doi:10.1016/j.jmb.2009.08.034 J. Mol. Biol. (2009) 393, 978–992

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Examining Polyglutamine Peptide Length: A Connection between Collapsed Conformations and Increased Aggregation

Robert H. Walters and Regina M. Murphy*

Department of Chemical and Biological Engineering, University of Wisconsin, Madison, 1415 Engineering Drive, Madison, WI 53706, USA

Received 13 April 2009; received in revised form 10 August 2009; accepted 13 August 2009
Available online 20 August 2009

Abnormally expanded polyglutamine domains in proteins are associated with several neurodegenerative diseases, of which the best known is Huntington's. Expansion of the polyglutamine domain facilitates aggregation of the affected protein, and several studies directly link aggregation to neurotoxicity. The age of onset of disease is inversely correlated with the length of the polyglutamine domain; this correlation motivates an examination of the role of the length of the domain on aggregation. In this investigation, peptides containing 8 to 24 glutamines were synthesized, and their conformational and aggregation properties were examined. All peptides lacked secondary structure. Fluorescence resonance energy transfer studies revealed that the peptides became increasingly collapsed as the number of glutamine residues increased. The effective persistence length was estimated to decrease from ~11 to ~7 Å as the number of glutamines increased from 8 to 24. A comparison of our data with theoretical results suggests that phosphate-buffered saline is a good solvent for Q8 and Q12, a theta solvent for Q16, and a poor solvent for Q20 and Q24. By dynamic light scattering, we observed that Q16, Q20, and Q24, but not Q8 or Q12, immediately formed soluble aggregates upon dilution into phosphate-buffered saline at 37 °C. Thus, Q16 stands at the transition point between good and poor solvent and between stable and aggregation-prone peptide. Examination of aggregates by transmission electron microscopy, along with kinetic assays for sedimentation, provided evidence indicating that soluble aggregates mature into sedimentable aggregates. Together, the data support a mechanism of aggregation in which monomer collapse is accompanied by formation of soluble oligomers; these soluble species lack regular secondary structure but appear morphologically similar to the sedimentable aggregates into which they eventually mature.

© 2009 Elsevier Ltd. All rights reserved.

Keywords: aggregation; dynamic light scattering; fluorescence resonance energy transfer (FRET); peptide conformation; polyglutamine

Edited by K. Kuwajima

Introduction

The "CAG" diseases are a collection of nine known neurodegenerative disorders, including Huntington's disease.[1] These disorders are causally linked to abnormally extended tracts of CAG codons in disease-specific genes. Expression produces a protein with an expanded polyglutamine (polyQ) domain; invariably, aggregates of the affected protein are found in neuronal inclusions.[2,3] Expansion of the polyQ domain beyond a critical length results in a disease phenotype; longer polyQ regions correspond to an earlier age of onset and a greater severity of symptoms.[4] The proteins involved in the CAG diseases have no sequence, compositional, or structural homologies, with the exception of the polyQ stretch,[5] which varies in both length and placement within the primary sequence of the various disease-associated proteins.[6]

*Corresponding author. E-mail address: regina@engr.wisc.edu.

Abbreviations used: GdmCl, guanidine hydrochloride; FRET, fluorescence resonance energy transfer; PBS(A), phosphate-buffered saline (with azide); SEC, size-exclusion chromatography; TEM, transmission electron microscopy; TFA, trifluoroacetic acid.

0022-2836/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.

The expanded polyQ domain facilitates aggregation of the affected protein, and several studies have directly linked aggregation to toxicity. For example, delivery of synthetic polyQ peptide aggregates to the nucleus was shown to cause cell death,[7] and the molecular chaperones Hsp40 and Hsp70 effectively decreased aggregation and toxicity in a zebrafish model of polyQ aggregation.[8] Onset of disease is more likely due to gain of toxic function rather than loss of normal function of the affected protein.[9] As one piece of evidence for this statement, addition of an extended polyQ tract into proteins unrelated to the known polyQ diseases produced polyQ-like aggregates and neurodegeneration in mice.[10] Some research suggests that large polyQ aggregates may be toxic,[11,12] but others have argued that a specific monomer conformation[13] or soluble oligomers[14] are responsible for cell death. Alternatively, the polyQ region may interact with other cell components, preventing them from fulfilling their functions and thus causing cell death. In particular, polyQ may interact with transcription factors, thereby disrupting transcriptional regulation,[15] as transcription factors have been found within nuclear inclusions.[16]

PolyQ peptides have been used in many studies as a tool for examining conformation and aggregation.[7,17–19] *In vitro*, longer polyQ domains aggregate faster.[20] Aggregation kinetic data were used to propose a nucleation–elongation model of polyQ peptide aggregation, in which the aggregation nucleus is a structured monomer in equilibrium with the bulk pool of disordered monomers; the equilibrium constant, though small, increases with increasing polyQ chain length.[20,21] However, recent work in our group demonstrated that the monomer loss data, used to support the hypothesis of a critical monomeric nucleus,[20] is insufficient to firmly establish a specific kinetic model, since a number of alternative kinetic mechanisms can capture the monomer loss data equally well.[22] Elucidation of the mechanism of aggregation requires a closer examination of monomer conformation and its relationship to aggregation.

PolyQ monomers lack regular secondary structural elements, as determined by both circular dichroism (CD) and NMR.[17,23–25] While polyQ monomers are disordered, they may still preferentially adopt collapsed or extended conformations. Theoretical studies have begun to probe the structure and conformation of polyQ peptides. Discontinuous molecular dynamics simulations revealed that, although some α and β backbone–backbone hydrogen bonds form in isolated polyQ monomers, there are not as many of these bonds as there are in a full α-helix or β-sheet.[26] Another molecular dynamics study concluded that the driving force for intermolecular associations in polyQ aggregation is governed by spontaneous conformational fluctuations within monomeric polyglutamine, and thus, aggregation is unlikely to follow a homogeneous nucleation mechanism with the monomer as the critical nucleus.[27]

Two experimental results have led to distinctly different conclusions about polyQ peptide conformation. A recent study, utilizing a tryptophan/cysteine contact quenching technique to measure the kinetics of loop closure, concluded that polyQ peptides adopt unusually extended conformations.[28] However, other researchers, using fluorescence correlation spectroscopy[29] to determine translational diffusion coefficients, reported that polyQ peptides adopt highly collapsed conformations. The clear discrepancy in these results warrants further investigation.

In this study, we systematically varied the length of polyQ peptides from 8 to 24 residues and examined the structure, conformation, and aggregation properties as a function of length. We employed fluorescence resonance energy transfer (FRET), which has previously been successfully applied to the study of the end-to-end distance of disordered peptides,[30–32] to measure whether polyQ peptides are extended or collapsed in aqueous solution. We also examined the aggregation kinetics using a combination of size-exclusion chromatography (SEC), laser light scattering, sedimentation, and electron microscopy.

Using FRET, we observed an increasing propensity for collapsed conformations as the number of glutamine residues is increased. Peptides containing 8 or 12 glutamines were relatively extended in solution and remained monomeric. Peptides containing 20 or 24 glutamines were relatively collapsed in solution and immediately associated into large soluble aggregates, which subsequently formed sedimentable aggregates whose sedimentation kinetics fit a sigmoidal curve. A comparison of the morphology of aggregates formed early and late in the aggregation process suggests that maturation of structure includes realignment within the aggregates themselves. Peptides containing 16 glutamines were intermediate in behavior; these formed large soluble aggregates that did not precipitate. The data support a mechanism of aggregation in which monomer collapse drives the formation of soluble oligomers. These soluble species lack regular secondary structure but appear morphologically similar to the sedimentable aggregates into which they eventually mature.

Results

Peptides of the type $K_2WQ_nXK_2$ were synthesized, with $n = 8$, 12, 16, 20, and 24. Flanking lysine residues were added to the polyQ core to increase solubility, consistent with previous work,[20] while tryptophan served as a FRET donor and for concentration determination. The N terminus was acetylated and the C terminus was amidated to eliminate the possibility of charge interactions of the termini. The residue X was either alanine or dansylated lysine. Peptides will be referred to subsequently as Q8, Q12, Q16, Q20, and Q24 when X = alanine and as Q8D, Q12D, Q16D, Q20D, and Q24D when X = dansylated lysine.

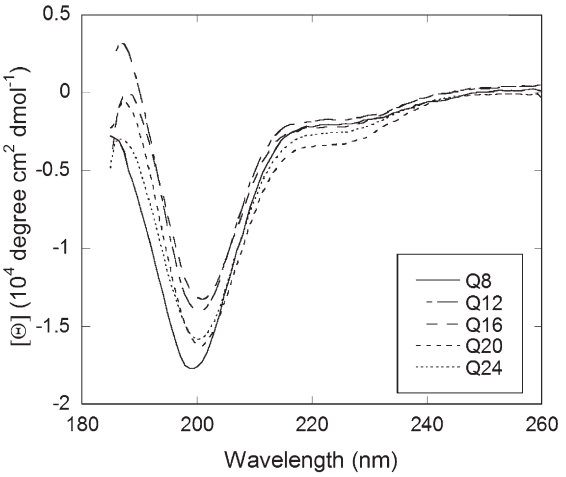


Fig. 1. CD spectra of polyQ peptides. Peptide stock solutions were diluted into a pH 7.4 phosphate buffer (10 mM buffer salts and 140 mM NaF) to a concentration of 20 μM peptide and filtered through a 0.45-μm membrane directly into a cuvette. Data were normalized to per-residue molar ellipticity.

Circular dichroism

Circular dichroic spectra were collected for solutions of Q8, Q12, Q16, Q20, and Q24. The spectra (Fig. 1) are consistent with that of a disordered peptide. Slight variations between the peptides are likely a consequence of tryptophan making up a different percentage of each peptide, as tryptophan can have an impact on CD spectra.[33] Spectra taken after 13 days (data not shown) are similar to those taken immediately after dilution into buffer. Our data are consistent with other CD results in the literature, indicating that monomeric polyQ peptides occupy a disordered state regardless of length.[17,24,34] NMR data indicated that all Q's in polyQ regions experience a similar chemical environment,[23] again supporting the conclusion that polyQ peptides under these conditions lack regular defined secondary structural elements.

FRET of polyQ monomers

We next sought to determine the average extension of the polyQ peptide in solution. CD spectra indicated that the peptides are disordered and lack regular secondary structure, but that does not necessarily mean that they are true random coils. We inserted a fluorescent donor (Trp) and acceptor (dansyl) at each end of the polyQ stretch. Dipole–dipole interactions between the fluorescence donor and acceptor allow the transfer of excited-state energy from donor to acceptor. The efficiency of energy transfer E is a function of the distance between donor and acceptor.[35] For a rigid material, E is related to the distance R between donor and acceptor as

$$E = \frac{R_0^6}{R_0^6 + R^6} \quad (1)$$

where R_0 is the Förster radius for the donor–acceptor pair. We assumed that $R_0 = 21.0$ Å for the Trp–dansyl pair based on literature data.[30,36,37] R_0 varies somewhat with peptides and solvents; for instance, Singh measured R_0 for dansyl-$(AGQ)_3$-W as 19.3 Å in water and 19.55 Å in 6 M guanidine hydrochloride (GdmCl).[28] We decided to use 21 Å since it is a more established value.

Data were collected with the peptides at 5 μM in buffer at physiological pH (7.4) and salt concentration (10 mM buffer salts plus 140 mM NaCl). The measured E was fit to Eq. (1) to obtain R; results are summarized in Fig. 2a. Generally, R increased as the number of glutamines increased. Limited experiments were conducted at lower peptide concentration (1 μM) and over a 2-week incubation time. We observed no change in R with concentration or with

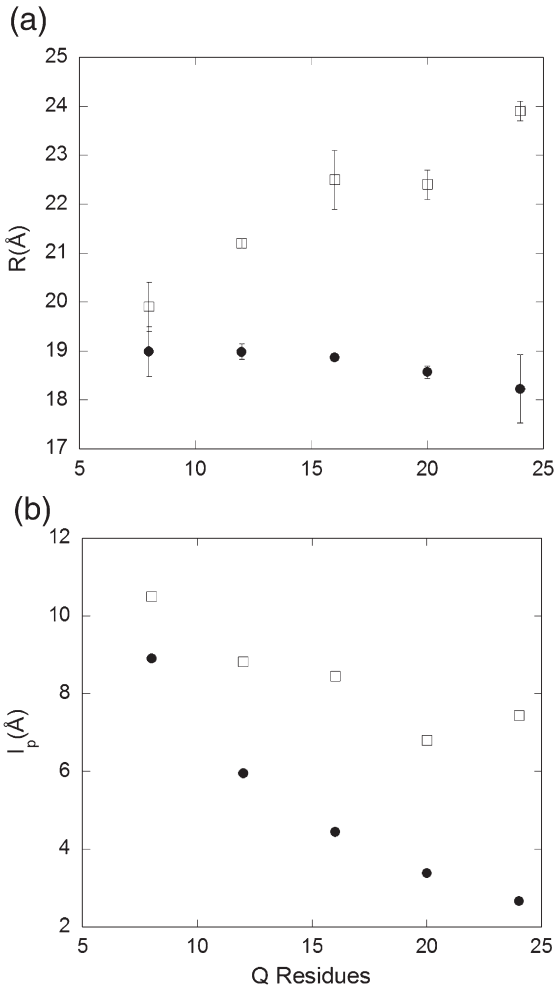


Fig. 2. Length of polyQ region determined by FRET. Fluorescence spectra of peptides with and without the fluorescence acceptor dansyl were collected and used to calculate the distance between the fluorophores R. Peptides were diluted to 5 μM in pH 7.4 PBSA (□) or pH 12 buffer (●). Readings were taken in duplicate, and error bars indicate the standard deviation of the measurement. (a) Average length of polyQ region determined by FRET. (b) Persistence length, l_p.

time (data not shown), suggesting that the measured R value is for monomers.

Equation (1) is true for homogeneous rigid materials. Two additional factors must be considered in analyzing FRET for disordered polypeptides: diffusion of the ends during the fluorescence lifetime of the donor and conformational heterogeneity. The intramolecular diffusion can be ignored if

$$D\tau_F/s^2 \ll 1 \tag{2}$$

where D is the intramolecular diffusion coefficient (relative diffusion of donor and acceptor), τ_F is the fluorescence lifetime of the donor, and s is the mean distance between donor and acceptor.[35] For peptides similar to ours containing 16 glutamines, $D \approx 2\times10^{-7}$ cm^2/s.[28] Fluorescence lifetimes for tryptophan depend somewhat on the peptide and the solvent but are typically less than 3 ns at pH 7.4.[38] From our data, we estimate $s \approx 20$ Å. Thus, $D\tau_F/s^2 \approx 0.02 \ll 1$; hence, it is reasonable to neglect diffusion during the fluorescence lifetime of the donor.

To address the issue of conformational heterogeneity, we assumed that the peptides populate a conformational ensemble with a probability distribution of donor–acceptor distances $P(r)$. The measured energy transfer is therefore an average $\langle E\rangle$ over the conformational distribution $P(r)$:

$$\langle E\rangle = \int_0^{L_c} E(r)P(r)\mathrm{d}r, \text{ with } E(r) = \left(1+\left(\frac{r}{R_0}\right)^6\right)^{-1} \tag{3}$$

and the measured R in Fig. 2a is an average distance between donor and acceptor:

$$R = R_0\left(\frac{1-\langle E\rangle}{\langle E\rangle}\right)^{1/6} \tag{4}$$

We considered several different models of polypeptide conformation in an effort to interpret the FRET data. If the behavior of the polypeptides could be modeled as freely jointed chains having a Gaussian distribution,[39] then

$$P(r) = 4\pi r^2\left(\frac{3}{2\pi n l_b^2}\right)^{3/2}\exp\left(-\frac{3r^2}{2nl_b^2}\right) \tag{5}$$

where l_b is the bond length and is taken to be 3.8 Å, the distance between repeating atoms in a polypeptide backbone.[40] The Gaussian chain model has been used to describe the conformational distribution of unfolded polypeptides in other studies, frequently with the bond length l_b as a fitted parameter; the model is better suited for longer chains.[41] We inserted the distribution given in Eq. (5) into Eq. (3) to calculate R_{FJ}, the average distance between donor and acceptor one would expect for a freely jointed chain with a bond length equal to one residue. Alternatively, if the polypeptides were modeled as fully extended rigid rods, then the donor–acceptor distance is simply $L_c = (n+1)l_b$, where n is the number of glutamines. (We neglect the displacement of the side chains from the backbone in this analysis.) The measured values of R fall between the freely jointed chain R_{FJ} and the fully extended rod L_c, as shown in Table 1. Thus, the behavior of our polyglutamine peptides appears to fall in between these two limiting cases.

Table 1. Measured R for polyQ peptides in PBSA, compared to peptide chain distances for a freely jointed chain (R_{FJ}), a fully extended chain (L_c), and a chain with excluded volume in a theta solvent (R_θ), calculated as described in the text

No. of Q	R (Å)	R_{FJ} (Å)	L_c (Å)	R_θ (Å)
8	19.9±0.5	13.3	34.2	17.1
12	21.2±0.1	15.2	49.4	20.5
16	22.5±0.6	16.6	64.6	23.5
20	22.4±0.3	17.7	79.8	26.1
24	23.9±0.2	18.6	95.0	28.5

Vitalis *et al.* recently completed molecular dynamics simulations of polyglutamine peptides (Ac-Q$_n$-Nm) and calculated average distances in a chain as a function of length and temperature.[27] They proposed that the average end-to-end distance for polyglutamine in a theta solvent could be estimated from $R_\theta = 5.7(n+1)^{0.5}$, with R_θ expressed in angstroms. We calculated R_θ from this equation as another point of comparison with the experimental data. This comparison indicates that Q8 and Q12 are more extended than one would expect for polyglutamine in a theta solvent ($R > R_\theta$), while Q20 and Q24 are more collapsed than one would expect in a theta solvent ($R < R_\theta$). This analysis suggests that phosphate-buffered saline (PBS) becomes less able to solvate polyQ peptides with increasing Q length and that the peptides become relatively more collapsed as the number of glutamines increases.

As an alternative analysis, we postulated that our polyQ peptides could be modeled as semiflexible (wormlike) chains.[28,42,43] The wormlike chain model has been used with reasonable success to describe the conformational distributions of unfolded polypeptides and flexible linkers within folded proteins.[41] We used an approximate analytical expression derived by Hyeon and Thirumalai for $P(r)$ as a function of L_c (the fully extended distance between donor and acceptor), L (the total length of the peptide from N to C terminus), and l_p (the persistence length, which is a measure of the extent to which the chain persists in the same direction).[42] In calculating L, we neglected the fact that the donor and acceptor are on side chains; we estimated that this neglect would affect the calculated value of l_p by less than 10% and would not change our conclusions. The wormlike chain model can account for polymer chains anywhere from a stiff rigid rod (high l_p) to a collapsed chain (low l_p). This distribution does not account for excluded volume effects, which become more important as L increases and as l_p decreases but should be negligible for peptides of

the length studied here.[41] We numerically integrated Eq. (3) using $P(r)$ from Hyeon and Thirumalai for each peptide length, as a function of l_p, calculated R per Eq. (4), and regressed the measured R value from the FRET data to find the least-squares best fit of l_p.[42] Results of these calculations are presented in Fig. 2b. There is a smooth decrease in the effective persistence length as the number of glutamines increases. Thus, the peptides become progressively more collapsed as the number of glutamines increases. This conclusion is similar to that obtained from the analysis summarized in Table 1.

Aggregation kinetics

We next used laser light scattering to test for the presence of aggregates in peptide solutions. No scattering above background was detected from solutions of Q24 at pH 3 over 24 h (data not shown), which we interpreted to indicate that the stock solution was unaggregated. Peptides (Q8, Q12, Q16, Q20, and Q24) were diluted from stock solutions into PBSA (PBS with azide) to a concentration of 20 μM and held at 37 °C. Both the average apparent hydrodynamic radius, R_{hz}, and the normalized scattered intensity at 90°, I_{norm}, were measured. With Q8 and Q12, we did not detect scattering intensities above background over the course of 24 h (not shown); therefore, $I_{norm} \sim 0$ and R_{hz} was not measurable. Thus, Q8 and Q12 did not aggregate to any measurable extent when diluted into PBSA.

For Q16, Q20, and Q24, scattered light was detectable and soluble aggregates formed essentially immediately after sample preparation. For these samples, the average apparent R_{hz} increased continuously over time and showed little variation with the length of the peptides, as shown in Fig. 3. Since R_{hz} is heavily weighted toward larger particles, this indicates that aggregate sizes for all three peptides are similar. R_{hz} reached ~100 nm after approximately 1 h; particles grew to an apparent radius of several hundred nanometers within 4 h (Fig. 3). For these three peptides, the total normalized scattering intensity, I_{norm}, increased over time for the first ~50 min (Fig. 4a). All samples extrapolate to zero intensity at time zero, supporting the notion that the peptides are not aggregated in the low pH stock solution and that dilution into PBSA at 37 °C initiates the aggregation process. I_{norm} increased with Q length: I_{norm} for Q20 was about 30% greater than that for Q16, and I_{norm} for Q24 was about 2 times that of Q20.

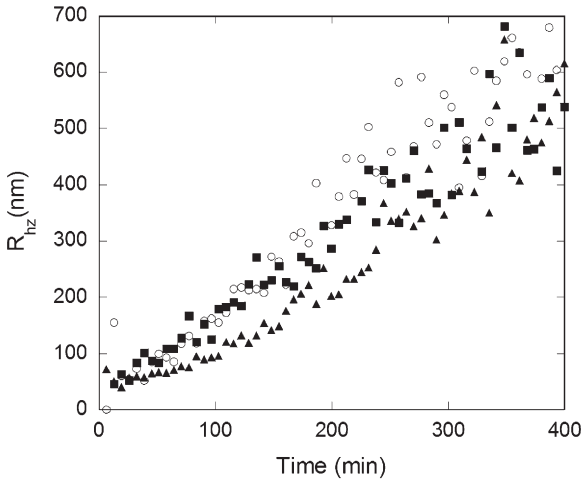


Fig. 3. Dynamic light-scattering hydrodynamic radius of polyQ peptides. Disaggregated peptide stock solution was diluted to a concentration of 20 μM in PBSA and then immediately filtered directly into a light-scattering cuvette and held at 37 °C. Apparent z-averaged hydrodynamic radius R_{hz}, as determined from cumulants analysis of autocorrelation data collected at 90° scattering angle. Data shown for Q16 (■), Q20 (○), and Q24 (▲). Q8 and Q12 did not scatter above background.

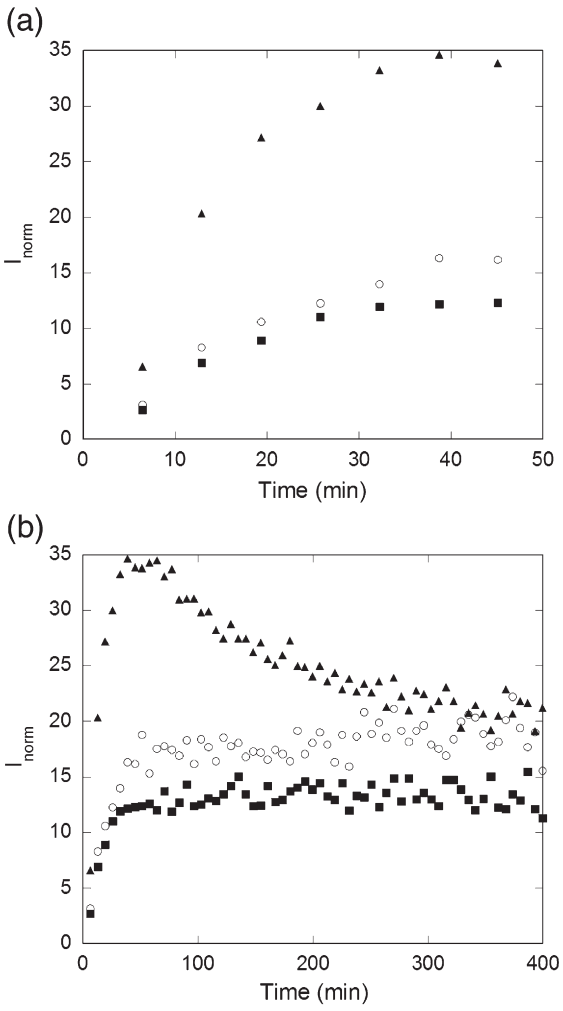


Fig. 4. Scattered light intensity of polyQ peptides. Data shown for Q16 (■), Q20 (○), and Q24 (▲). Q8 and Q12 did not scatter above background. (a) Intensity of scattered light at 90°. Scattering due to the solvent was subtracted, and results were normalized to the scattering intensity of toluene to account for changes in laser strength and aperture and to the mass concentration of peptide. (b) Total intensity of scattered light over a longer period of time.

$I_{norm} \propto \langle M \rangle_w P(90^\circ)$, where $\langle M \rangle_w = \sum w_i M_i$ is the weight-averaged molecular weight when w_i is the weight fraction of species with molecular weight M_i and $P(90°)$ is a particle scattering factor that depends on particle shape and size. Aggregate sizes (as measured by R_{hz}) are roughly equal for Q16, Q20, and Q24 (Fig. 3). Assuming that particle shapes are equivalent, then $P(90°)$ is roughly equal for all three peptides, and the differences in I_{norm} among the peptides likely indicate differences in $\langle M \rangle_w$. We assumed that there are only monomer and soluble aggregate populations during this time (see SEC data, below) and that the molecular weight of the aggregates is much greater than that of the monomers, or $M_{agg} \gg M_{mon}$. With these assumptions, $\langle M \rangle_w = \sum w_i M_i$, and increases in I_{norm} reflect an increase in the $\sum w_{agg} M_{agg}$ term. Because the average size R_{hz} is not much different for Q16, Q20, and Q24, as long as aggregate shapes are similar, then M_{agg} is similar for all three peptides, and therefore, the increase in I_{norm} with increasing Q length is primarily due to an increase in w_{agg}, the mass fraction of peptide that is incorporated into soluble aggregates.

At times longer than 1 h, I_{norm} plateaued for Q16 and Q20 and decreased for Q24. There are two possible explanations for these observations. First, as R_{hz} increases, $P(90°)$ decreases sharply. For example, at our conditions, $P(90°)$ for spheres = 0.9 if $R_h = 30$ nm but 0.27 if $R_h = 100$ nm. Thus, an increase in $\langle M \rangle_w$ could be offset by an equivalent decrease in $P(90°)$, leading to little change in I_{norm}. Under some circumstances, the decrease in $P(90°)$ could more than offset an increase in $\langle M \rangle_w$. Second, any loss of material due to precipitation could cause a decrease in I_{norm}.[34] Since we did not detect any visible precipitates at this early time in any of the samples, we suspect that the first explanation is the most likely.

Transmission electron microscopy

We prepared solutions of all five peptides in PBSA and examined them by transmission electron microscopy (TEM) 3 h after preparation. In agreement with the light-scattering data, we failed to see aggregates in the Q8 and Q12 samples. Aggregates were observed with both Q20 (Fig. 5a) and Q24 (Fig. 5b), although they were sparse, particularly in Q20. No aggregates were found in the Q16 sample. The gross morphology of Q20 and Q24 aggregates was similar, as both appeared as loose clusters of linear structures of similar size. However, there were some differences at a finer resolution. Q20 clusters contained many short, poorly defined rodlike structures. In contrast, Q24 aggregates showed a higher degree of order, with more clearly defined, longer rodlike structures. However, the rods themselves were not highly aligned.

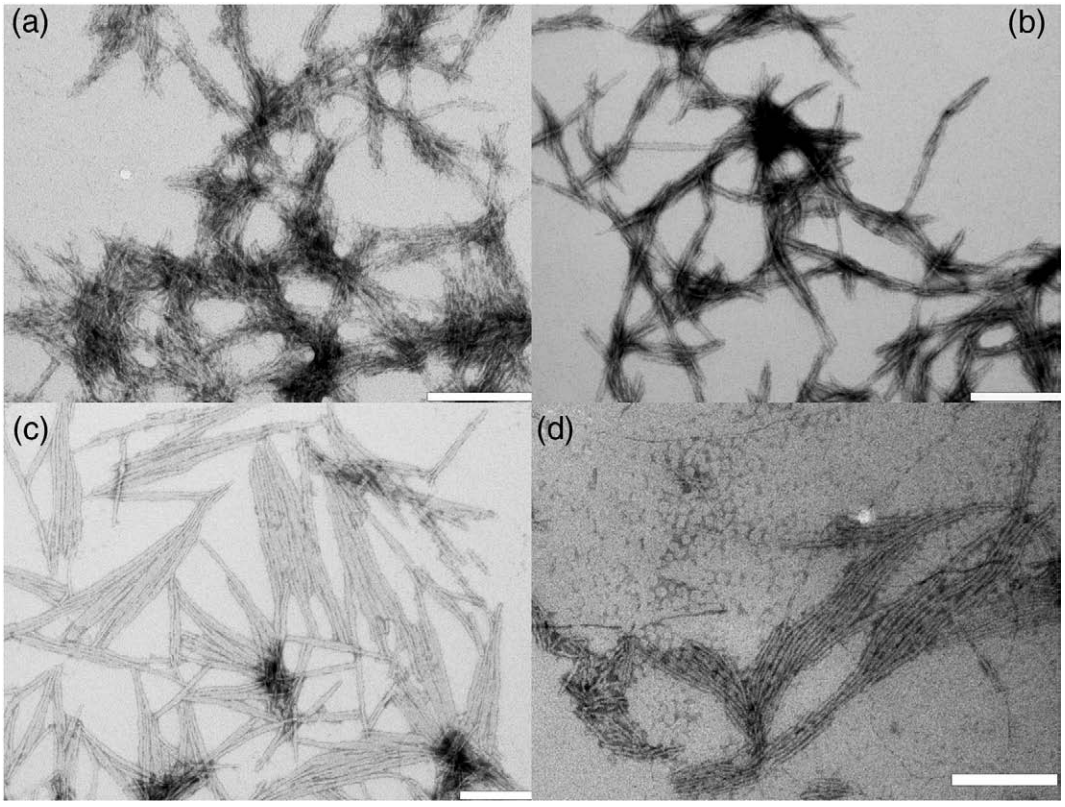


Fig. 5. TEM image of polyQ peptide aggregates. Samples were prepared at 20 μM in PBSA and then incubated at 37 °C prior to imaging. Images are representative of a large number of fields examined. The length of the white bar is 200 nm. (a) Q20, 3 h incubation. (b) Q24, 3 h incubation. (c) Q20, 40 days incubation. (d) Q16, 40 days incubation.

Solutions were incubated for 40 days and then imaged. Again, no aggregates were detected in the Q8 and Q12 samples. Aggregates were observed in Q16, Q20, and Q24; these were more plentiful than in the early time samples. Figure 5c shows a representative Q20 aggregate while Fig. 5d shows a Q16 aggregate. Morphologically similar aggregates were observed in Q24 samples (not shown). The mature aggregates shown in Fig. 5c and d are of similar cluster size but exhibit a more ordered appearance than the immature aggregates of Fig. 5a. For example, the aged Q20 aggregates are clearly fibrillar, with laterally aligned rods that are very well defined. Frequently, the ends of bundled fibrils contact the sides of a single rod, but it was rare to observe rods piled on top of one another, in contrast to the early aggregates. The aged aggregates of Q16 shown in Fig. 5d exhibit a high degree of ordered bundling of fibrils.

Sedimentation assay

A sedimentation assay was used to study kinetics of growth of insoluble aggregates. Results are shown in Fig. 6. The sedimentation assay utilizes the increased density of insoluble aggregates to determine the mass of peptide left in the supernatant, which includes both monomer and soluble aggregates. No sediment was observed in Q8 or Q12 samples even after 35 days of aggregation. This result is consistent with light-scattering and TEM data, as we were unable to detect aggregation in these peptides. With Q16, we did not measure any sedimentable aggregates even after 35 days, despite evidence from light scattering and TEM that aggregates exist. This indicates that either the aggregates present in Q16 remain soluble or the mass fraction of sedimentable aggregates in the sample is smaller than the error in this measurement (<~5%). Sedimentable aggregates clearly form in the case of Q20 and Q24, with Q24 aggregating more quickly and to a greater extent than Q20. Both Q20 and Q24 undergo a lag phase before sedimentable aggregates are detected. For Q20, the lag phase lasts approximately 5 days, while Q24 aggregates after 3 days.

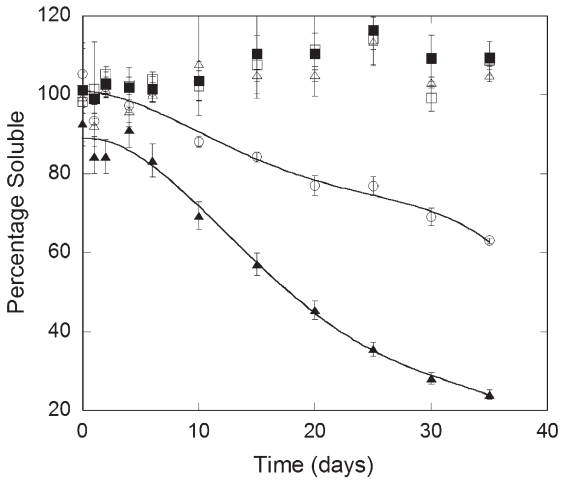


Fig. 6. Sedimentation kinetics of polyQ peptides. Peptides [Q8 (△), Q12 (□), Q16 (■), Q20 (○), and Q24 (▲)] were diluted into PBSA at 20 μM, aliquoted into tubes, and incubated at 37 °C. At the times indicated, a tube was centrifuged and the peptide concentration in the supernatant was determined using a bicinchoninic acid assay. Data are reported as the wt% peptide soluble, calculated from the ratio of the supernatant to the total concentration in an uncentrifuged sample. Lines are drawn as an aid to the eye.

SEC analysis

Freshly prepared peptide samples were injected onto a calibrated size-exclusion column (3–70 kDa range), and elution peaks were monitored by absorbance. In all cases, a single peak was observed, corresponding to a molecular weight slightly larger than expected based on column calibration. Q8 eluted at a molecular weight of 2420 (1854 expected), Q12 at 2939 (2366 expected), Q16 at 3181 (2879 expected), Q20 at 3892 (3391 expected), and Q24 at 5000 (3904 expected). No peaks eluting at higher molecular weights were detected. Identical samples were injected without the column in place, and the peak area of the eluting material was measured. The percentage of monomer recovered was calculated as the area of the monomer peak divided by the total peak area without the column. The percentage recovered decreased from ~90% for Q8 and Q12, to 83% for Q16, to ~76–78% for Q20 and Q24 on the fresh samples. The material that was not recovered in the monomer peak could include monomers that are adsorbed or otherwise lost elsewhere in the system, soluble oligomers that elute as a broad band rather than as a peak and therefore are not detected above baseline, and/or aggregates that adsorb to, or are trapped by, the column or filter and do not elute.

Repeated injections of Q16, Q20, and Q24 were made over a 35-day period (Table 2). The elution time of the monomer peak did not change (data not shown). The monomer peak area for each peptide and each day was normalized by the peak area without the column in place (% recovered) or the monomer peak area on day 0 (% initial). These data mirrored the trends observed with the sedimentation assay. Specifically, there was a lag time during which the monomer peak area did not change, and when the peak area did decrease, the change was detected earliest for Q24 and latest for Q16. Numerical values for monomer recovery at 35 days are consistent with the results from the sedimentation assay (Fig. 6).

Effect of pH on mean end-to-end distance and aggregation

As described earlier, we designed our peptides with two lysines flanking both ends of the polyglutamine stretch. Thus, at pH 7.4, the peptides have a net +4 charge. These lysines facilitate solubility but

Table 2. Mass fraction of polyQ peptide monomer measured by SEC

Time (days)	Q8		Q12		Q16		Q20		Q24	
	% rec	% initial	% rec	% initial	% rec	% initial	% rec	% initial	% rec	% initial
0	88	100	90	100	83	100	76	100	78	100
2	93	105	88	97	83	100	76	99	75	97
5	95	107	95	105	86	104	75	99	80	102
10	ND	ND	ND	ND	86	104	77	102	72	92
35	ND	ND	ND	ND	67	81	47	61	19	24

Peptides were diluted into PBSA at 20 μM and incubated at 37 °C, and solutions were analyzed by SEC. Percentage recovered (% rec) equals the peak area of the monomer divided by the peak area of the peptide injected without the column, times 100%. Percentage initial (% initial) equals the peak area of the monomer divided by the peak area of the monomer at 0 days, times 100%. ND, not determined.

likely also influence the conformational and aggregation properties through electrostatic repulsion. We diluted disaggregated peptide into buffer at pH 12, above the pK_a of the lysine side chain, where the lysines should be uncharged. FRET data were collected and analyzed to obtain values of R at pH 12. Neutralization of the lysine charges led to a small (~5%) decrease in R for Q8 relative to the results in PBS. The effect of pH was dramatically greater for the longer peptides, with R for Q24 decreasing by ~25% with the change from pH 7.4 to pH 12. In fact, at pH 12, R was essentially independent of peptide length (Fig. 2a).

We fitted the data to $P(r)$ for a wormlike chain as described earlier and found that the effective persistence length decreased from ~9 Å for Q8 to less than 3 Å for Q24 (Fig. 2b). Thus, the electrostatic repulsion provided by the charged lysines at neutral pH masks the intrinsic tendency for polyQ peptides to collapse, and the influence of the flanking residues becomes increasingly important as the Q length increases.

We examined peptide solutions at pH 12 by light scattering. For Q8 and Q12, although there was some scattering above background, the signal was weak and there was no growth in aggregate size or scattering intensity over a 4-h measuring period. For Q16 at pH 12, scattered intensity and growth kinetics were similar to that obtained in PBS. For Q20, scattered intensity was about 4 times stronger at pH 12 than in PBS and the growth kinetics were similar. We were unable to collect light-scattering data for Q24 because the peptide immediately formed insoluble precipitates upon dilution into pH 12 buffer. Thus, the extent of aggregation was accentuated by neutralization of lysines, but the Q16 still sat at the transition point separating non-aggregating and aggregating lengths of polyglutamine peptides.

Discussion

Two groups that have examined the conformation of polyQ peptides in aqueous solution came up with surprisingly different conclusions. Tryptophan–cysteine triplet quenching experiments led one group to conclude that polyQ peptides are unusually extended.[28] Based on fluorescence correlation spectroscopy, another group concluded that polyQ peptides are collapsed and that water is a surprisingly poor solvent.[29] There are two key differences between the peptides used by the two groups that may explain these discrepant results: length and charge. Singh and Lapidus worked with peptides of 4 to 16 Q's, while Crick *et al.* used peptides containing 15 to 53 Q's. Second, Singh and Lapidus worked with peptides with two flanking lysines at either end of the polyglutamine domain (KKWQ$_n$CKK), similar to our peptides, whereas Crick *et al.* used peptides of the form GQ$_n$C∗KK, where C∗ indicates a cysteine coupled to a fluorescent dye that carries two negative charges. Thus, the Singh peptide has a net charge of +4, whereas the Crick peptide contains both positive and negative charges but is overall neutral.

As an alternative approach to the conformational question, we used FRET to measure the mean end-to-end distance of monomeric polyQ peptides of Q lengths varying from 8 to 24 residues. We observed that the average R increased only slightly as the number of Q residues increased (Fig. 2). We interpreted the measured R by comparing calculated values for freely jointed chains with a bond length of one residue (R_{FJ}), fully extended (rigid) chains (L_c), or polyQ peptides in a theta solvent (R_θ) (Table 1). In PBS, R for all five peptides was greater than that for the freely jointed chain but significantly less than that for a fully extended chain. Q8 is relatively extended ($R/L_c = 0.58$, $R/R_\theta = 1.16$), whereas Q24 is relatively collapsed ($R/L_c = 0.25$, $R/R_\theta = 0.84$). This comparison led us to conclude that PBS is a good solvent for our shorter peptides (e.g., Q8) but a poorer solvent for longer peptides (e.g., Q24). The transition ($R \sim R_\theta$) occurs at Q16.

To look at the data in a different way, we also used a wormlike chain model to calculate an effective persistence length. The persistence lengths shown in Fig. 2b clearly trend downward with increasing polyQ length, from 11 Å for Q8 to 7 Å for Q24. Singh and Lapidus reported a persistence length of 13.5 Å for their series of short polyQ peptides.[28] Their value is similar to, but somewhat greater than, our value of 11 Å for Q8 in PBS; the difference could be explained by the fact that triplet quenching and FRET sample different parts of the conformational distribution and/or inaccuracies in the wormlike chain model. Thus, our data are consistent with that

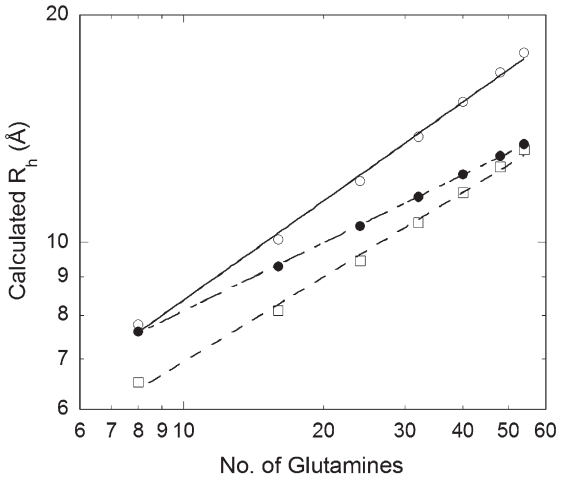


Fig. 7. Hydrodynamic radius of polyQ peptides from wormlike chain model. R_h was calculated from the Yamakawa–Fujii equations for a wormlike chain,[44] assuming a chain diameter of 4 Å and a contour chain length (in angstroms) of $3.8N = 3.8(N_Q + 6)$, where $N = N_Q + 6$ accounts for the number of glutamines in the peptide plus six flanking residues. Results are shown for $l_p = 5$ Å (□, $\nu = 0.37$), $l_p = 12$ Å (○, $\nu = 0.44$), and $l_p = 22\ N_Q^{-0.36}$ (●, $\nu = 0.30$).

of Singh and Lapidus: short polyQ peptides with flanking charged lysines are quite extended in aqueous buffer.

Using results from fluorescence correlation spectroscopy measures of translational diffusion times, Crick *et al.* reported that R_h of polyQ peptides varies as $N^{0.32}$ and therefore concluded that the peptides were highly collapsed and essentially globular.[29] We calculated the hydrodynamic radius for wormlike chains as a function of l_p and N, assuming a chain thickness of 4 Å and using analytical expressions derived by Yamakawa and Fujii.[44] We then fit the calculated values to $R_h = aN^{\nu}$ and found that $\nu = 0.33$ (compact sphere) only when the persistence length is very short ($l_p = 2$ Å). We observed $l_p < 3$ Å for Q24 at pH 12. Thus, our data are consistent with that of Crick *et al.*: longer polyQ peptides with no flanking charged residues are highly collapsed in aqueous buffer.

Our data indicated that l_p decreases with increasing peptide length. For further comparison with the Crick data, we fit a simple power law expression to l_p *versus* the number of Q's, N_Q, and obtained $l_p \sim 22\ N_Q^{-0.36}$. If we then use the variable l_p to calculate R_h as a function of polyQ peptide length and plot R_h *versus* N_Q, we find an apparent $\nu \sim 0.30$ (Fig. 7). In other words, a Q-length-dependent l_p flattens a plot of R_h on N_Q and leads to an estimate of a smaller ν than what one would obtain if one used a constant l_p. Without taking into account the variable persistence length, one would conclude that the polypeptides are more collapsed than they actually are. This is one additional possible contribution to the discrepancy between the Singh and Lapidus conclusion that polyQ is extended and the Crick *et al.* conclusion that polyQ is collapsed.

Our data show that charge (or lack thereof) of the flanking residues strongly influences the conformational properties of polyQ peptides, as illustrated by a comparison of FRET data taken at pH 7.4, where our peptides have a net +4 charge, and at pH 12, where there is no net charge. When the charge on the flanking lysine residues is neutralized by raising the pH to 12, Q8 remains relatively extended, with $R > R_\theta$ and $l_p \sim 9$ Å. Thus, even in the absence of electrostatic repulsion, we surmise that Q8 remains relatively well solvated. However, the effect of pH becomes dramatically stronger as the Q length increases. In the absence of electrostatic repulsion from the flanking lysines, R for Q24 approaches the freely jointed chain limit (R_{FJ}) and l_p decreases dramatically to <3 Å. These results demonstrate the strong intrinsic tendency of longer polyQ peptides towards collapse, as well as the major role that repulsive interactions from flanking residues play in mitigating collapse.

It is useful to compare our estimated l_p values to other data in the literature. A short persistence length, 3 Å, was reported for loops in folded proteins.[43] From FRET data reported by Tucker *et al.* for a 14-mer amphipathic peptide, Mastoparan X,[45] we also calculated a persistence length of 3 Å. More typically, unfolded proteins and polypeptides have persistence lengths ranging from 5 to 9 Å. For example, using single-molecule pulling experiments, l_p was found to be 6–7 Å for unfolded titin.[46] Similarly, Buscaglia *et al.* found $l_p = 6$–7 Å for $C(AGQ)_nW$, with $n = 1$–9.[47] We applied the wormlike chain analysis to FRET data reported in the literature and calculated $l_p = 6.3$, 6.5, and 9.9 Å for 14-mers of repeating AGQ, AGE, and AQE units, respectively.[32] A similar analysis of data from a study of prion repeats[30] containing 11, 21, and 31 total residues yielded estimates of $l_p = 5.6$, 6.8, and 5.1 Å, respectively, showing no dependence on the length of the segment, unlike our polyQ peptides. Schuler *et al.*[48] analyzed single-molecule FRET data for polyproline and found $l_p = 44$ Å, with a slight increase with increasing number of Pro residues. This result is not surprising given that polyproline is very stiff and is used as a "molecular ruler". In comparing these literature values to our results, we conclude that short polyQ peptides in relatively good solvents (e.g., Q8 in PBS) are more extended than typical unfolded peptides, while longer polyQ peptides in the absence of repulsive interactions in flanking residues (e.g., Q24 in pH 12) are highly collapsed.

We propose the following simple hypothesis to explain why polyQ peptides become more collapsed as the Q length increases. Glutamine is considered a hydrophilic polar side chain;[40] the amide side chain participates in hydrogen bonding with water as both a donor and an acceptor. With short polyQ peptides, we hypothesize that hydrogen bonding between side-chain amides and water is prevalent and that the peptide is well solvated and is therefore

somewhat more extended than the typical polypeptide. As the number of Q residues and the length increase, random contacts between distant parts of the chain become more likely, and the probability of forming intramolecular hydrogen bonds (either side chain–side chain, side chain–backbone, or backbone–backbone) increases. Formation of multiple intramolecular hydrogen bonds traps the collapsed conformation, leading to a decrease in the effective persistence length. Although the peptides become more collapsed as the number of glutamines increases, they do not develop any regular secondary structure (Fig. 1). Our data and hypothesis are consistent with several theoretical studies. For example, Marchut and Hall describe polyglutamine monomers at lower temperatures as collapsed random coils.[26] Vitalis *et al.* showed by atomistic simulations that polyglutamine has no preference for a specific, unique conformation.[27] A very recent theoretical study proposes that the total number of intramolecular hydrogen bonds, especially those involving side-chain groups, increases as Q length increases from 5 to 30 and that these collapsed peptides remain disordered rather than forming a regular β-sheet structure.[49]

It is important to note here that the FRET data could be skewed if the peptides were aggregated to a significant extent. Aggregation could place the fluorescence donors/acceptors of multiple peptides in close proximity to each other, causing significant error in the measurement of FRET efficiency and length of the polyQ region. This issue was controlled for by using a low (5 μM) concentration of polyQ peptides and taking measurements at room temperature. Further, R was found to be independent of peptide concentration or incubation time, indicating that aggregation is unlikely to influence the FRET data.

We next examined the aggregation properties as a function of Q length, using a variety of techniques. Previous work in our laboratory showed that early, soluble aggregates were detected in a polyQ peptide by laser light scattering long before sedimentable aggregates were found using other methods.[34] These results were confirmed in the present study. No soluble aggregates were detected for Q8 or Q12 in PBS by light scattering, and none were observed in TEM or by sedimentation assay. Q16 appears to straddle a transition point: we detected immediate formation of soluble aggregates by light scattering and found aggregates in aged samples but not fresh samples by TEM. SEC analysis suggested some loss of monomer for Q16 after 35 days, although there were no sedimentable aggregates. In PBS, Q20 and Q24 both immediately formed soluble aggregates, detectable by both light scattering and TEM. The higher scattered intensity for Q24 compared to Q20, coupled with the similar average size, suggests that the primary difference is the fraction of material incorporated into the soluble aggregates. Sedimentable aggregates and loss of monomer were detected in both samples but only after several days' incubation. The lag in appearance of insoluble aggregates is consistent with the lag phase observed by other groups for polyQ aggregation.[20,50] This lag phase has been cited as evidence of a nucleation–elongation-type mechanism for aggregate formation.[19–21] However, the sedimentation assay is only sensitive to insoluble aggregates; the light-scattering data of Figs. 3 and 4 show no lag time in formation of soluble aggregates. This result, along with previous work in our laboratory,[34] indicates that the lag phase is a lag in the accumulation of sedimentable aggregates, not a lag in the formation of aggregates themselves.

Our experimental results are summarized in Table 3. Interestingly, the polyglutamine length at which the transition between non-aggregating and aggregating peptides is observed, Q16, is also the length at which the mean end-to-end distance transitions from a good to a poor solvent ($R \sim R_\theta$). This result suggests that collapse is strongly correlated with aggregation.

Neutralization of the lysine charges by adjusting the pH to 12 affected both conformation and aggregation, with a greater effect as the Q length increased. Interestingly, R became essentially independent of Q length under these conditions. This result is consistent with a recent single-molecule study that showed that polyQ domains inserted into an unfolded protein were remarkably resistant to mechanical stretching and that the length of the polypeptide under stretching forces was independent of Q length.[51] For the shortest polyQ peptides, neutralization of the charge on the flanking residues has minor impact on R. This result suggests that short polyQ peptides are normally well solvated and prefer extended conformations, independent of the flanking residues. In contrast, in the absence of repulsive interactions, longer polyQ peptides have a

Table 3. Summary of experimental results

Peptide	R/R_θ (FRET)	Soluble aggregates (LS)?	Aggregates at 3 h (TEM)?	Aggregates at 40 days (TEM)?	Monomer loss at 35 days (SEC)?	Sedimentation at 35 days?
Q8	1.16	No	No	No	No	No
Q12	1.03	No	No	No	No	No
Q16	0.96	Yes	No	Yes	Yes (19%)	No
Q20	0.86	Yes	Yes	Yes	Yes (39%)	Yes (37%)
Q24	0.84	Yes	Yes	Yes	Yes (76%)	Yes (76%)

R/R_θ is the ratio of the measured average distance spanned between donor and acceptor R compared to that calculated for a polyQ peptide in a theta solvent (from Vitalis *et al.*[27]).

strong intrinsic tendency towards collapse. With neutralization of the charge on the flanking residues, both the extent of collapse (as determined by FRET) and the degree of aggregation (as determined by light scattering and observation of precipitates) become more strongly dependent on Q length. In other words, repulsive interactions of the flanking residues modulate the otherwise strong Q length dependence of both conformation and aggregation. Interestingly, the Q length at which a transition from minimal to strong aggregation occurs remains at Q16 at both neutral and basic pH.

Close examination of the TEM images (Fig. 5) leads to a number of observations that may shed further light on the mechanism of polyglutamine aggregation. First, the Q20 and Q24 aggregates formed after 3 h are roughly similar in size and shape—both appear as loose clusters of linear aggregates. However, the Q20 aggregates have a more amorphous appearance than Q24, with less sharply defined surfaces. The mature aggregates of Q20 take on a much more defined and highly ordered appearance compared to the early Q20 aggregates, while preserving the general shape and size of the clusters. We did not observe any of the more amorphous aggregates in the aged sample. These observations lead us to propose that the immature soluble aggregates are not "off-pathway" but undergo a slow internal conformational rearrangement over time that leads to the highly ordered mature aggregates. The fact that the early Q24 aggregates look more defined than the early Q20 aggregates could indicate that the Q24 peptide matures more quickly than the Q20 aggregate. We hypothesize that the slow internal conformational rearrangement involves making and breaking of multiple hydrogen bonds and dehydration, with the net effect being an increase in regular β-sheet structure, an increase in packing density, and a decrease in solubility of the aggregates. If true, then sedimentation is an indication of this conformational rearrangement within the preformed large aggregates. Our interpretation of the data is conceptually consistent with conclusions from a recent theoretical study.[49]

In the mature aggregates, the degree of order is striking, as the rodlike aggregates seem to have a clear preference for lateral alignment. Interestingly, the ends of the rods may have a similar property. In Fig. 5c, there are a number of locations where the ends of multiple rods contact the sides of another rod and seem to hold it into place. Thus, it is possible that once a certain number of contacts are made, be it through lateral alignment with another rod or with end-to-side contact with multiple rods, the relative positions of the rods become fixed. Additionally, it appears as though most of the aggregates lie in a single plane, with only a few examples of rods lying on top of each other. In other words, the conformational rearrangement leads to increased asymmetry of the rodlike structures, resulting in different properties on the top and bottom faces compared to those of the sides and ends.

We propose a schematic (Fig. 8) for Q-length-dependent aggregation that is consistent with our results. Short polyQ peptides occupy extended conformations while longer polyQ peptides adopt collapsed coil conformations; all lack regular secondary structure. Collapse is facilitated by a switch from peptide–solvent hydrogen bonding to intramolecular hydrogen bonding as Q length grows (phase 1). The amide side chain of glutamine allows for side chain–side chain and side–chain backbone hydrogen bonding, as well as backbone–backbone hydrogen bonding. Based on molecular dynamics simulations, all three types of hydrogen bonds are likely to occur in polyglutamine aggregates.[26,49] Water becomes a poorer solvent, and the collapsed polypeptides tend to self-associate (phase 2). The transition occurs at Q16. In phase 3, the aggregation-competent peptides rapidly associate into soluble oligomers that grow over time. The mass fraction of peptide incorporated into soluble oligomers is higher for Q24 than for Q16 or Q20, based on the increased scattering intensity (Fig. 4). These soluble aggregates lack regular secondary structure. However, at lower resolution, they possess some degree of order: by TEM, the early aggregates appear as loose clusters of rodlike particles. In phase 4, aggregates transition to fully mature, insoluble

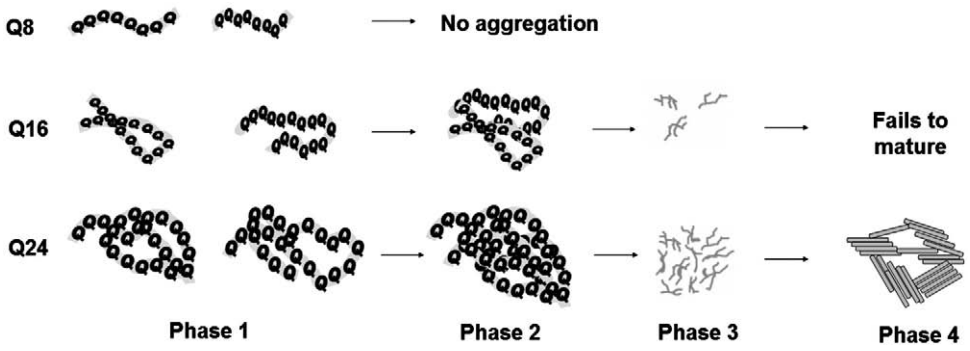


Fig. 8. Schematic illustrating the proposed kinetic pathway. Phase 1 consists of disordered peptides, with longer peptides showing more collapse. In phase 2, peptides with Q≥16 associate in a nonspecific manner, leading to the amorphous aggregates present in phase 3. The amount of aggregated material in phase 3 is proportional to the number of glutamines. In phase 4, Q24 forms fully mature, structured, insoluble aggregates, while Q16 failed to sediment in 35 days.

aggregates, characterized by a higher degree of order. The gross morphological similarities between the aggregates imaged after 3 h and those imaged after 40 days suggest that the soluble aggregates are precursors to the insoluble aggregates rather than "off-pathway". We hypothesize that maturation of the aggregates involves dehydration and an increase in density, thus leading to sedimentation. Over the time frame we investigated, Q20 and Q24 but not Q16 aggregates make this final transition.

Materials and Methods

Peptide synthesis and purification

All materials were from Fisher Scientific (Pittsburgh, PA) except where indicated. Peptides were synthesized using standard Fmoc solid-phase methods on a PerSeptive Biosystems Pioneer synthesizer. Alanine, glutamines with a Trityl side-chain protecting group, and lysines and tryptophan with a Boc side-chain protecting group were purchased from Novabiochem (Gibbstown, NJ). Dansylated lysine was purchased from Anaspec (San Jose, CA). The resin used was Fmoc-PAL-PEG-PS from Applied Biosystems (Foster City, CA). Resin sites were partially blocked with lysine-Boc to reduce on-bead aggregation. Extended cycles and double couplings were used to improve yield. The peptides were cleaved from the resin using a solution of 95% trifluoroacetic acid (TFA), 2.5% ethanedithiol (Fluka, Buchs, Switzerland), and 2.5% H_2O. Peptides were precipitated into cold *t*-butylmethyl ether and digested in aqueous solvent for 1 h before lyophilization. Crude peptide was pre-solubilized in a 1:1 solution of TFA and hexafluoroisopropanol (Fluka). This solution was evaporated under gentle N_2 flow, and the peptide was resuspended in water adjusted to pH 3 before purifying by reverse-phase HPLC on a Vydac C18 column. Peptide was eluted from the column with a linear gradient of acetonitrile and water with 0.1% TFA, starting at 5% and ending at 30% acetonitrile. The peak corresponding to the correct sequence was collected and lyophilized. Peptide identity was confirmed by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, which yielded molecular weights of 1854.0 (Q8, 1854.2 theoretical), 2366.2 (Q12, 2366.7 theoretical), 2879.5 (Q16, 2879.2 theoretical), 3390.7 (Q20, 3391.8 theoretical), and 3903.9 (Q24, 3904.3 theoretical). For the dansylated peptides, molecular weights were 2144.0 (Q8D, 2144.2 theoretical), 2656.0 (Q12D, 2656.7 theoretical), 3168.8 (Q16D, 2169.2 theoretical), 3680.8 (Q20D, 3681.7 theoretical), and 4194.0 (Q24D, 4194.2 theoretical).

Sample preparation

Lyophilized peptides were disaggregated using a protocol similar to that developed by others.[52] Briefly, peptide was incubated overnight in a 1:1 solution of TFA and hexafluoroisopropanol. Solvent was evaporated under a gentle flow of N_2, and the peptides were redissolved in water adjusted to pH 3 with TFA. Peptide stock solutions were filtered through a 0.45-μm filter, aliquoted, snap frozen in dry ice and ethanol, and stored at −80 °C. Prior to each experiment, a vial was thawed and centrifuged at 19,500 rcf for 30 min, and then the supernatant (top 75%) was removed and used immediately. Although the TFA used was the purest available (99.5% TFA for biochemistry, Acrōs Organics), evaporation of TFA was found to leave a residue that interfered with the absorbance of the peptide at 280 nm. Therefore, the concentration of peptide was determined using the fluorescence intensity of standard solutions of tryptophan in 6 M GdmCl. The concentration of tryptophan was determined by absorbance at 280 nm using an extinction coefficient of 5500 cm^{-1} M^{-1}. Tryptophan standards of 5 to 50 μM were used to produce a linear standard curve using the emission intensity at 350 nm with an excitation wavelength of 295 nm. Peptide concentrations were determined from the fluorescence intensity of a peptide stock sample diluted into 6 M GdmCl fit to the standard curve; the peptide concentration in the stock solution was typically ~200 μM. The concentration of the dansyl peptides was determined using a similar method, with the standard curve based on dansyl lysine (Anaspec), using an extinction coefficient at 330 nm of 4600 cm^{-1} M^{-1} to determine the stock concentration[53] and an excitation wavelength of 350 nm and an emission wavelength of 549 nm. Dansyl was used as the standard rather than tryptophan, as tryptophan fluorescence could be quenched by the dansyl group in these peptides.

Fluorescence resonance energy transfer

FRET requires both a fluorescence donor and an acceptor, where the emission spectrum of the donor overlaps the absorbance spectrum of the acceptor. Excitation of the fluorescence donor results in emission from the donor group. If the acceptor group is located in close proximity to the donor group, excited-state energy is transferred from the donor to the acceptor. The decrease in measured emission from the donor is used to calculate the distance between the donor and acceptor groups.[54]

Peptide samples for FRET experiments were diluted to 5 μM in pH 7.4 PBSA (0.01 M buffer salts, 0.14 M NaCl, and 3 mM sodium azide) or pH 12 potassium chloride buffer (0.01 M KCl and NaOH to adjust to pH 12). Limited additional data were obtained at 1 μM peptide to ascertain the effect of concentration. Fluorescence spectra for solutions of both Q_n and Q_nD peptides were obtained with a QuantaMaster Series spectrofluorometer (PTI, Inc., Birmingham, NJ), with excitation at 295 nm and emission spectra recorded from 310 to 550 nm. Measurements were made in duplicate and at steady state. An excitation wavelength of 295 nm was used to avoid energy transfer between two Trp groups that occurs with excitation at 280 nm.[55] FRET efficiency E was calculated as[54]

$$E = 1 - \frac{F_{DA}}{F_D} \tag{6}$$

where F_{DA} is the donor (tryptophan) fluorescence with the acceptor (dansyl) group present and F_D is the donor fluorescence with the acceptor group absent. The values of F_{DA} and F_D in Eq. (6) were determined by integrating the peak area of the emission scan, with background subtracted, from 330 to 370 nm.

Circular dichroism

Peptide stock solutions were diluted into phosphate buffer (10 mM K_2HPO_4/KH_2PO_4 and 140 mM NaF, pH 7.4) to a concentration of 20 μM peptide. Samples were filtered through a 0.45-μm membrane immediately before transfer to a 1-mm cell. CD spectra were collected

using an Aviv 202SF CD spectrometer at 37 °C immediately after filtration. Solvent spectra were collected and subtracted. An identical set of samples was produced as described above and incubated in sample tubes for 13 days at 37 °C before collecting CD spectra.

Transmission electron microscopy

Peptides at 20 μM were prepared in PBSA and incubated for 3 h or 40 days at 37 °C. A drop of sample was applied to a pioloform-coated grid and stained with methylamine tungstate stain and then imaged with a Philips CM120 scanning transmission electron microscope.

Sedimentation assay

Peptides were diluted into PBSA to a concentration of 20 μM, distributed into sample tubes in 100-μL aliquots, and incubated at 37 °C. At regular time intervals, a sample tube was withdrawn and centrifuged at 19,500 rcf for 30 min. The supernatant (top 75%) was removed, and the peptide concentration was determined in triplicate with a bicinchoninic acid (Pierce, Rockford IL) assay in the microplate format, using a BioTek EL 800 Universal Microplate Reader and following the manufacturer's protocol. All concentrations are reported as a percentage of the concentration of an uncentrifuged aliquot.

Laser light scattering

Samples were diluted into PBSA that had been filtered twice with a 0.22-μm filter and then immediately filtered again through a 0.45-μm filter directly into a clean light-scattering cuvette. The cuvette was placed in a 37 °C bath of the index-matching solvent decahydronaphthalene. Data were collected using a Malvern 4700c system and a Coherent Innova 90C argon ion laser operating at 488 nm wavelength. The autocorrelation function at 90° scattering angle was measured repeatedly over several hours. The z-averaged diffusivity was determined from the autocorrelation function by the method of cumulants. R_{hz}, the apparent z-averaged hydrodynamic radius, was calculated from $\langle D \rangle_z$ using the Stokes equation.[56] On the same samples and at the same time points, the total intensity of scattered light at 90°, $I_s(90°)$, was measured. Scattering of the solvent, $I_b(90°)$, was subtracted from $I_s(90°)$, and the result was normalized to the scattered intensity of toluene $I_{tol}(90°)$, to adjust for changes in laser power or aperture, and to the mass concentration of peptide (c_{tot}, mg/mL). Data are reported as a normalized intensity I_{norm}:

$$I_{norm} = \frac{I_s(90^\circ) - I_s(90^\circ)}{c_{tot} I_{tol}(90^\circ)} \quad (7)$$

I_{norm} is proportional to $\langle M \rangle_w$, the weight-averaged molecular weight, and $P(90°)$, the particle scattering factor at 90° scattering angle:

$$I_{norm} \propto \langle M \rangle_w P(90^\circ) \quad (8)$$

For small R_g, $P(90°)$ is independent of shape and

$$P(90^\circ)^{-1} \approx 1 + \frac{q^2 R_g^2}{3} \quad (9)$$

where R_g is the radius of gyration and q is the scattering vector. With our system, $q = 0.0242$ nm^{-1}; hence, this approximation is valid if $R_g < \sim 40$ nm. At larger particle size, $P(90°)$ depends on the particle shape and deviates from Eq. (9); analytical expressions for many particle shapes are available.[57]

Size-exclusion chromatography

Samples were prepared as described and incubated at 37 °C. Aliquots were removed at regular time intervals and analyzed by SEC on a Superdex 75 PC 3.2/30 column (Pharmacia), with a molecular mass range of 3–70 kDa, using a Waters 625LC system. The mobile phase (PBS, pH 7.4) flow rate was set to 0.1 mL/min, an overfilled 50 μL sample loop was used to apply a standard volume of sample, and peaks were detected by absorbance at 280 nm. The column was calibrated with a set of molecular weight protein standards. The weight fraction of a species eluting from the column at a specific elution time was determined by dividing the area of that peak by the total area of an identical injection without the column in place.

Acknowledgements

This work was supported by grant BES-0330537 from the National Science Foundation and grant 5T32 GM-08349 (R.H.W.) from the National Institutes of Health. We gratefully acknowledge the technical assistance of Dr. Gary Case, Dr. Randy Massey, and the University of Wisconsin Biotechnology Center mass spectrometry staff. CD data were obtained at the University of Wisconsin-Madison Biophysics Instrumentation Facility, which was established with support from the University of Wisconsin-Madison and grants BIR-9512577 (National Science Foundation) and S10 RR13790 (National Institutes of Health).

References

1. Bates, G. (2003). Huntingtin aggregation and toxicity in Huntington's disease. *Lancet*, **361**, 1642–1644.
2. Orr, H. T. (2001). Beyond the Qs in the polyglutamine diseases. *Genes Dev.* **15**, 925–932.
3. Wanker, E. E. (2000). Protein aggregation and pathogenesis of Huntington's disease: mechanisms and correlations. *Biol. Chem.* **381**, 937–942.
4. Gusella, J. F. & MacDonald, M. E. (2000). Molecular genetics: unmasking polyglutamine triggers in neurodegenerative disease. *Nat. Rev., Neurosci.* **1**, 109–115.
5. Albrecht, M., Golatta, M., Wullner, U. & Lengauer, T. (2004). Structural and functional analysis of ataxin-2 and ataxin-3. *Eur. J. Biochem.* **271**, 3155–3170.
6. Williams, A. J. & Paulson, H. L. (2008). Polyglutamine neurodegeneration: protein misfolding revisited. *Trends Neurosci.* **31**, 521–528.
7. Yang, W., Dunlap, J. R., Andrews, R. B. & Wetzel, R. (2002). Aggregated polyglutamine peptides delivered to nuclei are toxic to mammalian cells. *Hum. Mol. Genet.* **11**, 2905–2917.
8. Schiffer, N. W., Broadley, S. A., Hirschberger, T., Tavan, P., Kretzschmar, H. A., Giese, A. *et al.* (2007). Identification of anti-prion compounds as efficient

inhibitors of polyglutamine protein aggregation in a zebrafish model. *J. Biol. Chem.* **282**, 9195–9203.
9. Ross, C. A. (2002). Polyglutamine pathogenesis: emergence of unifying mechanisms for Huntington's disease and related disorders. *Neuron*, **35**, 819–822.
10. Ordway, J. M., Tallaksen-Greene, S., Gutekunst, C. A., Bernstein, E. M., Cearley, J. A., Wiener, H. W. *et al.* (1997). Ectopically expressed CAG repeats cause intranuclear inclusions and a progressive late onset neurological phenotype in the mouse. *Cell*, **91**, 753–763.
11. Zhang, X. Q., Smith, D. L., Merlin, A. B., Engemann, S., Russel, D. E., Roark, M. *et al.* (2005). A potent small molecule inhibits polyglutamine aggregation in Huntington's disease neurons and suppresses neurodegeneration in vivo. *Proc. Natl Acad. Sci. USA*, **102**, 892–897.
12. Weiss, A., Klein, C., Woodman, B., Sathasivam, K., Bibel, M., Regulier, E. *et al.* (2008). Sensitive biochemical aggregate detection reveals aggregation onset before symptom development in cellular and murine models of Huntington's disease. *J. Neurochem.* **104**, 846–858.
13. Takahashi, T., Kikuchi, S., Katada, S., Nagai, Y., Nishizawa, M. & Onodera, O. (2008). Soluble polyglutamine oligomers formed prior to inclusion body formation are cytotoxic. *Hum. Mol. Genet.* **17**, 345–356.
14. Shao, J. & Diamond, M. I. (2007). Polyglutamine diseases: emerging concepts in pathogenesis and therapy. *Hum. Mol. Genet.* **16**, R115–R123.
15. Truant, R., Atwal, R. S. & Burtnik, A. (2007). Nucleocytoplasmic trafficking and transcription effects of huntingtin in Huntington's disease. *Prog. Neurobiol.* **83**, 211–227.
16. Freiman, R. N. & Tjian, R. (2002). Perspectives: neurodegeneration—a glutamine-rich trail leads to transcription factors. *Science*, **296**, 2149–2150.
17. Chen, S., Berthelier, V., Yang, W. & Wetzel, R. (2001). Polyglutamine aggregation behavior in vitro supports a recruitment mechanism of cytotoxicity. *J. Mol. Biol.* **311**, 173–182.
18. Slepko, N., Bhattacharyya, A. M., Jackson, G. R., Steffan, J. S., Marsh, J. L., Thompson, L. M. & Wetzel, R. (2006). Normal-repeat-length polyglutamine peptides accelerate aggregation nucleation and cytotoxicity of expanded polyglutamine proteins. *Proc. Natl Acad. Sci. USA*, **103**, 14367–14372.
19. Chen, S. M., Berthelier, V., Hamilton, J. B., O'Nuallain, B. & Wetzel, R. (2002). Amyloid-like features of polyglutamine aggregates and their assembly kinetics. *Biochemistry*, **41**, 7391–7399.
20. Chen, S. M., Ferrone, F. A. & Wetzel, R. (2002). Huntington's disease age-of-onset linked to polyglutamine aggregation nucleation. *Proc. Natl Acad. Sci. USA*, **99**, 11884–11889.
21. Bhattacharyya, A. M., Thakur, A. K. & Wetzel, R. (2005). Polyglutamine aggregation nucleation: thermodynamics of a highly unfavorable protein folding reaction. *Proc. Natl Acad. Sci. USA*, **102**, 15400–15405.
22. Bernacki, J. P. & Murphy, R. M. (2009). Model discrimination and mechanistic interpretation of kinetic data in protein aggregation studies. *Biophys. J.* **96**, 2871–2887.
23. Masino, L., Kelly, G., Leonard, K., Trottier, Y. & Pastore, A. (2002). Solution structure of polyglutamine tracts in GST-polyglutamine fusion proteins. *FEBS Lett.* **513**, 267–272.
24. Altschuler, E. L., Hud, N. V., Mazrimas, J. A. & Rupp, B. (1997). Random coil conformation for extended polyglutamine stretches in aqueous soluble monomeric peptides. *J. Pept. Res.* **50**, 73–75.
25. Klein, F. A. C., Pastore, A., Masino, L., Zeder-Lutz, G., Nierengarten, H., Oulad-Abdeighani, M. *et al.* (2007). Pathogenic and non-pathogenic polyglutamine tracts have similar structural properties: towards a length-dependent toxicity gradient. *J. Mol. Biol.* **371**, 235–244.
26. Marchut, A. J. & Hall, C. K. (2007). Effects of chain length on the aggregation of model polyglutamine peptides: molecular dynamics simulations. *Proteins*, **66**, 96–109.
27. Vitalis, A., Wang, X. & Pappu, R. V. (2008). Atomistic simulations of the effects of polyglutamine chain length and solvent quality on conformational equilibria and spontaneous homodimerization. *J. Mol. Biol.* **384**, 279–297.
28. Singh, V. R. & Lapidus, L. J. (2008). The intrinsic stiffness of polyglutamine peptides. *J. Phys. Chem. B*, **112**, 13172–13176.
29. Crick, S. L., Jayaraman, M., Frieden, C., Wetzel, R. & Pappu, R. V. (2006). Fluorescence correlation spectroscopy shows that monomeric polyglutamine molecules form collapsed structures in aqueous solutions. *Proc. Natl Acad. Sci. USA*, **103**, 16764–16769.
30. Gustiananda, M., Liggins, J. R., Cummins, P. L. & Gready, J. E. (2004). Conformation of prion protein repeat peptides probed by FRET measurements and molecular dynamics simulations. *Biophys. J.* **86**, 2467–2483.
31. Moglich, A., Joder, K. & Kiefhaber, T. (2006). End-to-end distance distributions and intrachain diffusion constants in unfolded polypeptide chains indicate intramolecular hydrogen bond formation. *Proc. Natl Acad. Sci. USA*, **103**, 12394–12399.
32. Soranno, A., Longhi, R., Bellini, T. & Buscaglia, M. (2009). Kinetics of contact formation and end-to-end distance distributions of swollen disordered peptides. *Biophys. J.* **96**, 1515–1528.
33. Fasman, G. D. (1996). *Circular Dichroism and the Conformational Analysis of Biomolecules.*Plenum Press, New York, NY; 737 pp.
34. Lee, C. C., Walters, R. H. & Murphy, R. M. (2007). Reconsidering the mechanism of polyglutamine peptide aggregation. *Biochemistry,*, **46**, 12810–12820.
35. Lakowicz, J. R. (1999). *Principles of Fluorescence Spectroscopy*, 2nd ed. Springer, New York, NY; pp. 1–20.
36. Dunn, B. M., Pham, C., Raney, L., Abayasekara, D., Gillespie, W. & Hsu, A. (1981). Interaction of alpha-dansylated peptide inhibitors with porcine pepsin—detection of complex-formation by fluorescence energy-transfer and chromatography and evidence for a 2-step binding scheme. *Biochemistry*, **20**, 7206–7211.
37. Wu, P. G. & Brand, L. (1994). Resonance energy transfer—methods and applications. *Anal. Biochem.* **218**, 1–13.
38. Gudgin, E., Lopez-Delgado, R. & Ware, W. R. (1981). The tryptophan fluorescence lifetime puzzle. A study of decay times in aqueous solution as a function of pH and buffer composition. *Can. J. Chem.* **59**, 1037.
39. Flory, P. J. (1953). *Principles of Polymer Chemistry.* Cornell Univ. Press, Ithaca, NY; pp. 404–407.
40. Creighton, T. E. (1983). *Proteins: Structures and Molecular Principles.*W. H. Freeman and Company, New York, NY; pp. 4–10.
41. Zhou, H. X. (2004). Polymer models of protein stability, folding, and interactions. *Biochemistry*, **43**, 2141–2154.
42. Hyeon, C. & Thirumalai, D. (2006). Kinetics of interior loop formation in semiflexible chains. *J. Chem. Phys.* **124**, 104905.

43. Zhou, H. X. (2001). Loops in proteins can be modeled as worm-like chains. *J. Phys. Chem., B*, **105**, 6763–6766.
44. Yamakawa, H. & Fujii, M. (1973). Translational friction coefficient of wormlike chains. *Macromolecules*, **6**, 407–415.
45. Tucker, M. J., Oyola, R. & Gai, F. (2005). Conformational distribution of a 14-residue peptide in solution: a fluorescence resonance energy transfer study. *J. Phys. Chem., B*, **109**, 4788–4795.
46. Kellermayer, M. S. Z., Smith, S. B., Granzier, H. L. & Bustamante, C. (1997). Folding–unfolding transitions in single titin molecules characterized with laser tweezers. *Science,*, **276**, 1112–1116.
47. Buscaglia, M., Lapidus, L. J., Eaton, W. A. & Hofrichter, J. (2006). Effects of denaturants on the dynamics of loop formation in polypeptides. *Biophys. J.* **91**, 276–288.
48. Schuler, B., Lipman, E. A., Steinbach, P. J., Kumke, M. & Eaton, W. A. (2005). Polyproline and the "spectroscopic ruler" revisited with single-molecule fluorescence. *Proc. Natl Acad. Sci. USA*, **102**, 2754–2759.
49. Vitalis, A., Lyle, N. & Pappu, R. V. (2009). Thermodynamics of β-sheet formation in polyglutamine. *Biophys. J.* **97**, 303–311.
50. Nagai, Y., Inui, T., Popiel, H. A., Fujikake, N., Hasegawa, K., Urade, Y. *et al.* (2007). A toxic monomeric conformer of the polyglutamine protein. *Nat. Struct. Mol. Biol.* **14**, 332–340.
51. Dougan, L., Li, J., Badilla, C. L., Berne, B. J. & Fernandez, J. M. (2006). Single homopolypeptide chains collapse into mechanically rigid conformations. *Proc. Natl Acad. Sci. USA*, **106**, 12605–12610.
52. Chen, S. M. & Wetzel, R. (2001). Solubilization and disaggregation of polyglutamine peptides. *Protein Sci.* **10**, 887–891.
53. Bobkov, A. A., Muhlrad, A., Pavlov, D. A., Kokabi, K., Yilmaz, A. & Reisler, E. (2006). Cooperative effects of cofilin (ADF) on actin structure suggest allosteric mechanism of cofilin function. *J. Mol. Biol.* **356**, 325–334.
54. dos Remedios, C. G. & Moens, P. D. J. (1999). Resonance energy transfer in proteins. In (Andrews, D. L. & Demidov, A. A., eds), pp. 1–64, Chichester, UK, Wiley.
55. Weber, G. & Shinitzk, M. (1970). Failure of energy transfer between identical aromatic molecules on excitation at long wave edge of absorption spectrum. *Proc. Natl Acad. Sci. USA*, **65**, 823–830.
56. Shen, C. L., Fitzgerald, M. C. & Murphy, R. M. (1994). Effect of acid predissolution on fibril size and fibril flexibility of synthetic beta-amyloid peptide. *Biophys. J.* **67**, 1238–1246.
57. Burchard, W. (1983). Static and dynamic light-scattering from branched polymers and bio-polymers. *Adv. Polym. Sci.* **48**, 1–124.