

Polyglutamine Homopolymers Having 8–45 Residues Form Slablike β -Crystallite Assemblies

Deepak Sharma,¹ Leonid M. Shinchuk,^{1†‡} Hideyo Inouye,¹ Ronald Wetzel,² and Daniel A. Kirschner^{1*}

¹Biology Department, Boston College, Chestnut Hill, Massachusetts

²Graduate School of Medicine, University of Tennessee Medical Center, Knoxville, Tennessee

ABSTRACT At least nine inherited neurodegenerative diseases, including Huntington's, are caused by poly(L-glutamine) (polyGln, polyQ) expansions > 35–40 repeats in widely or ubiquitously expressed proteins. Except for their expansions, these proteins have no sequence homologies, and their functions mostly remain unknown. Although each disease is characterized by a distinct pathology specific to a subset of neuronal cells, the formation of neuronal intranuclear aggregates containing protein with an expanded polyQ is the hallmark and common feature to most polyQ disorders. The neurodegeneration is thought to be caused by a toxic gain of function that occurs at the protein level and depends on the length of the expansion: Longer repeats cause earlier age of onset and more severe symptoms. To address whether there is a structural difference between polyQ having < 40 versus > 40 residues, we undertook an X-ray fiber diffraction study of synthetic polyQ peptides having varying numbers of residues: Ac-Q₈-NH₂, D₂Q₁₅K₂, K₂Q₂₈K₂, and K₂Q₄₅K₂. These particular lengths bracket both the range of normalcy (9–36 repeats) and the pathological (45 repeats), and therefore could be indicative of the structural changes expected in expanded polyQ domains. Contrary to expectations of different length-dependent morphologies, we accounted for all the X-ray patterns by slablike, β -sheet structures, ~20 Å thick in the β -chain direction, all having similar monoclinic lattices. Moreover, the slab thickness indicates that K₂Q₄₅K₂, rather than forming a water-filled nanotube, must form multiple reverse turns. *Proteins* 2005;61:398–411.

© 2005 Wiley-Liss, Inc.

Key words: amyloid; Huntington's disease; X-ray diffraction; β -crystallite; β -helix

INTRODUCTION

At least nine neurological disorders, including Huntington's disease, result from expansion of the nucleotide triplet repeat (CAG)_n that codes for polyglutamine (polyGln; polyQ) tracts within different proteins,^{1,2} and all these diseases show an autosomal dominant pattern of inheritance except for X-linked spinal bulbar muscular atrophy (SBMA, or Kennedy's disease). The affected proteins, which show no sequence homology except for the polyQ stretch, are ubiquitously expressed but mostly have un-

known functions, although in some cases a role has been and/or is being inferred from experimental observations. Each disease is characterized by a distinct pathology for a specific subset of neuronal cells, but the polyQ disorders have several common features.² The onsets of these progressive neurodegenerative diseases occur typically in midlife and their duration is 10–20 years.^{2,3} A toxic gain of function at the protein level is thought to underlie the neurodegeneration. Proteins that have longer repeats correlate with an earlier age of onset and more severe symptoms. The threshold for all diseases excluding spinocerebellar ataxia type 6 (SCA6) is a repeat of 35–40 residues.²

The histopathological feature that is common to most polyQ disorders is the formation of neuronal intranuclear aggregates containing protein with the expanded polyQ repeat. Observations suggest that these diseases could share a common mechanism of pathogenesis in which the polyQ aggregates affect cellular function and eventually cause neuronal death. This hypothesis is supported by several studies of animal models and of tissue culture, which demonstrate that the intranuclear inclusions are toxic for neuronal cells,⁴ and by a number of additional observations.^{5,6} However, whereas neuronal death does correlate with the length of the polyQ extension and the amount of diffuse forms of huntingtin protein, it does not correlate with the formation of inclusion bodies.⁷ Possible unifying hypotheses include the suggestion that the true toxic species may be the aggregation-prone monomeric state, or that the most toxic aggregates are small, difficult to detect oligomers rather than the large, more visible inclusions.⁵ While the mechanism by which the mutated polyQ proteins induce neurodegeneration is not unambiguously established, the presence of protein aggregates nonetheless remains a common feature shared not only by most

Grant sponsor: Alzheimer's Association/T. L. L. Temple Foundation Discovery Award (to D. A. Kirschner). Grant sponsor: Boston College. Grant sponsor: NIH-NIA; Grant number: R03-AG19322 (to R. Wetzel).

†Current address: Department of Physical Medicine and Rehabilitation, Boston Medical Center, Boston, MA 02118

‡Some of the research was carried out while an undergraduate at Boston College.

*Correspondence to: Daniel A. Kirschner, Department of Biology, Higgins Hall 510, Boston College, 140 Commonwealth Avenue, Chestnut Hill, MA 02467-3811. E-mail: kirschnd.bc.edu

Received 1 February 2005; Accepted 29 March 2005

Published online 19 August 2005 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.20602

polyQ disorders but also by other neurodegenerative diseases, such as Alzheimer's disease, Parkinson's disease, prion disease, and amyotrophic disease.^{8–11} The typical finding that ubiquitin, chaperones, and components of the proteasomal apparatus are present in the disease-specific protein aggregates suggests that the aggregation process is related to protein misfolding.¹ In the case of polyQ expansion diseases, additional evidence for protein misfolding is the identification of antibodies that can distinguish between expanded tracts and normal ones,¹² although this result may also be due to a linear lattice effect and not a folding change.¹³ Thus, characterizing the misfolding and aggregation that is induced by expansion may prove important for clarifying the pathogenic mechanism.

In vivo interactions with other proteins or cellular components may mediate aggregation of polyQ-containing proteins; however, in vitro experiments indicate clearly that aggregation is self-driven, depending on protein concentration and expansion length with the same length threshold as that in tissue culture and in human pathologies.¹⁴ Thus, the aggregation process may be initiated by a conformational change induced by the expanded polyQ stretch, such as in promoting nucleus formation.¹⁵ Relatively few studies have focused on the structural aspects of polyQ diseases, or on the structures of polyQ stretches and of the proteins that contain them (as recently reviewed¹⁶). Perutz et al.¹⁷ proposed that polyQ chains could form antiparallel β -strands held together by hydrogen bonds between side-chain and main-chain amides, and the β -strands could then be assembled into β -sheets or β -barrels through a network of hydrogen bonds, giving rise to "polar zippers." Additional types of assemblies that have been proposed for polyQ include parallel and antiparallel β -sheet, π -helix, and μ -helix.¹⁸ Thakur and Wetzel¹⁹ used mutagenesis and aggregation kinetics to conclude that the aggregates consist of an alternating series of extended chain and turn elements, with the optimal extended chain length being seven or eight glutamines.

The insolubility of synthetic polyQ peptides in water, which has hindered experimental studies on their folding and assemblies, has been countered by flanking the glutamine stretch with charged residues and by treatment with disaggregating solvents. For example, studies on peptide $D_2Q_{15}K_2$ by circular dichroism (CD), electron microscopy (EM), and electron and X-ray diffraction led to the proposal that this peptide folds into hydrogen-bonded hairpins and aggregates to yield tightly packed β -sheets.²⁰ By contrast, characterization of 9-mer and 17-mer polyQ peptides flanked by sequences rich in alanine and lysine indicate a random coil conformation, and led to the notion that longer polyQ stretches may adopt different conformations.²¹ CD and solid phase Fourier transform infrared spectroscopy (FTIR) studies of polyQ stretches that were synthesized with intervening histidines to simulate the protein involved in SCA1 (i.e., Q_8HQHQ_8 , $Q_{10}HQHQ_{10}$, and $Q_8HQ_4HQ_8$) indicated aggregates rich in β -sheet rather than random coil; at low concentrations, these polyQ peptides were reported to form intramolecular β -hairpins, and at higher concentrations, to aggregate

through formation of large intermolecular β -sheet structures.²² The ability of added charged residues to provide transient solubility to polyQ peptides is limited, however. For peptides as short as Q_{15} , pretreatment of peptides with disaggregating solvents is required to generate completely aggregate-free solutions.²³ Assemblies of polyQ sequences containing flanking pairs of lysine residues (i.e., $K_2Q_nK_2$, for $5 \leq n \leq 44$) are predominantly random coil in conformation; however, when $n > 37$ glutamine residues, there is a much greater tendency to form amyloid-like aggregates.²⁴

The neuronal cell death that occurs in CAG extension disease correlates with the accumulation of insoluble fibrils containing polyglutamine stretches > 40 residues. With respect to their molecular organization, the fibrils have been proposed to form a 20-Å-diameter nanotubular structure^{25,26} composed of a β -helical array with 20 residues per turn, and in which the structural unit is two residues in the β -conformation. Thus, a β chain of 40 residues encompasses two helical turns that are held together by hydrogen bonding between amides of the glutamine backbone, as well as those of the side-chains. A polyglutamine stretch of 40 residues, therefore, could form a stable nucleus for fibril elongation in the H-bonding direction.^{25,26} The β -helix model does not, however, account for all of the observed X-ray reflections from aggregates of shorter polyQ stretches (e.g., $polyQ_{15}$ ^{20,25–27}), whereas the antiparallel β -sheet polar-zipper model²⁰ does. To determine whether there is a structural difference between the assemblies formed by polyglutamine stretches having > 40 versus < 40 residues, and to elucidate the molecular mechanism of self-aggregation of polyQ having different lengths, we describe here our analysis of X-ray diffraction patterns from the assemblies formed by a range of polyQ peptides (AcQ_8NH_2 , $D_2Q_{15}K_2$, $K_2Q_{28}K_2$, and $K_2Q_{45}K_2$) under lyophilized, vapor-hydrated, and solubilized/dried conditions. We report that all polyQ samples gave similar diffraction patterns that could be accounted for by slablike β -crystallites. This analysis suggests that $polyQ_{45}$ forms multiple reverse turns rather than a continuous β -helix.

MATERIALS AND METHODS

Peptides and Preparation of Diffraction Samples

Samples: We examined four different peptides: AcQ_8NH_2 (lyophilized; material obtained from Biopolymers Laboratory, Brigham and Women's Hospital, Boston, MA), $D_2Q_{15}K_2$, $K_2Q_{28}K_2$, and lyophilized monomer $K_2Q_{45}K_2$. The latter three peptides were synthesized at the Keck Biotechnology Center (Yale University) and used for X-ray diffraction without further purification. The peptide disaggregation and aggregation protocols have been reported previously.^{23,24,28} For simplicity, we refer to these peptides and their assemblies as Q8, Q15, Q28, and Q45. The flanking charged residues that were introduced into Q15, Q28, and Q45 ensured their solubility. Whereas these residues are not in the natural polypeptides and could therefore affect their folding here, we consider this unlikely for reasons cited in the Discussion section (see below).

The peptides were analyzed under three different conditions: lyophilized, vapor-hydrated, and solubilized then dried/concentrated. For the first condition, the lyophilized peptide was gently packed into a thin-walled glass capillary (0.7 mm diameter; Charles A. Supper Co.; South Natick, MA) to form a 1-mm disk. To maintain a dry environment inside the tube, a small piece of desiccant (anhydrous CaSO_4) was placed in the mouth of the capillary, which was then sealed at both ends. For vapor hydration, peptide was gently packed in the capillary as above, and then a small volume of water was placed in the mouth of the tube, which was then sealed. The sample was left at room temperature for 7–10 days to allow equilibration of the peptide against 100% relative humidity. For preparing solubilized/dried peptide assemblies, lyophilized peptide was first dissolved in ultrapure 18 megaOhm ($\text{M}\Omega$) water (U.S. Filter; Lowell, MA) at 10 mg/mL. All peptides dissolved readily in water. The solutions were briefly vortexed to ensure complete mixing, then centrifuged at 16,000 g for 10 min to sediment any precipitate. The supernatant was then slowly drawn into siliconized, thin-walled glass capillary tubes (0.7 mm outer diameter; Charles A. Supper Co., South Natick, MA) with the aid of a slight vacuum. The capillary tubes were sealed at the narrow end and placed in a 2-Tesla permanent magnet²⁹ (Charles A. Supper Co.), which can promote fibril alignment by diamagnetic anisotropy.^{30–32} The wide end of the capillary was sealed with wax through which a pinhole was punched using a hot needle. The peptide solution was then allowed gradually to dry under ambient temperature and humidity. The dehydration was monitored by periodic observation using a polarizing microscope. When the peptide solutions had dried to a small, uniform disk, the capillary tubes were removed from the magnetic field and transferred to the sample holder for analysis by X-ray diffraction.

For some samples the lyophilized peptides were first placed inside the wide mouth of the capillary, and then water was added. The solubilized peptide gel was then pushed into the thinner end of the tube. The peptide solution was dried under ambient conditions while in the magnetic field (as described above). The diffraction pattern from peptide Q15 gave an oriented fiber pattern with the cylindrical or rotation axis parallel to the H-bonding direction of the β -sheet.

X-Ray Diffraction Experiment

Data collection

X-ray diffraction patterns were obtained using nickel-filtered, double-mirror focused $\text{CuK}\alpha$ radiation from an Elliott GX-20 rotating anode X-ray generator (GEC Avionics; Hertfordshire, UK) with a 200 μm focal spot, operated at 35 kV and 35 mA. A helium tunnel was placed in the X-ray path to reduce air scatter. Patterns were recorded on Kodak DEF films (Rochester, NY) with exposure time of 55–160 h. The known Bragg spacing of calcite, 3.035 Å,³³ was used to calibrate the specimen-to-film distance (66.0 or 88.0 mm). The Bragg spacings of reflections were measured from densitometer tracings of the X-ray films,

which were digitized at either 85 or 115 μ with a ScanMaker 5 Scanner (Microtek Laboratory, Inc., 3715 Doolittle Drive, Redondo Beach, CA 90278-1226) using the Macintosh program ScanWizard PPC 3.2.2. The eight-bit readout was converted to optical density units by a characterization curve derived from a calibrated step-wedge (Kodak Photographic Step Tablet No. 2; Eastman Kodak Company, Rochester, NY). After subtraction of polynomial-fit background curves, the resulting intensity curves were fit with Gaussian profiles³⁴ to determine integral areas and widths of the X-ray reflections. The integral width of the incident direct beam was in the range of 242–455 μm (corresponding to 0.0018–0.0034 Å^{−1}).

Rapid recording of X-ray diffraction

Measurements at room temperature from peptides Q15 and Q45 (monomer) were also conducted using the Oxford Diffraction Xcalibur PX Ultra System (Oxford Diffraction Ltd.; 130A Baker Avenue, Concord, MA 01742) located in the laboratory of Dr. Andrew Bohm (Department of Biochemistry, Tufts University, Boston, MA). The $\text{CuK}\alpha$ X-ray beam was generated using an Enhance Ultra, which is a sealed tube-based system incorporating confocal multilayer optics. The X-ray beam was monochromated and the $\text{K}\beta$ component was removed by means of the double bounce within the confocal optic. The X-ray beam was focused to 0.3 mm \times 0.3 mm (full-width at half-maximum at detector position). A two-dimensional (2D) Onyx charge-coupled device (CCD) detector (Oxford Diffraction Inc., Concord, MA) was placed either 75 or 100 mm from the sample position, covering the scattering range of Bragg spacing 2.8–63 Å. The sample-to-detector distance was calibrated by a spherical ylid crystal (molecular formula $\text{C}_{10}\text{H}_{10}\text{SO}_4$) or a cubic alum crystal. The active range of the detector was 165 mm, and the 2D image (1024 \times 1024 pixels; in 2 \times 2 binning) was collected using the software package CrysAlis [CrysAlis CCD and RED, version 171 (2004), Oxford, UK] and stored in the compressed image format IMG. The output readout, which was stored on a hard disk, was linear with X-ray intensity to $\sim 1.3 \times 10^5$. Exposure times were 360 s.

Analysis of X-Ray Diffraction Data

Powder diffraction patterns (which are circularly symmetric) were analyzed by measuring the intensity along a radial strip on the film. The public domain NIH Image program (developed at the U.S. National Institutes of Health and available at <http://rsb.info.nih.gov/nih-image/>) facilitated this quantitation. The observed intensities are related to the structure amplitudes by Lorentz and polarization factors. The Lorentz factor L in powder diffraction is given by $\lambda/4\sin\theta$, where λ is X-ray wavelength and θ is half of the scattering angle. As the spherically distributed intensity I_s is measured along the radial direction on flat film, the structure factor F is related to I_s by $I_s = F^2LP/2\pi R = F^2P/4\pi R^2$, where $R = 2\sin\theta/\lambda$ and the polarization factor $P = (1 + \cos^2 2\theta)/2$.³⁵

The lattice constants and indices were searched by comparing the observed and calculated spacings using

different combinations of indices within a particular range of lattice constants and indices. The final values of lattice constants were determined by linear least-squares fitting.

Fourier synthesis iteration and molecular modeling in powder diffraction: The procedure for fiber or powder diffractions has been described in detail.^{27,35,36} In brief:

1. We first choose the initial phase model that defines the atomic fraction coordinates (x_j, y_j, z_j) for the j th atom in the unit cell. The initial model here includes the C α , C β , and peptide bonds of the β -backbone, which was built using Swiss-PDB viewer.³⁷
2. Next, we calculate the structure factors $F_{calc}(hkl)$ with given ranges of h, k, l for the observed unit cell according to $F_{calc}(hkl) = \sum_j f_j \exp(i2\pi(hx_j + ky_j + lz_j))$, where f_j is the atomic factor. We use space group P1 (which assumes no symmetry) and calculate the structure factors for both positive and negative indices.
3. We then combine the observed structure amplitude $|F_{obs}(hkl)|$ and calculated phase $\phi_{calc}(hkl)$. For powder diffraction the observed intensity is given as the spherically distributed intensity $I_s(R)$ as a function of radial component of spherical reciprocal coordinates R . Because multiple indices may contribute to each peak in powder diffraction, the observed structure amplitudes for the specified indices are calculated by dividing the observed intensity in proportion to the calculated intensities from the model. The observed structure amplitudes are therefore given by $|F_{obs}(hkl)|^2 = C(hkl)I_s(R)/LP$, where $C(hkl) = |F_{calc}(hkl)|^2/\sum |F_{calc}(hkl)|^2$, and LP is the Lorentz polarization factor (see above). The indices hkl for the summation satisfy the condition $|R_{hkl}^{-1} - R^{-1}| < \Delta d$, where R_{hkl} is the radial component of the Bragg spacing in spherical coordinates, and Δd is a defined critical value between the observed and calculated Bragg spacings (0.1 Å for the current broad reflections). The relative deviation between the observed and calculated structure amplitudes is given as $R_{obs-amp} = \sum ||F_{obs}| - |F_{calc}|| / \sum |F_{obs}|$.
4. Finally, we use the observed amplitudes and calculated phases to calculate the electron density map with XtalView³⁸ and fit the density map with a molecular model. With the new atomic model, we then go back to step 2.

Molecular Modeling

The stereochemical quality of the model was examined using WHAT IF,³⁹ which is available online in the Biotech Validation Suite for Protein Structures at <http://biotech.ebi.ac.uk:8400/>. The molecular model was displayed and manipulated using MOLSCRIPT,⁴⁰ XtalView,³⁸ RAS-MOL,⁴¹ and Swiss-PDB Viewer.³⁷

RESULTS AND INTERPRETATION

β -Sheet Structures Account for the X-Ray Diffraction Patterns

The diffraction patterns of all peptides (Figs. 1 and 2, and Table I) showed Bragg reflections that were accounted for by a monoclinic lattice of $a = \sim 9.6$ Å, $b = \sim 7.1$ Å, $c =$

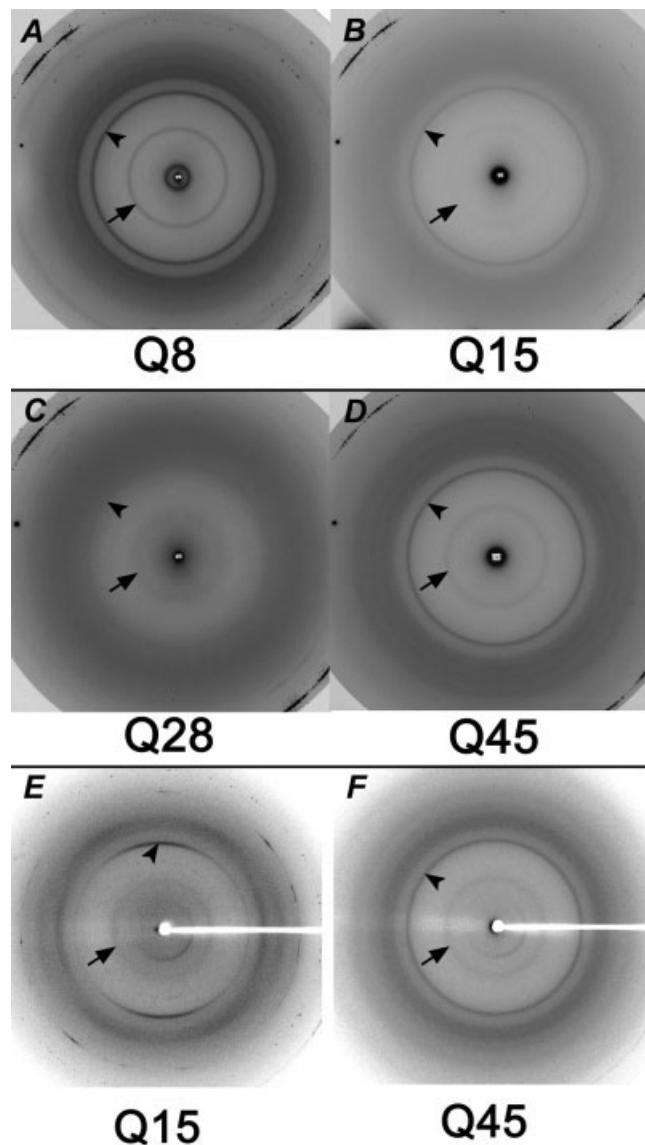


Fig. 1. X-ray diffraction patterns from a series of polyQ peptides. (A) Ac-Q₈-NH₂, lyophilized; (B) D₂Q₁₅K₂, vapor-hydrated; (C) K₂Q₂₈K₂, vapor-hydrated; (D) Q₄₅, concentrated by drying preformed aggregates; (E) D₂Q₁₅K₂; and (F) Q₄₅, prepared by solubilizing the lyophilized peptides at the mouth of the capillary tube with added water, and then pushing the gel into the capillary. The arrowhead and arrow indicate the positions of the 4.8 Å H-bonding and 8.2 Å intersheet reflections that are detected, though with different intensities, in all the samples. Patterns (A) through (D) were recorded for 48–96 h, using flat film and the X-ray beam generated by a rotating anode and focused using double mirror optics, and (E) and (F) were recorded for 6 min using the Oxford Diffraction system (Tufts University, Department of Biochemistry).

~ 8.2 Å and $\gamma = \sim 95^\circ$. These spacings most likely arise from a β -sheet structure, defined by axes a , b , and c along the H-bonding, polypeptide chain, and intersheet stacking directions, respectively.³⁵ While the characteristic (001), (200), and (210) reflections of β -sheets were evident, the polyQ peptides were unique in giving a strong 4.1 Å reflection indexed by (002) and/or (201), a 2.8 Å reflection indexed as (003), and a novel doublet at 3.5 Å (020) and 3.3 Å (021).

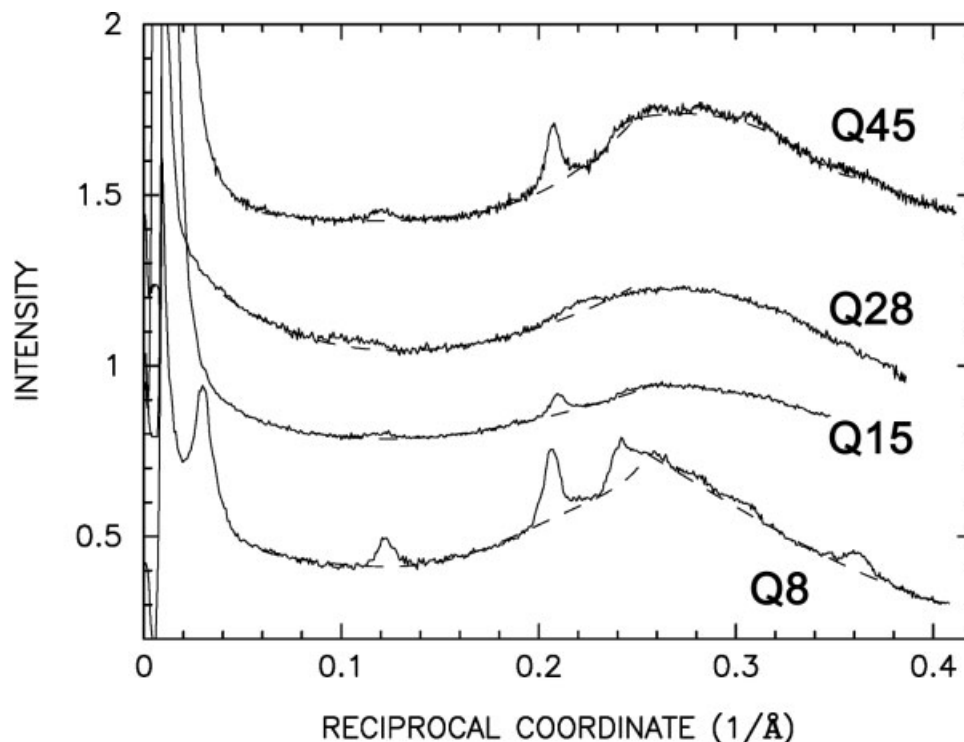


Fig. 2. The intensity distribution for Q45 (dried), Q28 (vapor-hydrated), Q15 (vapor-hydrated), and Q8 (lyophilized) as derived from digitizing the X-ray diffraction patterns recorded on film. The intensity was measured from the linear scan, converted to the optical density scale, and plotted as a function of reciprocal coordinate in \AA^{-1} . The intensity was not corrected for the Lorentz and polarization factors. The intensity scale shown is for the polyQ8, and the other spectra were displaced vertically for clarity. Dashed lines show the background curves as approximated by a combination of polynomial curves.

Underlying Structure Is Slablike

Among the different peptides examined here (Table I), peptide Q15 assemblies^{20,26} showed a fiber pattern with a cylindrical (rotation) axis along the equator. In agreement with the reported measurement,²⁰ ours showed meridional reflections at 8.4 \AA , 4.8 \AA , 4.2 \AA , and 2.8 \AA , an off-meridional reflection at 3.9 \AA , and equatorial reflections at 3.6 \AA and 3.2 \AA . The meridional reflections were indexed two dimensionally as ($h0l$), and the equatorial ones as (020) and (021) of a monoclinic unit cell (with $a = 9.6$ \AA , $b = 7.2$ \AA , $c = 8.4$ \AA , and $\gamma = 93^\circ$). The cylindrical (rotation) axis was in the b -direction of the β -sheet. That the ($h0l$) reflections are sharper than the ($0k0$) reflection indicates that the size of the coherent domain in the directions of a and c is larger than that in the b -direction. Therefore, the Q15 assembly is slablike, and the slabs are stacked in the b -direction with a 30- \AA period as indicated by the equatorial low-angle reflection. If Q15 formed an extended and straight β chain with a ~ 3.5 - \AA rise per residue, its total length should be ~ 50 \AA , which is considerably larger than the stacking period. Therefore, the peptide likely forms a reverse turn in the middle of Q15, and the N- and C-termini are positioned in close proximity to one another.

A similar β -crystallite model has been recently proposed⁴² for $D_2Q_{15}K_2$.^{25,26} In our nomenclature, the reported unit cell is orthorhombic, where $a = 9.50$ \AA , $b = 6.95$ \AA , and $c = 16.60$ \AA . Note that their c -dimension is twice as

large as the intersheet distance, as they assigned a (201) index to the very weak 4.5 \AA reflection. With the orthorhombic unit cell, unlike the monoclinic cell in our analysis, the agreement between the observed and calculated spacing for the 3.6 \AA (020) reflection is poor.

The powder patterns for Q8 and Q45 (Figs. 1 and 2, Table II) showed nearly identical Bragg spacings as the ones in the oriented Q15 sample^{20,26} (Tables I and II). From the linear least-squares optimization using the observed spacings, the best lattice constant was monoclinic, giving $\gamma = 96^\circ$ for Q8 and 94° for Q45 (Table I). The low-angle reflection at 35 \AA in Q8 is the first intensity maximum that arises from the interference between slabs. Absence of higher orders of the stacking period indicates that the stacking array is not highly ordered, but rather liquidlike (see below⁴³). The vapor-hydrated Q15, Q28, and dried Q45 on the other hand showed strong forward scattering and no interference peak, which indicates that their slabs are not stacked. The (001) reflection at ~ 10 \AA for vapor-hydrated Q28 was larger than that of the other peptides by 2 \AA , suggesting that the intersheet space in Q28 is expanded by the positive-charged and bulky lysine residues at the N- and C-termini.

For the oriented fiber pattern for Q15 (Fig. 1 and Table I), the cylindrical (rotation) axis was in the direction of H-bonding. Thus, the (200) 4.8- \AA reflection was on the meridian, the (210) 3.8- \AA was off-meridional, and the (001)

TABLE I. Summary of X-Ray Diffraction Spacings by Different Polyglutamine Peptide Assemblies

AcQ ₈ NH ₂ Lyophilized	D ₂ Q ₁₅ K ₂ V.H.	D ₂ Q ₁₅ K ₂ S.D.	D ₂ Q ₁₅ K ₂ ²⁰ (see footnote)	K ₂ Q ₂₈ K ₂ V.H.	Q ₄₅ Preformed fibrils dried	Q ₄₅ S.D.
$a = 9.73, b = 7.14,$ $c = 8.16, \gamma = 95.7^\circ$			$a = 9.6,$ $b = 7.2,$ $c = 8.4,$ $\gamma = 93^\circ$		$a = 9.66,$ $b = 7.10,$ $c = 8.33,$ $\gamma = 94.0^\circ$	
34.9 Cvs	Forward scatter		31*	No low-angle	Forward scatter	
8.14 Cm (001)	8.32 Cm (001)	4.81 Ms (200)	8.4 Mw (001)	9.95 Cm (001)	8.31 Cm (001)	4.81 Cs
4.84 Cs (200)	4.77 Cs (200)	3.88 M'm (210)	4.8 Ms (200)	4.62 Cm (200)	4.82 Cs (200)	4.05 Cw
4.13 Cs (002) (201)	4.06 Cm (002) (201)	8.29 Ew (001)	4.2 Mm (002)(201)		4.15 Cw (201)(002)	3.53 Cw
3.83 Cw (210)		4.14 Em (002)	3.9** M'm (210)		3.86 Cw (210)	3.07 Cw
3.55 Cvw# (020)		3.58 Ew (020)	2.8 Mvw (003)		3.54 Cw# (020)	
3.29 Cvw\$ (021)		3.16 Cvw	3.6 Em# (020)		3.23 Cw\$ (021)	
2.76 Cw (003)			3.2 Em\$ (021)		2.83 Cvw (003)	

The Bragg spacings Å are summarized for the observed diffraction patterns (Fig. 1) and for the Q15 peptide of Perutz et al.²⁰ Samples were either lyophilized, vapor-hydrated (V.H.), solubilized then dried (S.D.), or solubilized in the capillary mouth and then pushed into the narrow bore section; the Perutz et al.²⁰ pattern was obtained from “wet particles spun into a quartz capillary.” The monoclinic lattice constants, a , b , c , and γ are indicated, where a is along the H-bonding direction, b is the chain direction, and c is along the intersheet direction. The major Miller indices are indicated in parenthesis (see also the overlapping indices in Table II). The orientation and intensity are indicated qualitatively by *C*, circular; *E*, equatorial; *M*, meridional; *M'*, off-meridional; *vs*, very strong; *s*, strong; *m*, medium; *w*, weak; and *vw*, very weak. The Q15 pattern of Perutz et al.²⁰ has similar Bragg spacings as that recently reported.²⁶ The 3.9 Å** was measured directly from a copy of the pattern kindly provided by Dr. J. T. Finch. The broad low-angle reflection at 31 Å* was also indicated previously.^{25,26} All peptides gave similar Bragg spacings. #, \$: The circular wide-angle reflections near 3.6 Å and 3.2 Å spacings may correspond to those observed in the oriented fiber pattern of D₂Q15K₂.²⁰

TABLE II. Observed and Calculated Bragg Spacings (Å) for Q8 and Q45 Powder Patterns

Q8			Q45		
$a = 9.73 (0.07)$ $b = 7.14 (0.06)$ $c = 8.16 (0.03)$ $\gamma = 95.7^\circ (1.1)$			$a = 9.66 (0.06)$ $b = 7.10 (0.06)$ $c = 8.33 (0.03)$ $\gamma = 94.0^\circ (1.1)$		
Miller index	d_{obs}	d_{calc}	Miller index	d_{obs}	d_{calc}
001	8.14	8.16	001	8.31	8.33
200	4.84	4.84	200	4.82	4.82
201	4.13	4.16	201	4.15	4.17
002		4.08	002		4.17
211	3.83	3.73	210	3.86	3.86
210		3.83			
020	3.55	3.55	020	3.54	3.54
021	3.29	3.26	021	3.23	3.26
120		3.29	120		3.23
003	2.76	2.72	003	2.83	2.78

The observed indices include (hkl), ($hk-l$) and the respective Friedel pairs. Goodness of fit between observed and calculated spacings is 0.39% for Q8 and 0.41% for Q45. The odd h order (120) contributes to the weak intensity of the 3.29 Å spacing for Q8 and 3.23 Å for Q45.

8.3-Å, (002) 4.1-Å, and (020) 3.5-Å reflections were all on the equator. These indices were consistent with the ones given above from the fiber pattern with the cylindrical (rotation) axis parallel to the β -chain direction and from the powder pattern.

Low-Angle Reflections of Q8, Q15, and Q45 Provide Measure of Slab Thickness

Assemblies of peptide Q8 showed a strong reflection at 35 Å Bragg spacing (Figs. 1 and 2). This intensity maximum arises from the one-dimensional (1D) interference function $Z(R)$ with period d , and the subsequent reflection

likely arises from the Fourier transform $F(R)$ of the slab of thickness a , where R is the reciprocal coordinate in the direction normal to the slab surface. Here the intensity $I(R)$ is written as

$$I(R) = |F(R)|^2 Z(R),$$

where

$$F(R) = a \text{sinc}(\pi a R),$$

$$Z(R) = \frac{1 - [H(R)]^2}{1 + [H(R)]^2 - 2[H(R)]\cos(2\pi d R)},$$

TABLE III. Analysis of the Low-Angle Scattering

Peptide	Data sets	d (Å)	Δ (Å)	a (Å)	R (%)
Q8	3	30.7 (0.9)	7.3 (1.1)	17.0 (2.0)	14 (7)
Q15	3			19.0 (2.0)	18 (5)
Q45	9			22.6 (3.1)	16 (6)

Q8 showed an interference peak at d (Å), while Q15 and Q45 showed strong forward scattering. Paracrystalline theory was used to define the lattice period d and disorder Δ . The structure amplitude was calculated from the slab of thickness a . The relative error between the observed and calculated intensity (R -factor) was measured by systematically changing the variables. The number of data sets is indicated, and the mean error is shown in parenthesis.

and

$$H(R) = \exp(-2\pi^2 R^2 \Delta^2),$$

which is the Fourier transform of the Gaussian probability $h(r)$ for the nearest lattice point, and Δ is the deviation of the $h(r)$.^{34,44,45} To fit the calculated intensity to the observed one, the approximate values of parameters d , Δ , and a were first derived (Table III). The $[F(R)]^2$ minima occur at n/a , where n is a nonzero integer. Assuming that the first zero intensity is at the edge of the first reflection, then the thickness of the slab should be about 24 Å. Assuming that the second intensity maximum at $R_2 = (15 \text{ Å})^{-1}$ is the intensity maximum of $[F(R)]^2$, then the thickness a is 23 Å (from $aR_2 = 3/2$). Both assumptions gave similar a values. As the first intensity maximum refers to the interference peak of the stacking, the period d is about 35 Å.

Assuming that the loss of Bragg reflections arising from the interference function may be characterized by the region where the intensity maxima do not exceed the background by more than 20%,^{44,46} then we can derive the lattice disorder parameter Δ from $X\Delta = 0.4$, where X is the reciprocal coordinate at which the highest order reflection is barely detected. Since the second order of 35 Å was not observed clearly, the X value was taken as $(17 \text{ Å})^{-1}$. Therefore, $\Delta \approx 7 \text{ Å}$. The best values for $d = 30 \text{ Å}$, $\Delta = 7 \text{ Å}$ and $a = 15 \text{ Å}$ were obtained by searching for the minimum residual (24%) between the normalized observed intensity and the normalized calculated intensity according to the above equations. A plot of the R -factor as a function of slab thickness for specified period and its variation showed a single minimum (Fig. 3). To account for the spherical distribution of the scattered intensity, the calculated intensity was divided by R^2 (Fig. 4). The optimized stacking period d was slightly shifted to 30 Å from the measured peak at 35 Å due to the influence of the structure amplitude. Carrying out the same analysis for both the right and left sides of the scattering curves gave averages of period $d = 30.7 \text{ Å} (\pm 0.9)$, $\Delta = 7.3 \text{ Å} (\pm 1.1)$, and $a = 17.0 \text{ Å} (\pm 2.0)$, with a mean error of 14.2% (± 6.8) (Table III).

Samples of vapor-hydrated Q15 and solubilized/dried Q45 (monomer) (Figs. 1 and 2), unlike the lyophilized Q8 sample, showed strong forward scattering, indicating a lack of interference between the slablike structures. If the thickness of the slab is along the direction normal to the surface of the slab, then the Fourier transform gives

the value along this direction and is zero along the directions parallel to the surface. Given the reciprocal coordinate R , the structure factor $F(R)$ is given as described above. The first zero [i.e., where $F(R) = 0$] is at $R = 1/a$. The best a value for the slab thickness was derived by changing a systematically and searching for the minimum relative error between the observed and calculated intensities (Fig. 5). Plotting the R -factor as a function of slab thickness gave a single minimum (Fig. 5, inset). For Q15, the mean value of a was 19.0 Å (± 2.0), with the averaged R -factor 17.8% ($\pm 4.6\%$); and for Q45, the mean value of a was 22.6 Å (± 3.1), with $R = 15.5\%$ ($\pm 6.2\%$) (Table III).

Electron Density Map Revealed Side-Chain Conformation

The diffraction from lyophilized peptide Q8 showed concentric rings of Bragg reflections, which were indexed by a monoclinic cell of $a = 9.73 \text{ Å}$, $b = 7.14 \text{ Å}$, $c = 8.16 \text{ Å}$ and $\gamma = 95.7^\circ$, where a , b , and c are in the directions of the H-bonding, β -chain, and intersheet stacking, respectively. The structure amplitudes of the reflections were measured and compared with the calculated ones derived from the structural model. The electron density map was calculated by combining the structure amplitudes from the observed intensity and the phase angles from the model (Fig. 6, left). By fitting the molecular structure to the electron density map, a new atomic model was built. The agreement between the model structure and the observed intensity was estimated by calculating the R -factor (i.e., the relative error between the observed and calculated structure amplitudes). The initial phase model was a β -chain backbone built using the Swiss-PDB Viewer. Due to the geometrical restriction of the unit cell along the β -chain direction, the β -chains were nearly fully extended, and thus the sheet arrangement was likely to be antiparallel. The initial phase model gave an R -factor of 47% for Q8, and 65% for Q45. The electron density map, as expected, showed a higher electron density at the peptide backbone position, but also gave peaks between the β -chains (Fig. 6, right). Including the glutamine side-chains in a bent conformation reduced the R -factors to 24% and 35% for Q8 and Q45, respectively.

DISCUSSION

To account for the X-ray diffraction patterns recorded for $D_2Q_{15}K_2$, Perutz et al.^{20,26} proposed the “polar zipper”

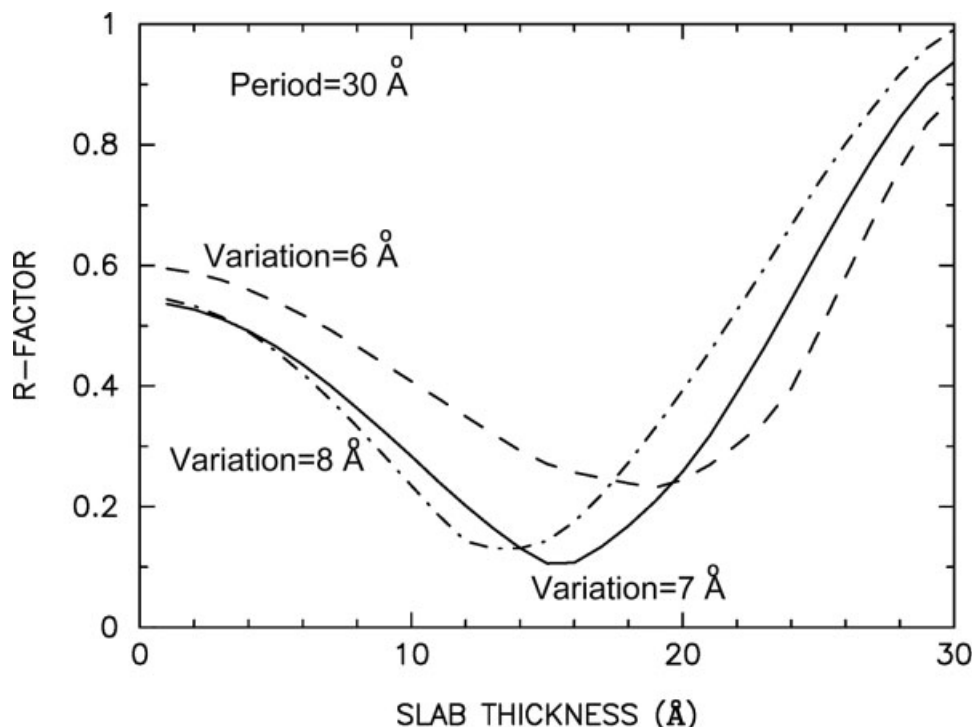


Fig. 3. Dependence of the agreement between the observed and calculated intensities (R -factor) on different slab thicknesses. Shown are the 30 Å period of slab stacking, and three different lattice variations from 6 Å to 8 Å.

model, in which not only the β -chain backbone but also the glutamine side-chains are H-bonded, but later replaced this model with a β -helix model (or “water-filled nanotube”).²⁵ In the current article, we have analyzed the X-ray diffraction patterns from synthetic polyQ peptides having glutamine lengths of 8, 15, 28, and 45 residues, and with different flanking residues at the N- and C-termini [i.e., acetyl/amide, D_2/K_2 and K_2/K_2 (Fig. 1 and Table I). Our objective was to determine the structural consequence of polyglutamine stretches having > 40 residues. Analysis of the X-ray patterns showed that polyglutamine peptides with vastly different numbers of residues all form slablike structures. The Q8 peptide gave a slab thickness of 17 Å, indicating that the slab is composed of five glutamine residues. Thus, three glutamine residues are likely disordered and located in the 14-Å-wide interslab space. Since the slab thickness of Q45 is 23 Å, seven residues are in the crystallite domain. Assuming four residues comprise the turn domain indicates that each Q45 molecule forms four β -chains of seven residues each, a result consistent with a previous theoretical model.^{47,48} This appears to be consistent also with the structural implication based on the kinetic results that peptides consisting of four Q9 or Q10 stretches interspersed with proline-glycine elements undergo spontaneous aggregation as efficiently as the Q45 sequence.¹⁹

The peptide samples we used contain flanking residues—either positive-charged lysines or negative-charged aspartates. While we observed a slightly expanded intersheet distance for Q28 having K/K residues at the N- and

C-termini (likely due to electrostatic repulsion), all samples showed a crystalline β -sheet structure for the polyQ domain. This suggests that the charged flanking residues do not impact the β -chain folding of the polyglutamine, but rather that the local H-bonding between the peptide backbone and glutamine side-chains are the major factors. This conclusion is consistent with the findings, based on experiments using short peptides (e.g., VYK for tau PHF,⁴⁹ LVFF for A β amyloid,^{50,51} NFGSVQ for medin,⁵² and DFNKF for calcitonin⁵³) that amyloid cores are usually composed of short sequences. Thus, the major determinants in amyloid formation are likely to be local interactions involving H-bonding and aromatic residues between short sequences.^{54,55}

One interesting implication of the trend for increasing slab thickness (and hence increasing extended chain lengths traversing the β -sheet) with increasing polyglutamine repeat length is that the aggregate structures for longer repeat lengths tend to feature both more extended chain elements per molecule and also longer extended chains per element. Thus, both thermodynamic and kinetic considerations may explain why longer polyglutamines undergo a more rapid spontaneous aggregation.^{14,24} The thermodynamic stability of the aggregation nucleus is the dominant determinant of nucleation kinetics, which in turn determines the kinetics of spontaneous aggregation. It has already been shown that the trend of increasing spontaneous aggregation rates with increasing polyglutamine repeat length is associated with greater stability of the aggregation nucleus.¹⁵ The optimal length

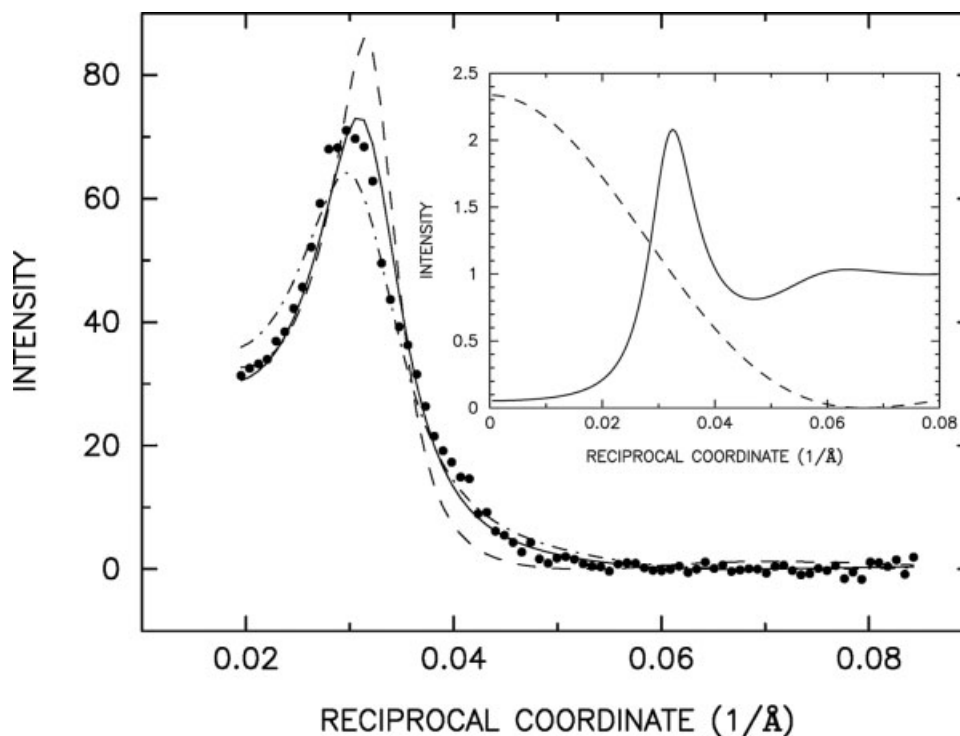


Fig. 4. Observed low-angle X-ray scattering intensity of Q8 and the calculated intensity for different lattice variations. The observed intensity at low angle (\bullet) was derived after background subtraction (here approximated by a straight line). The calculated intensities were derived according to paracrystalline theory for a slab periodicity of 30 Å, and for three different slab thicknesses and lattice variations: **solid line**, 15 Å and 7 Å; **dashed line**, 19 Å and 6 Å; and **dash-dotted line**, 14 Å and 8 Å. The calculated curves were corrected for the powder Lorentz type geometric factor to compare with the observed curve, and normalized so that the area under each curve was 1. The agreement between the observed and the calculated intensities for different slab thickness is shown in Figure 3. As the measured integral beam width (0.003 Å^{-1}) was negligible, the smearing effect due to finite width of the direct beam was not considered. **Inset**: the scattered intensity of a 15-Å-thick slab (**dashed line**) and the paracrystalline interference function (**solid line**) where the lattice has a 30 Å period and 7 Å variation. (Lorentz correction not applied.)

of an extended polypeptide chain in a β -sheet appears to be in the vicinity of seven residues.⁵⁶ Therefore, if the width of the β -sheet in the mature aggregate is replicated in the aggregation nucleus, the longer extended chains present in the antiparallel sheet would be expected to stabilize the nucleus. Kinetics may also play a role. In seeded elongation reactions, polyGln repeat sequences exhibit modest increases in elongation rate with increasing repeat length.²⁴ This may be due, at least in part, to a more rapid turn formation in hairpins that feature longer extended chains. Since increases in elongation kinetics of the nucleus and nascent aggregates can also contribute to observed nucleation kinetics (Bhattacharyya and Wetzel, unpublished observations), this factor might also impact observed rates of spontaneous aggregation.

Comparison With Other Models

As indicated above, the polar zipper model was proposed from X-ray diffraction analysis of the $D_2Q_{15}K_2$ peptide.²⁰ For the oriented pattern, the meridional reflections were interpreted as follows: The strong 4.8 Å reflection is from the hydrogen bonding distance, and the 8.4 Å, 4.2 Å, and 2.8 Å reflections are the second, fourth, and sixth orders,

respectively, of a 16.8 Å repeat in the intersheet direction. The equatorial reflections were interpreted as follows: The 3.6 Å reflection is from a fully extended chain or is the higher order of a long repeat, and the 3.2 Å reflection is the axial repeat for a β pleated sheet. By contrast with their interpretation, we took the 8.4 Å reflection as the lattice constant along the intersheet direction, as no odd order reflections of a 16.8 Å repeat were observed. The broad off-meridional 3.9 Å reflection, which was present but not accounted for by them, we assigned as the (210) reflection, which is characteristic of β -crystallites.³⁵ Moreover, the fiber pattern from this peptide is similar to the oriented fiber patterns from A β 15-28, A β 17-28, and A β 18-28³⁵ in which the cylindrical (rotation) axis is along the chain direction (b axis), and not along the more typical H-bonding direction (a axis).

The more recently proposed water-filled nanotube model^{25,26} is a reinterpretation of the original diffraction data for $D_2Q_{15}K_2$. The 4.8 Å reflection, which is in the fibril direction, is the H-bonding distance between the β -chains in successive turns of a β -helical array, where there are 20 residues per turn. The 20 Å diameter of the tube formed by the main-chain is derived from the 30 Å outer diameter

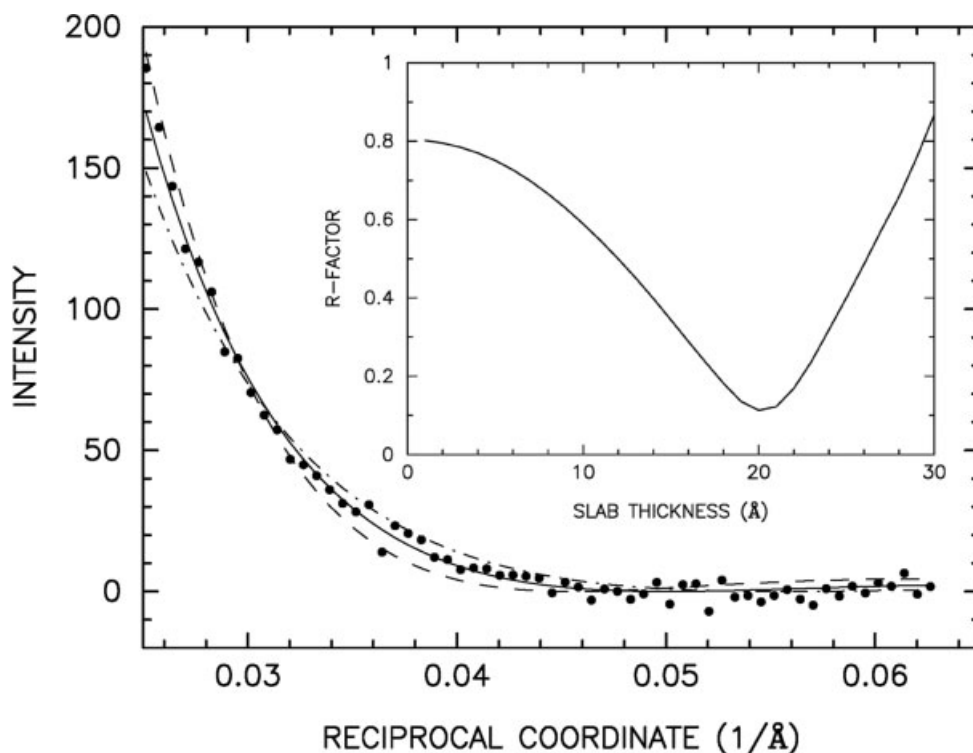


Fig. 5. Low-angle X-ray scattering of Q45 and calculated intensity for the slab model. The observed intensity at low-angle (●) was derived after background subtraction (approximated as a straight line). A 20 Å-thick slab (solid line) best fit the observed intensity. Two models having slightly different slab thickness are shown: 18 Å (dash and dot) and 22 Å (dashed line). Inset: Dependency of relative error between the observed and calculated intensity (*R*-factor) on slab thickness.

(which accounts for the low-angle reflection) and the 5 Å extent of the glutamine side-chain [i.e., $30 \text{ Å} - (5 \text{ Å} \times 2) = 20 \text{ Å}$]. Breaking with their earlier polar zipper model, the authors no longer interpreted the series of 8.4 Å reflections as coming from intersheet stacking, but rather from some unknown origin. The absence of a 10 Å reflection, which had been expected for the intersheet spacing, was interpreted as indicating that the β -sheets are not stacked, but rather that a single β -sheet constitutes the thin wall of the tubular helix. The significant and appealing structural consequence of the nanotube model is that two turns with a total of 40 residues are held together by H-bonds between amides of successive turns, and this forms a stable nucleus for growth of the helical fiber.

Since the rise per one turn is the same as the H-bonding distance ($\sim 4.8 \text{ Å}$), and the asymmetric unit of the discrete helix includes two residues in a β conformation, then the selection rule for this helix ($l/c = n/P + m/h$, where c is the period along the fiber direction, P is pitch, and h is rise per unit) becomes $l = n + 10m$, because the radius is 10 Å , $c = P = 4.8 \text{ Å}$, and $h = 4.8 \text{ Å}/10$. This array for their model predicts that layer lines should be observed at every $(4.8 \text{ Å})^{-1}$. The first layer line at 4.8 Å spacing gives the intensity maximum that arises from the J_1 Bessel function at $R = (35 \text{ Å})^{-1}$ (see Appendix). In fact, the intensity distribution calculated from the model (Fig. 7) reproduces the observed 4.8 Å reflection, and the distance between the helices accounts for the low-angle equatorial intensity maximum at 30 Å ; how-

ever, it does not account for other reflections—the 3.6 and 3.2 Å equatorials, and the 8.3 Å meridional.

A different helical structure for polyQ, one that acts as an ion channel, has also been proposed.⁵⁷ In this case, the rise per residue $h = 0.81 \text{ Å}$ and the pitch $P = 5 \text{ Å}$; thus, the period along the fiber axis c is 20 Å . The selection rule for this helical symmetry is $l = 4n + 25m$. If the helix is made of 40 residues and longer, the length of the helix becomes 32 Å ($= 40 \times 0.8$), which is comparable to the thickness of a biological membrane. This model suggests a correlation between the polyQ length and potential channel formation by Q45 with possible pathological effects; however, the model does not account for the observed diffraction patterns of polyQ aggregates including the strong 4.8 Å reflection.

Atomic Model of Amyloid Fibril: Parallel and Antiparallel Conformations

Models at atomic resolution for the folding of β chains in various amyloid fibrils (e.g., A β and prion PrP peptides) are often contradictory, notwithstanding their extensive characterization by structural biophysical techniques (e.g., X-ray diffraction,^{27,35,45,58–61} solid state NMR,⁶² FTIR,^{63,64} and vibrational Raman optical activity⁶⁵). FTIR of A β 1-40 shows two absorptions at about 1620 cm^{-1} and 1690 cm^{-1} , which are characteristic of an antiparallel β -sheet conformation^{66,67}; however, solid-state NMR of A β 1-40 and also of A β 10-35 deduce parallel β -sheets.^{68,69} A similar parallel arrangement has been determined for A β 1-40 and A β 1-42

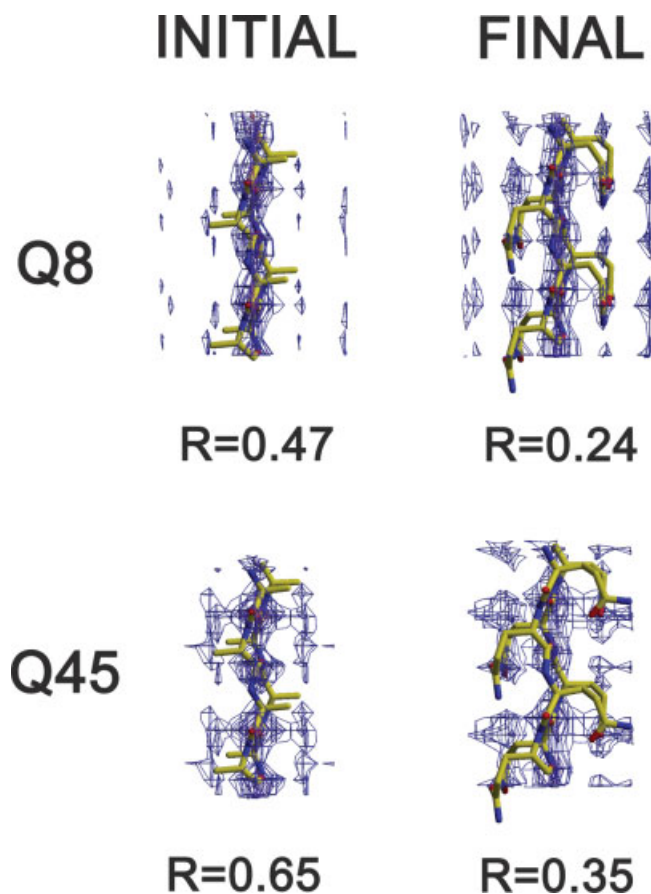


Fig. 6. Electron density profile and skeletal model of Q8 and Q45 assemblies. The unit cell for both peptides was monoclinic, and the lattice constants were $a = 9.73 \text{ \AA}$, $b = 7.14 \text{ \AA}$, $c = 8.16 \text{ \AA}$, and $\gamma = 95.7^\circ$ for Q8, and $a = 9.66 \text{ \AA}$, $b = 7.10 \text{ \AA}$, $c = 8.33 \text{ \AA}$, and $\gamma = 94.0^\circ$ for Q45. The a , b , and c axes are in the directions of hydrogen bond, chain, and intersheet stacking of the β -sheet structure, respectively. The initial map was calculated by using the observed structure amplitudes and the phases derived from the peptide chain backbone of antiparallel β structure including N, C $_{\alpha}$, C', O, C $_{\beta}$. The initial R -factor indicates the relative error between the observed intensity and the calculated one as derived from the initial phase model. The initial map gave high electron densities as expected from the β backbone structure, but also gave additional density peaks that were not accounted for by the initial model. These peaks likely arise from the glutamine side-chains. A new model with side-chains was therefore built on the initial map. The new model gave new phase information, and with the observed structure amplitudes gave the final electron density map. The final R -factor indicates the relative error between the observed intensity and calculated one for the final atomic model. The R -factors (here indicated as fractional value rather than percent) are indicated. The side-chains nearly overlapped in the hydrogen bonding direction. This difference in side-chain conformation and disorder likely accounts for the differences in observed intensity.

by site-directed spin labeling.⁷⁰ Whether there is a turn within the A β 10-35 sequence is also controversial. Solid-state NMR shows from measurements on interstrand distances that the entire sequence of 10-35 forms an extended β -sheet (i.e., no turns⁶⁹); however, molecular dynamics simulation,⁷¹ as well as solid-state NMR⁶² and disulfide cross-linking⁷² studies on A β 1-40 fibrils all indicate a turn within this sequence.

X-ray diffraction patterns of prion analogs provide evidence for an array or lattice structure of β -crystallites,^{45,73}

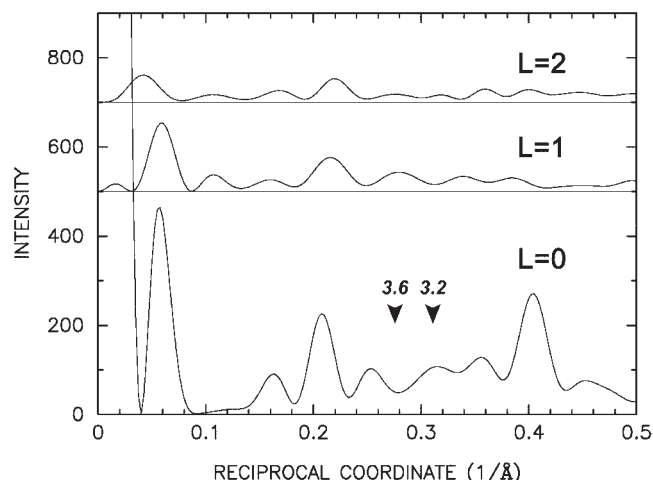


Fig. 7. Diffracted intensity calculated from the helical nanotube model (Perutz et al.²⁵). The intensities along the equator, first layer line at $(4.8 \text{ \AA})^{-1}$ and second layer line at $(2.4 \text{ \AA})^{-1}$ are shown. The equatorial intensity has a strong low-angle maximum at 0.06 \AA^{-1} arising from a 10 \AA -radius helical tubule, and intensity maxima at 0.2 \AA^{-1} and 0.4 \AA^{-1} arising from the distances between the glutamine side-chains. The positions of the intensity maxima at 3.2 \AA and 3.6 \AA that are observed in the diffraction patterns reported previously^{20,25,26} and here are not satisfactorily reproduced in the intensity calculated from the model, nor is the 8.4 \AA meridional spacing.

which is similar to the β -sheet assembly of poly(L-lysine)^{35,74,75} and is also indicated by vibrational Raman optical activity.⁶⁵ EM images⁷⁶ and FTIR studies^{63,64} indicate an antiparallel β -sheet structure for assemblies of prion peptides; however, other studies suggest instead a parallel β -helical structure.^{77,78}

For polyQ peptides, contradictory results have also been reported: *parallel* β -sheets according to IR (for Q78/ataxin-3 fibrils⁷⁹) and structural considerations⁸⁰; and antiparallel β -sheets according to CD and FTIR (for Q22²²). Additional support for an antiparallel arrangement comes from kinetics analyses of Q45 analogs consisting of four Q9 or Q10 segments interspersed with L-ProGly or D-ProGly.^{19,48} The Pro-Gly pairs in these synthetic peptides disfavor extended chain while tolerating β -turn formation (while the D-ProGly sequences favor turn formation) such that the H-bonded β -chains are arranged in an antiparallel conformation.^{19,81,82} Our current X-ray results on polyQ peptides showed that the β -chain arrangement is likely to be antiparallel, particularly as the $\sim 3.5 \text{ \AA}$ rise per unit for the β -chain indicates that the chains are extended, and such β -chains are arranged in an antiparallel conformation by H-bonding as originally proposed by Pauling and Corey.⁸³ Most of the observed reflections can be indexed by even orders of h . Therefore, the a -axis unit length can be approximated as 4.8 \AA , which is one-half the expected unit cell dimension for an antiparallel β -sheet structure. The absence of odd orders has previously been noted for poly-L-alanine⁸⁴ and β keratin,⁸⁵ whereas IR data have favored the antiparallel β -sheet. This apparent contradiction has been resolved by introducing the notion that the crystal structure is statistical with neighboring sheets⁸⁴ or

that there is disorder in sheet packing.⁸⁵ Similar disorders may arise in the polyQ β -sheets.

Amyloidogenic proteins or polypeptides may fold and assemble as either parallel or antiparallel β -sheets. Fibrils of A β 1-40, for example, are folded as parallel β -sheets in which the hydrophilic N-terminal domain is on the surface of the protofilament, and the hydrophobic C-terminal domain defines the protofilament core.^{27,62} Shorter fragments of A β can form antiparallel β -sheets.³⁵ By contrast with the well-defined domain distribution of hydrophilic and hydrophobic side-chains in A β 1-40, the polyQ homopolymer lacks hydrophobic side-chains, instead having H-bond donors and acceptors along the entire chain. Based on the close agreement (low *R*-factors) between the observed and calculated spacings (given the defined unit cells) and the self-consistency of the modeling, our data suggest that the polyQ polypeptide most likely folds as an antiparallel β -sheet.

In summary, our X-ray diffraction study showed that polyQ peptides having different residue numbers can assemble into antiparallel slablike structures, and therefore indicates that Q45 needs to form multiple reverse turns. The specific pathological effect of polyQ having more than 40 residues in CAG extension disease may be related to the rapid accumulation of insoluble fibrils of polyglutamine due to the thermodynamic and kinetic consequences of longer polyQ repeats exhibiting more reverse turns. The significance of reverse turns in initiating or nucleating fibril assembly has been suggested recently from analysis at atomic resolution of powder diffraction patterns of A β 31-35⁶⁰ and by single crystallography of synthetic peptides containing D-Pro.⁸²

ACKNOWLEDGMENTS

We thank Drs. Andrew Bohm, Joseph Toth, and Gretchen Meinke (Department of Biochemistry, Tufts University) for graciously granting us access to their X-ray diffraction facility, and kind assistance in using the Oxford Diffraction Xcalibur PX Ultra. We thank also an anonymous referee for valuable comments on the folding of β -sheets, and Ms. Abby Gross for editorial assistance.

REFERENCES

- Cummings CJ, Zoghbi HY. Fourteen and counting: Unraveling trinucleotide repeat diseases. *Hum Mol Genet* 2000;9:909–916.
- Bates GP, Benn C. The polyglutamine diseases. In: Bates GP, Harper PS, Jones L, editors. *Huntington's disease*. Oxford, UK: Oxford University Press; 2002. p 429–472.
- Zoghbi HY, Orr HT. Glutamine repeats and neurodegeneration. *Annu Rev Neurosci* 2000;23:217–247.
- Zoghbi HY, Orr HT. Polyglutamine diseases: protein cleavage and aggregation. *Curr Opin Neurobiol* 1999;9:566–570.
- Michalik A, Van Broeckhoven C. Pathogenesis of polyglutamine disorders: aggregation revisited. *Hum Mol Genet* 2003;12(Suppl 2):R173–R186.
- Bates G. Huntingtin aggregation and toxicity in Huntington's disease. *Lancet* 2003;361:1642–1644.
- Arrasate M, Mitra S, Schweitzer ES, Segal MR, Finkbeiner S. Inclusion body formation reduces levels of mutant huntingtin and the risk of neuronal death. *Nature* 2004;431:805–810.
- Martin JB. Molecular basis of the neurodegenerative disorders [published erratum]. *N Engl J Med* 1999;340:1970–1980.
- Paulson HL. Protein fate in neurodegenerative proteinopathies: Polyglutamine diseases join the (mis)fold. *Am J Hum Genet* 1999;64:339–345.
- Kaytor MD, Warren ST. Aberrant protein deposition and neurological disease. *J Biol Chem* 1999;274:37507–37510.
- Taylor JP, Hardy J, Fischbeck KH. Toxic proteins in neurodegenerative disease. *Science* 2002;296:1991–1995.
- Trottier Y, Lutz Y, Stevanin G, Imbert G, Devys D, Cancel G, Saudou F, Weber C, David G, Tora L, Agid Y, Brice A, Mandel J-L. Polyglutamine expansion as a pathological epitope in Huntington's disease and four dominant cerebellar ataxias. *Nature* 1995;378:403–406.
- Bennett MJ, Huey-Tubman KE, Herr AB, West AP Jr, Ross SA, Bjorkman PJ. A linear lattice model for poly-glutamine in CAG expansion diseases. *Proc Natl Acad Sci USA* 2002;99:11634–11639.
- Scherzinger E, Sittler A, Schweiger K, Heiser V, Lurz R, Hasenbank R, Bates GP, Lehrach H, Wanker EE. Self-assembly of polyglutamine-containing huntingtin fragments into amyloid-like fibrils: implications for Huntington's disease pathology. *Proc Natl Acad Sci USA* 1999;96:4604–4609.
- Chen S, Ferrone F, Wetzel R. Huntington's disease age-of-onset linked to polyglutamine aggregation nucleation. *Proc Natl Acad Sci USA* 2002;99:11884–11889.
- Masino L, Pastore A. Glutamine repeats: structural hypotheses and neurodegeneration. *Biochem Soc Trans* 2002;30:548–551.
- Perutz MF, Staden R, Moens L, DeBaere I. Polar zippers. *Curr Biol* 1993;3:249–253.
- Lathrop RH, Casale M, Tobias DJ, Marsh JL, Thompson LM. Modeling protein homopolymeric repeats: possible polyglutamine structural motifs for Huntington's disease. *Proc Int Conf Intell Syst Mol Biol* 1998;6:105–114.
- Thakur AK, Wetzel R. Mutational analysis of the structural organization of polyglutamine aggregates. *Proc Natl Acad Sci USA* 2002;99:17014–17019.
- Perutz MF, Johnson T, Suzuki M, Finch JT. Glutamine repeats as polar zippers: their possible role in inherited neurodegenerative diseases. *Proc Natl Acad Sci USA* 1994;91:5355–5358.
- Altschuler EL, Hud NV, Mazrimas JA, Rupp B. Random coil conformation for extended polyglutamine stretches in aqueous soluble monomeric peptides. *J Pept Res* 1997;50:73–75.
- Sharma D, Sharma S, Pasha S, Brahmachari SK. Peptide models for inherited neurodegenerative disorders: conformation and aggregation properties of long polyglutamine peptides with and without interruptions. *FEBS Lett* 1999;456:181–185.
- Chen S, Wetzel R. Solubilization and disaggregation of polyglutamine peptides. *Protein Sci* 2001;10:887–891.
- Chen S, Berthelie V, Yang W, Wetzel R. Polyglutamine aggregation behavior in vitro supports a recruitment mechanism of cytotoxicity. *J Mol Biol* 2001;311:173–182.
- Perutz MF, Finch JT, Berriman J, Lesk A. Amyloid fibers are water-filled nanotubes. *Proc Natl Acad Sci USA* 2002;99:5591–5595.
- Perutz MF, Pope BJ, Owen D, Wanker EE, Scherzinger E. Aggregation of proteins with expanded glutamine and alanine repeats of the glutamine-rich and asparagine-rich domains of Sup35 and of the amyloid β -peptide of amyloid plaques. *Proc Natl Acad Sci USA* 2002;99:5596–5600.
- Inouye H, Kirschner DA. Alzheimer's β amyloid: insights into fibril formation and structure from Congo red binding. In: Harris JR, Fahrenholz F, editors. *Alzheimer's disease: cellular and molecular aspects of amyloid*. Series: Subcellular Biochemistry, Vol. 38. New York: Springer-Verlag; 2005. p 203–224.
- Wetzel R. Protein folding and aggregation in the expanded polyglutamine repeat diseases. In: Buchner J, Kiefhaber T, editors. *The protein folding handbook*. New York: Wiley; 2004. p 1170–1214.
- Oldenbourg R, Phillips WC. Small permanent magnet for fields up to 2.6 T. *Rev Sci Instr* 1986;57:2362–2365.
- Glucksman MJ, Hay RD, Makowski L. X-ray diffraction from magnetically oriented solutions of macromolecular assemblies. *Science* 1986;231:1273–1276.
- Pauling L. Diamagnetic anisotropy of the peptide group. *Proc Natl Acad Sci USA* 1979;76:2293–2294.
- Worcester DL. Structural origins of diamagnetic anisotropy in proteins. *Proc Natl Acad Sci USA* 1978;75:5475–5477.
- Graf DL. Crystallographic tables for the rhombohedral carbonates. *Am Mineralogist* 1961;46:1283–1316.

34. Inouye H, Karthigasan J, Kirschner DA. Membrane structure in isolated and intact myelins. *Biophys J* 1989;56:129–137.
35. Inouye H, Fraser PE, Kirschner DA. Structure of β -crystallite assemblies formed by Alzheimer β -amyloid protein analogues: analysis by X-ray diffraction. *Biophys J* 1993;64:502–519.
36. Inouye H, Kirschner DA. X-ray diffraction analysis of scrapie prion: intermediate and folded structures in a peptide containing two putative α -helices. *J Mol Biol* 1997;268:375–389.
37. Guex N, Diemand A, Peitsch MC. Protein modelling for all. *Trends Biochem Sci* 1999;24:364–367.
38. McRee DE. "XtalView/Xfit"—a versatile program for manipulating atomic coordinates and electron density. *J Struct Biol* 1999;125:156–165.
39. Vriend G. WHAT IF: a molecular modeling and drug design program. *J Mol Graph* 1990;8:52–56.
40. Kraulis PJ. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J Appl Crystallogr* 1991;24:946–950.
41. Sayle RA, Milner-White EJ. RASMOL: biomolecular graphics for all. *Trends Biochem Sci* 1995;20:374–376.
42. Sikorski P, Atkins E. New model for crystalline polyglutamine assemblies and their connection with amyloid fibrils. *Biomacromolecules* 2005;6:425–432.
43. Inouye H, Bond J, Baldwin MA, Ball HL, Prusiner SB, Kirschner DA. Structural changes in a hydrophobic domain of the prion protein induced by hydration and by Ala \rightarrow Val and Pro \rightarrow Leu substitutions. *J Mol Biol* 2000;300:1285–1298.
44. Vainshtein BK. Diffraction of X-rays by chain molecules. Amsterdam: Elsevier; 1966. 414 p.
45. Inouye H, Kirschner DA. X-ray fibre diffraction analysis of assemblies formed by prion-related peptides: polymorphism of the heterodimer interface between PrP^C and PrP^{Sc}. *Fibre Diffraction Rev* 2003;11:102–112.
46. Inouye H. X-ray scattering from a discrete helix with cumulative angular and translational disorders. *Acta Crystallogr A* 1994;50:644–646.
47. Starikov EB, Lehrach, Wanker E. Folding of oligoglutamines: a theoretical approach based upon thermodynamics and molecular mechanics. *J Biomol Struct Dyn* 1999;17:409–427.
48. Ross CA, Poirier MA, Wanker EE, Amzel M. Polyglutamine fibrillogenesis: the pathway unfolds. *Proc Natl Acad Sci USA* 2003;100:1–3.
49. Goux WJ, Kopplin L, Nguyen AD, Leak K, Rutkowsky M, Shanmuganandam VD, Sharma D, Inouye H, Kirschner DA. Twisted and straight filaments formed from short segments of tau protein lacking β -turn regions. *J Biol Chem* 2004;279:26868–26875.
50. Inouye H, Kirschner DA. Refined fibril structures: the hydrophobic core in Alzheimer's amyloid β -protein and prion as revealed by X-ray diffraction. In: Bock GR, Goode JA, editors. *Nature and origin of amyloid fibrils*. CIBA Foundation Symposium No. 199. Chichester, UK: Wiley; 1996. p 22–39.
51. Tjernberg LO, Tjernberg A, Bark N, Shi Y, Ruzsicska BP, Bu ZM, Thyberg J, Callaway DJE. Assembling amyloid fibrils from designed structures containing a significant amyloid β -peptide fragment. *Biochem J* 2002;366:343–351.
52. Reches M, Gazit E. Amyloidogenic hexapeptide fragment of medin: homology to functional islet amyloid polypeptide fragments. *Amyloid* 2004;11:81–89.
53. Reches M, Porat Y, Gazit E. Amyloid fibril formation by pentapeptide and tetrapeptide fragments of human calcitonin. *J Biol Chem* 2002;277:35475–35480.
54. Gazit E. A possible role for pi-stacking in the self-assembly of amyloid fibrils. *FASEB J* 2002;16:77–83.
55. Makin OS, Atkins E, Sikorski P, Johansson J, Serpell LC. Molecular basis for amyloid fibril formation and stability. *Proc Natl Acad Sci USA* 2005;102:315–320.
56. Stanger HE, Syud FA, Espinosa JF, Girit I, Muir T, Gellman SH. Length-dependent stability and strand length limits in antiparallel β -sheet secondary structure. *Proc Natl Acad Sci USA* 2001;98:12015–12020.
57. Monoi H, Futaki S, Kugimiya H, Minakata H, Yoshihara K. Poly-L-glutamine forms cation channels: relevance to the pathogenesis of the polyglutamine disease. *Biophys J* 2000;78:2892–2899.
58. Blake C, Serpell L. Synchrotron X-ray studies suggest that the core of the transthyretin amyloid fibril is a continuous β -sheet helix. *Structure* 1996;4:989–998.
59. Malinchik SB, Inouye H, Szymowski KE, Kirschner DA. Structural analysis of Alzheimer's β (1–40) amyloid: protofilament assembly of tubular fibrils. *Biophys J* 1998;74:537–545.
60. Bond JP, Deverin SP, Inouye H, El-Agnaf OM, Teeter MM, Kirschner DA. Assemblies of Alzheimer's peptides A β 25–35 and A β 31–35: reverse-turn conformation and side-chain interactions revealed by X-ray diffraction. *J Struct Biol* 2003;141:156–170.
61. Makin OS, Serpell LC. The structure of amyloid. *Fibre Diffraction Rev* 2004;12:29–35.
62. Tycko R. Insights into the amyloid folding problem from solid-state NMR. *Biochemistry* 2003;42:3151–3159.
63. Silva RAGD, Barber-Armstrong W, Decatur SM. The organization and assembly of a β -sheet formed by a prion peptide in solution: an isotope-edited FTIR study. *J Am Chem Soc* 2003;125:13674–13675.
64. Sokolowski F, Modler AJ, Masuch R, Zirwer D, Baier M, Lutsch G, Moss DA, Gast K, Naumann D. Formation of critical oligomers is a key event during conformational transition of recombinant Syrian hamster prion protein. *J Biol Chem* 2003;278:40481–40492.
65. McColl IH, Blanch EW, Gill AC, Rhie AGO, Ritchie MA, Hecht L, Nielsen K, Barron LD. A new perspective on β -sheet structures using vibrational Raman optical activity: From poly(L-lysine) to the prion protein. *J Am Chem Soc* 2003;125:10019–10026.
66. Hilbich C, Kisters-Woike B, Reed J, Masters CL, Beyreuther K. Aggregation and secondary structure of synthetic amyloid β A4 peptides of Alzheimer's disease. *J Mol Biol* 1991;218:149–163.
67. Fraser PE, Nguyen JT, Inouye H, Surewicz WK, Selkoe DJ, Podlisny MB, Kirschner DA. Fibril formation by primate, rodent, and Dutch-hemorrhagic analogues of Alzheimer amyloid β -protein. *Biochemistry* 1992;31:10716–10723.
68. Antzutkin ON, Balbach JJ, Leapman RD, Rizzo NW, Reed J, Tycko R. Multiple quantum solid-state NMR indicates a parallel, not antiparallel, organization of β -sheets in Alzheimer's β -amyloid fibrils. *Proc Natl Acad Sci USA* 2000;97:13045–13050.
69. Lynn DG, Meredith SC. Review: model peptides and the physicochemical approach to β -amyloids. *J Struct Biol* 2000;130:153–173.
70. Torok M, Milton S, Kaye R, Wu P, McIntire T, Glabe CG, Langen R. Structural and dynamic features of Alzheimer's A β peptide in amyloid fibrils studied by site-directed spin labeling. *J Biol Chem* 2002;277:40810–40815.
71. Ma B, Nussinov R. Stabilities and conformations of Alzheimer's β -amyloid peptide oligomers (A β 16–22, A β 16–35, and A β 10–35): sequence effects. *Proc Natl Acad Sci USA* 2002;99:14126–14131.
72. Shivaprasad S, Wetzel R. An intersheet packing interaction in A β fibrils mapped by disulfide crosslinking. *Biochemistry* 2004;43:15310–15317.
73. Nguyen JT, Inouye H, Baldwin MA, Fletterick R, Cohen FE, Prusiner SB, Kirschner DA. X-ray diffraction of scrapie prion rods and PrP peptides. *J Mol Biol* 1995;252:412–422.
74. Shmueli U, Traub W. An X-ray diffraction study of poly-L-lysine hydrochloride. *J Mol Biol* 1965;12:205–214.
75. Padden FJ Jr, Keith HD, Giannoni G. Single crystals of poly-L-lysine. *Biopolymers* 1969;7:793–804.
76. DeMarco ML, Daggett V. From conversion to aggregation: protofibril formation of the prion protein. *Proc Natl Acad Sci USA* 2004;101:2293–2298.
77. Wille H, Michelitsch MD, Guenebaut V, Supattapone S, Serban A, Cohen FE, Agard DA, Prusiner SB. Structural studies of the scrapie prion protein by electron crystallography. *Proc Natl Acad Sci USA* 2002;99:3563–3568.
78. Govaerts C, Wille H, Prusiner SB, Cohen FE. Evidence for assembly of prions with left-handed β -helices into trimers. *Proc Natl Acad Sci USA* 2004;101:8342–8347.
79. Bevivino AE, Loll PJ. An expanded glutamine repeat destabilizes native ataxin-3 structure and mediates formation of parallel β -fibrils. *Proc Natl Acad Sci USA* 2001;98:11955–11960.
80. Kajava AV, Baxa U, Wickner RB, Steven AC. A model for Ure2p prion filaments and other amyloids: the parallel superpleated β -structure. *Proc Natl Acad Sci USA* 2004;101:7885–7890.
81. Inouye H, Bond JE, Deverin SP, Lim A, Costello CE, Kirschner DA. Molecular organization of amyloid protofilament-like assembly of betabellin 15D: Helical array of β -sandwiches. *Biophys J* 2002;83:1716–1727.
82. Aravinda S, Harini VV, Shamala N, Das C, Balaram P. Structure and assembly of designed β -hairpin peptides in crystals as models for β -sheet aggregation. *Biochemistry* 2004;43:1832–1846.
83. Pauling L, Corey RB. Configurations of polypeptide chains with

- avored orientations around single bonds: two new pleated sheets. Proc Natl Acad Sci USA 1951;37:729–740.
84. Arnott S, Dover D, Elliott A. Structure of β -poly-L-alanine: refined atomic co-ordinates for an anti-parallel β -pleated sheet. J Mol Biol 1967;30:201–208.
 85. Fraser RDB, MacRae TP, Parry DAD, Suzuki E. The structure of β -keratin. Polymer 1969;10:810–826.
 86. Cochran W, Crick FHC, Vand V. The structure of synthetic polypeptides: I. The transform of atoms on a helix. Acta Crystallogr 1952;5:581–586.
 87. Franklin RE, Klug A. The splitting of layer lines in X-ray fibre diagrams of helical structures: application to tobacco mosaic virus. Acta Crystallogr 1955;8:777–780.

APPENDIX

Testing the β -Helical Nanotube Model of Polyglutamine²⁵

Comparing the intensity calculated from the model to the observed diffraction pattern for Q15 provided a test of this model. As shown in the analysis of betabellin 15D,⁸¹ a discrete helix can be approximated as a tubular structure having constant density. Along the equator the intensity $I(R)$ of a tubular structure is given as a function of the radial component in cylindrical coordinates (R). $I(R)$ is proportional to the product of the structure amplitude and the interference function $Z(R)$ as $|F(R)|^2 Z(R)$. For a tubular cylinder of radius r_0 , $F(R) = J_0(2\pi r_0 R)$. After cylindrical averaging the interference function $Z(R) = \sum_j \sum_k J_0(2\pi r_{jk} R)$, where r_{jk} is the distance between the vector r_j and vector r_k in cylindrical coordinates.³⁵ The interference function never becomes zero, and gives the first intensity maximum at $R = 7.037/(2\pi a)$, where a is the intercylinder distance. If there is no interference between cylinders, then $Z(R) = 1$. The first intensity maximum, in this case, corresponds to the intensity maximum of the structure amplitude (i.e., $2\pi r_0 R = 3.770$). If the observed low-angle reflection at 31 Å spacing arises from the first intensity maximum of the interference term, then the intercylinder distance a is 34.7 Å. If the tubular radius is 10 Å, the structure amplitude gives the first intensity maximum at 17 Å, and first zero at 26 Å (see Fig. 7).

According to the nanotube model,²⁵ the selection rule for a discrete helix is given by $l/c = n/P + m/h$, where c is the period along the fiber direction, P is pitch, h is rise per unit, $c = P = 4.8$ Å, and $h = 4.8$ Å/10; then for a helical radius $r_h = 10$ Å, the selection rule becomes $l = n + 10m$. This model predicts that layer lines will be observed at every $(4.8 \text{ Å})^{-1}$ in the fiber direction. As the 4.8-Å reflection corresponds to the J_1 Bessel function, then this reflection will occur at $R = 1.8/(2\pi r_0) = (35 \text{ Å})^{-1}$. An antiparallel β -sheet is expected to give a 9.4-Å reflection

along the hydrogen bonding direction, but this model predicts the absence of the 9.4-Å layer line. Thus, favoring the model would be the absence of the 9.4-Å reflection, and also a split of the 4.8-Å reflection; however, experimental proof may be difficult to achieve because of the fibril disorientation.

Fourier Transform of the β Helix

A helical array of polyglutamine was constructed in the following manner. First, the Cartesian coordinates (x, y, z) of the two residues in β conformation were defined. Second, the center of gravity of the molecule (x_0, y_0, z_0) was determined. Third, the coordinates (x, y, z) were modified according to $(x - x_0 + r_h, y - y_0, z - z_0)$, where r_h is the helical radius. Fourth, the Cartesian coordinates were translated to cylindrical coordinates (r, ϕ, z), where $r^2 = (x - x_0 + r_h)^2 + (y - y_0)^2$ and $\phi = \arccos[(x - x_0 + r_h)/r_h]$. Fifth, a series of coordinates was calculated by choosing sequential integer n ($r, \phi + 2\pi n h/P, z + n h$). Finally, the cylindrical coordinates were converted back to Cartesian coordinates. The modeled fibril structure in the helical array was derived using the parameters $r_h = 10$ Å, $c = P = 4.8$ Å, and $h = 0.48$ Å. The size of the helix was calculated to be 30 Å, which is similar to the observed low-angle reflection.

The cylindrically averaged intensity of the discrete helix array was calculated according to the following equation.⁸¹ If the disorder parameters are neglected, the helical intensity reduces to the one for the ideal helix^{86,87}:

$$\langle I(R, \Phi, l/c) \rangle_\Phi = \sum_n (A_n^2 + B_n^2), \text{ where}$$

$$A_n = \sum_j f_j J_n(2\pi r_j R) \cos[n(\pi/2 - \phi_j) + 2\pi l z_j / c],$$

$$B_n = \sum_j f_j J_n(2\pi r_j R) \sin[n(\pi/2 - \phi_j) + 2\pi l z_j / c].$$

Here, the structure factor is written as

$$F(R, \Phi, l/c) = \sum_n G_n \exp[in(\Phi + \pi/2)], \text{ where}$$

$$G_n(R) = \sum_j f_j J_n(2\pi r_j R) \exp[i[n(\phi_j - \pi/2) + 2\pi l z_j / c]].$$

The layer line number l is related to the Bessel term n according to the helical selection rule $l = un + vm$, where all variables are integers. The atomic coordinate j refers to the atoms in the asymmetric unit.