Classification: Biological Sciences; Biochemistry

Huntingtin exon 1 fibrils feature an interdigitated β-hairpin-based polyglutamine core.

Short title: Assembly of the huntingtin exon1 fibril core

Cody L. Hoop[a,1,2], Hsiang-Kai Lin[a,1], Karunakar Kar[a,b,3], Gábor Magyarfalvi[c], Jonathan M. Lamley[d], Jennifer C. Boatz[a], Abhishek Mandal[a], Józef R. Lewandowski[d], Ronald Wetzel[a,b], Patrick C.A. van der Wel[a,4]

Author affiliations:

[a] Department of Structural Biology, University of Pittsburgh School of Medicine, Biomedical Science Tower 3, 3501 Fifth Ave, Pittsburgh, PA 15260, USA.

[b] Pittsburgh Institute for Neurodegenerative Diseases, University of Pittsburgh School of Medicine, Biomedical Sciences Tower 3, 3501 Fifth Avenue, Pittsburgh, PA 15260, USA

[c] Laboratory of Molecular Spectroscopy, Institute of Chemistry, Eötvös University, PO Box 32, H-1518, Budapest 112, Hungary.

[d] Department of Chemistry, University of Warwick, Coventry CV4 7AL, United Kingdom.

1 C.L.H. and H.K.L. contributed equally to this work.

2 Current address: Department of Chemistry and Chemical Biology, Rutgers, Piscataway, NJ, USA.

3 Current address: Indian Institute of Technology Jodhpur, Jodhpur, Rajasthan, India.

4 Corresponding author: Patrick van der Wel; University of Pittsburgh School of Medicine, Department of Structural Biology; 3501 Fifth Ave, BST3 room 2044, Pittsburgh PA 15260; phone: +1(412)-383-9896; email: vanderwel@pitt.edu.

Keywords:

Solid-state NMR; Huntington's Disease; Amyloid disease; Protein aggregation; Amyloid

Abstract:

Polyglutamine expansion within the exon1 of huntingtin leads to protein misfolding, aggregation, and cytotoxicity in Huntington's Disease. This incurable neurodegenerative disease is the most prevalent member of a family of CAG repeat expansion disorders. Although mature exon1 fibrils are viable candidates for the toxic species, their molecular structure and how they form have remained poorly understood. Using advanced magic angle spinning solid state NMR, we directly probe the structure of the rigid core that is at the heart of huntingtin exon1 fibrils and other polyglutamine aggregates, via measurements of long-range intra- and inter-molecular contacts, backbone and side chain torsion angles, relaxation measurements, and calculations of chemical shifts. These reveal the presence of β-hairpin-containing β-sheets that are connected through interdigitating extended side chains. Despite dramatic differences in aggregation behavior, huntingtin exon1 fibrils and other polyglutamine-based aggregates contain identical β-strand-based cores. Prior structural models, derived from X-ray fiber diffraction and computational analyses, are shown to be inconsistent with the solid-state NMR results. Internally, the polyglutamine amyloid fibrils are co-assembled from differently structured monomers, which we describe as a type of 'intrinsic' polymorphism. A stochastic polyglutamine-specific aggregation mechanism is introduced to explain this phenomenon. We show that the aggregation of mutant huntingtin exon1 proceeds via an intramolecular collapse of the expanded polyglutamine domain, and discuss the implications of this observation for our understanding of its misfolding and aggregation mechanisms.

Significance

Huntington's Disease is a devastating and incurable inherited neurodegenerative disease. Like at least eight other diseases, its primary genetic cause is the CAG repeat expansion in a specific gene. Mutant huntingtin protein undergoes misfolding and aggregation, causing degeneration of neurons through as-yet poorly understood mechanisms. Attempts to characterize the implicated protein deposits have until now had limited success. We present our structural studies of mutant huntingtin-derived protein deposits by advanced solid-state NMR spectroscopy. We determine the essential structural features of the fibrils' rigid core, which is shown to feature intramolecular β-hairpins tied together via interdigitating extended side chains. These structural insights have direct implications for the mechanism by which the mutant protein misfolds and selfassembles.

The misfolding and aggregation of proteins is a common, but as-yet poorly understood, cause for human disease. One family of protein misfolding diseases involves the expansion of CAG repeats in specific genes (1). Beyond a threshold value, increasing CAG repeat lengths correlate to decreasing age of pathological onset and increasing toxicity. The most prevalent example is Huntington's Disease (HD): an incurable neurological disorder that impacts motor and cognitive abilities and is ultimately fatal. In HD, the expansion affects a polyglutamine (polyQ) domain near the N-terminus of the huntingtin protein (htt), causing protein misfolding, N-terminal fragmentation, and aggregation. Much attention has focused on the N-terminal fragment coinciding with Htt exon1 (Fig. 1A), as it is generated in vivo and induces HD-like disease pathology in mouse models (2, 3). Htt exon1 is known to misfold and self-assemble via a series of aggregated species, including spherical oligomers, protofibrils, mature amyloid fibrils, and large fibril clusters (3-6). Although some studies suggested that large inclusions are non-toxic and may be protective (7), recent work also revealed the presence in cells of smaller amyloid aggregates of exon1 that are not visible in normal fluorescence microscopy (8) and specific toxic cellular events triggered by inclusions (9). Despite their potential importance, only limited atomic resolution structural data are available on the fibrillar aggregates that are formed by htt exon1 or other polyQ-based proteins or peptides (10). Although it is generally accepted that the mature fibrils feature antiparallel β-sheets, there continue to be conflicting models not only for their fibrillar structure, but also for the specifics of the multi-stage pathway by which they are formed. For instance, intramolecular polyQ-based β-hairpins are both proposed to be either nuclei that initiate the rapid formation of β-hairpin-based fibrils (11), or semi-stable monomeric or oligomeric species that at least transiently resist progression to fibrils (12). Concrete structural data on the fibril's internal structure are essential pre-requisites for a truly molecular understanding of the way in which mutant htt exon1 and other polyQ disease proteins misfold and aggregate.

Magic-angle-spinning (MAS) solid-state NMR (ssNMR) spectroscopy has developed into an essential tool for the determination of amyloid fibril structure (13, 14). We previously used MAS ssNMR to elucidate the domain structure of aggregated htt N-terminal fragments, in which we identified the rigid amyloid core (15, 16). This polyQ-based amyloid core was found to have an unusual spectral signature that is also shared by other polyQ aggregates (11, 15, 17). However, until now, no ssNMR-based structural constraints on the htt exon1 amyloid core were available. Here we present structural ssNMR measurements on the exon1 amyloid core and compare it to polyQ peptide fibrils. We show that two kinds of β-strands make up the core assembly, where they engage in intimate intra-protein interactions. Interactions between β-sheets involve steric-zipper-like side-chain interdigitation (18), based on side-chain torsion angle measurements and other structural and dynamic constraints. Implications of the β-hairpin-based core structure for the misfolding and aggregation pathways followed by expanded polyQ domains in context of htt exon1 and beyond are discussed.

RESULTS

The β-sheet based huntingtin exon1 fibril core.

Mutant htt exon1 with an expanded 44-residue polyQ domain was expressed as a maltose-binding protein (MBP) fusion construct (Fig. 1A) (5, 16). Cleavage with factor Xa releases exon1, which first forms oligomeric aggregates (Fig. 1C), and later amyloid-like fibrils (Fig. 1D-E). The MAS ssNMR spectrum in Figure 2A shows the Gln ^{13}C signals of the rigid amyloid core of uniformly ^{13}C and ^{15}N (U-^{13}C,^{15}N) labeled exon1 fibrils, which have a width of ~15 nm (Fig. 1F). Colored lines mark the Gln peaks of two types ('a' and 'b') of rigid core residues present in equal amounts. Signals from the partly mobile flanking domains outside the amyloid core (15, 16) are not marked. Despite reports of temperature-dependent polymorphism (19), no difference is seen between the polyQ core signals of fibrils formed at room temperature and 37 °C (Fig. S1). A much smaller set of signals (type 'c') is observed for Gln outside the amyloid core (16, 17) (Fig. S1). In "backbone walk" spectra we correlate the ^{15}N and ^{13}C signals of sequential residues (Fig. 2B,C). The intra- and inter-residue correlation spectra are identical (Fig. 2D), which implies that each Gln is preceded and followed by residues of the same type. No evidence of direct 'a'-'b' connections is observed. Thus, the exon1 fibril core must contain some combination of two structurally distinct types of uninterrupted polyQ tracts, each of which is exclusively comprised of one of two distinct Gln conformers that differ in their NMR signals (Fig. 2E).

Fig. S1 shows that these NMR signals are indistinguishable from those of "simple" polyQ fibrils without htt flanking domains (11, 15-17). Previous studies have noted that these chemical shifts are indicative of β-sheet rather than α-helical or random coil structure. However, some argue that polyQ aggregation may involve a non-standard secondary structure, known as α-sheet, which may be indistinguishable from β-sheet by its NMR shifts (20, 21). To test this, we did chemical-shift-independent torsion angle measurements (22) and find unambiguous evidence for a β-sheet conformation. We used "NCCN" experiments that probe the relative orientation of N_i-C_{i^α} and C_i'-N_{i+1} dipolar coupling vectors, and thus report on $|\psi|$ (the magnitude of the ψ torsion angle; Fig. S2) (23). We labeled two sequential Gln within a polyQ peptide, where both feature the amyloid core ssNMR signals (Fig. S1H,I). We obtained NCCN measurements for each of these signals (Fig. 3A,B; Fig. S2), and find that type 'b' Gln have $|\psi| = 152 \pm 2°$. This unambiguously contradicts an α-sheet-based assembly in which residues would occupy the α and Lα regions (Fig. 3C). The NCCN measurement and chemical shift analysis (24) agree on the β-strand conformation (Fig. 3C), similar to prior ssNMR studies of amyloid structure (25, 26). The type 'a' Gln have a different NCCN signal from type 'b' Gln (Fig. 3A). It also fits a β-sheet structure, although in this case not uniquely so. We conclude that it reflects a β-strand structure based on agreement from chemical shift analysis (Fig. 3C) and the fact that the two conformers co-assemble in the amyloid core (see below). Thus, the doubled ssNMR peaks of the htt exon1 fibril core and other polyQ amyloids are due to two equally populated, but structurally distinct β-strand types.

The htt exon1 core β-strands form intramolecular β-hairpins.

We probed for interactions between these two types of β-strands by recording a ^{13}C-^{13}C spectrum with a longer ^{13}C-^{13}C polarization exchange time (Fig. 4A,B). In this spectrum with 250 ms proton-driven spin diffusion (PDSD) mixing we observe many strong cross-peaks between the

two sets of Gln peaks. Both in the literature (27, 28), and in reference experiments (Fig. S3A), this mixing time allows for detectable signal transfer over up to ~7 Å. Many of the "inter-form" (i.e., between 'a' and 'b' conformers) peaks are even visible at shorter PDSD mixing times (Fig. S3D). We see extensive transfer between backbones, from backbone to side chains and *vice versa,* which implies that the two β-strand types are in intimate contact with each other and form a "composite" amyloid core (29).

An in-depth structural analysis of these data is difficult, since in fully labeled amyloid fibrils (Fig. 4A,B), the peaks reflect both interactions within a particular protein as well as those between different proteins. Thus, we prepared fibrils from a mix of ^{13}C-only and ^{15}N-only labeled exon1. Here, one expects that ^{13}C-^{13}C polarization transfer between ^{13}C sites *within* each ^{13}C-labeled monomer will be unchanged, but ^{13}C-^{13}C contacts *between* proteins will be suppressed. Using fibrils containing 26% ^{13}C- and 74% ^{15}N-labeled protein we repeat the ^{13}C-^{13}C experiment (Fig. 4C). Even with the four-fold ^{13}C dilution, cross-peaks between the 'a'/'b' conformers are still observed. This can be seen clearly in 1D slices (Fig. 4D-F) comparing the fully labeled and mixed fibrils, both for polarization transfer from backbone (circled C^{α} & C') or side chain carbons (C^{s}; circled). Peaks reflecting transfer across to the other β-strand (e.g. 'a' to 'b' or vice versa) are specifically marked. For more quantitative insights, we measured the 2D peak intensities, and found the inter-form peaks to reflect 2-8 % of the originating ^{13}C signal (Fig. 4G; Fig. S3B). In analogous data on a reference sample of known structure, such peak volumes correspond to distances up to 6.5 Å (Fig. S3A). Based on the cross-β X-ray pattern of polyQ (Fig. 5A-B and refs. (30, 31)) the backbones of β-sheets are separated by 8.4 Å. Then, the observed backbone-to-backbone interactions cannot be contacts between β-sheets, but must reflect interactions within a β-sheet. More precisely, they must occur between neighboring β-strands that are 4.7 Å apart (Fig. 5). Given that these are isotopically diluted fibrils, these neighboring 'a' and 'b' strands must be part of the same protein monomer (Fig. S3A), forming an intramolecular β-hairpin (Fig. 5C).

The similar cross-peak patterns in ^{13}C-^{13}C spectra of fully- and mixed-labeled fibrils shows that the interactions within and between proteins must be similar. To selectively probe for the latter category of contacts we looked for the dominant ^{13}C-^{15}N contacts in the ^{13}C/^{15}N mixed fibrils, using a NHHC experiment (Fig. S4A, (32)). We found much stronger peaks for the side chain nitrogens than for the backbones, which indicates that backbone-backbone contacts occur more commonly *within* proteins, whilst side chains mediate extensive interactions *between* proteins.

Interdigitation of the extended and rigid Gln side chains.

Among the dominant side chain N^{ε} contacts in the NHHC spectrum, we see close contact between one protein's side chain N^{ε} and the backbone C^{α} of another protein (Fig. S4A). The ^{13}C-^{13}C spectra also show strong peaks between the end of side chains (C^{s}) and backbone carbons C^{α} and C', in particular in the absence of isotopic dilution (Fig. 4; Fig. S3D). These data fit well to an interdigitated steric-zipper interface that places the side chain termini of one β-sheet close to the backbone of a neighboring β-sheet (18, 29). We tested for other characteristic features of steric-zippers. First, we adapted a "HCCH" experiment previously used to measure a molecular torsion angle in rhodopsin's retinal (33) to measure side chain torsion angles in polyQ amyloid with one or two U-^{13}C,^{15}N Gln. We obtained χ_2-sensitive HCCH data that were indistinguishable for the two β-strand types (Fig. 3E; Fig. S5H) and show that $\chi_2 = 180 \pm 15°$ in both β-strands. This is inconsistent with bent side chains (Fig. S6A; (34)) but fits a steric zipper (Fig. 5D). We

also obtained χ_1-dependent HCCH data, which were found to significantly differ between the 'a' and 'b' Gln (Fig. 3F; Fig. S5G), showing that 'a' and 'b' β-strands differ in their side chain structure. The best fitting χ_1 angles are -65° for conformer 'a' and 55° for conformer 'b', although reduced χ^2 analysis cannot exclude other rotamers (Fig. S5). We probed polyQ side chain motion via ^{15}N R_1, ^{15}N $R_{1\rho}$, ^{13}C R_1, and ^{13}C $R_{1\rho}$ relaxation measurements, in part enabled by ultrafast MAS (35, 36) (Fig. 3G-I; Fig. S5I). This showed the side chains to have a rigidity characteristic of their confinement in a steric zipper amyloid core (18, 29). Thus, the polyQ backbone in the amyloid core relaxes as slowly as the very rigid backbone of the crystalline globular protein GB1 (e.g. Fig. 3H,I). Remarkably, compared with Gln side chains in GB1 crystals, the side chains in the polyQ amyloid core exhibit relaxation similar to that of the backbone (Fig. 3H,I), suggesting a lack of interaction with the solvent consistent with a "dry" interface. Altogether, the ssNMR data all point to an interdigitated steric-zipper-like structure, as illustrated in Fig. 5D.

Discussion

The htt exon1 polyQ amyloid core features β-sheets interacting via steric zippers.

Using MAS ssNMR we examined the structure of the polyQ amyloid core of htt exon1 fibrils and other polyQ aggregates. The exon1 amyloid core features equal populations of two specific types of Gln signals that are identically reproduced in amyloid fibrils formed by shorter N-terminal htt fragments and polyQ model peptides (Fig. S1). MAS ssNMR torsion angle measurements showed that these doubled signals are not from α-sheets, but reflect two β-strand types with different backbone and side chain χ_1 torsion angles. They do share an identical χ_2 angle (of 180°), having extended side chains. On the basis of polyQ's cross-β parameters and various complementary ssNMR measurements, we found these side chains to interdigitate to form a steric-zipper interface between β-sheets.

The antiparallel polyQ β-sheets contain intramolecular β-hairpins.

The ^{13}C-^{13}C spectra showed that the 'a' and 'b'-type β-strands are in intimate contact, even when mixed with a majority of protein without ^{13}C labels. Thus, the contacts represent interactions *within a single* protein, rather than *between different* proteins. These interactions include intimate (<6.5 Å) interactions between the backbones of the two β-strands, which cannot be between β-sheets (which are farther apart; Fig. 5A-B). Thus, these intimate backbone-to-backbone distances must occur between neighboring polyQ β-strands in the same β-sheet. Given the lack of long loops or turns (see below), the only way to see such contacts within a single polypeptide chain is in the form of β-hairpins. To the best of our knowledge this is the first direct evidence that β-hairpins are a prominent building block of the amyloid core of htt exon1 fibrils.

A β-hairpin-based structure implies that the fibrils must contain Gln in β-turns in addition to those forming β-strands. As noted above, the ssNMR signals of the exon1 fibril core are so strongly dominated by the β-sheet amyloid signals, that it is hard to analyze the ssNMR signals from the turn regions. This is compounded by likely structural heterogeneity in these turns, as predicted by our mechanistic model introduced below and observed in polyQ with widely-studied (11, 12, 37) β-hairpin-stabilizing mutations (see Fig. S7). The 'c'-type Gln are the best candidates for the turn structure (16), and they constitute at most ~10% of the total Gln signal. This is qualitatively similar for polyQ amyloid lacking htt's flanking domains (17). If

approximately 90% of a 44-residue exon1 polyQ domain forms β-strands, and considering that β-turns usually span four or more residues (38), then the implication is that perhaps just a single turn region occurs per protein. Kinetic studies indicate that polyQ segments down to 26 residues in length aggregate via a β-hairpin-based monomeric nucleus, suggesting a minimal β-strand length of ~11 residues (11, 39). Although 20-residue β-strands might seem unusually long, they fit easily within the exon1 fibril width (~15 nm; Fig. 1F).

The polyQ amyloid core building block.

The obtained backbone torsion angles allow the construction of two β-strands with slightly different backbone conformations that are able to align with each other and form hydrogen bonds in an antiparallel fashion. Since residues *within each β-strand* have identical chemical shifts, it is most likely that all residues within each β-strand have the same torsion angles (in both backbone and side chain). The side chain dihedral angles, relaxation and ^{13}C-^{13}C correlation constraints were then used to construct the model shown in Fig. 5D. It features two different, but structurally compatible β-strands. The model uses the best-fit χ_1 angles, which (as discussed above) are not unique solutions in the absence of other constraints. The obtained extended side chains are able to form an interdigitated interface between β-sheets consistent with the observed ^{13}C-^{13}C and ^{13}C-^{15}N contacts, as well as the (relatively short) β-sheet repeat distance of polyQ (Fig. 5A). The residues form a Gln "ladder" within each β-sheet, which was assumed to set the (as yet unconstrained) χ_3 angle to allow hydrogen bonding between stacked side chains.

Chemical shift signature of the polyQ amyloid structure.

The ssNMR signals of polyQ amyloid are highly unusual and seemingly unique (11). A good structural model should rationalize this ssNMR signature. The most striking feature is that C^{β} and C^{γ} carbons have highly unusual chemical shifts (11). Gln in a few globular proteins reproduce some of the shifts (17), but these residues are typically surface exposed, dynamic, and sample widely varying side chain conformations. We instead submitted different polyQ models to ab initio ^{13}C chemical shift calculations. The absence of side chain motion (Fig. 3G-I) and lack of aromatic residues render the polyQ amyloid core particularly amenable to this kind of analysis, at least for non-hydrogen-bonding carbons. We calculated ^{13}C chemical shifts for reference compounds and different polyQ models (18, 31, 34, 40, 41) (Tables S3-S4). The latter fail to reproduce the experimental ssNMR results (Fig. S6), with the exception of our ssNMR-derived model (Fig. 5D). It predicts identical shifts throughout each β-strand, atypically small chemical shift differences between C^{β} and C^{γ} in both β-strands, and C^{β}/C^{γ} shifts for strand 'a' that are several ppm higher than those of strand 'b'. Thus, our model reproduces the shift patterns and rationalizes the presence of two distinct ssNMR signals at equal intensities.

PolyQ's intrinsic peak doubling explained by a stochastic assembly mechanism.

In spite of the agreement between this model and our data, the model does not explain an intriguing feature of the polyQ signature: the peak doubling is also seen when just a *single residue* is labeled (Fig. 6A, Fig. S1) (11, 15, 16). Thus, in half of the proteins in the sample this Gln residue is present in an 'a'-type β-strand, and in the other half it is part of a 'b'-type β-strand. We propose that this is a universal feature of polyQ amyloids, and reflects an aggregation mechanism that is stochastic during nucleation, fibril elongation or both. The β-hairpin-based fibril structure implies that elongation must involve β-hairpin formation. During elongation incoming proteins add to the exposed β-strands of the fibril ends. Fig. 6B schematically shows

an exposed ‘b’-type β-strand. A section of the incoming polyQ domain then must form the other β-strand-type (i.e. the ‘a’-type) as it binds, since same-to-same interactions are not allowed due to constraints on pairing of hydrogen bonding in the backbone and side chains (Fig. 5E). Although this step may thus appear deterministic, stochastic assembly arises from the degeneracy of the polyQ sequence: different sections of the incoming polyQ domain are equally capable of being the initial point of interaction. This is schematically visualized in Fig. 6B for a polyQ domain forming a single β-hairpin. A single labeled Gln near the N-terminus randomly ends up in either of the two β-strand configurations, leading to that one residue showing both peaks at a 1:1 ratio, as observed experimentally.

We hypothesize that even the formation of the initial elongation-capable structure could be stochastic. Mechanistic studies suggest β-hairpin formation to be critical in the nucleation process for long polyQ (11, 39). We propose that this could involve formation of not one specific β-hairpin structure, but rather the stochastic formation of structurally related, but nonetheless different β-hairpins. They would be structurally related by always ending up with the complementary ‘a’ and ‘b’ β-strands, while being different in the ways that those strands are arranged. An N-terminal type-‘a’ β-strand can be combined with a more C-terminal type-‘b’ strand, or vice versa (Fig. 6C left/right), with likely minimal energetic or kinetic differences.

Thus, we propose mutant htt exon1 and other polyQ peptides to follow a stochastic assembly mechanism that is a universal feature of polyQ, independent of aggregation kinetics and sequence context (Fig. S1). A prior study (17) reported that $D_2Q_{15}K_2$ fibrils showed the same two signals, but argued that individual Gln were present in either one or the other conformer, but not in both. We were intrigued by the potential implication that single-residue peak doubling would depend on the length of the polyQ segment, and thus might point toward a structural or mechanistic rationale for the polyQ threshold phenomenon. When we measured $D_2Q_{15}K_2$ fibrils with a single labeled Gln, however, we observed the doubled peak pattern (Fig. S1F), showing no evidence of a change in the stochastic assembly process as a function of polyQ length.

Fibril polymorphism in polyQ amyloid

Other work has argued for polymorphism in the structure of aggregated polyQ, with potential implications for aggregate toxicity (19). The ability for a single sequence to form polymorphic fibrils is common for amyloid-forming proteins, and ssNMR chemical shifts are the gold standard for detecting the underlying structural differences (14). However, polyQ fibrils always feature the same pattern of chemical shifts (Fig. S1), reflecting the described combination of two β-strand configurations. One might expect that a purely Gln-based sequence could fulfill the fundamental architecture of an interdigitating β-hairpin-based assembly in different ways, but this appears not to be the case. We hypothesize that the origins for this must be in the nucleation process that initiates the amyloid formation. This event dictates the structure of the initial assembly that is faithfully extended and reproduced during elongation (Fig. 6). As examined in a recent molecular dynamics study, different β-strand-based polyQ structures have distinct propensities for initiating the aggregation process (42). We propose therefore that the particular structure that we observe in the polyQ amyloid core would be uniquely capable of nucleated elongation. It is not immediately obvious from the current data what makes this conformation so unique, but we hope that the structural insights enabled by ssNMR will facilitate computational and experimental explorations of this issue.

Conclusion.

We have shown by ssNMR that the amyloid core of htt exon1 fibrils and other polyQ aggregates feature β-sheets that interact via interdigitation of side chains. We described how specific β-hairpin structures are present in the htt exon1 fibrils, as well as a stochastic aggregation mechanism of expanded polyQ. These new insights greatly enhance our understanding of polyQ misfolding and aggregation, and provide support for mechanistic studies that have pinpointed β-hairpin formation to be a pivotal event in the aggregation process.

Materials and Methods

Fibril sample preparation

Fibrillar htt exon1 with a 44-residue polyQ domain (Fig. 1A; Table S1) was prepared following protocols similar to those reported (16), with several modifications described in the Supporting Materials. Site-specifically labeled peptide fibrils (Table S1) were prepared according to previously reported protocols (11, 43), with more details provided in the Supporting Materials.

Transmission electron microscopy

TEM was performed on mature U-^{13}C,^{15}N htt exon1 fibrils and on htt exon1 samples harvested during aggregation (see Extended Methods online). Samples were negatively stained with 1% (w/v) uranyl acetate. Imaging at 6,500-15,000-fold magnification was done using a Technai T12spirit transmission electron microscope (FEI; Hillsboro, OR) operating at 120 kV and equipped with an UltraScan 1000 CCD camera (Gatan; Pleasanton, CA).

MAS ssNMR spectroscopy

MAS ssNMR experiments were performed using Bruker spectrometers operating at 600 and 800 MHz ^{1}H Larmor frequencies, using 1.3 and 3.2 mm MAS NMR probes. Assignments were performed using 2D ^{13}C-^{13}C and ^{15}N-^{13}C assignment measurements described in the Supporting Materials. Distance constraints were obtained using proton-driven spin diffusion (PDSD) and NHHC experiments (32). Backbone and side chain torsion angles were measured using "NCCN" (23, 26) and "HCCH"-style dipolar recoupling measurements (33) (Figs. S2, S5). ^{15}N $R_{1\rho}$, ^{13}C $R_{1\rho}$ and ^{13}C R_1 relaxation rates were measured at 60 kHz MAS, and ^{15}N R_1 rates were measured at 20 kHz MAS (29, 35, 36). Further details can be found in the Supporting Materials.

Chemical shift calculations

Ab initio calculations were carried out on different Gln-containing candidate structures, using the PQS program (44). Density functional GIAO shielding calculations were performed with the B3LYP functional and the Ahlrichs TZP basis set (45). Absolute shieldings were corrected as described (46). Reference calculations were carried out on an amyloidogenic peptide of known structure (18, 40). Additional details are provided in the Supporting Materials.

Acknowledgments

We thank Mike Delk and Jochem Struppe for technical assistance, and acknowledge funding from the University of Pittsburgh and the National Institutes of Health (grants R01 GM112678 and AG019322 to P.v.d.W. and R.W., R01 GM099718 to R.W., T32 GM088119 to C. H.), BBSRC (grant BB/L022761/1) and EPSRC (grant EP/L025906/1) to J.R.L and EPSRC to J.M.L., and grant UL1 RR024153 from the National Center for Research Resources (NCRR). Molecular graphics were prepared with UCSF Chimera, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIGMS P41-GM103311).

References

1. Bates GP & Benn C (2002) The polyglutamine diseases. *Huntington's Disease*, eds Bates GP, Harper P, & Jones L (Oxford), 3rd Ed Vol 45, pp 429-472.
2. Mangiarini L*, et al.* (1996) Exon 1 of the HD gene with an expanded CAG repeat is sufficient to cause a progressive neurological phenotype in transgenic mice. *Cell* 87(3):493-506.
3. DiFiglia M*, et al.* (1997) Aggregation of huntingtin in neuronal intranuclear inclusions and dystrophic neurites in brain. *Science* 277(5334):1990-1993.
4. Scherzinger E*, et al.* (1997) Huntingtin-encoded polyglutamine expansions form amyloid-like protein aggregates in vitro and in vivo. *Cell* 90(3):549-558.
5. Poirier MA*, et al.* (2002) Huntingtin spheroids and protofibrils as precursors in polyglutamine fibrilization. *J. Biol. Chem.* 277(43):41032-41037.
6. Legleiter J*, et al.* (2009) Monoclonal antibodies recognize distinct conformational epitopes formed by polyglutamine in a mutant huntingtin fragment. *J. Biol. Chem.* 284(32):21647-21658.
7. Arrasate M, Mitra S, Schweitzer ES, Segal MR, & Finkbeiner S (2004) Inclusion body formation reduces levels of mutant huntingtin and the risk of neuronal death. *Nature* 431(7010):805-810.
8. Duim WC, Jiang Y, Shen K, Frydman J, & Moerner WE (2014) Super-resolution fluorescence of huntingtin reveals growth of globular species into short fibers and coexistence of distinct aggregates. *ACS Chem. Biol.* 9(12):2767-2778.
9. Liu K-Y*, et al.* (2015) Disruption of the nuclear membrane by perinuclear inclusions of mutant huntingtin causes cell-cycle re-entry and striatal cell death in mouse and cell models of Huntington's disease. *Hum. Mol. Genet.* 24(6):1602-1616.
10. Wetzel R & Mishra R (2014) Order, disorder, and conformational flux in the structural biology of Huntington's Disease. *Huntington's Disease*, eds Bates GP, Tabrizi S, & Jones L (Oxford University Press), 4th Ed.
11. Kar K*, et al.* (2013) β-hairpin-mediated nucleation of polyglutamine amyloid formation. *J. Mol. Biol.* 425(7):1183-1197.
12. Zhang QC*, et al.* (2011) A compact beta model of huntingtin toxicity. *J. Biol. Chem.* 286(10):8188-8196.
13. Comellas G & Rienstra CM (2013) Protein structure determination by magic-angle spinning solid-state NMR, and insights into the formation, structure, and stability of amyloid fibrils. *Annu. Rev. Biophys.* 42(1):515-536.
14. Tycko R (2014) Physical and Structural Basis for Polymorphism in Amyloid Fibrils. *Protein Sci* 23(11):1528-1539.
15. Sivanandam VN*, et al.* (2011) The aggregation-enhancing huntingtin N-terminus is helical in amyloid fibrils. *J. Am. Chem. Soc.* 133(12):4558-4566.
16. Hoop CL*, et al.* (2014) Polyglutamine amyloid core boundaries and flanking domain dynamics in huntingtin fragment fibrils determined by solid-state nuclear magnetic resonance. *Biochemistry* 53(42):6653-6666.
17. Schneider R*, et al.* (2011) Structural characterization of polyglutamine fibrils by solid-state NMR spectroscopy. *J. Mol. Biol.* 412(1):121-136.
18. Sawaya MR*, et al.* (2007) Atomic structures of amyloid cross-beta spines reveal varied steric zippers. *Nature* 447(7143):453-457.
19. Nekooki-Machida Y*, et al.* (2009) Distinct conformations of in vitro and in vivo amyloids of huntingtin-exon1 show different cytotoxicity. *Proc. Natl. Acad. Sci. U.S.A.* 106(24):9679-9684.
20. Armen RS, Bernard BM, Day R, Alonso DOV, & Daggett V (2005) Characterization of a possible amyloidogenic precursor in glutamine-repeat neurodegenerative diseases. *Proc. Natl. Acad. Sci. U.S.A.* 102(38):13433-13438.

21. Hayward S & Milner-White EJ (2011) Simulation of the β- to α-sheet transition results in a twisted sheet for antiparallel and an α-nanotube for parallel strands: Implications for amyloid formation. *Proteins* 79(11):3193-3207.
22. Hong M (2006) Torsion Angle Determination by Solid-State NMR. *Modern Magnetic Resonance*, ed Webb GA (Springer), pp 723-729.
23. Costa PR, Gross JD, Hong M, & Griffin RG (1997) Solid-state NMR measurement of Psi in peptides: a NCCN 2Q-heteronuclear local field experiment. *Chem. Phys. Lett.* 280(1-2):95-103.
24. Shen Y, Delaglio F, Cornilescu G, & Bax A (2009) TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR* 44(4):213-223.
25. Jaroniec CP*, et al.* (2004) High-resolution molecular structure of a peptide in an amyloid fibril determined by magic angle spinning NMR spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* 101(3):711-716.
26. van der Wel PCA, Lewandowski JR, & Griffin RG (2010) Structural characterization of GNNQQNY amyloid fibrils by magic angle spinning NMR. *Biochemistry* 49(44):9457-9469.
27. De Paëpe G, Lewandowski JR, Loquet A, Böckmann A, & Griffin RG (2008) Proton assisted recoupling and protein structure determination. *J. Chem. Phys.* 129(24):245101.
28. Li J & van der Wel PCA (2013) Spinning-rate encoded chemical shift correlations from rotational resonance solid-state NMR experiments. *J. Magn. Reson.* 230:117-124.
29. Lewandowski JR, van der Wel PCA, Rigney M, Grigorieff N, & Griffin RG (2011) Structural complexity of a composite amyloid fibril. *J. Am. Chem. Soc.* 133(37):14686-14698.
30. Perutz MF, Pope BJ, Owen D, Wanker EE, & Scherzinger E (2002) Aggregation of proteins with expanded glutamine and alanine repeats of the glutamine-rich and asparagine-rich domains of Sup35 and of the amyloid beta-peptide of amyloid plaques. *Proc. Natl. Acad. Sci. U.S.A.* 99(8):5596-5600.
31. Sikorski P & Atkins E (2005) New model for crystalline polyglutamine assemblies and their connection with amyloid fibrils. *Biomacromolecules* 6(1):425-432.
32. Lange A, Luca S, & Baldus M (2002) Structural constraints from proton-mediated rare-spin correlation spectroscopy in rotating solids. *J. Am. Chem. Soc.* 124(33):9704-9705.
33. Feng X*, et al.* (1997) Direct determination of a molecular torsional angle in the membrane protein rhodopsin by solid-state NMR. *J. Am. Chem. Soc.* 119(29):6853-6857.
34. Sharma D, Shinchuk LM, Inouye H, Wetzel R, & Kirschner DA (2005) Polyglutamine homopolymers having 8-45 residues form slablike beta-crystallite assemblies. *Proteins* 61(2):398-411.
35. Lewandowski JR*, et al.* (2010) Measurement of site-specific ^{13}C spin−lattice relaxation in a crystalline protein. *J. Am. Chem. Soc.* 132(24):8252-8254.
36. Lewandowski JR, Sass HJ, Grzesiek S, Blackledge M, & Emsley L (2011) Site-specific measurement of slow motions in proteins. *J. Am. Chem. Soc.* 133(42):16762-16765.
37. Thakur AK & Wetzel R (2002) Mutational analysis of the structural organization of polyglutamine aggregates. *Proc. Natl. Acad. Sci. U.S.A.* 99(26):17014-17019.
38. Chou PY & Fasman GD (1977) β-turns in proteins. *J. Mol. Biol.* 115(2):135-175.
39. Kar K, Jayaraman M, Sahoo B, Kodali R, & Wetzel R (2011) Critical nucleus size for disease-related polyglutamine aggregation is repeat-length dependent. *Nat. Struct. Mol. Biol.* 18(3):328-336.
40. van der Wel PCA, Lewandowski JR, & Griffin RG (2007) Solid-state NMR study of amyloid nanocrystals and fibrils formed by the peptide GNNQQNY from yeast prion protein Sup35p. *J. Am. Chem. Soc.* 129(16):5117-5130.
41. Perutz MF, Finch JT, Berriman J, & Lesk A (2002) Amyloid fibers are water-filled nanotubes. *Proc. Natl. Acad. Sci. U.S.A.* 99(8):5591-5595.
42. Miettinen MS, Knecht V, Monticelli L, & Ignatova Z (2012) Assessing Polyglutamine Conformation in the Nucleating Event by Molecular Dynamics Simulations. *J. Phys. Chem. B* 116(34):10259-10265.

43. O'Nuallain B*, et al.* (2006) Kinetics and thermodynamics of amyloid assembly using a high-performance liquid chromatography-based sedimentation assay. *Methods Enzym.* 413:34-74.
44. Baker J*, et al.* (2009) Quantum chemistry in parallel with PQS. *J. Comp. Chem.* 30(2):317-335.
45. Schäfer A, Horn H, & Ahlrichs R (1992) Fully optimized contracted Gaussian basis sets for atoms Li to Kr. *J. Chem. Phys.* 97(4):2571-2577.
46. Magyarfalvi G & Pulay P (2003) Assessment of density functional methods for nuclear magnetic resonance shielding calculations. *J. Chem. Phys.* 119(3):1350-1357.
47. Marley J, Lu M, & Bracken C (2001) A method for efficient isotopic labeling of recombinant proteins. *J. Biomol. NMR* 20(1):71-75.
48. Harris RK*, et al.* (2008) Further conventions for NMR shielding and chemical shifts (IUPAC Recommendations 2008). *Magn. Reson. Chem.* 46(6):582-598.
49. Takegoshi K, Nakamura S, & Terao T (2001) ^{13}C-^{1}H dipolar-assisted rotational resonance in magic-angle spinning NMR. *Chem. Phys. Lett.* 344(5-6):631-637.
50. Bennett AE, Rienstra CM, Auger M, Lakshmi KV, & Griffin RG (1995) Heteronuclear decoupling in rotating solids. *J. Chem. Phys.* 103(16):6951-6958.
51. Hohwy M, Rienstra CM, & Griffin RG (2002) Band-selective homonuclear dipolar recoupling in rotating solids. *J. Chem. Phys.* 117:4973-4987.
52. Lee M & Goldburg WI (1965) Nuclear-magnetic-resonance line narrowing by a rotating rf field. *Phys. Rev.* 140(4A):A1261.
53. Yao XL & Hong M (2004) Structure Distribution in an Elastin-Mimetic Peptide (VPGVG)3 Investigated by Solid-State NMR. *J. Am. Chem. Soc.* 126(13):4199-4210.

Figure Legends

Figure 1. *Huntingtin exon1 construct design and fibril formation. (A) Sequence of the employed MBP-htt exon1 fusion protein. The exon1 sequence, factor Xa cleavage site and location of the fibrils' rigid amyloid core (16) are indicated. (B-E) Negatively stained TEM as a function of time after factor Xa release of unlabeled exon1. (B) Uncleaved htt exon1 MBP fusion protein. (C) Oligomers observed 1 h after cleavage. (D) By 3 hours, fibrils have begun to form. (E) After 25 hours, fibrils have grown and oligomers are no longer visible on the grid. (F) Mature [U-*^{13}*C,*^{15}*N] labeled fibrils prepared for ssNMR. Scale bars are 200 nm.*

Figure 2. *MAS NMR on the polyQ core of mature huntingtin exon1 fibrils prepared at room temperature. (A) 2D ^{13}C-^{13}C spectrum shows the two sets of Gln peaks (type 'a' and 'b') that account for the rigid amyloid core (red and blue lines). (B) An intra-residue NCACX spectrum connects ^{13}C signals to their own backbone ^{15}N. (C) An inter-residue NCOCX spectrum connects the ^{13}C signals to the ^{15}N backbone shift of the next Gln. (D) Overlay of B and C. (E) The identical NCACX and NCOCX spectra show that connected Gln always have the exact same chemical shifts. No direct backbone connections between 'a' and 'b' are observed.*

Figure 3. *PolyQ amyloid structure and dynamics by MAS NMR. (A,B) Intersections (circled) of experimental NCCN data (horizontal lines) with the theoretical dependence on ψ (black line). Shaded areas indicate the standard error. (C) Consensus of experimental backbone angles for polyQ amyloid, based on chemical shift analysis (diamonds) and ψ-angle measurements (colored lines & shaded areas) for conformers 'a' (red) and 'b' (blue). (D) Simulated HCCH curves for distinct χ_2 side chain angles. (E) Experimental χ_2-sensitive HCCH data (C^{β}/C^{γ}) for conformers 'a' (red diamonds) and 'b' (blue triangles), along with the theoretical $\chi_2 = 180°$ curve (solid line). (F) χ_1-sensitive HCCH data, with simulated curves for $\chi_1 = -65°$ and 55° (lines). (G) $R_{1\rho}$ and R_1 ^{13}C relaxation for backbone and side chains of a Gln in polyQ amyloid, measured at 60 kHz MAS. (H,I) ^{15}N R_1 and $R_{1\rho}$ values for Gln in polyQ amyloid (right) and in GB1 protein crystals (left), showing a striking difference for the side chain N^{ε}.*

Figure 4. *Intra- and inter-molecular β-strand-β-strand interactions within the htt exon1 core. (A) Extended (250 ms) ^{13}C-^{13}C mixing 2D spectrum on exon1 fibrils with complete ^{13}C labeling. (B) Enlargement showing strong cross-peaks between the 'a' and 'b' Gln conformers. (C) Analogous 2D spectrum on diluted (26 %) ^{13}C labeled fibrils (see Methods), which also shows significant cross-peaks between the 'a' and 'b' signals. (D-F) 1D slices extracted from the 2D spectra for polarization transferred from C^{α}, C' and C^{δ} carbons of each Gln conformer (circled labels). Peaks due to transfer to the other β-strand type are marked (color-coded labels). These spectra were obtained at 800 MHz (^{1}H) and with 13 kHz MAS. (G) Normalized volumes for cross peaks between the 'a' and 'b' strands in the mixed fibrils (panel (C)).*

Figure 5. *PolyQ amyloid structure. (A) X-ray powder diffraction on $K_2Q_{31}K_2$ fibrils shows the cross-β dimensions of polyQ amyloid. (B) The cross-β dimensions reflect repeat distances between β-strands (4.7 Å) and between β-sheets (8.4 Å). (C) The intra-protein ^{13}C-^{13}C contacts between 'a' and 'b' Gln backbones are too short to occur between sheets, and are therefore between β-strands (color-coded by type) within a β-sheet. (D) β-strand structures that fulfill the torsion angle constraints, close proximity of side chain C^δ to the backbone, and allow hydrogen bonding of both backbones and side chains. Extended side chains form a steric zipper interface to allow the 8.4 Å sheet-to-sheet distance. (E) The 'a' and 'b' strands are mutually compatible, but 'a'-'a' or 'b'-'b' interactions are not possible.*

Figure 6. *Stochastic polyQ β-sheet assembly mechanism. (A) Peak doubling is seen for a single labeled Gln in the amyloid core of a fibrillar htt N-terminal fragment (15). The specific Gln (Q19) is distributed between both β-strand types, shown with their schematic ssNMR spectra. (B) Schematic fibril core containing β-sheets with 2n β-strands (top). Elongation maintains the alternating β-strand pattern, but can be initiated by different segments of the polyQ domain, which then causes the N-terminal 'labeled' Gln (circles) to end up in both β-strand types. (C) During nucleation a 'a'-'b' β-strand assembly can be formed in two ways, yielding related but distinct β-hairpin structures. The circles mark a single "labeled" Gln nearer the N-terminus.*

Figure 1

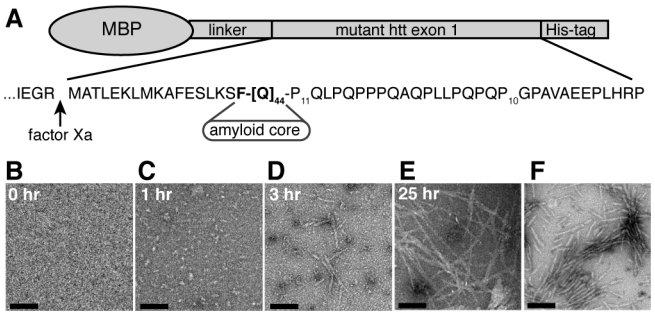


Figure 2

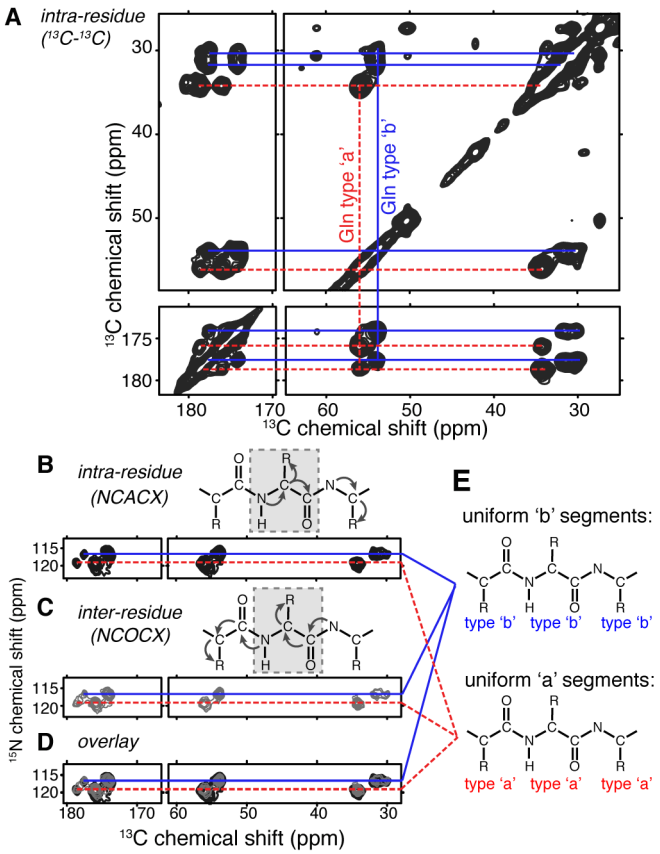


Figure 3

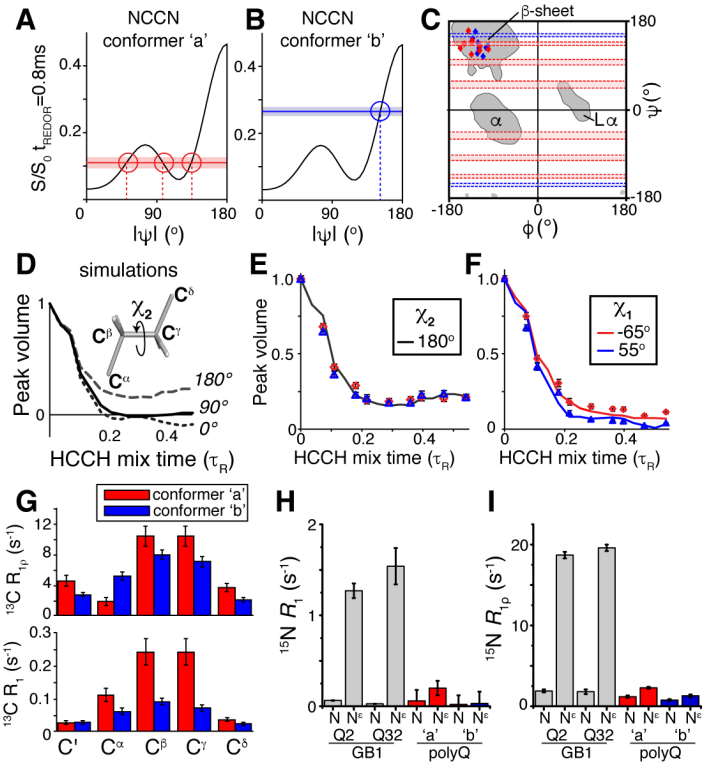

A
NCCN
conformer 'a'
B
NCCN
conformer 'b'
C
β-sheet
α
Lα
ψ (°)
φ (°)
S/S0 tREDOR=0.8ms
|ψ| (°)
D
simulations
Peak volume
HCCH mix time (τR)
180°
90°
0°
E
χ2
180°
F
χ1
-65°
55°
G
conformer 'a'
conformer 'b'
13C R1ρ (s-1)
13C R1 (s-1)
C'
Cα
Cβ
Cγ
Cδ
H
15N R1 (s-1)
I
15N R1ρ (s-1)
N
Nε
Q2
Q32
'a'
'b'
GB1
polyQ

Figure 4

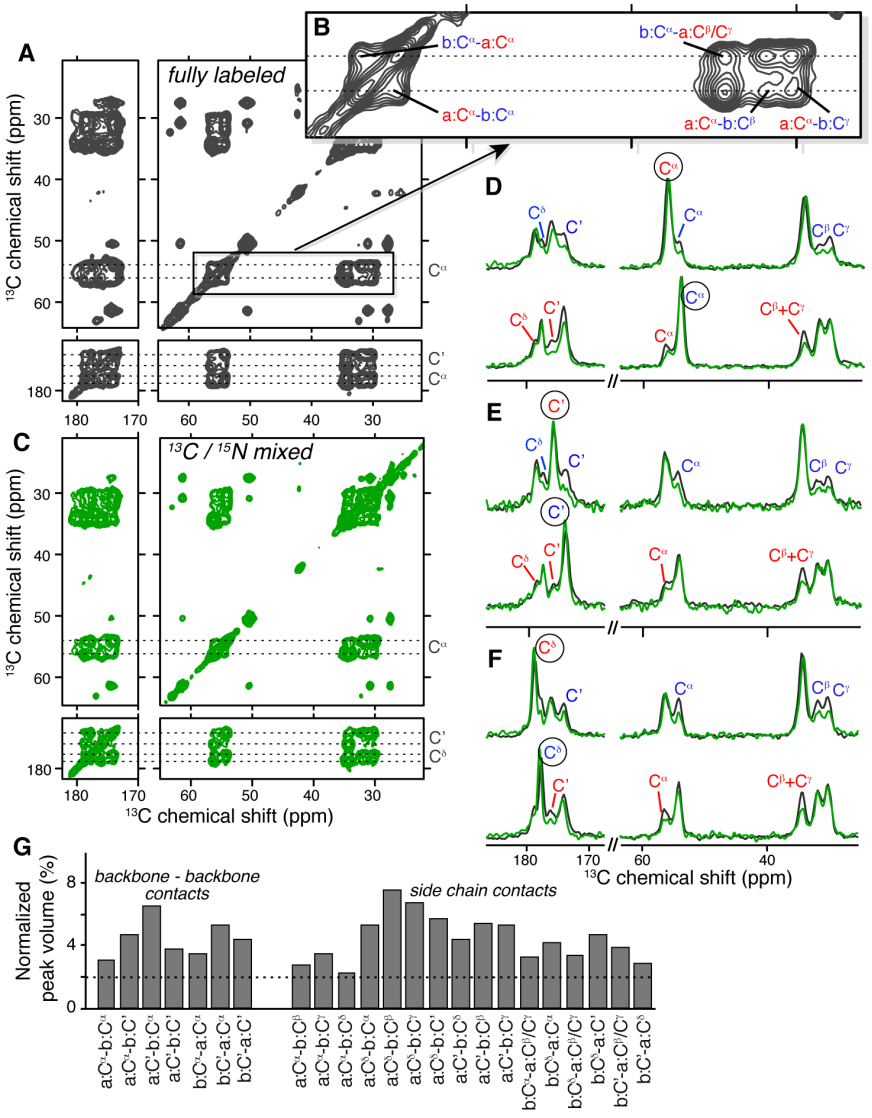

A
fully labeled
B
b:Cα-a:Cα
a:Cα-b:Cα
b:Cα-a:Cβ/Cγ
a:Cα-b:Cβ
a:Cα-b:Cγ
C
13C / 15N mixed
13C chemical shift (ppm)
D
E
F
G
Normalized peak volume (%)
backbone - backbone contacts
side chain contacts

Figure 5:

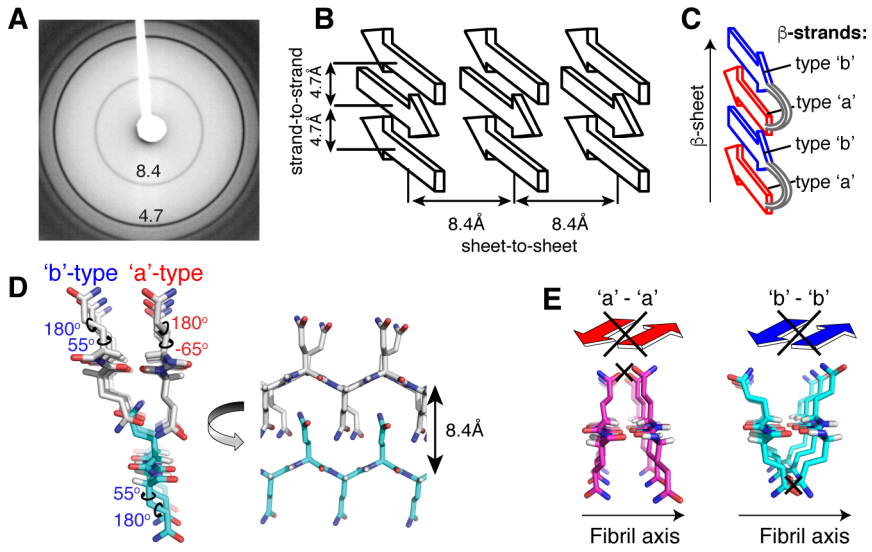
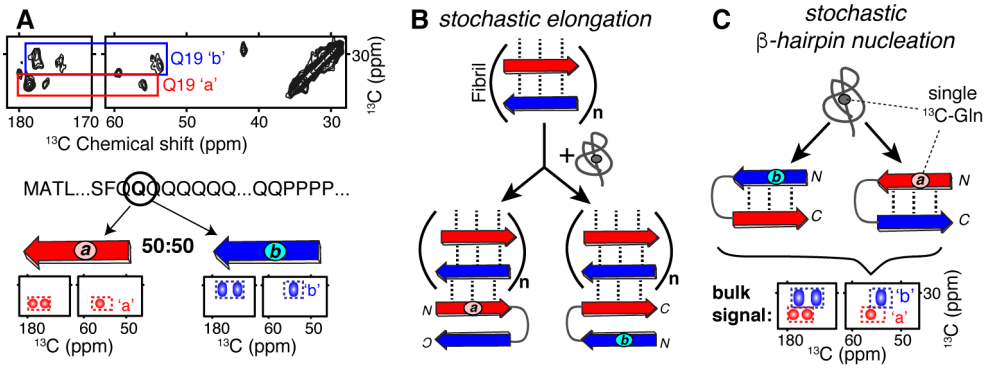


Figure 6:



Supporting Materials

for

Huntingtin exon 1 fibrils feature an interdigitated β-hairpin-based polyglutamine core.

Cody L. Hoop[a], Hsiang-Kai Lin[a], Karunakar Kar[a,b], Gábor Magyarfalvi[c], Jonathan M. Lamley[d], Jennifer C. Boatz[a], Abhishek Mandal[a], Józef R. Lewandowski[d], Ronald Wetzel[a,b], Patrick C.A. van der Wel[a]

Author affiliations:

[a] Department of Structural Biology, University of Pittsburgh School of Medicine, Biomedical Science Tower 3, 3501 Fifth Ave, Pittsburgh, PA 15260, USA.

[b] Pittsburgh Institute for Neurodegenerative Diseases, University of Pittsburgh School of Medicine, Biomedical Sciences Tower 3, 3501 Fifth Avenue, Pittsburgh, PA 15260, USA

[c] Laboratory of Molecular Spectroscopy, Institute of Chemistry, Eötvös University, PO Box 32, H-1518, Budapest 112, Hungary.

[d] Department of Chemistry, University of Warwick, Coventry CV4 7AL, United Kingdom.

* Correspondence to: vanderwel@pitt.edu.

Extended Methods

Preparation of htt exon1 fibrils

Htt exon1 with a 44-residue polyQ domain was expressed in *Escherichia coli* as a fusion protein featuring an N-terminal maltose binding protein (MBP) solubility tag (Fig. 1a), following similar protocols as described previously (16). A codon optimized version of the htt exon1 gene was synthesized by Genscript Inc. (Piscataway, NJ), and sub-cloned into a pMAL-c2x plasmid using the EcoRI and HindIII restriction sites. Labeled fusion protein was obtained by overexpression in M9 minimal media supplemented with U-^{13}C-D-glucose and/or ^{15}N-ammonium chloride (Cambridge Isotope Laboratories, Tewksbury, MA), following published protocols (16, 47). The intact fusion protein was purified as described (16). The MBP solubility tag was removed with factor Xa protease (5, 16) (Promega, Madison, WI), such that it generates the exact N-terminal sequence specified by mammalian mRNA (Fig. 1A; (16)). Upon release, the htt exon1 undergoes aggregation at room temperature. Protein aggregation for the preparation of MAS NMR samples was induced by factor Xa cleavage of either only the U-^{13}C,^{15}N-labeled protein, or mixtures of U-^{13}C-labeled and U-^{15}N-labeled protein. The obtained fibrils were washed with deionized water or PBS buffer prior to experimental characterization. For the time-dependent TEM studies, 0.55 μg of factor Xa was added to 10 μL (7 μg htt exon 1) aliquots of the unlabeled fusion protein. PBS buffer was added in place of factor Xa for the 0 time point. After aggregating undisturbed for a set amount of time, samples were diluted 2 x into PBS buffer. Samples (5 μL) were applied to freshly glow-discharged carbon coated 400 mesh size copper grids, adsorbed for 30 seconds, and negatively stained with 1% (w/v) uranyl acetate. Excess sample and stain was blotted with filter paper between steps, and grids were allowed to air dry.

Preparation of site-specifically labeled peptide fibrils

Site-specifically labeled polyQ peptides were prepared by solid-phase peptide synthesis (Table S1), using Fmoc- and side-chain-protected ^{13}C, ^{15}N-labeled amino acids from Cambridge Isotope Laboratories (Andover, MA) and Isotec (Sigma-Aldrich, St. Louis, MO). $D_2Q_{15}K_2$ with a single U-^{13}C,^{15}N-labeled Gln in position 6 and an acetylated N-terminus was synthesized by Anaspec (Fremont, CA). Other peptides were synthesized by the Small Scale Synthesis facility of the Keck Biotechnology Resource Laboratory of Yale University. The peptide sequences and labeling patterns are in Table S1. Peptides were obtained crude, purified in-house, and disaggregated prior to fibril formation, and allowed to aggregate in PBS buffer (pH 7.4) at 37 °C (11, 43).

Solid-state NMR spectroscopy

For MAS NMR, the mature fibrils were packed into 3.2 mm or 1.3 mm zirconia MAS rotors (Bruker Biospin, Billerica, MA and CortecNet, Voisins-le-Bretonneux, France) using home-built ultracentrifugal sample packing tools. An ultrafast MAS sample was pelleted in a Beckmann L8-70M ultracentrifuge running at ~100,000 g for 6 h. Other MAS NMR samples were packed in a Beckman Coulter Optima L-100 XP ultracentrifuge with SW-32 Ti rotor at up to 130,000 x g. The samples were kept unfrozen and hydrated at all times, facilitated in part by use of silicone-based spacers in the 1.3 mm samples. The MAS ssNMR experiments were performed with a wide bore Bruker Avance I NMR spectrometer operating at 600 MHz ^{1}H Larmor frequency (14.1 T) and using a 3.2 mm Efree HCN (Bruker Biospin) MAS NMR probe, unless stated otherwise. The sample temperature was controlled using a constant flow of cooled gas. Spectra were processed and analyzed with NMRPipe, Sparky, and CCPNMR/Analysis software.

External referencing to 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) (for ^{13}C) was done indirectly via the ^{13}C signals of adamantane (48). For all spectra shown, additional experimental details can be found in Table S5. Intra-residue ^{13}C assignments were based on 2D ^{13}C-^{13}C experiments obtained with ^{1}H-^{13}C cross polarization (CP) and dipolar assisted rotational resonance (DARR) (49) ^{13}C-^{13}C mixing as indicated. Typically, 83 kHz two pulse phase modulation (TPPM) (50) ^{1}H decoupling was applied during acquisition and evolution, and MAS spinning rates were 8-13 kHz. For several samples the side chain ^{13}C assignments were confirmed using 2D single quantum - double quantum (SQ-DQ) experiments with the ^{13}C-^{13}C DQ coherence generated via 1.2-1.8 ms $SPC5_3$ mixing (51), as indicated. 1D and 2D ^{15}N-^{13}C double CP-based experiments, including NCACX, NCOCX, and CONCX measurements, were used to make intra- and inter-residue assignments of ^{15}N and ^{13}C chemical shifts.

MAS ssNMR distance measurements

Short-range carbon proximities in the htt exon1 fibrils were probed via ^{13}C-^{13}C recoupling experiments using short DARR and proton-driven spin diffusion (PDSD) mixing at 10 and 13 kHz (Table S5). Variable-mixing time 2D PDSD experiments were performed at 13 kHz MAS, on 600 and 800 MHz spectrometers, using both fully U-^{13}C,^{15}N-labeled exon1 fibrils and ^{13}C/^{15}N mixed exon1 fibrils. The PDSD data shown in Fig. 4 were obtained on an isotopically mixed sample in which the ^{13}C-labeled protein constituted 26 mol-% of the total protein, based on absorption and mass spectrometry measurements. For a qualitative analysis of approximate ^{13}C-^{13}C distances we performed corresponding PDSD experiments on isotopically diluted (10 mol-% U–^{13}C,^{15}N-labeled) N-acetyl-Val-Leu crystals, prepared as previously described (28). Peak volumes were measured and normalized as previously described (28): peak fitting was performed in the Sparky program, using 2D Gaussian line shapes. Normalized peak volumes were calculated by dividing the (cross) peak volumes at non-zero mixing times by the peak intensities of the corresponding diagonal peaks in analogous 2D spectra obtained in absence of ^{13}C-^{13}C mixing. Intermolecular NHHC distance measurements (32) were performed at 800 MHz (^{1}H frequency) on fibrils prepared from a mixture of ^{13}C labeled and ^{15}N-labeled htt exon1 (molar ratio 60:40 based on mass spectrometry analysis). ^{1}H-^{15}N and ^{15}N-^{1}H CP contact times were 400 μs, ^{1}H-^{13}C CP contact time was 200 μs, and the ^{1}H-^{1}H mixing time was set to 250 μs, at a 13 kHz MAS rate. Other experimental conditions are listed in Table S5.

MAS ssNMR torsion angle measurements

Chemical shift-based backbone torsion angle analysis (diamonds in Fig. 3C) was performed with the TALOS+ software package (24). Backbone torsion angle experiments were performed via ^{13}C SQ-DQ NCCN experiments (Fig. S2A), as previously described (23, 26). The NCCN experiments were applied to the two sequentially labeled Gln in the polyQ core of [U-^{13}C,^{15}N-Q8,Q9] pG-Q_{22} fibrils. SPC5 recoupling was used to generate the C^{α}-C' DQ coherence, with a 10 kHz MAS rate, 50 kHz ^{13}C rf power, and 100 kHz CW decoupling. The DQ signal was dephased under the ^{15}N-^{13}C dipolar couplings, using variable-time rotational echo double resonance (REDOR) mixing with 50 kHz ^{15}N and ^{13}C rf powers for the respective π pulses. The obtained experimental data were corrected for contributions due to unlabeled residues. Side-chain torsion angle experiments were performed on site-specifically labeled synthetic peptide aggregates, featuring either one or two labeled Gln, as indicated. The employed pulse sequence is shown in Fig. S5A, and is adapted from prior work (33). SPC5 was used to establish DQ coherence between the side chain carbons. Variable-time Lee-Goldburg (LG) irradiation on ^{1}H allowed for concomitant decoupling of ^{1}H-^{1}H interactions, while enabling the ^{1}H-^{13}C dipolar couplings (52).

By gradually incrementing the LG mixing time, χ_1 and χ_2-dependent HCCH dephasing curves were obtained. To extract backbone and side chain torsion angles from the NCCN and HCCH experimental results, numerical simulations of the respective pulse sequences were performed with the SPINEVOLUTION program (www.SpinEvolution.com). These simulations were done on multi-core workstations or a departmental 48-core high-performance linux computer. The coordinate system used for the NCCN simulation was N_i-C^{α}_i-C'_i-N_{i+1}, where the N_i-C^{α}_i distance was 1.33 Å, the C^{α}_i-C'_i distance was 1.52 Å, the C'_i-N_{i+1} distance was 1.46 Å, the angle defined by N_i-C^{α}_i-C'_i was 114.1°, and the angle defined by C^{α}_i-C'_i-N_{i+1} was 117.8°. For a given torsion angle, the signal intensity of the ^{13}C spins was simulated with an increasing REDOR period. Such dephasing curves were simulated for an array of torsion angles between -180° and 180°. The HCCH simulations included five or six spins: H^{α}-C^{α}-C^{β}-H^{β}_2 (for χ_1) or H^{β}_2-C^{β}-C^{γ}-H^{γ}_2 (for χ_2). H-C bond distances were 1.13 Å, the C-C distance was 1.33 Å, and the angles defined H-C-C were 109.5°. For a given torsion angle, the ^{13}C signal intensity was recorded as a function of increasing LG period up to the length of one rotor period. Interference curves were simulated for an array of HCCH torsion angles between -180° to 180°. To fit the data, reduced χ^2 was calculated for each of these curves relative to the experimental data, and a reduced χ^2 test was used to find torsion angles that fit within a 90% confidence interval. The experimental error for each data point was estimated by taking the noise level in the spectra relative to the Gaussian integration of the peak and was considered in the fitting algorithm.

MAS ssNMR relaxation measurements

^{15}N longitudinal R_1 relaxation was measured for backbone and side chain ^{15}N sites using a series of ^{1}H-^{15}N CP experiments incorporating a ^{15}N R_1 relaxation period and 83 kHz TPPM ^{1}H decoupling during acquisition, analogous to earlier work (29). These measurements were performed at 22 kHz MAS and 600 MHz ^{1}H frequency, conditions where systematic MAS-dependent measurements had shown ^{15}N-^{15}N spin diffusion to be effectively suppressed.^{15}N $R_{1\rho}$, ^{13}C $R_{1\rho}$ and ^{13}C R_1 relaxation rates were measured at a MAS frequency of 60 kHz using a Bruker 1.3 mm triple resonance probe in a Bruker Avance II+ spectrometer operating at a ^{1}H Larmor frequency of 600 MHz. Sample cooling was applied using a Bruker BCU-X variable temperature unit such that the internal sample temperature was held at 296 ± 1 K, measured from the chemical shift of water protons with respect to an internal DSS reference. Relaxation rates were measured from peak intensities in a series of 1D spectra, with maximum spin-lock pulse lengths of 300 ms (^{15}N $R_{1\rho}$), 160 ms (^{13}CO $R_{1\rho}$) and 80 ms ($^{13}C_{aliphatic}$ $R_{1\rho}$) and maximum relaxation delay times of 15 s (^{13}CO R_1) and 5 s ($^{13}C_{aliphatic}$ R_1). For each scan, initial ^{1}H magnetization was prepared with a 100 kHz 90° pulse, and subsequently transferred to ^{13}C/^{15}N via adiabatic DQ CP of 2.0 ms (^{15}N), 1.4 ms (^{13}C') or 1.0 ms ($^{13}C_{aliphatic}$) with nutation frequencies of ~50 kHz (^{1}H) and 10 kHz (^{13}C/^{15}N). Spin-lock nutation frequencies for $R_{1\rho}$ experiments were set to 17 kHz (36), while for R_1 experiments the 90° pulses either side of the incremented relaxation delay were set to 100 kHz (^{13}C) or 83 kHz (^{15}N). During acquisition, low-power (15 kHz) slpTPPM (35) ^{1}H decoupling was applied.

Chemical shift calculations

Ab initio calculations were carried out using version 4.0 of the PQS program package (Parallel Quantum Solutions; Fayetteville, AR) (44). Density functional GIAO shielding calculations were performed with the B3LYP functional and the Ahlrichs TZP basis set (45). Absolute shieldings were corrected as described previously (46). As a simple structural model we used a

N-acetyl-Gln-$NHCH_3$ diamide with backbone and sidechain torsion angles constrained to specific values. All other geometry parameters were optimized before shielding calculations. As a reference system, an analogous shielding calculation was carried out for amyloidogenic peptide GNNQQNY in its monoclinic crystal form (18). For these calculations the torsion angles of a single peptide were constrained based on the X-ray structure. The experimental MAS NMR chemical shifts (40) were used as a benchmark when converting calculated NMR shieldings to chemical shifts. The linear transformation of shielding values that gave the best agreement over all carbon shifts of GNNQQNY was used in all subsequent chemical shift calculations. To test the validity of the single Gln residue model, we calculated shieldings and shifts also for tripeptides built from both the 'a' and 'b' Gln conformers. The side chain carbon shifts of the single residue model and the central residue of the tripeptide remain within 1 ppm giving an indication of the reliability of the model.

X-ray powder diffraction

PolyQ amyloid fibrils were prepared from synthetic, unlabeled peptide $K_2Q_{31}K_2$, as described previously (15). The hydrated fibrils were packed into a glass capillary (0.7 mm diameter) using a syringe, after which the capillary was sealed with wax. X-ray powder diffraction data were measured at room temperature using a Rigaku Saturn 944 CCD camera (Tokyo, Japan), following one minute of beam exposure using a Rigaku FR-E generator (2 kW, spot size 0.07 mm) as the X-ray source. Diffraction data was processed using Structure Studio software from Rigaku.

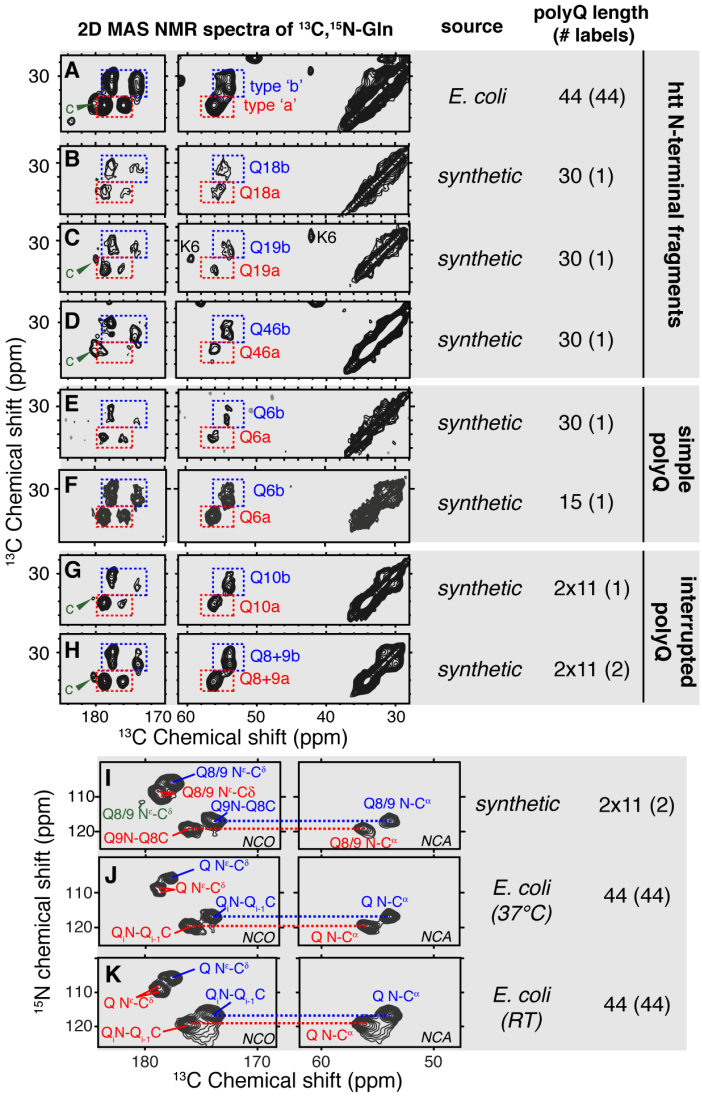


Figure S1. **MAS NMR spectra for $^{13}C,^{15}N$-labeled Gln within different polyQ amyloid cores.** (A-H) 2D ^{13}C-^{13}C DARR spectra on: (A) Uniformly $^{13}C,^{15}N$-labeled htt exon1, expressed as MBP fusion protein in *E. coli* and aggregated at room temperature. (B) First Gln in $htt^{NT}Q_{30}P_{10}K_2$ fibrils. (C) Second Gln in $htt^{NT}Q_{30}P_{10}K_2$ fibrils. (D) Penultimate Gln in $htt^{NT}Q_{30}P_{10}K_2$ fibrils. (E) Single Gln in [U-$^{13}C,^{15}N$-Q6] $K_2Q_{30}K_2$ fibrils. (F) Single Gln in [U-$^{13}C,^{15}N$-Q6] $D_2Q_{15}K_2$ fibrils. (G) Single Gln in [U-$^{13}C,^{15}N$-Q10] PG-Q_{22} fibrils. (H) Two sequential Gln in [U-$^{13}C,^{15}N$-Q8Q9] pG-Q_{22} fibrils. (I-K) ^{15}N-^{13}C 2D spectra on: (I) Two sequential Gln in [U-$^{13}C,^{15}N$-Q8Q9] pG-Q_{22} fibrils. (J) Uniformly $^{13}C,^{15}N$-labeled htt exon1 fibrils formed at 37 °C. (K) Uniformly $^{13}C,^{15}N$-labeled htt exon1 fibrils formed at room temperature. Further information on the experimental and sample conditions can be found in Table S5. Panels B-D, F, and G are adapted from previously reported data (11, 15, 16), with permission from Elsevier and the American Chemical Society. The 'a', 'b', and 'c' Gln conformers are marked in red, blue, and green respectively.

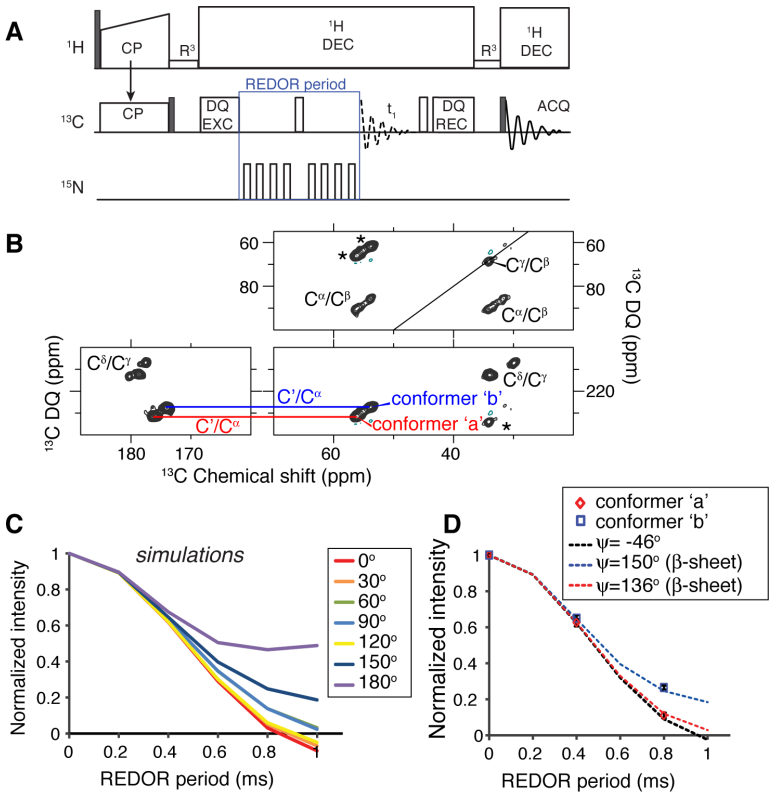


Figure S2. MAS NMR backbone torsion angle measurements of polyQ amyloid core residues. *(A) Schematic of the employed pulse sequence. The REDOR period is indicated, during which the ^{15}N-^{13}C dipolar interactions are recoupled. Abbreviations: ACQ. = acquisition; CP = cross polarization; DQ. EXC = double-quantum excitation; DQ REC = double quantum reconversion; ^{1}H DEC = ^{1}H decoupling (here: TPPM); R^3 = rotary resonance recoupling during the z-filters. Grey and white bars are π/2 and π pulses, respectively. For further details see ref. (26). (B) INADEQUATE-style SQ-DQ 2D spectra afford site-specific resolution in the NCCN experiments on sequential labeled polyQ core residues of [U-^{13}C,^{15}N-Q8,Q9]-pG-Q_{22} fibrils. The backbone C'-C^{α} correlations for the two Gln core conformers (highlighted with red/blue markers and connecting horizontal lines) were integrated as a function of the REDOR dephasing time. (C) Numerical simulations of the NCCN experiment, showing the different data curves for different ψ angles, as indicated. (D) Overlay of experimental data on the best-fitting curves (see also Fig. 3). Error bars are shown, but are similar in size to the data markers.*

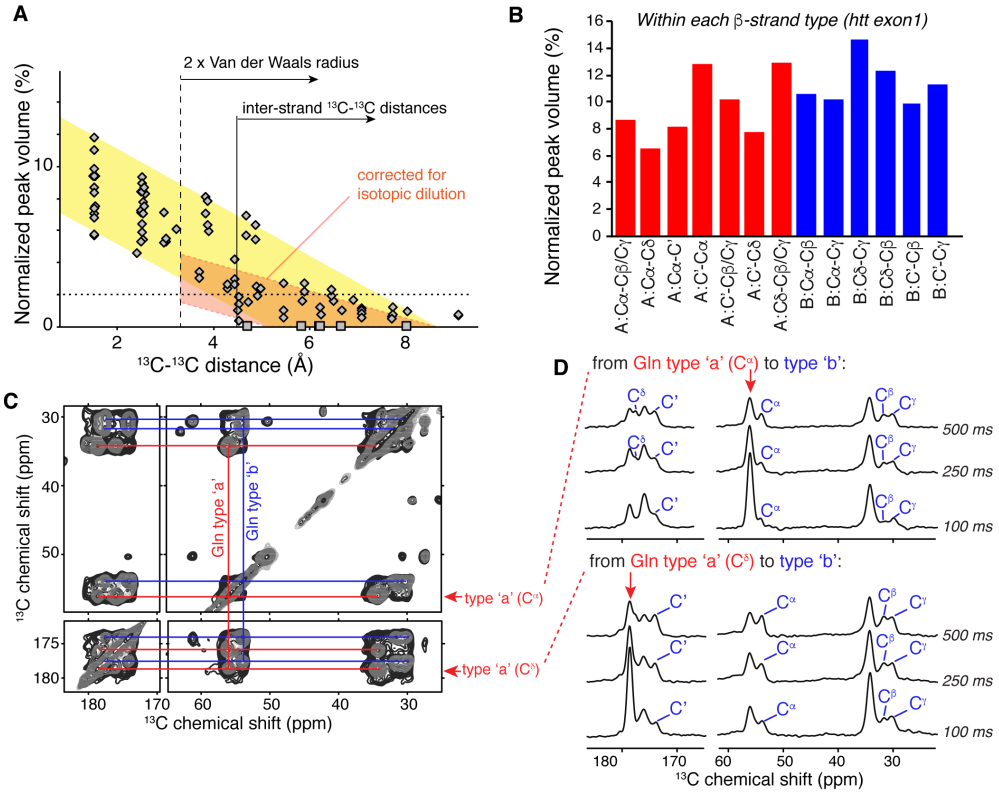


Figure S3. Long-range ^{13}C-^{13}C distance measurements. *(A) Normalized peak volumes from a 250 ms PDSD 2D spectrum acquired on 10%-U-^{13}C,^{15}N-labeled N-acetyl-valine-leucine crystals. The x-axis shows the corresponding intra-molecular ^{13}C-^{13}C distance in the known crystal structure. Diamonds reflect observed peaks, whilst squares show inter-atomic distances of less than 8 Å for which no detectable peaks are observed. The yellow area indicates the significant vertical spread, which is a known feature of these kinds of ^{13}C-^{13}C transfer experiments (28). The pale red area shows an estimation of the normalized (apparent) transfer intensity for intermolecular contacts upon a 25% isotopic dilution, which causes each ^{13}C-labeled β-strand to have only a 50% likelihood of having a ^{13}C labeled neighboring strand either above or below. Note that intermolecular contacts (e.g. between β-strands) have to be at least twice the Van der Waals radius (vertical dashed line). The solid vertical line indicates the minimal distance expected for carbons of neighboring, hydrogen-bonded β-strands (approximately ≥4.5 Å; (18)). The horizontal dashed line indicates the lowest peak volume observable above the noise in the htt exon1 fibril data of Fig. 4 and Fig. S3B. Thus, cross-peaks among backbone carbons of β-strands in different proteins are expected to have a normalized volume no more than ~4% after the extensive isotopic dilution, with most peaks expected to be indistinguishable from the noise. (B) Integrated cross-peaks reflecting local contacts within each of the two Gln conformers (i.e. within each β-strand type) in the htt exon1 fibril data from Fig. 4. The horizontal dashed line indicates the lowest observable peak size above the noise. (C) Overlay of short- and long ^{13}C-^{13}C mixing spectra for the U-^{13}C,^{15}N-labeled htt exon1's polyQ core, with 50- (gray) and 500-ms (black) PDSD mixing at 600 MHz (^{1}H). (D) 1D slices from a series of 2D PDSD ^{13}C-^{13}C spectra with variable ^{13}C-^{13}C mixing times showing the polarization transfer from C^{γ} and C^{δ} of type-'a' Gln to indicated atoms of Gln type 'b' (blue labels). The spin diffusion ^{13}C-^{13}C mixing time is shown (in ms). Note the clear cross-peaks between the 'a' and 'b' conformers even at 100 ms PDSD mixing.*

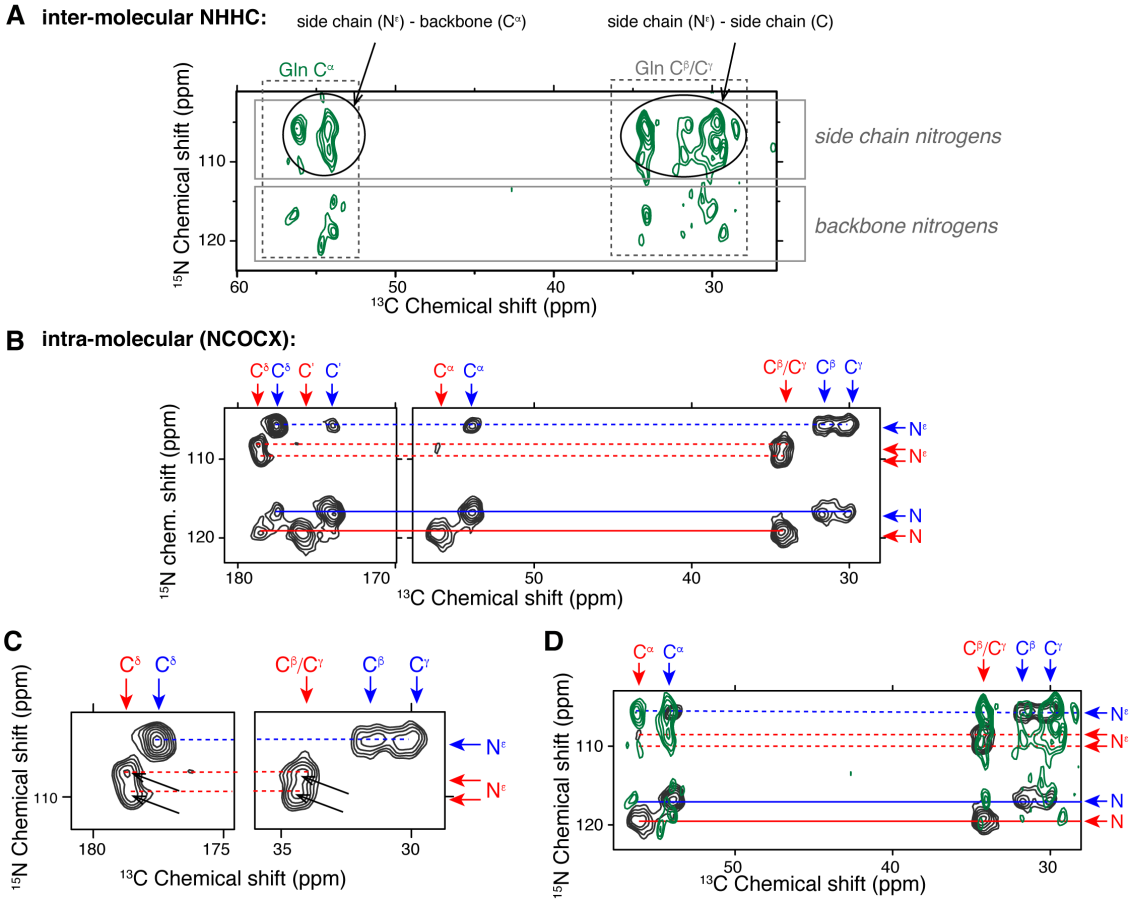


Figure S4. Inter-protein $^{13}C/^{15}N$ contacts. *(A) 2D NHHC spectrum with 250 μs ^{1}H-^{1}H transfer time on aggregated co-mixed ^{15}N- and ^{13}C-only labeled htt exon1 (40:60 molar ratio). Given that this a mixed $^{13}C/^{15}N$ sample, the observed peaks must reflect the interactions between different proteins. The cross peaks between backbones of different proteins are close to the noise level, and are clearly much weaker than those from side chains nitrogens (circled groups of peaks). Strong cross-peaks between the side chain N^{ε} of one protein and the backbone C^{α} of other proteins imply the presence of intermolecular steric zipper interfaces: interdigitation of side chains allows proteins in opposing β-sheets to bring their N^{ε} nitrogens and C^{α} carbons in close proximity. (B) To illustrate the chemical shift assignments and allow direct comparison to the intra-molecular cross-peak pattern we reproduce here the NCOCX data from Fig. 2 (now including the side chain nitrogen (N^{ε}) peaks). Colored arrows indicate ^{15}N and ^{13}C chemical shifts of the type-'a' (red) and type-'b' (blue) polyQ conformers. Solid lines connect the backbone nitrogen cross-peaks; dashed lines connect side chain nitrogen peaks. (C) Enlarged sections from (B) showing a doubling of the $^{15}N^{\varepsilon}$ peak of the type-'a' side chain (black arrows), which likely contributes to the weaker type-'a' (compared to type-'b') cross-peaks in the NHHC spectrum. (D) Overlay of the NHHC (green) and NCOCX (black) spectra, allowing a direct comparison. Colored markers indicate the ^{15}N and ^{13}C chemical shift assignments of the type-'a' (red) and type-'b' (blue) polyQ conformers.*

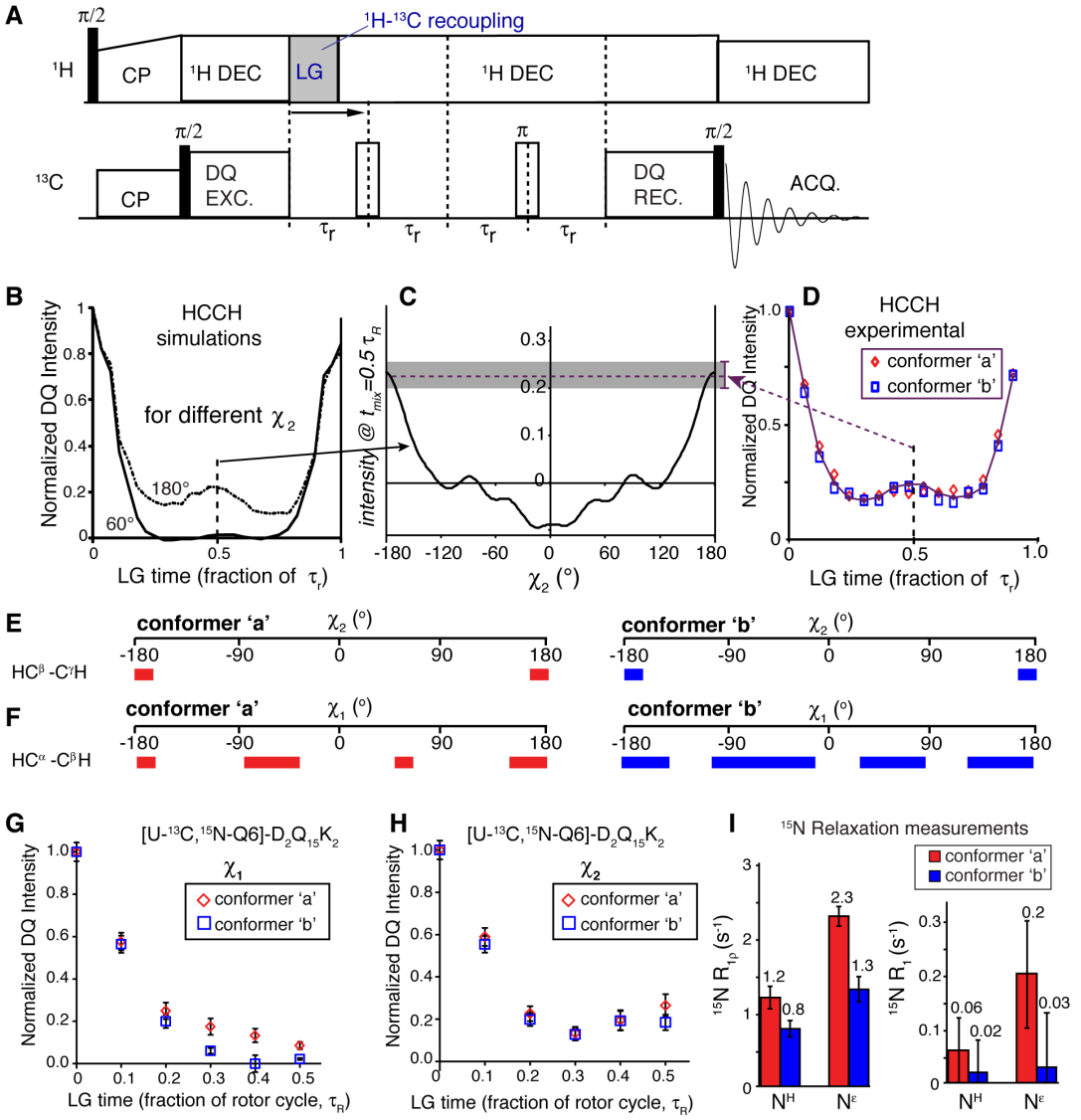


Figure S5. MAS NMR side chain torsion angle analysis of polyQ core Gln residues. *(A) Schematic of the pulse sequence employed for HCCH torsion angle measurements. Abbreviations as in Figure S2, with LG = Lee-Goldburg, and τ_r = time for a full rotation of the spinning sample (rotor period). (B) Numerical simulation of HCCH curves for bent (Fig. S6A) and straight side chains, having χ_2 = 60 and 180°. (C) Dependence of mid-point intensity ($t_{mix} = \tau_r/2$) on the χ_2 angle. (D) Full experimental HCCH curves for Gln conformers 'a' (red) and 'b' (blue), showing the normalized DQ-filtered C^{β}/C^{γ} intensity as a function of LG time. Data acquired on the polyQ core residues Q8 and Q9 (both U-^{13}C,^{15}N-labeled) in pG-Q_{22} fibrils. (E, F) Colored bars indicate χ_1 and χ_2 angle constraints from the HCCH measurements for Gln conformer 'a' (left) and 'b' (right). (G, H) χ_1 and χ_2 HCCH measurements on $D_2Q_{15}K_2$ fibrils featuring a single labeled Gln (U-^{13}C,^{15}N-Q6). Conformers 'a' and 'b' are shown in red and blue, respectively. Aside from lower signal/noise, these data on a single labeled residue reproduce the results on the two-residue-labeled sample (Fig. 3E,F). (I) ^{15}N $R_{1\rho}$ and R_1 relaxation rates for the backbone and side chain of both Gln conformers of U-^{13}C,^{15}N-labeled Q10 in PG-Q_{22} fibrils, measured at 60 and 22 kHz MAS, respectively.*

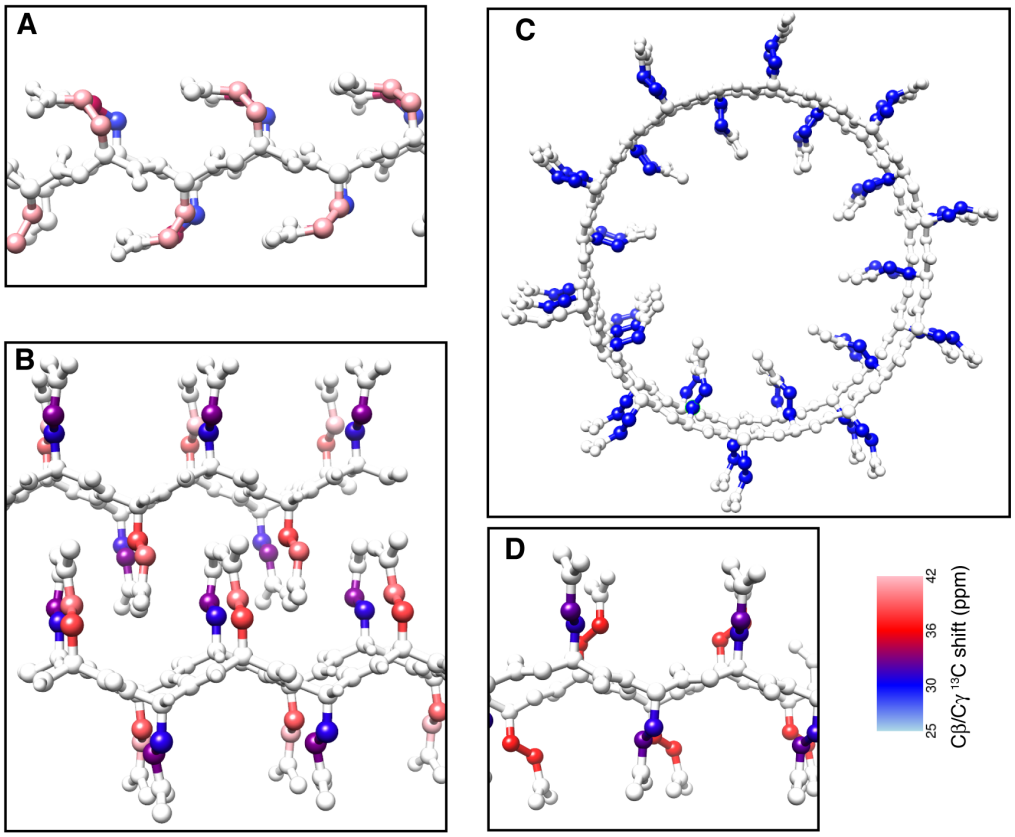


Figure S6. Calculated side chain ^{13}C shifts for polyQ structural models. *The Gln conformations from different polyQ structural models were used for ab initio chemical shift calculations. The non-hydrogen-bonding side chain Cβ/Cγ sites are color coded according to the calculated ab initio chemical shift values (see scale bar at right; Table S3). Note that by MAS NMR there should be two types of β-strands, which uniformly feature Cβ/Cγ shifts near 34 ppm or near 30-31 ppm, respectively. (A) X-ray based structure for slab-like aggregates, predicting two types of internally uniform β-strands (34). Conceptually this model fits the dehydrated nature of the polyQ amyloid core (15-17), but neither the calculated Cβ/Cγ shifts nor the χ_2 angles are consistent with our MAS NMR constraints. (B) X-ray-derived β-hairpin-based model (31), which correctly contains intramolecular β-hairpins and features interdigitating side chains. However, it is incorrect, since subsequent residues within the β-strands differ in structure and are calculated to yield alternating Cβ/Cγ shifts for odd/even residues (see color coding). (C) The Perutz nanotube model (41) would imply extensive solvent access for the Gln side chains, which is inconsistent with side chain dynamics and water access MAS NMR data (15, 17). Moreover, in ab initio analyses it fails to predict the observed signal doubling. (D) MAS NMR derived structure (see main text). The* ab initio *calculated Cβ/Cγ shifts match the experimentally observed pattern, since they are the same for all residues within each β-strand. These graphics were prepared with the UCSF Chimera program, using atomic coordinates kindly supplied by the respective authors.*

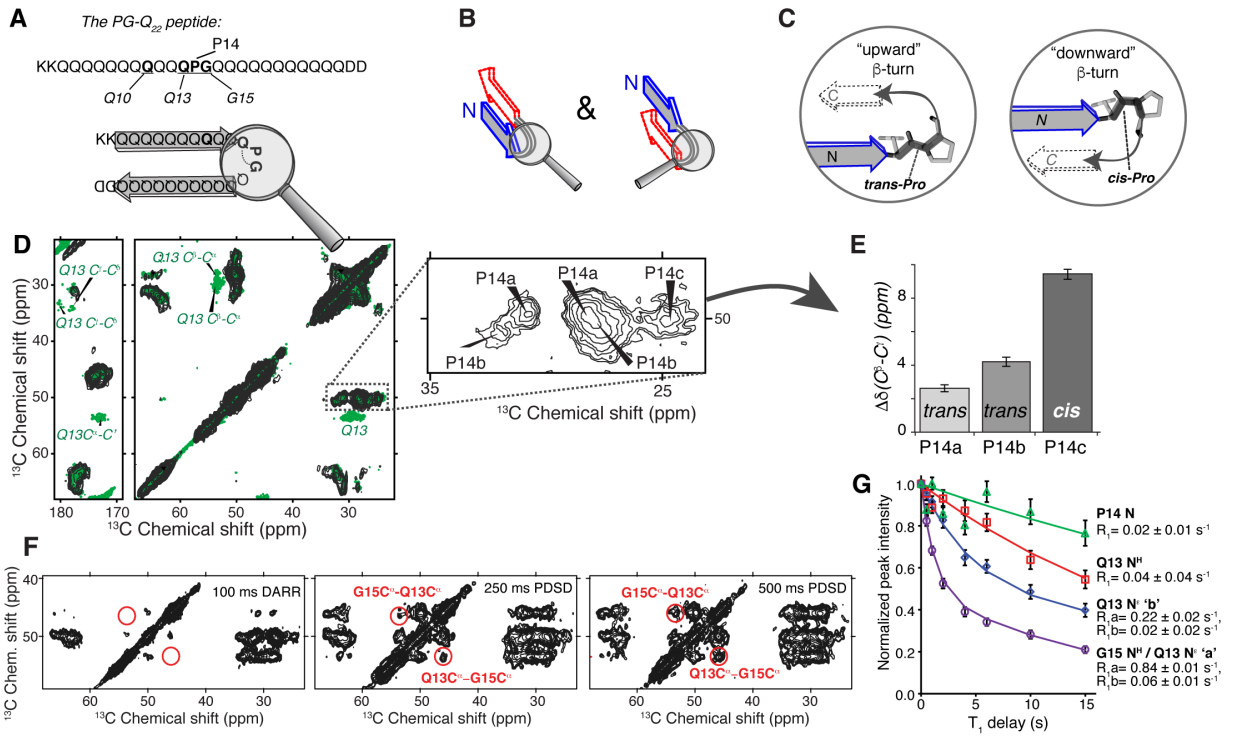


Figure S7. Polymorphic β-turn structures in Pro-Gly interrupted polyQ amyloid. *(A) Sequence of the PG-Q_{22} polyQ peptide with a central Pro-Gly insertion that is expected to stabilize a β-hairpin structure and end up in a β-turn outside the amyloid core (11, 12, 37). ^{13}C,^{15}N-labeled residues are shown in bold and underlined, of which Q10 shows the typical doubled polyQ core signal (Fig. S1G). (B) Two schematic structures of a peptide with its N-terminal half forming a type-'b' β-strand (blue) and its C-terminal half forming a type-'a' β-strand (red). Despite featuring an identically structured β-strand assembly, these structures differ in the conformation of the intervening turn regions (which here are ^{13}C,^{15}N-labeled). (C) Schematics showing how this could be accomplished by either trans-Pro in an "upward" β-II turn or cis-Pro in a "downward" β-VI turn. (D) 2D ^{13}C-^{13}C spectra for fibrils with labeled Q13/P14/G15 (green), or P14/G15 only (grey). Peaks only visible in green are from Q13. These Q13 signals are heterogeneous, do not match the shifts typical of the amyloid core (e.g. Q10; Fig. S1G), but are instead similar to the minor population of type 'c' Gln seen in htt exon1 and other polyQ fibrils (Table S2). Three types of Pro signals are observed (see enlarged inset), indicating the presence of different coexisting structures. (E) The C^{β} - C^{γ} shift difference ($\Delta\delta(C^{\beta}$-$C^{\gamma})$) of the Pro signals indicates that both trans- and cis-Pro are present, matching the illustrations in panel C. Whilst limited sample size, peak overlap, and general heterogeneity prevent a rigorous structural analysis, we note that the cis-Pro conformation fulfills the features of β-VI turns (which characteristically have a cis-Pro as part of the turn) (38). The trans-Pro could reflect various turn structures, but appear to fit well to β-II turns, based on structural and dynamic constraints (below), torsion angle analysis (24, 38), and chemical shift similarities to previously ssNMR-studied Pro-Gly β-II turns (53). (F) 2D ^{13}C-^{13}C spectra on the Q13/P14/G15-labeled fibrils shows cross-peaks between Q13C^{α} and G15C^{α} (red circles), for 250 and 500 ms PDSD mixing (i.e. distance < 6.5 Å). This $i \rightarrow i+2$ C^{α}-C^{α} distance is marked as a dashed line in the bottom scheme of panel (A). (G) ^{15}N R_1 relaxation at 22 kHz MAS, with fit curves (solid lines). Where appropriate, double exponential fits are indicated, e.g. for the overlapping signals of G15 N and Q13 N^{ε}. The backbone has slow relaxation, and does not appear to be flexible. The Gln side chain shows increased mobility, consistent with a solvent-accessible location outside the amyloid core, as expected for these β-turns as they always involve four or more residues (38).*

Table S1. - Sequences, labeling schemes, and amounts of isotopically labeled MAS NMR samples. *In all cases, the indicated proteins and peptides were studied as mature amyloid-like fibrils.*

Name (labeled sites)	Labeling details	Sequence [a)]	Sample size (mg)
Htt exon1, 37 °C fibrils [b)]	U-^{13}C,^{15}N	See Figure 1	2
Htt exon1, room temperature	U-^{13}C,^{15}N	See Figure 1	4.4
Htt exon1, room temperature	U-^{15}N/U-^{13}C mixed; 40:60	See Figure 1	4.0
Htt exon1, room temperature	U-^{15}N/U-^{13}C mixed; 74:26	See Figure 1	4.8
$D_2Q_{15}K_2$ (Q6)	U-^{13}C, ^{15}N-[Q6]	$D_2QQQQQ_{11}K_2$	5
PG-Q_{22} (Q13/P14/G15)	U-^{13}C,^{15}N-[Q13, P14, G15]	$K_2Q_{10}\underline{QPG}Q_{11}D_2$	10
PG-Q_{22} (P14/G15)	U-^{13}C,^{15}N-[P14, G15]	$K_2Q_{10}Q\underline{PG}Q_{11}D_2$	7
pG-Q_{22} (Q8/Q9)	U-^{13}C, ^{15}N-[Q8, Q9]	$K_2QQQQQ\underline{QQ}QQpGQ_{11}K_2$	9

a) Underlined residues are labeled as indicated. Lower-case 'p' = D-Pro.
b) See ref. (16).

Table S2. ^{13}C and ^{15}N chemical shift assignments for polyQ amyloid core residues in this study. *Chemical shifts have errors of ±0.1-0.3 ppm unless otherwise stated, and are referenced to dilute aqueous DSS.*

Residue(s)	C'	Cα	Cβ	Cγ	Cδ	N	Nε
Htt exon1 fibrils (room temperature)							
Gln 'a'	175.9	56.1	34.2	34.2	178.6 / 178.6	119.4	108.3 / 109.8
Gln 'b'	174.1	53.9	31.7	30.2	177.6	116.9	105.7
Gln 'c'	--	--	--	33.9	180.3	--	--
$D_2Q_{15}K_2$ fibrils							
Q6a	175.9	56.1	34.2	34.1	178.8	--	--
Q6b	173.9	53.9	31.7	30	177.5	--	--
Q6c	--	--	--	33.8	180.4	--	--
$K_2Q_{11}pGQ_{11}K_2$ fibrils (pG-Q_{22})							
Q8a	176.1	56.3	34.2	34.2	178.8	119.1	108.5
Q8b	173.9	53.9	31.7	30	177.5	116.9	105.9
Q8c	--	--	--	33.7	180.3	--	111.4
Q9a	176.1	56.3	34.2	34.2	178.8	119.1	108.5
Q9b	173.9	53.9	31.7	30	177.5	117.1	105.9
Q9c	--	--	--	33.7	180.3	--	111.4
$K_2Q_{11}PGQ_{11}D_2$ fibrils (PG-Q_{22}) [a)]							
Q10a	175.8	55.8	34	33.9	178.5	--	--
Q10b	173.8	53.6	31.5	29.9	177.5	--	--
Q10c	--	--	--	33.2	180	--	--
Q13a	173.1	53.5	29.8 +/- 0.5	32.5	177.6	119	104.2
Q13b	172.8	53.7	30.3	33.6	180.1	119	109.1
P14a	177.4	66.2	31.2	28.3	50.2	135.6	
P14b	176.6	64	31.9	27.8	50.1	133.5	
P14c	175.1	63.2	34.1	24.9	50.1	131.8	
G15a	174	46				111	
G15b	174	45.9				112.2	
G15c	174	46.1				107	

a) Q10 chemical shifts from ref. (11).

***Table S3*. Predicted side chain ^{13}C chemical shifts for polyQ amyloid structural models.** *Ab initio calculations were performed for distinct Gln side chain conformations present in the indicated model structures, kindly provided as PDB files by the respective authors. Some of the listed dihedral angles reflect averaged values of similarly structured residues present in the supplied coordinate files.*

Source	Conformer	χ_1, χ_2, χ_3 (°)	C^β	C^γ	Citation
X-ray-based slab-like model	β-strand 1	-175, -89, -28	28.8	34.3	(34)
	β-strand 2	-91, 88, -150	41.1	41.0	
X-ray-based β-hairpin model	β-strand 1, conformer 1	-110, -167, 160	38.8	41.6	(31)
	β-strand 1, conformer 2	69, -174, -162	31.0	32.8	
	β-strand 2, conformer 1	-120, 172, -162	38.1	39.5	
	β-strand 2, conformer 2	64, 167, -152	30.9	32.8	
Water-filled nanotube		76, -150, -33	29.9	29.7	(41)
MAS NMR	β-strand 'a'	-60, 180, 150	36.4	36.2	This work
	β-strand 'b'	55, 180, -150	31.1	32.3	

Table S4. Experimental solid-state ^{13}C isotropic chemical shifts for the side chains of the central Gln of crystalline GNNQQNY compared to shift values calculated from the respective structures. *Gln residue Q4 is involved in steric zippers in the parallel in-register β-sheet-based dehydrated core of amyloid-like GNNQQNY nanocrystals (18). Experimental MAS NMR shifts were from ref. (40). The* ab initio *calculations were performed on a single residue model (see Extended Methods) or the full peptide, as indicated, using torsion angles from the published crystal structures (18). The tabulated chemical shift values are in ppm, and are relative to dilute aqueous DSS. The experimental (Exp.) and calculated (Calc.) shifts are shown, along with their difference (Diff.). Larger deviations are observed for carbon sites involved in hydrogen bonding interactions (C^δ).*

Sample	Ab initio model	C^β			C^γ			C^δ		
		Exp.	Calc.	Diff.	Exp.	Calc.	Diff.	Exp.	Calc.	Diff.
monoclinic crystals	single residue	32.4	33.6	+1.2	34.8	35.1	+0.3	179.2	175.7	-3.5
	heptapeptide	32.4	32.7	+0.3	34.8	34.7	-0.1	179.2	180.8	+1.6
orthorhombic crystals	single residue	32.6	34.0	+1.4	34.7	36.4	+1.7	179.7	175.9	-3.8

Table S5. Detailed experimental conditions of the MAS NMR experiments. *Abbreviations: NS, number of scans per t_1 point; Temp., temperature of cooling gas; Spec., Spectrometer field strength; MAS, magic angle spinning rate; RD, recycle delay; t_1 evol., number and length (in μs) of t_1 evolution increments (including imaginary); TPPM, ^1H decoupling power during evolution and acquisition (using two-pulse phase modulation scheme). Sample identifiers: Htt exon1 (RT): htt exon1 fibrils prepared at room temperature; Htt exon1 (37 °C): U-^{13}C,^{15}N-labeled exon1 fibrils prepared at 37 °C (16); $D_2Q_{15}K_2$ (Q6): $D_2Q_{15}K_2$ fibrils featuring U-^{13}C,^{15}N-labeling of residue 6; others as indicated in Table S1.*

Figure	Sample	Experiment	Spec. (MHz)	NS	Temp (K)	MAS (kHz)	RD (s)	TPPM (kHz)	t_1 evol. (μs)	Mixing (ms)
2A,S3C	U-^{13}C,^{15}N htt exon1 (RT)	PDSD 2D	600	96	275	13	2.8	95	526x33.11	50
2B,D	U-^{13}C,^{15}N htt exon1 (RT)	NCACX 2D	600	512	275	13	2.8	95	74x410.76	20
2C,D, S4B-D	U-^{13}C,^{15}N htt exon1 (RT)	NCOCX 2D	600	512	275	13	2.8	95	74x410.76	50
3A,B	pG-Q_{22} (Q8,9)	NCCN	600	256	276	10	3	83	100x40.00	0.8
3E,F	pG-Q_{22} (Q8,9)	HCCH	600	1536	276	6	3	83		
3G	PG-Q_{22} (Q10)	^{13}C $R_{1\rho}$	600	2048	296	60	2	15		
3G	PG-Q_{22} (Q10)	^{13}C R_1	600	2048	296	60	2	15		
3H, S5I	PG-Q_{22} (Q10)	^{15}N R_1	600	790	275	22	6	83		
3I, S5I	PG-Q_{22} (Q10)	^{15}N $R_{1\rho}$	600	10,240	296	60	2	15		
4A,B, D-F	U-^{13}C,^{15}N htt exon1 (RT)	PDSD 2D (+ 1D slices)	800	32	275	13	3	83	496x25.0	250
4C, D-F	^{13}C/^{15}N-mixed htt exon1 (RT)	PDSD 2D (+ 1D slices)	800	64	275	13	3	83	520x25.0	250
S1A	U-^{13}C,^{15}N htt exon1 (RT)	DARR 2D	600	128	275	10	2.6	95	448x35.6	25
S1E	$K_2Q_{30}K_2$ (Q6)	DARR 2D	600	96	275	10	3	83	550x22.07	25
S1F	$D_2Q_{15}K_2$ (Q6)	DARR 2D	600	128	275	10	2.8	83	512x38.95	8
S1H	pG-Q_{22} (Q8,9)	DARR 2D	600	96	276	10	2.8	83	400x36.78	8
S1I	pG-Q_{22} (Q8,9)	NCO	600	256	276	10	2.8	83	98x200.0	NA
S1J	U-^{13}C,^{15}N htt exon1 (37°C)	NCO	600	352	275	10	2.6	86	74x273.84	NA
S1J	U-^{13}C,^{15}N htt exon1 (37°C)	NCA	600	352	275	10	2.6	86	74x273.84	NA
S1K	U-^{13}C,^{15}N htt exon1	NCO	600	192	275	13	2.8	95	55x410.76	NA

Figure	Sample	Experiment	Spec. (MHz)	NS	Temp (K)	MAS (kHz)	RD (s)	TPPM (kHz)	t_1 evol. (μs)	Mixing (ms)
S1K	(RT) U-^{13}C,^{15}N htt exon1	NCA	600	192	275	13	2.8	95	55x410.76	NA
S2B	(RT) pG-Q_{22} (Q8,Q9)	NCCN S0	600	256	276	10	3.0	83	100x40.00	0.4
S3A	[10% U-^{13}C,^{15}N] N-Ac-VL	PDSD 2D data	600	16	275	13	2.8	83	480x70.00	250
S3C	U-^{13}C,^{15}N htt exon1	PDSD 2D	600	96	275	13	2.8	95	526x33.11	500
S3D	(RT) U-^{13}C,^{15}N htt exon1	1D slices from PDSD 2D spectra	600	96	275	13	2.8	95	526x33.11	100, 250, 500
S4A	(RT) ^{13}C/^{15}N-mixed exon1	2D NHHC	800	2560	275	13	2.5	80	40x224.18	0.25
S5G,H	(RT) $D_2Q_{15}K_2$ (Q6)	HCCH	600	2048	275	6	2.7	83		
S7D	PG-Q_{22} (QPG)	DARR 2D	600	344	275	10	3	83	334x30.00	10
S7D	PG-Q_{22} (PG)	DARR 2D	600	88	275	10	3	83	334x30.00	8
S7F	PG-Q_{22} (QPG)	PDSD 2D	600	104	275	8	2.5	83	332x30.00	500
S7F	PG-Q_{22} (QPG)	PDSD 2D	600	104	275	8	2.5	83	332x30.00	250
S7F	PG-Q_{22} (QPG)	DARR 2D	600	104	275	8	2.5	83	332x30.00	100
S7G	PG-Q_{22} (QPG)	^{15}N R_1	600	1024	275	22	6	83		