Amyloid-like Features of Polyglutamine Aggregates and Their Assembly Kinetics†

Songming Chen, Valerie Berthelier, J. Bradley Hamilton,‡ Brian O'Nuallain, and Ronald Wetzel*

Graduate School of Medicine, University of Tennessee Medical Center, 1924 Alcoa Highway, Knoxville, Tennessee 37920

Received September 7, 2001; Revised Manuscript Received March 13, 2002

ABSTRACT: The repeat length-dependent tendency of the polyglutamine sequences of certain proteins to form aggregates may underlie the cytotoxicity of these sequences in expanded CAG repeat diseases such as Huntington's disease. We report here a number of features of various polyglutamine (polyGln) aggregates and their assembly pathways that bear a resemblance to generally recognized defining features of amyloid fibrils. PolyGln aggregation kinetics displays concentration and length dependence and a lag phase that can be abbreviated by seeding. PolyGln aggregates exhibit classical β-sheet-rich circular dichroism spectra consistent with an amyloid-like substructure. The fundamental structural unit of all the in vitro aggregates described here is a filament about 3 nm in width, resembling the protofibrillar intermediates in amyloid fibril assembly. We observed these filamentous structures either as isolated threads, as components of ribbonlike sheets, or, rarely, in amyloid-like twisted fibrils. All of the polyGln aggregates described here bind thioflavin T and shift its fluorescence spectrum. Although all polyGln aggregates tested bind the dye Congo red, only aggregates of a relatively long polyGln peptide exhibit Congo red birefringence, and this birefringence is only observed in a small portion of these aggregates. Remarkably, a monoclonal antibody with high selectivity for a generic amyloid fibril conformational epitope is capable of binding polyGln aggregates. Thus, polyGln aggregates exhibit most of the characteristic features of amyloid, but the twisted fibril structure with Congo red birefringence is not the predominant form in the polyGln repeat length range studied here. We also find that polyGln peptides exhibit an unusual freezing-dependent aggregation that appears to be caused by the freeze concentration of peptide and/or buffer components. This is of both fundamental and practical significance. PolyGln aggregation is revealed to be a highly specific process consistent with a significant degree of order in the molecular structure of the product. This ordered structure, or the assembly process leading to it, may be responsible for the cell-specific neuronal degeneration observed in Huntington's and other expanded CAG repeat diseases.

There are eight hereditary diseases, including Huntington's disease, in which the genetic expansion of a CAG repeat in an open reading frame leads to neurodegeneration (*1*). The neurotoxicity in these diseases appears to derive primarily from a toxic gain of function associated with the expansion of the CAG_N-encoded polyglutamine (polyGln) repeat. Although not uniformly accepted, the most attractive mechanistic hypothesis linking CAG repeat expansion to toxicity involves the tendency of longer polyGln sequences, regardless of protein context, to form insoluble aggregates (*2*−*6*). The potential role of polyGln aggregation in the disease mechanism makes it particularly important to better understand the process and products of the self-assembly of these sequences.

Previous studies demonstrated the tendency of polyGln sequences to form insoluble aggregates. Perutz and coworkers showed that a chemically synthesized Q_{15} peptide, under certain conditions, aggressively aggregates to form β-sheet-rich structures appearing as short filaments in electron micrographs (*7*). Wanker and colleagues showed by electron microscopy that proteins containing longer polyGln peptides, generated by recombinant expression, aggregate in a polyGln repeat length-dependent manner to form amyloid, or amyloid-like, fibrils (*8*, *9*). On this basis, it might be justified to include the expanded polyglutamine diseases into the growing family of amyloid diseases known to affect either the brain (*10*) or organs in the periphery (*11*, *12*). However, amyloid fibrils exhibit a number of defining characteristics (*11*), and there is no general agreement on what the minimum criteria should be for classifying an aggregate as amyloid nor on how to classify aggregates that meet some criteria but fail others. Because the structures and properties of polyGln aggregates may well define their cytotoxic activity (*13*), we have undertaken an examination of the structures of polyGln aggregates by a number of experimental procedures.

With the recent development of protocols for solubilizing chemically synthesized polyGln peptides of repeat lengths up to at least Q_{50} (*14*), it has become possible to conduct controlled studies of the aggregation process and to characterize the structures of both monomeric peptides and aggregates. In this report we present data demonstrating that polyGln aggregation is an ordered process resembling in many respects amyloid fibril formation in both assembly kinetics and aggregate structure. At the same time, we find that polyGln peptides can give rise to a number of different

† Supported in part by the Lindsay Young Alzheimer's Disease Research Gift Fund, the Reuben Louise Cates Mount Research Endowment, and grants from the Hereditary Disease Foundation.

* To whom correspondence should be addressed.

‡ Present address: RHeoGene, 706 Forest St., Charlottesville, VA 22903.

10.1021/bi011772q CCC: $22.00
Published on Web 05/17/2002

aggregate morphologies, some of which, while ordered, do not resemble amyloid in all respects. For instance, we describe an unusual freeze concentration-dependent aggregation that generates an end-product aggregate that more resembles the protofibrils described as intermediates in Aβ amyloid fibril assembly.

MATERIALS AND METHODS

General Materials and Methods. All peptides were obtained by custom solid-phase syntheses from the Keck Biotechnology Center at Yale University (http://info.med.yale.edu/wmkeck/). Flanking sets of charged residues were included to enhance general solubility, with a general attempt to ensure that peptides possess significant charge in the neutral pH region. Structures and purities of polyGln peptides were confirmed by mass spectrometry. Solid-phase synthesis of long polyGln tracks can generate significant levels of deletion peptides. For example, MS analysis revealed that the peptide synthesized as Q_{30} was predominantly 29 glutamines in length, with minor amounts of Q_{27}, Q_{28}, and Q_{30}, giving a weight-average repeat length of Q_{28}. For most of the experiments described here, this heterogeneous product was used. Previously, we found no significant difference in the aggregation kinetics between an unpurified Q_{30} mixture and a major component purified from that mixture (*15*).

Wild-type Aβ(1−40) was synthesized and purified by the Keck Center. Amyloid fibrils were prepared from Aβ as described previously (*16, 17*).

Circular dichroism (CD)[1] spectroscopy was performed on an OLIS RSM CD spectrophotometer. Solutions or suspensions of 0.4 mg/mL disaggregated or aggregated polyGln peptide, in 10 mM Tris−TFA, pH 7.2, were placed in a 1 mm path-length cylindrical cuvette. Ellipticity was measured with a 1 nm bandwidth and 1 nm step size at 37 °C. Accumulation time at each wavelength was variable, adjusted by the OLIS software to ensure a constant amount of light delivered to the detector at each wavelength. Buffer baselines were subtracted from the spectra. The spectra reported here are representative of an average of six runs, smoothed after averaging using a fast Fourier transform method using Origin 6.0 software (Microcal Software, Inc.). The data were expressed as mean residue molar ellipticities [Θ].

Light scattering measurements were performed by transferring 300 μL aliquots of aggregation reactions into a 4 mm × 4 mm quartz fluorescence cuvette and reading the apparent fluorescence in a Perkin-Elmer LS50B luminescence spectrophotomer, in which both the emission and excitation wavelengths were set at 450 nm, and both slit widths were set at less than 2 nm.

For electron microscopy, aggregates were fixed to mica grids, negatively stained with a 0.25% potassium phosphotungstate solution, and analyzed by transmission electron microscopy on a Hitachi H-600 electron microscope. Electron microscopy was performed at the microscopy center in the University of Tennessee Division of Biology.

Solubilization, Disaggregation, and Quantitation of PolyGln Peptides. Disaggregation and solubilization were performed as described (*14*). Thus, peptides were dissolved in a 50:50 mixture of TFA (Pierce, Rockford, IL) and HFIP (Sigma, St. Louis, MO) at 0.5 mg/mL, vigorously agitated, and incubated for at least 30 min at room temperature. The volatile solvents were evaporated under a stream of argon in a fume hood, and the peptide residue was solubilized to a concentration of 0.5 mg/mL in water adjusted to pH 3 with TFA. The exact peptide concentration was determined by reverse-phase high-performance liquid chromatography (RP-HPLC) on a Series 1100 chromatograph (Agilent, Palo Alto, CA) using a Zorbax SB-C3 column. An aliquot of the solubilized peptide was injected onto the column and analyzed on a 0−50% (v/v) acetonitrile gradient with 0.05% TFA applied at a rate of 2% per minute. The peak area obtained at 215 nm was converted to micrograms of peptide using a weight-based standard curve[2] previously established with the peptide $K_2Q_{15}K_2$, whose concentration was determined independently by amino acid composition analysis (Commonwealth Biotechnologies, Inc., Richmond, VA). The HPLC analysis was conducted immediately after the solubilization step to minimize the waiting time of the disaggregated concentrated peptides. Solutions of peptides should not be stored prior to completing subsequent manipulations as described below.

Measurements of nonaggregated polyGln levels during aggregation reactions were obtained by transferring a 300 μL aliquot of the reaction mixture into an Eppendorf tube and centrifuging at 14000 rpm (20800*g*) for 30 min at 25 °C. An aliquot of 270 μL was carefully drawn from the supernatant and diluted with 30 μL of aqueous 0.5% TFA. This sample was analyzed by HPLC as described above. The area of the polyGln peak was determined by integration, using the Agilent software, and converted to micrograms of polyGln using the peak area from a Q_{15} standard curve. A 40 μM sample of $K_2Q_{42}K_2$ in PBS spiked with 0.05% TFA was stable for 4 h in an autosampler vial before peptide recovery began to decrease due to aggregation in the vial (data not shown).

Preparation of Polyglutamine Aggregates. Aggregates were prepared from polyGln peptides solubilized and disaggregated as described above. Aggregates grown at 37 °C were prepared as described previously (*13*) by incubation of appropriate concentrations of polyGln peptide in PBSA and monitoring by light scattering, ThT, or HPLC, as described below. Other aggregates were prepared by the freeze-concentration method as described in this paper. In this method, a 10 μM solution in PBSA of disaggregated polyGln peptide was snap-frozen in liquid N_2 and then incubated at −20 °C for 48 h before thawing at room temperature. Mature aggregates were harvested by centrifugation at 20800*g* for 30 min at 37 °C. The pellet was resuspended in PBSA at various concentrations depending on intended use. To develop sonicated aggregates for use as seeds, aggregates were suspended to a final concentration

[1] Abbreviations: TFA, trifluoroacetic acid; HFIP, 1,1,1,3,3,3-hexafluoro-2-propanol; CD, circular dichroism; CR, Congo red; PBS, phosphate-buffered saline; PBSA, PBS plus 0.05% sodium azide; ThT, thioflavin T.

[2] Since absorbance of polypeptides at 215 nm is dominated by the peptide bond, and since the molecular weights of polyGln sequences increase in proportion to the number of peptide bonds in the sequence, we can use this standard curve to determine the weight concentrations of polyGln peptides of different lengths.

of 180 μM (monomer equivalent) and sonicated on ice with a probe sonicator for six successive pulses of 30 s each.

Thioflavin Binding Experiments. Thioflavin binding-dependent fluorescence was routinely assessed on kinetic time points as follows. Immediately after the light scattering measurement (see above), a 10 μL aliquot of a 2.5 mM stock of ThT was added, with stirring, to the aliquot of the aggregation reaction in the fluorescence cuvette, and the fluorometer was reset to an excitation wavelength of 450 nm (slit width, 5 nm) and an emission wavelength of 489 nm (slit width, 10 nm). Data were collected over a 10 min period and time-averaged using the Perkin-Elmer software.

Characterization of the binding of ThT to polyGln aggregates was carried out by ThT titration as described previously (*18*, *19*). To analyze the data, we followed the procedure of Naiki et al. (*18*). Because of a number of assumptions that are made in this procedure, the binding constants generated are considered only approximations of the K_d value; however, these numbers are useful in being able to make a number of comparisons among different aggregates. Since the fluorescence of the aggregate-bound ThT is linear from 0 to 80 μM aggregates (monomer equivalents) for all peptides (data not shown), we can use such curves to determine the concentration of the ThT-bound aggregate. Scatchard plot analysis was used to generate the K_d values. As described in Results, the K_d value obtained in this way for Aβ(1−40) agreed very well with values obtained by others using similar approaches.

Congo Red Staining and Birefringence. Congo red staining was conducted using a modification of the protocol of Puchtler et al. (*20*). Thus, 5 μL of a 0.1% fresh alkaline solution of Congo red (Sigma) was added to 200 μL of a suspension of polyGln aggregates (concentration equivalent to 10 μM monomeric polyGln). The sample was centrifuged at 20800*g* for 10 min and the pellet washed with H_2O to remove the excess CR. This wash step was repeated three times after which the pellet was resuspended in 20 μL of H_2O. Two drops of this CR-bound polyGln aggregate suspension was allowed to dry on a glass microscope slide for 5 min. Birefringence was determined with an Olympus microscope (40×) equipped with a polarizing stage.

Monoclonal Antibody Binding to Aggregates. The day before an experiment, high-binding microtiter wells (COSTAR) were coated with protein aggregates suspended in PBSA (100 ng per well) by incubating the uncovered plates overnight at 37 °C in an incubator oven. Plates were washed three times with PBSA containing 0.05% Tween 20. These are the standard washing conditions used throughout this protocol. [Essentially quantitative nonspecific binding of test proteins to the plastic was confirmed by a recovery experiment (*21*).] Wells were blocked using 300 μL per well of 1% gelatin in PBSA at 37 °C for 1 h. A tissue culture supernatant containing WO1 was diluted 1:2 in PBS containing 0.05% sodium azide, then added to each well (100 μL per well), and incubated as before. Plates were washed three times and then incubated with biotinylated goat anti-mouse antibody (Sigma). The binding of antibody to each modified protein was then detected in triplicate by time-resolved fluorescence using a streptavidin−europium detection system (*22*).

RESULTS

Formation of Aggregates. We designed polyGln peptides to include pairs of charged residues to improve general solubility (*7*, *13*−*15*, *23*). Different combinations of charged residues were explored. As reflected in their yields in 37 °C aggregation reactions (not shown), we found that the peptide $D_2Q_{15}K_2$ aggregated considerably better than did $DK_2Q_{15}DK_2$, which in turn aggregated more readily than did $K_2Q_{15}K_2$. The variation in yields is presumably a direct consequence of differences in the critical concentration, C_r, which reflects the position of the aggregation equilibrium (*24*). We surmise that the apparent enhanced thermodynamic stability of the $D_2Q_{15}K_2$ peptide, with respect to the other Q_{15} polyGln peptides, might be due either to a nonspecific structural effect, because of the low net charge of this peptide at pH 7, or to a specific structural effect, because of pairing of opposite charges contributing to antiparallel alignment of peptides to facilitate β-sheet formation.

We also tested whether flanking charged residues are required for kinetic solubility, as follows. A Q_{37} molecule with no flanking charged residues was obtained. Although this peptide was found to dissolve in our normal TFA/HFIP protocol, the peptide after removal of volatile solvent by an argon stream exhibited only limited solubility at pH 3 (2.7 μM by HPLC analysis). The peptide aggregates extremely rapidly when adjusted to pH 7, in contrast to a similar peptide with flanking Lys residues. At the same time, the aggregates of this peptide exhibit similar filamentous structure by electron microscopy to aggregates of polyGln peptides containing flanking residues (not shown). Thus, flanking charged residues appear to greatly enhance polyGln solubility and to slow polyGln aggregation but do not appear to qualitatively alter aggregate morphology compared to a simple polyGln peptide. By choosing as a constant sequence context flanking pairs of Lys residues for most of the peptides described here, we feel we have been able to extend the repeat length range over which we can control and analyze polyGln aggregation, without compromising the significance of the results.

Processing chemically synthesized polyGln peptides by the protocol described in the Materials and Methods section produces a disaggregated peptide solution that exhibits low light scattering and ThT fluorescence and that does not pellet out when centrifuged at high speed. Figure 1 shows that the soluble polyGln peptide obtained in this way exhibits a random coil circular dichroism spectrum (*25*). When a PBSA solution of this soluble peptide is incubated at 37 °C, the peptide aggregates. Figure 2 shows the aggregation kinetics in PBSA and 37 °C of a 20 μM solution of the peptide $K_2Q_{28}K_2$ that had been disaggregated by our standard solvent treatment but had not been further disaggregated by ultracentrifugation. The figure shows that aggregation begins within the first 20 h and that the progress of the aggregation reaction is very similar whether monitored by Rayleigh light scattering, ThT fluorescence, or solubility (HPLC analysis of centrifugation supernatants). Figure 3 shows the 37 °C aggregation kinetics for a 20 μM PBSA solution of the same ($K_2Q_{28}K_2$) peptide after it had been disaggregated by both TFA/HFIP treatment and ultracentrifugation. The lag phase now increases to 4−8 days compared to that seen in Figure 2, strongly suggesting that ultracentrifugation removes mi-

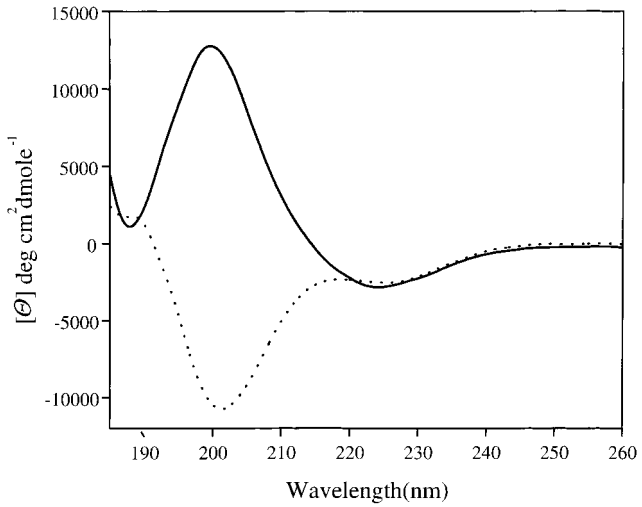

FIGURE 1: Circular dichroism spectra of $K_2Q_{42}K_2$ either freshly solubilized (—) or aggregated at 37 °C (---) in 10 mM Tris−TFA buffer, pH 7.0. Spectra were collected and processed as described in Materials and Methods.

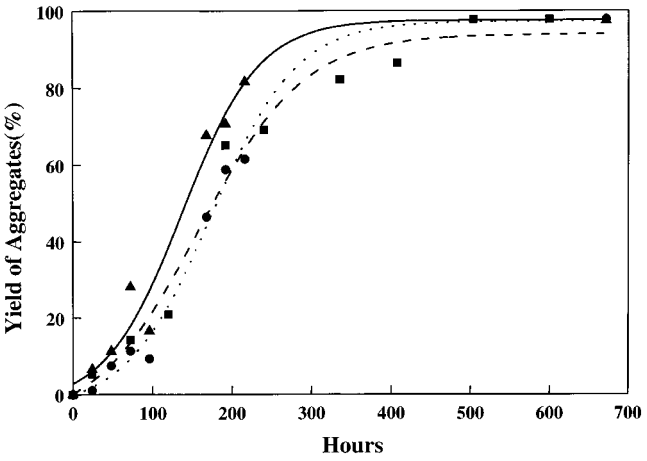

FIGURE 2: Time course of aggregation of a 20 μM solution of $K_2Q_{28}K_2$ in PBSA at 37 °C, as monitored by light scattering (···), ThT fluorescence (---), and solubility (—). The peptide was prepared by the TFA/HFIP disaggregation protocol described in Materials and Methods but was not subjected to ultracentrifugation.

croparticulate aggregates that can seed aggregate growth. The figure also shows that when previously prepared aggregates of the same peptide are added to the reaction at $t = 0$ at 1 wt % of total peptide in solution, the reaction is rigorously accelerated with complete elimination of the lag phase. This confirms that polyGln aggregation can be seeded. Seeding is typical of the growth of amyloid fibrils, where it has been interpreted as evidence supporting a nucleated growth assembly mechanism (*24*).

In the course of our initial studies with polyGln peptides we discovered that they readily aggregate during storage in frozen aqueous buffers. Figure 4 shows that a solution of a polyGln peptide undergoes facile aggregation when incubated in the frozen state at −10 °C but not when incubated in the liquid state at the same temperature. In a similar experiment at −20 °C, we found that aggregates form from a polyGln solution whether it is snap-frozen in liquid nitrogen before being allowed to equilibrate at −20 °C or is allowed to slowly freeze at this temperature. These and other features suggest that stimulation of aggregation in the frozen state is due to the process of freeze concentration (see Discussion).

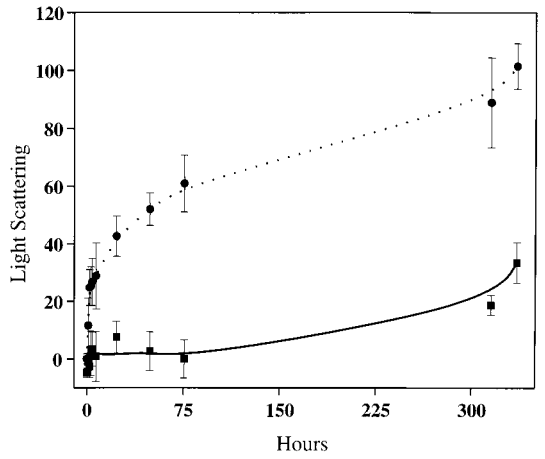

FIGURE 3: Aggregation of a solution of 20 μM $K_2Q_{28}K_2$ in pH 7.4 PBSA, as monitored by light scattering, either with (---) or without (—) seeding. Seeding was with 1 wt % of $K_2Q_{28}K_2$ aggregates prepared by aggregation in frozen solution at −20 °C followed by sonication in PBS (six pulses of 30 s each). The soluble peptide was prepared by a disaggregation protocol that included both treatment with TFA/HFIP and ultracentrifugation, as described in Materials and Methods.

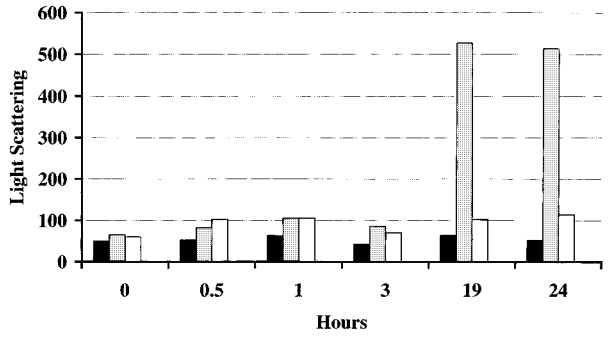

FIGURE 4: Time course of aggregation of 10 μM $K_2Q_{37}K_2$ in PBS, as monitored by light scattering, with samples incubated under various conditions: solid black bars, 37 °C; open bars, liquid state (supercooled) at −10 °C; gray bars, sample snap frozen in liquid nitrogen and then incubated at −10 °C (sample remained frozen at 24 h). ThT readings of reaction time points showed similar trends.

In agreement with previous reports (*7*), we find that polyGln aggregates are rich in β-sheet. Figure 1 shows that the circular dichroism spectra of a polyGln peptide before and after aggregation exhibit a classical random coil to β-sheet transition (*26*). The quality of the spectrum of the aggregate suspension is surprisingly good, since normally aggregates give poor CD spectra due to loss of incident light via light scattering; the ability of the lamp to continue to deliver photons to the multiplier tubes even at low wavelengths presumably reflects the relatively low scattering efficiency of the aggregates and hence, presumably, their relatively small particle size.

Dye Binding Properties of Aggregates. Although light scattering and solubility changes are expected to occur in any reaction involving the formation of large aggregates, the ability of the polyGln aggregation shown in Figure 2 to respond to ThT is not typical for all protein aggregation reactions and is consistent with an amyloid-like structure. Table 1 shows the apparent binding constants of ThT for a variety of polyGln aggregates, in comparison to amyloid fibrils formed from the Alzheimer's peptide Aβ. The data show that polyGln peptides of different lengths and/or with

Table 1: Fluorescent Properties of Thioflavin T Stained Aggregates

peptide component[a]	K_d (μM)[b]	fluorescence yield (arbitrary units/μM peptide)
$D_2Q_{15}K_2$	2.7	50
$DK_2Q_{15}DK_2$	7.2	25
$K_2Q_{15}K_2$	11.9	25
$K_2Q_{20}YK_2$	5.6	27
$K_2Q_{28}K_2$	16.1	12.9
$K_2Q_{33}K_2$	16.0	14.2
$K_2Q_{37}K_2$	14.9	31.2
$K_2Q_{37}K_2$[c]	2.8	41.3
Aβ(1−40)	0.5	200
Aβ(1−42)	5.7	70

[a] Except where indicated, polyGln aggregates were prepared at 37 °C at concentrations of peptide ranging from 50 to 80 μM as described in Materials and Methods. [b] Apparent K_d values were obtained as described in Materials and Methods. [c] Aggregates prepared in frozen solution as described in Materials and Methods.

different flanking residues all bind ThT with approximately the same strength, in the 3−16 μM range. These apparent binding constants are of the same order of magnitude as binding constants obtained in this laboratory and elsewhere for amyloid fibrils of various forms of Aβ. Thus, Table 1 shows that we obtained a K_d for Aβ(1−40) of 0.5 μM. This agrees well with the values of 0.9 μM (*27*) and 2.0 μM (*19*) obtained previously; the slight differences may reflect the different pH values that these two previous determinations were conducted at (pH 8.5 and 6.0, respectively) or may reflect other differences, such as the source of Aβ peptide and fibril growth conditions. The physical meaning of the different apparent binding constants within the polyGln series of aggregates is not clear. It is interesting that, among the three aggregates shown in Table 1 that all have a repeat length of Q_{15}, there is a trend of stronger ThT binding as the net positive charge on the peptide decreases. Since ThT carries a net positive charge, this may reflect a role for destabilizing (like) ionic interactions in ThT binding. At the same time, since ThT binds well to aggregates of polyGln peptides flanked only by Lys residues, we conclude that nonionic interactions must play a dominant role in ThT binding to polyGln aggregates. Table 1 also shows that, for polyGln sequences of intermediate length (Q_{28} to Q_{37}) in a common context of flanking Lys residues, binding constants for ThT are very similar. However, aggregates formed from the same peptide ($K_2Q_{37}K_2$) by growth either at 37 °C or at −20 °C in frozen solution exhibit apparent binding constants that differ by almost an order of magnitude.

Table 1 also shows that the fluorescent yields obtained for ThT bound to various polyGln aggregates are all in the same range and that all of these are less strong by factors ranging from 2 to 15 than the corresponding values for Aβ fibrils. The significance of this result is not clear, but it is possible that this information will be important in understanding how ThT binds to amyloid and amyloid-like aggregates.

Despite several indications that polyGln aggregates possess an amyloid substructure, the amyloid diagnostic dye Congo red does not confer on these aggregates the strong, uniform birefringence that would be expected for a homogeneous preparation of amyloid fibrils (*28*). In fact, aggregates grown in PBSA from polyGln peptides up to a repeat length of Q_{42} in our hands exhibit no birefringence in polarized microscopy

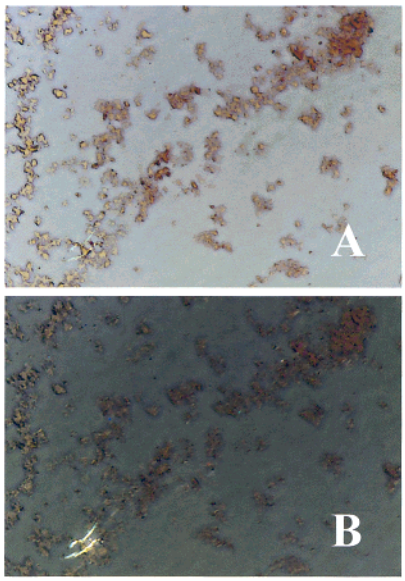

FIGURE 5: Nonpolarized (A) and polarized (B) light microscopy of Congo red stained Q_{47} aggregates.

(data not shown). We only observed birefringence consistent with amyloid fibril formation with an aggregate prepared from the Q_{47} peptide. Figure 5A shows that CR-treated Q_{47} aggregates exhibit a uniformly red tint when viewed under a nonpolarized light microscope. Figure 5B shows the same field, at the same magnification, under polarized light. It can be seen that only a small portion of the sample is birefringence. The lack of uniform, strong CR birefringence shown in Figure 5 does not seem to be due to any inability of the polyGln aggregates to bind CR. As noted above, aggregates treated with CR take on a red cast indicative of CR binding (Figure 5A). Furthermore, quantitative flow cytometry studies using CR fluorescence (*29*) show that polyGln aggregates made from either short or long repeat lengths bind CR in a manner similar to that of amyloid fibrils (V. Berthelier and R. Wetzel, unpublished). These experiments suggests that, while all polyGln aggregates bind Congo red, the long-range order necessary for birefringence is only generated, among the peptides we tested, in aggregates of Q_{47} peptides. The lack of birefringence of the bulk of the Q_{47} aggregates does not by itself address whether these aggregates have ordered or amorphous structures at atomic resolution.

Structures of Aggregates. We observe several different morphologies of polyGln aggregate by electron microscopy. All exhibit a significant underlying order at EM resolution (Figure 6). In analogy to previous observations (*7*), aggregates of a Q_{15} peptide grown at 37 °C appear as short, flexible filaments with diameters in the 3 nm range (Figure 6A). Other aggregated forms of polyGln appear to involve higher order assembly of filaments similar to those seen in Figure 6A. Peptides with longer polyGln repeat lengths grown at 37 °C in PBSA form mostly broad ribbons 10−40 nm in width, which appear to be comprised of parallel-aligned filaments (Figure 6B, Q_{20}; Figure 6C, Q_{37}). Figure 6C shows that longer peptides incubated at 37 °C produce somewhat wider ribbons. Figure 6D shows that when aggregates of longer peptides are prepared in frozen solution, they do not form ribbons but rather form the kind of short, flexible filaments shown in Figure 6A, either as a dense meshwork or as fine, gossamer threads. As shown in Figure

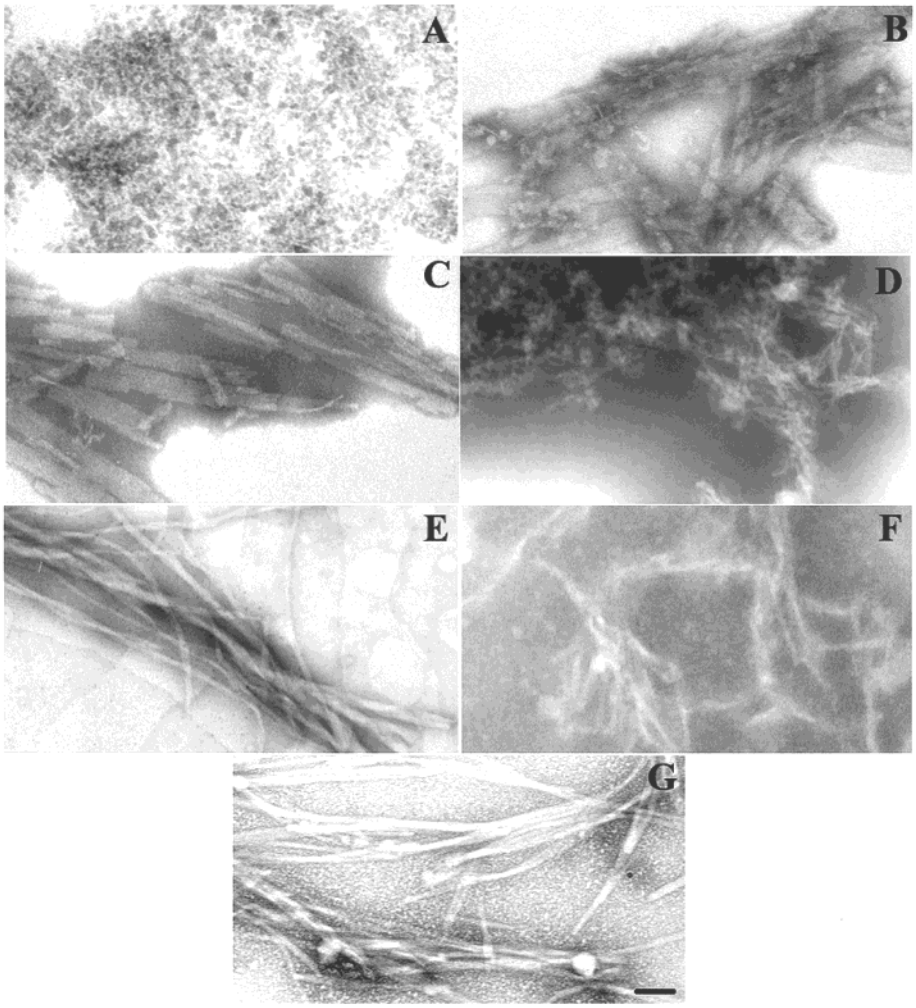

FIGURE 6: Electron micrographs of various protein aggregates: (A) $D_2Q_{15}K_2$ aggregates grown in PBS at 37 °C; (B) $K_2Q_{20}K_2$ aggregates grown in PBS at 37 °C; (C) $K_2Q_{37}K_2$ aggregates grown in PBS at 37 °C; (D) $K_2Q_{37}K_2$ aggregates grown in PBS in the frozen state at −20 °C; (E) fluoroscein-tagged $K_2Q_{42}K_2$ aggregates grown in PBS at 37 °C; (F) $K_2Q_{42}K_2$ aggregates grown in Tris-HCl, pH 7.2, at 37 °C; (G) Aβ(1−40) fibrils grown in PBS at 37 °C. Images were collected on a Hitachi 500 electron microscope on samples adsorbed onto mica grids and stained with 0.05% KPTA. The bar represents a distance of 50 nm.

6E, longer polyGln peptides (above Q_{32}) also can form amyloid-like fibrils. The peptide used to make the aggregates shown in 6E contains an N-terminal fluoroscein for use in experiments not described here; in our work with such peptides, we have seen no indication that the fluoroscein group modifies the morphologies of the aggregates formed (W. Yang and R. Wetzel, unpublished results). A more uniform aggregate preparation that resembles amyloid fibrils in all respects is formed when a Q_{42} peptide is aggregated in Tris-HCl, pH 7.2 (no NaCl) at 37 °C (Figure 6F). As a comparison, Figure 6G shows amyloid fibrils of Aβ(1−40) prepared in our laboratory.

It must be emphasized that while most of the images in Figure 6 are representative of the majority of grids observed in the electron microscopy of these different aggregates, the image shown in Figure 6E is very much a minor form of the Q_{42} aggregates, while the predominant form is of ribbons such as those shown in Figure 6B,C. It remains possible that aggregates grown in PBS under different growth conditions might exhibit a more consistent amyloid structure; except for the polyGln aggregation conducted in Tris buffer described above, no attempt was made to explore other growth conditions. The various morphological forms of aggregates that can be formed by polyGln suggest a delicate balance of forces at work in the formation of the higher ordered structure of these aggregates. After the β-sheet-rich, fundamental filament is formed, it appears that it can superassemble in a number of ways, depending on conditions.

PolyGln aggregates also exhibit an antibody epitope associated with the amyloid folding motif. Recently, this laboratory has isolated and characterized monoclonal antibodies (MAbs) capable of binding to Aβ amyloid fibrils (the immunogen) and not to the soluble, monomeric form of the peptide. Surprisingly, these antibodies also bind to a variety of other amyloid fibrils made in vitro but not to the corresponding native states of the amyloid precursors, nor to various kinds of naturally occurring or artificial protein aggregates (*21*). Here we prepared aggregates of the peptide $K_2Q_{42}K_2$ in PBSA at 37 °C, and these aggregates, as well as monomeric Q_{42} peptide, were immobilized onto plastic microtiter plates and their binding to WO1 was compared to Aβ(1−40). We found that, while neither monomeric Aβ nor monomeric Q_{42} bound measurable WO1, an equal weight of Q_{42} aggregate binds about 40% as much WO1 as do Aβ(1−40) fibrils (data not shown). Under these conditions, negligible binding occurs to nonamyloid protein aggregates (*21*). Surprisingly, WO1 does not bind to polyGln peptides aggregated in frozen aqueous solution (not shown). These

experiments show that polyGln aggregates grown at 37 °C possess a conformational epitope that is shared in common with a number of amyloid fibrils but is not found in other protein aggregates.

DISCUSSION

The tendency of polyGln sequences to aggregate into highly insoluble precipitates is, at first appearances, surprising. Most protein aggregation phenomena, such as irreversible thermal denaturation (*30*, *31*) and off-pathway aggregation during protein folding (*32*−*34*), are traditionally viewed as being driven by large contributions from the interactions of hydrophobic residues exposed on the surfaces of partially folded, or unfolded, states (*35*). Furthermore, highly amyloidogenic peptides such as Aβ, part of which is derived from the transmembrane sequence of its precursor protein APP (*36*), contain a significant portion of hydrophobic residues. In contrast, the glutamine residue, while uncharged, contains a polar side chain amide group whose potential for H-bonding would suggest a propensity for interactions with solvent water that would minimize any net stabilization from H-bonded self-interactions. Perutz addressed this conundrum by proposing a new class of peptide−peptide interactions called the "polar zipper", in which self-interactions are reinforced by the collective strength of cooperative H-bonding involving side chain amide groups (*7*). While this model has a number of attractive features, the molecular basis of polyGln aggregate structure remains unknown, in part because of the intrinsic difficulty of determining high-resolution structures of large, insoluble aggregates. In this paper we report a number of aspects of the assembly pathways and structures of polyGln aggregates, including some comparisons to bona fide amyloid fibrils composed of the Alzheimer's disease peptide Aβ.

Like the Aβ monomer (*37*, *38*), freshly solubilized polyGln appears to have no stable structure in solution, exhibiting a classical random coil CD spectrum (Figure 1). There does not appear to be a lower energy form of a monomeric polyGln peptide that can be significantly populated in neutral pH solution. The only structural transition undergone in the CD spectrum of an incubated sample of polyGln is a conversion into β-sheet, which occurs precisely in parallel with the onset of aggregation (S. Chen and R. Wetzel, submitted for publication). There is no information on the structure of even short polyGln sequences within the context of larger proteins. Although it has been speculated that these sequences may serve as association domains, in analogy to leucine zippers (*7*), the normal role of polyGln sequences in proteins is unknown. The role of these sequences in normal biology may be subtle, or nil, since homologous proteins from different organisms contain polyGln sequences of widely differing lengths [for example, human TATA binding protein has a 38 Gln repeat while the mouse, chicken, and viper versions, which otherwise are highly homologous to the human protein, have Gln repeats of only 13, 5, and 3, respectively (Swiss Protein Database)].

Like Aβ peptides, polyGln peptides incubated in PBS at 37 °C undergo an aggregation reaction characterized by a lag phase followed by rapid aggregate formation that continues until monomer is largely depleted from solution (Figures 2 and 3). As in Aβ fibril formation, the addition of a previously formed aggregate to a solution of polyGln reduces or entirely eliminates the lag phase in 37 °C aggregate growth (Figure 3). These similarities in aggregation kinetics suggest that polyGln aggregation is an ordered process.

PolyGln peptides, of all lengths tested, form aggregates when incubated in frozen aqueous buffers at intermediate freezing temperatures. While freezing is required for this enhancement of polyGln aggregation (Figure 4), aggregation does not occur if the frozen solution is incubated at −80 °C (data not shown). These results are best explained by the phenomenon of freeze concentration, in which frozen aqueous solutions above the eutectic points of the solutes contain liquid micropools possessing highly elevated solute concentrations (*39*, *40*). Such elevated solute concentrations, in turn, can promote protein aggregation, as seen in the freeze concentration-mediated inactivation of enzymes (*39*, *40*). The role of freeze concentration of NaCl and phosphate in this process is supported by our observations that polyGln aggregation at 37 °C increases with increasing NaCl concentration and that frozen polyGln solutions at −20 °C have a greater tendency to aggregate in phosphate buffer than in Tris-HCl buffer (not shown). At the same time, we also observe aggregate formation of polyGln peptides dissolved in pure water, suggesting that the eutectic properties of polyGln peptides and their concentration dependence of aggregation may also contribute to this phenomenon. Whatever the details of the physical factors contributing to this effect, we expect that more complex, natural proteins and protein fragments containing polyGln sequence elements may also be susceptible to aggregation induced by freezing, striking a note of caution for researchers who store protein samples frozen at −20 °C. The ability of freezing at −20 °C to encourage aggregate formation in polyGln peptides is in contrast to solutions of Aβ(1−40), which do not exhibit any detectible aggregation in our hands under these conditions.

As shown in Figure 6, the polyGln peptide aggregates formed in frozen solutions exhibit a morphology that is substantially different from aggregates formed at 37 °C. Freeze concentration aggregates resemble the protofibrils observed in Aβ fibril assembly (*41*), while aggregates formed at 37 °C resemble ribbons composed of parallel-aligned filaments and occasional straight filaments. It is not clear what it is about these different assembly conditions that determines aggregate morphology nor how any of these aggregated forms may relate to aggregates of polyGln protein fragments formed in vivo.

The circular dichroism spectrum (Figure 1) indicates that polyGln aggregates are dominated by β-sheet. Although this is consistent with an amyloid-like structure, many kinds of protein aggregates are dominated by β-sheet structure. The ability of polyGln aggregates to produce an amyloid-like response to ThT and the similarity in apparent ThT binding constants of Aβ amyloid and polyGln aggregates shown in Table 1 are consistent with an amyloid-like substructure even in polyGln aggregates that do not exhibit true amyloid fibril morphology in the EM. More details on the β-sheet arrangement within the polyGln aggregate and how this arrangement might differ among the various morphological forms of polyGln aggregates (Figure 6) may be elucidated by X-ray fiber diffraction studies.

In contrast to the ThT binding experiments, only weak birefringence is observed when polyGln aggregates are exposed to the classical defining test for amyloid, Congo red birefringence. Furthermore, only aggregates of the Q_{47} peptide exhibit any birefringence in our hands (Figure 5). Aggregates grown in PBS from polyGln peptides in the length range Q_{20} to Q_{42} produced no birefringence in repeated experiments (data not shown). Since all polyGln aggregates tested are able to bind CR (V. Bertherlier and R. Wetzel, unpublished), the lack of birefringence observed for these polyGln aggregates is consistent with an absence of the long-range order required for observation of birefringence. It cannot be ruled out, however, that some polyGln aggregates do not possess the kind of repeat structures required to elicit birefringence. Although Congo red remains a very useful probe for amyloid deposits in tissue pathology (*28*), it appears to be highly unsatisfactory as a molecular probe for amyloid structure (*42*).

Additional evidence for amyloid-like structure in the fundamental folding pattern of polyGln aggregates lies in the ability of an amyloid-specific monoclonal antibody, WO1, to bind to polyGln aggregates. This antibody is capable of binding to amyloid fibrils from a variety of proteins, including Aβ, islet amyloid polypeptide, immunoglobulin light chain, β_2-microglobulin, and transthyretin, but exhibits no significant binding to the native/monomeric states of these proteins. Specificity for the amyloid folding motif is suggested by the lack of binding to either amorphous protein aggregates, gelatin, or collagen (*21*). As described in the Results section, WO1 binds to polyGln aggregates but not to immobilized polyGln monomers. The nature of the WO1 epitope on these aggregates is yet to be established. Until the epitope is better defined, the full significance of WO1 binding to polyGln aggregates cannot be assessed. At the same time, the ability of WO1 to bind to amyloid fibrils as well as polyGln aggregates, but not to other protein aggregates, suggests that at least some polyGln aggregates may share a specific three-dimensional structural feature with other amyloid fibrils. In the future, conformational antibodies such as WO1 may serve as measures of the underlying folded structure within protein aggregates that will be more discerning than ThT and Congo red binding.

In summary, on the basis of a number of macroscopic properties, it can be inferred that polyGln aggregates possess a fundamental repeat structure at the atomic level that resembles the common folding motif of amyloid fibrils formed by many proteins, including Aβ. These aggregates form according to nucleation-dependent growth kinetics, are rich in β-sheet, and bind to ThT and alter its fluorescence spectrum in the same way as do Aβ amyloid fibrils. All polyGln aggregates exhibit some kind of ordered structure in the EM that involves a filamentous structure or substructure of 2−3 nm diameter. PolyGln aggregates also bind the pan-amyloid MAb WO1. All polyGln aggregates uniformly bind the Congo red molecule. At the same time, polyGln aggregates appear to only rarely adopt a form that exhibits the classic defining features of amyloid fibrils: twisted fibrillar morphology in the EM and CR birefringence. Whether the formation, and efficiency of formation, of bona fide amyloid fibrils by polyGln peptides has any significance for the human expanded CAG repeat diseases remains to be seen. HD brain tissue stained with thioflavin or Congo red does not exhibit classic amyloid signals (*43*). At the same time, aggregates in HD brain extracts do exhibit Congo red birefringence (*44*), as do aggregates made in vitro from recombinant polyGln protein material (*8*, *44*).

Despite the uncertainty surrounding the role of polyGln aggregates exhibiting the classic amyloid architecture in human pathology, the *amyloid-like substructure* that seems to exist in all forms of polyGln aggregates may be important to the cytotoxicity of expanded CAG repeat diseases. Recently, an attractive hypothesis for the mechanism of polyGln neurotoxicity has emerged, which postulates that aggregates of expanded polyGln peptides have the ability to recruit other polyGln-containing proteins within the cell and by so doing deprive the local environment of their beneficial activities (*4*, *6*, *13*, *44*−*48*). The polyGln aggregates described here clearly have the ability to be elongated in a variety of experimental settings (see, for example, Figure 3) and are also capable of recruiting polyGln peptides of much shorter length once aggregation has been initiated (*13*). We have presented data elsewhere suggesting that, in the context of this recruitment hypothesis, the most toxic aggregates would be expected to be small, exhibiting a high ratio of growth surface to total mass (*13*). Regardless of whether the recruitment−sequestration hypothesis proves correct, biophysical analysis of the process and products of polyGln aggregation would seem essential to the development of a complete understanding of the pathology of the expanded CAG repeat diseases.

ACKNOWLEDGMENT

We thank Dr. Rudi Hrnic for help with the Congo red staining protocol, Angela Williams for the Aβ amyloid fibrils, and Dustin Charles for initial studies on ThT binding to polyGln aggregates. We also acknowledge John Dunlap for the electron micrographs and Richard Andrews for help with the flow cytometer.

REFERENCES

1. Cummings, C. J., and Zoghbi, H. Y. (2000) *Hum. Mol. Genet. 9*, 909−916.
2. Zoghbi, H. Y., and Orr, H. T. (1999) *Curr. Opin. Neurobiol. 9*, 566−570.
3. Ross, C. A., Wood, J. D., Schilling, G., Peters, M. F., Nucifora, F. C., Jr., Cooper, J. K., Sharp, A. H., Margolis, R. L., and Borchelt, D. R. (1999) *Philos. Trans. R. Soc. London, Ser. B 354*, 1005−1011.
4. Preisinger, E., Jordan, B. M., Kazantsev, A., and Housman, D. (1999) *Philos. Trans. R. Soc. London, Ser. B 354*, 1029−1034.
5. Wanker, E. E. (2000) *Biol. Chem. 381*, 937−942.
6. McCampbell, A., and Fischbeck, K. H. (2001) *Nat. Med. 7*, 528−530.
7. Perutz, M. F., Johnson, T., Suzuki, M., and Finch, J. T. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 5355−5358.
8. Scherzinger, E., Lurz, R., Turmaine, M., Mangiarini, L., Hollenbach, B., Hasenbank, R., Bates, G. P., Davies, S. W., Lehrach, H., and Wanker, E. E. (1997) *Cell 90*, 549−558.
9. Scherzinger, E., Sittler, A., Schweiger, K., Heiser, V., Lurz, R., Hasenbank, R., Bates, G. P., Lehrach, H., and Wanker, E. E. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 4604−4609.
10. Martin, J. B. (1999) *N. Engl. J. Med. 340*, 1970−1980.
11. Sipe, J. D. (1992) *Annu. Rev. Biochem. 61*, 947−975.
12. Falk, R. H., Comenzo, R. L., and Skinner, M. (1997) *N. Engl. J. Med. 337*, 898−909.
13. Chen, S., Berthelier, V., Yang, W., and Wetzel, R. (2001) *J. Mol. Biol. 311*, 173−182.

14. Chen, S., and Wetzel, R. (2001) *Protein Sci. 10*, 887−891.
15. Berthelier, V., Hamilton, J. B., Chen, S., and Wetzel, R. (2001) *Anal. Biochem. 295*, 227−236.
16. Kheterpal, I., Zhou, S., Cook, K. D., and Wetzel, R. (2000) *Proc. Natl. Acad. Sci. U.S.A. 97*, 13597−13601.
17. Kheterpal, I., Williams, A., Murphy, C., Bledsoe, B., and Wetzel, R. (2001) *Biochemistry 40*, 11757−11767.
18. Naiki, H., Higuchi, K., Hosokawa, M., and Takeda, T. (1989) *Anal. Biochem. 177*, 244−249.
19. LeVine, H., III (1993) *Protein Sci. 2*, 404−410.
20. Puchtler, H., Sweat, F., and Levine, M. (1962) *J. Histochem. Cytochem. 10*, 355−364.
21. O'Nuallain, B., and Wetzel, R. (2002) *Proc. Natl. Acad. Sci. U.S.A. 99*, 1485−1490.
22. Diamandis, E. P. (1988) *Clin. Biochem. 21*, 139−150.
23. Altschuler, E. L., Hud, N. V., Mazrimas, J. A., and Rupp, B. (1997) *J. Pept. Res. 50*, 73−75.
24. Harper, J. D., and Lansbury, P. T., Jr. (1997) *Annu. Rev. Biochem. 66*, 385−407.
25. Woody, R. W. (1996) in *Circular Dichroism and Conformational Analysis of Biomolecules* (Fasman, G. D., Ed.) pp 25−67, Plenum, New York.
26. Tilstra, L., and Mattice, W. L. (1996) in *Circular dichroism and the conformational analysis of macromolecules* (Fasman, G. D., Ed.) pp 261−283, Plenum, New York.
27. Naiki, H., and Nakakuki, K. (1996) *Lab. Invest. 74*, 374−383.
28. Westermark, G. T., Johnson, K. H., and Westermark, P. (1999) *Methods Enzymol. 309*, 3−25.
29. Wall, J., and Solomon, A. (1999) in *Methods in Enzymology: Amyloid, Prions and other Protein Aggregates* (Wetzel, R., Ed.) pp 460−466, Academic Press, San Diego.
30. Tanford, C. (1968) *Adv. Protein Chem. 23*, 121−282.
31. Wetzel, R., Perry, L. J., Mulkerrin, M. G., and Randall, M. (1990) in *Protein Design and the Development of New Therapeutics and Vaccines; Proceedings of the Sixth Annual Smith, Kline and French Research Symposium* (Poste, G., and Hook, J. B., Eds.) pp 79−115, Plenum, New York.
32. London, J., Skrzynia, C., and Goldberg, M. E. (1974) *Eur. J. Biochem. 47*, 409−415.
33. Mitraki, A., and King, J. (1989) *Biotechnology 7*, 690−697.
34. King, J., Fane, B., Haase-Pettingell, C., Mitraki, A., Villafane, R., and Yu, M.-H. (1990) in *Protein Folding: Deciphering the Second Half of the Genetic Code* (Gierasch, L. M., and King, J., Eds.) pp 225−240, American Association for the Advancement of Science, Washington, DC.
35. Fink, A. L. (1998) *Folding Des. 3*, R9−R23.
36. Selkoe, D. J. (1994) *Annu. Rev. Neurosci. 17*, 489−517.
37. Gursky, O., and Aleshkov, S. (2000) *Biochim. Biophys. Acta 1476*, 93−102.
38. Zhang, S., Iwata, K., Lachenmann, M. J., Peng, J. W., Li, S., Stimson, E. R., Lu, Y., Felix, A. M., Maggio, J. E., and Lee, J. P. (2000) *J. Struct. Biol. 130*, 130−141.
39. Franks, F. (1985) *Biophysics and biochemistry at low temperatures*, Cambridge University Press, Cambridge.
40. Franks, F. (1993) in *Protein Biotechnology: Isolation, Characterization and Stabilization* (Franks, F., Ed.) pp 489−531, Humana Press, New York.
41. Harper, J. D., Wong, S. S., Lieber, C. M., and Lansbury, P. T. (1997) *Chem. Biol. 4*, 119−125.
42. Khurana, R., Uversky, V. N., Nielsen, L., and Fink, A. L. (2001) *J. Biol. Chem. 276*, 22715−22721.
43. Duyckaerts, C., Durr, A., Cancel, G., and Brice, A. (1999) *Acta Neuropathol. (Berlin) 97*, 201−207.
44. Huang, C. C., Faber, P. W., Persichetti, F., Mittal, V., Vonsattel, J. P., MacDonald, M. E., and Gusella, J. F. (1998) *Somat. Cell Mol. Genet. 24*, 217−233.
45. McCampbell, A., Taylor, J. P., Taye, A. A., Robitschek, J., Li, M., Walcott, J., Merry, D., Chai, Y., Paulson, H., Sobue, and Fischbeck, K. H. (2000) *Hum. Mol. Genet. 9*, 2197−2202.
46. Nucifora, F. C., Jr., Sasaki, M., Peters, M. F., Huang, H., Cooper, J. K., Yamada, M., Takahashi, H., Tsuji, S., Troncoso, J., Dawson, V. L., Dawson, T. M., and Ross, C. A. (2001) *Science 291*, 2423−2428.
47. Wetzel, R., Chen, S., Berthelier, V., and Yang, W. (2001) *Soc. Neurosci. Abstr. 27*, 578.12.
48. La Spada, A. R., Fu, Y., Sopher, B. L., Libby, R. T., Wang, X., Li, L. Y., Einum, D. D., Huang, J., Possin, D. E., Smith, A. C., Martinez, R. A., Koszdin, K. L., Treuting, P. M., Ware, C. B., Hurley, J. B., Ptacek, L. J., and Chen, S. (2001) *Neuron 31*, 913−927.

BI011772Q