Kurutz, J. W., and Meredith, S. C. (2009) Mechanism of cis-inhibition of polyQ fibrillation by polyP: PPII oligomers and the hydrophobic effect. *Biophys. J. 97*, 2295−2305.

(13) Hoffner, G., and Djian, P. (2014) Monomeric, Oligomeric and Polymeric Proteins in Huntington Disease and Other Diseases of Polyglutamine Expansion. *Brain Sci. 4*, 91−122.

(14) Nekooki-Machida, Y., Kurosawa, M., Nukina, N., Ito, K., Oda, T., and Tanaka, M. (2009) Distinct conformations of in vitro and in vivo amyloids of huntingtin-exon1 show different cytotoxicity. *Proc. Natl. Acad. Sci. U.S.A. 106*, 9679−9684.

(15) Pieri, L., Madiona, K., Bousset, L., and Melki, R. (2012) Fibrillar α-Synuclein and Huntingtin Exon 1 Assemblies Are Toxic to the Cells. *Biophys. J. 102*, 2894−2905.

(16) Wetzel, R. (2012) Physical chemistry of polyglutamine: Intriguing tales of a monotonous sequence. *J. Mol. Biol. 421*, 466−490.

(17) Schneider, R., Schumacher, M. C., Mueller, H., Nand, D., Klaukien, V., Heise, H., Riedel, D., Wolf, G., Behrmann, E., Raunser, S., Seidel, R., Engelhard, M., and Baldus, M. (2011) Structural characterization of polyglutamine fibrils by solid-state NMR spectroscopy. *J. Mol. Biol. 412*, 121−136.

(18) Sivanandam, V. N., Jayaraman, M., Hoop, C. L., Kodali, R., Wetzel, R., and Van Der Wel, P. C. A. (2011) The aggregation-enhancing huntingtin N-terminus is helical in amyloid fibrils. *J. Am. Chem. Soc. 133*, 4558−4566.

(19) Hoop, C., Lin, H.-K., Kar, K., Hao, Z., Poirier, M., Wetzel, R., and Van Der Wel, P. C. A. (2014) Polyglutamine Amyloid Core Boundaries and Flanking Domain Dynamics in Huntingtin Fragment Fibrils Determined by Solid-State Nuclear Magnetic Resonance. *Biochemistry 53*, 6653−6666.

(20) Michalek, M., Salnikov, E. S., Werten, S., and Bechinger, B. (2013) Membrane interactions of the amphipathic amino terminus of huntingtin. *Biochemistry 52*, 847−858.

(21) Kim, M. W., Chelliah, Y., Kim, S. W., Otwinowski, Z., and Bezprozvanny, I. (2009) Secondary structure of Huntingtin amino-terminal region. *Structure 17*, 1205−1212.

DOI: 10.1021/acs.biochem.5b00281
Biochemistry 2015, 54, 3942−3949

(22) Bugg, C. W., Isas, J. M., Fischer, T., Patterson, P. H., and Langen, R. (2012) Structural features and domain organization of huntingtin fibrils. *J. Biol. Chem. 287*, 31739−31746.
(23) Fodale, V., Kegulian, N. C., Verani, M., Cariulo, C., Azzollini, L., Petricca, L., Daldin, M., Boggio, R., Padova, A., Kuhn, R., Pacifici, R., Macdonald, D., Schoenfeld, R. C., Park, H., Isas, J. M., Langen, R., Weiss, A., and Caricasole, A. (2014) Polyglutamine- and temperature-dependent conformational rigidity in mutant huntingtin revealed by immunoassays and circular dichroism spectroscopy. *PLoS One 9*, e112262.
(24) Marley, J., Lu, M., and Bracken, C. (2001) A method for efficient isotopic labeling of recombinant proteins. *J. Biomol. NMR 20*, 71−75.
(25) Takegoshi, K., Nakamura, S., and Terao, T. (2001) ^{13}C-^{1}H dipolar-assisted rotational resonance in magic-angle spinning NMR. *Chem. Phys. Lett. 344*, 631−637.
(26) Baldus, M., Petkova, A. T., Herzfeld, J., and Griffin, R. G. (1998) Cross polarization in the tilted frame: Assignment and spectral simplification in heteronuclear spin systems. *Mol. Phys. 95*, 1197−1207.
(27) Chen, L., Olsen, R. A., Elliott, D. W., Boettcher, J. M., Zhou, D. H., Rienstra, C. M., and Mueller, L. J. (2006) Constant-time through-bond ^{13}C correlation spectroscopy for assigning protein resonances with solid-state NMR spectroscopy. *J. Am. Chem. Soc. 128*, 9992−9993.
(28) Comellas, G., Lemkau, L. R., Zhou, D. H., George, J. M., and Rienstra, C. M. (2012) Structural intermediates during α-synuclein fibrillogenesis on phospholipid vesicles. *J. Am. Chem. Soc. 134*, 5090−5099.
(29) Harris, R. K., Becker, E. D., Menezes, S. M. C., De Granger, P., Hoffman, R. E., and Zilm, K. W. (2008) Further Conventions for NMR Shielding and Chemical Shifts (IUPAC Recommendations 2008): International Union of Pure and Applied Chemistry Physical and Biophysical Chemistry Division. *Magn. Reson. Chem. 46*, 582−598.
(30) Siemer, A. B., Arnold, A. A., Ritter, C., Westfeld, T., Ernst, M., Riek, R., and Meier, B. H. (2006) Observation of highly flexible residues in amyloid fibrils of the HET-s prion. *J. Am. Chem. Soc. 128*, 13224−13228.
(31) Lu, J.-X., Qiang, W., Yau, W.-M., Schwieters, C. D., Meredith, S. C., and Tycko, R. (2013) Molecular structure of β-amyloid fibrils in Alzheimer's disease brain tissue. *Cell 154*, 1257−1268.
(32) Wang, Y., and Jardetzky, O. (2002) Probability-based protein secondary structure identification using combined NMR chemical-shift data. *Protein Sci. 11*, 852−861.
(33) Candel, A. M., Conejero-Lara, F., Martinez, J. C., van Nuland, N. A. J., and Bruix, M. (2007) The high-resolution NMR structure of a single-chain chimeric protein mimicking a SH3-peptide complex. *FEBS Lett. 581*, 687−692.
(34) Sone, M., Yoshimizu, H., Kurosu, H., and Ando, I. (1994) Side-chain conformation of poly(l-proline) form II in the crystalline state as studied by high-resolution solid-state ^{13}C NMR spectroscopy. *J. Mol. Struct. 317*, 111−118.
(35) Graf, R., Spiess, H. W., Floudas, G., Butt, H.-J., Gkikas, M., and Iatrou, H. (2012) Conformational Transitions of Poly(L-proline) in Copolypeptides with Poly(γ-benzyl-L-glutamate) Induced by Packing. *Macromolecules 45*, 9326−9332.
(36) Babin, V., Roland, C., and Sagui, C. (2011) The α-sheet: A missing-in-action secondary structure? *Proteins: Struct., Funct., Bioinf. 79*, 937−946.
(37) Gehman, J. D., Paulson, E. K., and Zilm, K. W. (2003) The influence of internuclear spatial distribution and instrument noise on the precision of distances determined by solid state NMR of isotopically enriched proteins. *J. Biomol. NMR 27*, 235−259.
(38) Pawar, A. P., DuBay, K. F., Zurdo, J., Chiti, F., Vendruscolo, M., and Dobson, C. M. (2005) Prediction of "aggregation-prone" and "aggregation-susceptible" regions in proteins associated with neurodegenerative diseases. *J. Mol. Biol. 350*, 379−392.
(39) Margittai, M., and Langen, R. (2006) Side chain-dependent stacking modulates tau filament structure. *J. Biol. Chem. 281*, 37820−37827.
(40) Williamson, T. E., Vitalis, A., Crick, S. L., and Pappu, R. V. (2010) Modulation of Polyglutamine Conformations and Dimer Formation by the N-Terminus of Huntingtin. *J. Mol. Biol. 396*, 1295−1309.
(41) Mchaourab, H. S., Lietzow, M. A., Hideg, K., and Hubbell, W. L. (1996) Motion of spin-labeled side chains in T4 lysozyme. Correlation with protein structure and dynamics. *Biochemistry 35*, 7692−7704.
(42) Isas, J. M., Langen, R., Haigler, H. T., and Hubbell, W. L. (2002) Structure and dynamics of a helical hairpin and loop region in annexin 12: A site-directed spin labeling study. *Biochemistry 41*, 1464−1473.
(43) Ruff, K. M., Khan, S. J., and Pappu, R. V. (2014) A Coarse-Grained Model for Polyglutamine Aggregation Modulated by Amphipathic Flanking Sequences. *Biophys. J. 107*, 1226−1235.
(44) Wegmann, S., Medalsy, I. D., Mandelkow, E., and Müller, D. J. (2013) The fuzzy coat of pathological human Tau fibrils is a two-layered polyelectrolyte brush. *Proc. Natl. Acad. Sci. U.S.A. 110*, E313−E321.
(45) Radhakrishnan, A., Vitalis, A., Mao, A. H., Steffen, A. T., and Pappu, R. V. (2012) Improved atomistic Monte Carlo simulations demonstrate that poly-L-proline adopts heterogeneous ensembles of conformations of semi-rigid segments interrupted by kinks. *J. Phys. Chem. B 116*, 6862−6871.

DOI: 10.1021/acs.biochem.5b00281
Biochemistry 2015, 54, 3942−3949