Aggregation Behavior of Chemically Synthesized, Full-Length 
Huntingtin Exon1

Bankanidhi Sahoo,† David Singer,‡ Ravindra Kodali,† Thole Zuchner,‡ and Ronald Wetzel‡*†

†Department of Structural Biology and Pittsburgh Institute for Neurodegenerative Diseases, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania 15260, United States
‡Institute of Bioanalytical Chemistry, Faculty of Chemistry and MineroLOGY, Center for Biotechnology and Biomedicine, Leipzig University, Leipzig 04103, Germany

ABSTRACT: Repeat length disease thresholds vary among the 10 expanded polyglutamine (polyQ) repeat diseases, from about 20 to about 50 glutamine residues. The unique amino acid sequences flanking the polyQ segments are thought to contribute to these repeat length thresholds. The specific portions of the flanking sequences that modulate polyQ properties are not always clear, however. This ambiguity may be important in Huntington’s disease (HD), for example, where in vitro studies of aggregation mechanisms have led to distinctly different mechanistic models. Most in vitro studies of the aggregation of the huntingtin (HTT) exon1 fragment implicated in the HD mechanism have been conducted on inexact molecules that are imprecise either on the N-terminus (recombinantly produced peptides) or on the C-terminus (chemically synthesized peptides). In this paper, we investigate the aggregation properties of chemically synthesized HTT exon1 peptides that are full-length and complete, containing both normal and expanded polyQ repeat lengths, and compare the results directly to previously investigated molecules containing truncated C-termini. The results on the full-length peptides are consistent with a two-step aggregation mechanism originally developed based on studies of the C-terminally truncated analogues. Thus, we observe relatively rapid formation of spherical oligomers containing from 100 to 600 HTT exon1 molecules and intermediate formation of short protofibril-like structures containing from 500 to 2600 molecules. In contrast to this relatively rapid assembly, mature HTT exon1 amyloid requires about one month to dissociate in vitro, which is similar to the time required for neuronal HTT exon1 aggregates to disappear in vivo after HTT production is discontinued.

Huntington’s disease1–2 (HD) is one of 10 known expanded CAG repeat diseases,3,4 autosomal dominant genetic disorders in which polyglutamine (polyQ) sequences above characteristic repeat lengths in specific disease proteins trigger neurodegeneration. The molecular mechanism of HD has not been elucidated, but the consensus is that polyQ expanded forms of N-terminal fragments, typified by the translation product of the huntingtin (HTT) gene’s first exon, are susceptible to aberrant folding behaviors that somehow interfere with normal neuronal function or survival. This protein fragment, which has become known as HTT exon1, can be generated in the cell either by proteolytic fragmentation of the full-length HTT protein5 or by translation of an alternatively spliced version of HTT mRNA.6 Animal models based on overexpression of HTT exon1 exhibit robust neurodegenerative phenotypes.7

Regardless of polyQ repeat length, HTT exon1 itself is not stably folded but rather is an amalgam of three sequence-defined segments (Figure 1), each of which exhibits aspects of intrinsically disordered protein (IDP) behavior. The N-terminal 16 or 17 amino acid segment, HTTNT, is disordered in solution in the monomeric state10 but takes on α-helical structure when it self-associates11 and when it interacts with membranes.12 The central polyQ segment exists in the monomer in an energetically favored compact coil state9,13–15 consisting of fluctuating short segments of coil, α, and β structure in which glutamine side chain amides spend significant time H-bonded to main chain amides groups.16 The C-terminal proline-rich domain (PRD), consisting of runs of P10 and P11 alternating with short segments of mixed, Pro-rich sequence, favors polyproline type II (PPII) structure but is not expected to exhibit a single, strongly favored conformation.9,15 Whether and how these three disordered sequence elements interact with each other, through the peptide backbone and through space, to define or modulate the solution properties of HTT exon1 is a challenging problem that is only beginning to be explored.9,17,18

Under the umbrella of opinion that HD is a protein misfolding disease, there is a wide range of ideas about the nature of the toxic misfolded species.3–5 Some research has been interpreted to indicate that repeat expansion favors the
time-dependent formation of an alternatively folded state of polyQ in which the average monomer in solution possesses β-hairpin structure. These ideas are hard to reconcile, however, with our understanding of the energetic features of the polyQ monomer conformational landscape, and there are reasonable alternative interpretations of the supporting data. In contrast, other results have been interpreted to implicate the repeat length and time-dependent formation of some kind of aggregated state of HTT exon1 as potential toxic entities. These ideas are consistent with the well-established existence of large HTT exon1 containing inclusions in neurons of affected individuals and in cell and animal models, and in the similarities between the polyQ repeat length dependences of disease risk and age-of-onset and in vitro and in vivo aggregation behavior. Although compelling evidence has been presented that the large inclusions that appear late in the cellular aggregation time course are more likely to be protective than toxic, it is now clear that smaller amyloid fibrils as well as more difficult to detect nonamyloid aggregates are also generated in cell and animal models. The undiminished feasibility of the aggregation hypothesis makes it imperative that we continue to work to understand molecular mechanisms of aggregation and how different segments of HTT exon1 act and interact to propel or retard these mechanisms.

Data on HTT exon1 self-assembly mechanisms and products have been generated using molecules from two different types of sources. One source generates essentially full-length HTT exon1 made in bacterial or mammalian cell culture, using a fusion protein approach that helps retain the aggregation-prone product in solution in the cell but leads to compromises in the integrity of the HTT exon1 N-terminus when the fusion partner is removed. The other major approach is to use peptides obtained by solid phase peptide synthesis, when the fusion partner is removed. The other major approach is to use peptides obtained by solid phase peptide synthesis, which, however, is limited by the length of peptides that can be conveniently synthesized. For this reason, most synthetic peptide studies have been constrained to relatively short (no longer than the low 40s) polyQ repeat lengths and truncated C-termini. Thus, the choice of the source of HTT exon1 analogues for in vitro studies of self-assembly and toxicity is guided in part by one’s perception of the relative importance of maintaining absolute sequence integrity in the N-terminus vs the C-terminus.

Previous studies using chemically synthesized HTT exon1 fragments led to development of a detailed mechanism (Figure 2) for how peptides in this class self-assemble. The presence of a functioning HTT sequence leads to a dominant pathway (A) featuring early formation of oligomers (b, c, d) consisting of superassembled tetramers that are themselves held together by bundling of α-helical HTT segments (green cylinders). Oligomer formation leads to very high local concentrations of compact coil polyQ segments (orange), which favors stochastic amyloid nucleation events (e) that trigger fibril elongation (f). This mechanism is supported by the identification of early nonamyloid oligomers held together by HTTNT interactions, and the observation of well-behaved tetramers and octamers requiring the presence of HTTNT11, the ability of HTTNT with or without added polyQ to form α-helix in a concentration-dependent manner, and the ability of peptides consisting only of the HTTNT segment to inhibit nucleation by coassembling into mixed oligomers that reduce the local polyQ concentration (Figure 2i). Interestingly, when such inhibitors compromise the HTTNT-mediated pathway (A), the nucleation mechanism of polyQ amyloid reverts to the less efficient pathway (B) favored by simple polyQ peptides. In contrast to the mechanism shown in Figure 2, studies based on recombinant HTT exon1 led to a proposed alternative mechanism in which non-β oligomers are assembled primarily via polyQ interactions. It is not clear whether the stark differences in these mechanistic proposals and the data on which they are based stem from the structural differences between the HTT exon1-like molecules being studied, their synthetic source, the detailed methods used to initiate and monitor reactions, or other factors. Previously, some of us reported the solid phase chemical synthesis of full-length HTT exon1 molecules containing repeats of 23 and 42 glutamine residues. The availability of these well-defined molecules served as the basis for the direct comparative studies reported here, allowing us to cleanly focus on the question of the relative importance of sequences within the proline-rich domain (PRD) downstream from the first 10 prolines. The results confirm a number of key points in the mechanism devised previously based on work with C-terminally truncated HTT exon1 peptides, suggesting that for many

Figure 2. Mechanisms of polyglutamine amyloid assembly. Monomeric HTT exon1 analogues (a) can assemble into polyQ-core amyloid by two mechanisms, nominally in competition. Pathway A. Monomers assemble into tetramers via concerted α-helix formation and bundling of the N-terminal 16–17 amino acid segment HTTNT (green) (b–d). Non-β oligomers assemble via poorly understood self-association of tetramers. This reversible oligomerization brings the polyQ chains (orange) close together in space at a very high local concentration, facilitating sporadic amyloid nucleation (e), which leads to amyloid elongation (f) into fibrils. The proline rich C-terminal segment (black) tends to favor PPII conformations throughout. Pathway B. If pathway A is compromised, for example, by coassembly with HTTNT peptides (i) (see text), amyloid nucleation can still occur by the classical nucleated growth polymerization type mechanism previously described for simple polyQ sequences. Reproduced from ref 46. Copyright 2012 American Chemical Society. dx.doi.org/10.1021/bi500300c
studies C-terminally truncated HTT exon1 analogues can provide information relevant to the behavior of full-length HTT exon1. The results also provide new information on the role of the PRD on aggregate assembly and stability.

## EXPERIMENTAL PROCEDURES

**Materials and General Methods.** Preparation of purified full-length HTT exon1 peptides has been described. Other synthetic peptides were obtained from the Keck Biotechnology Center at Yale University and were purified by reverse phase HPLC as described. Mass spectrometry evaluation of the purified peptides, which run as a single peak in analytical reverse phase HPLC, gave absolute purities in the 70–90% range. In addition, if peptides containing only single deletions or insertions of glutamine or proline are included, purities rise to the following levels: HTT\(^{NTQ23P10K2}\), 97%; HTT exon1-Q23, 94%; HTT\(^{NTQ42P10K2}\), 86%; HTT exon1-Q23, 93%.

Peptides were disaggregated in a mixture of trifluoroacetic acid and hexafluoroisopropanol immediately prior to initiating aggregation kinetics analyses, as described previously. Determination of the equilibrium concentration of monomer in amyloid formation reactions \((C_m)\) was accomplished by using the HPLC sedimentation assay to determine monomer concentrations at each time. Reverse reactions were initiated by diluting part of a nearly complete aggregation reaction in PBS and incubating, as described previously. FTIR spectroscopy was conducted on isolated aggregates, on an ABB Bomem FTIR spectrometer, as described. Dynamic light scattering measurements were conducted on aggregation reaction time points on a Wyatt DynaPro, as described. Negative stain electron microscopy was conducted on aggregation reaction time points using a Tecnai T12 microscope, as described. Analysis of dimensions of selected particles in the EM was done with the help of NIH ImageJ software (http://image.nih.gov/ij/).

## RESULTS

### Aggregation Kinetics

To examine the roles of the polyQ flanking sequences in HTT exon1-like peptides, we studied the peptide sequences shown in Figure 1, where the polyQ repeat length is either 23 or 42. The spontaneous aggregation kinetics curves of PBS solutions of rigorously disaggregated samples of these peptides are shown in Figure 3. As described previously, K\(_2\)Q\(_42\)K\(_2\) (green star) aggregates imperceptibly at 40 \(\mu\)M even after 1000 h (Figure 3A). In contrast, aggregation of HTT\(^{NTQ23K2}\) (green square) at a 10-fold lower concentration of 4 \(\mu\)M is essentially complete after 75 h. As previously described, the addition of a P\(_{10}\) sequence to generate HTT\(^{NTQ23P10K2}\) slows aggregation in comparison to HTT\(^{NTQ23K2}\), so that it takes about 1000 h for a 13 \(\mu\)M solution of this peptide to aggregate to completion (red filled circle). Importantly, the aggregation kinetics time course of an 11 \(\mu\)M solution of full-length HTT exon1-Q23 (black circle) is very similar to that of HTT\(^{NTQ23P10K2}\). The slight difference in shape, which appears to be within the error bars of the kinetics measurements (Figure 3A), may be due to a combination of differences in nucleation efficiency, elongation rate constants, and critical concentration (see below).

Similar results were obtained with the Q42 versions of these peptides. The simple polyQ peptide K\(_2\)Q\(_{42}\)K\(_2\) aggregates relatively slowly at 4.8 \(\mu\)M (Figure 3B, green star; note different time scale compared with panel A). Although we did not examine a HTT\(^{NTQ42P10K2}\) molecule, we previously reported that the peptide HTT\(^{NTQ42K2}\) aggregates to completion at 5 \(\mu\)M within about 3 h. In analogy to the Q42 series, addition of a P\(_{10}\) sequence to generate HTT\(^{NTQ42P10K2}\) extends the aggregation time course so that a 6.2 \(\mu\)M reaction nears completion at about 70 h (Figure 3B, red filled circle). The aggregation of a 5.5 \(\mu\)M solution of full-length HTT exon1-Q42 (black circle) is somewhat slower than that of HTT\(^{NTQ42P10K2}\), but their relationship is qualitatively similar to the Q23 series, in that both Q42 versions of HTT exon1 aggregate faster than K\(_2\)Q\(_{42}\)K\(_2\) and more slowly than HTT\(^{NTQ23P10K2}\). Interestingly, there are
some quantitative differences between the Q23 and Q42 series. First, the overall impact of the N- and C-terminal flanking sequences on aggregation appears to be greater for Q23 than for Q42. This seems to be less due to the impact of the HTTNT sequence, which is strong for both polyQ repeat lengths, and more due to the ability of PRD sequences to work against the rate enhancement by HTTNT. Second, the presence of the full PRD sequence, in comparison to only a P10 sequence, has a measurable impact on Q42 kinetics but not Q23 kinetics. This effect might be due to reductions in non-β oligomer (Figure 2b,c,d) stability, in the efficiency of nucleation within this oligomer (Figure 2e), or in nucleus or fibril elongation (Figure 2f). As suggested by Figure 2, the preferred aggregation mechanism of HTT exon1-like peptides involves the initial formation of oligomeric intermediates, which have been demonstrated in vitro from incubation of both recombinant full-length HTT exon141 and chemically synthesized versions lacking full-length PRDs.10

Electron microscopy examination of early assembly intermediates of chemically synthesized, full-length HTT exon1-Q23 shows the existence of spherical oligomers (Figure 4D) and protofibril or fibril-like structures (Figure 4E) very similar in size and appearance to those described previously for the initial stages of the aggregation of an HTTNTQ23P10K2 peptide.10 Nonamyloid (ThT-negative) oligomer formation has also been observed previously in the initial stages of HTTNTQ20P10K2 aggregation.10 For a more direct comparison, we show here that spherical oligomers are also formed in the early hours of incubation of HTTNTQ23P10K2 (Figure 4A). Nonamyloid oligomers are also present in early incubation time points of HTT exon1-Q23 incubation (Figure 4G). Interestingly, the HTT exon1-Q42 oligomers exhibit qualitatively different staining properties in the EM. Despite this staining difference, we believe the oligomers formed early by both polyQ repeat length versions of full-length HTT exon1 are structurally similar, because both aggregation reactions exhibit similar sensitivities to inhibition by HTTNT peptides (Figure 2; see below).

Dynamic light scattering (DLS) analysis of early time points of these incubation reactions are shown in Figure 5. The DLS curves of aliquots of the reaction mixtures at t = 0 exhibit amplitudes in the range of 1.05 or less and are indistinguishable from PBS buffer (not shown). The DLS time courses give particle size progressions as expected from the EM analysis, with rough agreement in absolute values at different time points. Thus, for HTT exon1-Q23 at 5 h, EM shows a range of particle sizes centered at 14 and 24 nm in diameter (Figure 4 legend), while at the same time point DLS gives a homogenous suspension of particles with ~5 nm diameter.
(2.6 nm hydrodynamic radius). At 30 h, EM shows a uniform population of protofibrils of ∼17 nm in diameter and 50–130 nm in length, while DLS gives a hydrodynamic radius of 80 nm for a filament length in the 100–200 nm range, in good agreement with the EM analysis. HTT exon1-Q42 incubated for only 30 min yields scatter from a diverse population of particles that cannot be easily fit by the software (Figure 5B). This is typical of amyloid assembly reactions, in which time points associated with amyloid fibrils give strong scattering in DLS that fails to give an interpretable correlation.

Recently, we described the ability of isolated HTTNT peptides to act as inhibitors of the nucleation phase of the amyloid formation of HTT exon1-like peptides.48 This ability appears to be due to the coassembly of the HTTNT peptides with HTT exon1 during oligomer formation, by virtue of mixed helical bundle formation as shown in Figure 2i. This dilutes the average local concentration of polyQ chains within the mixed oligomers, reducing nucleation efficiency. Once nucleation occurs, however, amyloid formation proceeds unabated.48 Since we found that HTTNT in trans has no effect on simple polyQ aggregation,48 we can use the sensitivity of an aggregation reaction to HTTNT inhibition as a test of mechanism.

In fact, we found that, to different extents, both of the chemically synthesized, full-length HTT exon1 peptides used in this study are susceptible to HTTNT inhibition (Figure 3). HTT exon1-Q42 is especially susceptible, showing only about 10% aggregation in the presence of a 1.5-fold molar excess of HTTNT (□) at 1000 h, a time when the peptide without inhibitor has essentially aggregated to completion (Figure 3A, ■). Mirroring their overlapping aggregation curves, HTTNTQ23P10K2 exhibits inhibition kinetics (red open circle) very similar to the inhibition of HTT exon1-Q31 (Figure 3A, □). HTTNT inhibition of Q42 HTT exon1-like peptides is somewhat less effective than for the Q31 peptides. At a time when the uninhibited reaction of HTTNTQ23P10K2 (red filled circle) is 50% aggregated (about 11 h), it is about 30% aggregated in the presence of a 1.5-fold molar excess of HTTNT (Figure 3B, red open circle). Full-length HTT exon1-Q42 is more susceptible to inhibition, so in the presence of a 1.5-fold HTTNT it has aggregated only negligibly at 25 h (Figure 3B, □), when the uninhibited reaction is 50% aggregated (Figure 3B, ■). Consistent with the proposed mechanism of HTTNT inhibition involving coassembly, the t = 0 time point of a mixture of HTTNT plus HTT exon1-Q42 exhibits more scattering intensity than the t = 0 time point of either component alone (Figure 5B). At the same time, oligomer size grows modestly within the lag phase of the inhibited reaction, giving an interpretable DLS curve after 30 min incubation (Figure 5B), consistent with a buildup of larger, HTTNT-mediated oligomers. As noted above, the same amount of HTT exon1-Q42 incubated without HTTNT for the same time gives more scattering and larger scattering particles (Figure 5B). As noted above, the sensitivity of HTT exon1-Q42 aggregation to inhibition by HTTNT peptides suggests that, as shown for other HTT exon1 analogues,48 the amyloid nucleation pathway must involve a required intermediate featuring HTTNT-based, nonamyloid assembly. That is, this pathological repeat length occurs, however, amyloid formation proceeds unabated.48 Since we found that HTTNT in trans has no effect on simple polyQ aggregation,48 we can use the sensitivity of an aggregation reaction to HTTNT inhibition as a test of mechanism.

Aggregate Structure. We assessed the impact of the PRD on polyQ amyloid structure by electron microscopy, dynamic light scattering, FTIR, fibril stability measurements, and cross-seeding experiments. By EM, the mature fibrils produced by both the Q31 (Figure 4F) and Q42 (Figure 4I) versions of chemically synthesized full-length HTT exon1 exhibit similar morphologies, both to each other and to fibrils from both chemically synthesized model peptides of the HTTNTQ23P10K2 design (Figure 4B) and from recombinantly produced, nearly exact full-length HTT exon1 peptides.49 One unexpected result is that at an intermediate time point in the aggregation of the chemically synthesized HTT exon1 peptides, the Q31 version, but not the Q42 version, exhibits a remarkably uniform population of very short fibrils with lengths in the 50–130 nm range (Figure 4E). Interestingly, these particles have somewhat larger diameters than mature fibrils in the EM, on the order of 17.2 nm compared with the fibril diameter of 12.5 nm (Figure 4). It is possible that these structures are related to the protofibrils often observed at intermediate times in the assembly of other amyloids. As in the case of these other amyloids, the role of the structures shown in Figure 4E in the assembly of mature amyloid fibrils is not clear. As observed previously,48 the mature fibrils of polyQ peptides containing HTT flanking sequences are substantially different in EM morphology from the amyloid produced by simple polyQ peptides (Figure 4C).

The size distribution of HTT exon1 aggregation products is of some interest, because of ongoing efforts to identify the toxic molecular species that is presumably populated in a time- and repeat length-dependent manner in the pathogenic mechanism of HD.9 In particular, it has become important to gauge the size range of aggregates in the context of the limited ability of fluorescence microscopy to detect aggregates smaller than inclusions (see Discussion). We therefore calculated the number of HTT exon1 monomers that could be packed into some of the smaller aggregates we observed from incubation of HTT exon1-Q23, using an estimate for the density of a folded protein of 1.37 g/cm3.57 For the spherical oligomers observed at 5 h by EM (Figure 4D), we calculated that the oligomers contain from 123 to 623 molecules, while the short fibrils or protofibrils observed at 30 h (Figure 3E) contain from 530 to 2740 molecules.

Despite the EM differences between HTT exon1 fibrils and simple polyQ fibrils, there is substantial evidence that the polyQ amyloid cores of these various aggregates are largely identical. For example, the second derivative FTIR spectra of the truncated and full-length HTT exon1 aggregates are essentially identical to the spectrum of a simple polyQ peptide amyloid (Figure 6). In particular, the three bands characteristic of simple polyQ amyloid [1605 cm−1 (glutamine side chain N–H bending), 1625 cm−1 (β-sheet), and 1659 cm−1 (glutamine side chain C=O stretch)] dominate all of the amyloid fibrils derived from HTT exon1-related peptides. In addition, both for Q31 and for Q42 versions, the FTIR spectra of the full HTT exon1 and HTTNTQ23P10K2 aggregates are superimposable (Figure 6). The small differences that exist between spectra, located in the broad region from 1610 to 1640 cm−1 that is normally associated with β-structure,58 recur in other FTIR spectra of polyQ aggregates without correlating with any obvious primary sequence feature, be it polyQ repeat length or the presence or absence of flanking sequences. We think it is likely that the fine structure sometimes observed in this region, as seen in the spectra for the Q31 peptide aggregates in Figure 6, is related to better signal in these spectra due to the availability of larger amounts of sample.

A powerful measure of the relatedness of two amyloid systems is the ability of the fibrils of one to “cross-seed” the
elargon of monomers of the other.\textsuperscript{59} In general, the efficiency of seeding elongation is thought to depend on the structural compatibility between the monomer and the fibril seed. We found excellent “cross-talk” between fibrils of HTT\textsuperscript{NTQ23P10K2} and HTT\textsuperscript{exon1-Q23}. Thus, in an experiment with equal concentrations of freshly disaggregated HTT\textsuperscript{exon1-Q23} monomers, the addition of 20\% by weight of amyloid fibrils gives essentially the same enhanced aggregation kinetics for seeds of either HTT\textsuperscript{exon1-Q23} (red filled circle) and HTT\textsuperscript{NTQ23P10K2} (blue triangle) amyloid, when compared with the spontaneous, unseeded aggregation reaction (■) (Figure 7A). Similarly, with freshly disaggregated HTT\textsuperscript{NTQ23P10K2} monomers, 20\% by weight of HTT\textsuperscript{NTQ23P10K2} amyloid (red filled circle) and HTT\textsuperscript{exon1-Q23} amyloid (blue triangle) gives identical aggregation stimulation compared with monomer alone (Figure 7B, ■). We take this to indicate a strong similarity in the polyQ amyloid cores of these two peptide fibrils. Another method for characterizing amyloid structures is by the stabilities of the fibrils against dissociation.\textsuperscript{60} Unless the monomers within a fibril are chemically cross-linked, fibrils should generally be capable of dissociating to monomers in native buffer until an equilibrium position is reached\textsuperscript{60} and this position is a measure of fibril stability with values that can be quite robust.\textsuperscript{51} Fibril stabilities as estimated by $C_r$ determinations were previously successfully used to conduct several comparative studies of amyloid stability, including a set of polymorphic amyloid fibrils derived from the same $A\beta_40$ sequence,\textsuperscript{62} fibrils generated from simple polyQ of different repeat lengths,\textsuperscript{51} and fibrils generated from polyQ molecules with or without $\beta$-hairpin encouraging mutations.\textsuperscript{53} Here we determined the equilibrium positions, expressed as the concentration of monomer at equilibrium (the $C_r$ value), for various Q\textsubscript{23} versions of HTT\textsuperscript{exon1} and compared the results with the previously published\textsuperscript{56} value for K\textsubscript{2Q23K2}. When HTT\textsuperscript{exon1-Q23} is incubated at $\sim$9 $\mu$M, it aggregates slowly over a period of weeks, reaching an equilibrium concentration of monomer after about one month (Figure 8A, ■). When this final aggregation reaction is diluted to reduce the concentration of monomer and the diluted reaction is further incubated, the monomer concentration increases as fibrils dissociate (red square), until equilibrium is reestablished, also after a period of about one month (Figure 8A, inset). The mean of the equilibrium positions measured for the association and dissociation directions for this peptide is 0.44 ± 0.13 $\mu$M. This $C_r$ value, the concentration of monomer below which aggregation is thermodynamically disallowed, is a convenient measure of fibril stability; the lower $C_r$, the more stable the fibril. A similar analysis of HTT\textsuperscript{NTQ23P10K2} (Figure 8B) gives a $C_r$ value of 0.28 ± 0.11. Thus, fibrils of full-length HTT\textsuperscript{exon1} and of the C-terminally truncated version are of essentially equal stability. Unfortunately, we could not obtain the corresponding values for the Q\textsubscript{42} peptides, since these appear to be below our level of detection.

We obtained two additional $C_r$ values. Previously we reported a value of 3.0 $\mu$M for the $C_r$ of K\textsubscript{2Q23K2} amyloid.\textsuperscript{56} In the present study, we determined a $C_r$ of $\leq$0.1 $\mu$M for HTT\textsuperscript{NTQ23K2} (the actual value may be lower, but the limited sensitivity of our methods preclude measuring concentrations below 0.1 $\mu$M for these peptides). These $C_r$ values allow us to calculate a number of important $\Delta\Delta G$ values or limits. By comparing the values for K\textsubscript{2Q23K2} and HTT\textsuperscript{NTQ23K2}, we obtain a positive contribution of the HTT\textsuperscript{NT} sequence to amyloid stability of $\geq$2.2 kcal/mol. This is consistent with previous data.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure6.png}
\caption{FTIR spectra of isolated final aggregates. The second derivative spectra of aggregates collected, when the aggregation reaction was judged complete, at the following times: K\textsubscript{2Q23K2} (300 h); HTT\textsuperscript{NTQ23P10K2} (1000 h); HTT\textsuperscript{exon1-Q23} (1000 h); HTT\textsuperscript{NTQ23P10K2} (70 h); and HTT\textsuperscript{exon1-Q42} (70 h).}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure7.png}
\caption{Seeded aggregation kinetics. Monomer alone (■), self-seeding (red filled circle), and cross-seeding (blue triangle): (A) 11 $\mu$M HTT\textsuperscript{exon1-Q23} monomer alone (■) or seeded with 20% by weight of amyloid fibrils of either HTT\textsuperscript{exon1-Q23} (red filled circle) or HTT\textsuperscript{NTQ23P10K2} (blue triangle); (B) 13 $\mu$M HTT\textsuperscript{NTQ23P10K2} monomer alone (■) or seeded with 20% by weight of amyloid fibrils of either HTT\textsuperscript{NTQ23P10K2} (red filled circle) or HTT\textsuperscript{exon1-Q23} (blue triangle). All data points have error bars based on two measurements.}
\end{figure}
suggested that HTTNT remains α-helical in assembled fibrils and appears to self-associate in that context. 


discussn, our data are consistent with the hypothesis that the PRD appears to make a negligible difference in this stability effect, with the P10 sequence being sufficient to represent essentially the full fibril destabilizing action of the PRD.

One interesting aspect of the C1 measurement that generated the above stability values is the dissociation kinetics of HTT exon1 amyloid. Previously we reported very slow dissociation kinetics for simple polyQ amyloid fibrils, with Kd/Q1-Q2 fibrils requiring about 2 months to dissociate to equilibrium. We show here that HTT exon1 analogues feature a similarly slow dissociation rate, despite the presence of flanking sequences that modulate the thermodynamic stability of the fibrils. Thus, while Aβ40 amyloid dissociates to equilibrium in PBS at 37°C over a period of only 24 h,60 it takes about 4 weeks for HTT exon1-Q23 (or HTT NTQ23P10K2) fibrils to dissociate to equilibrium (Figure 8). These new data suggest that the slow kinetics are relatively unaffected by flanking sequences and therefore likely be tied to the structure of the polyQ amyloid core. Interestingly, aggregates of an HTT exon1-Q94 protein expressed in the brains of a mouse model required 3 weeks to dissociate after HTT exon1 expression was turned off.66 The authors reported data supporting a role of the ubiquitin proteasome system in the clearance of these aggregates (as well as monomers).64 Together with those findings, our data are consistent with the hypothesis that the dissociation rate of aggregates in the mouse brain is driven primarily by the biophysical properties of the HTT exon1 amyloid and not by cellular factors. The role of the proteasome may be simply to help remove monomeric HTT exon1 from the thermodynamic system, which would tend to drive dissociation by virtue of Le Chatelier’s principle.

One issue that continues to plague the HD field is our ignorance of the potential size range of aggregated proteins in the cell. Although evidence has been presented that the large inclusions observed in HD cells and cell models are more likely to be protective than toxic,53 these data have no bearing on a possible role for any smaller aggregated species that might exist in the cell.55 In fact, the inclusions easily observed in HD brain tissue and cell models tend to be about 5 μm in diameter,
corresponding roughly to an HTT exon1 content of $10^9$ molecules, and this allows for a huge range of smaller aggregate sizes that will presumably be more difficult to detect in cells or cell extracts. In fact, super-resolution fluorescence microscopy was recently used to detect individual amyloid fibrils in a cell model of HD.  

We calculate these fibrils to have monomer contents in the range of $10^5$ monomers per aggregate.  

In the present paper, we show that well-characterized full-length HTT exon1 peptides aggregated \textit{in vitro} are capable of populating even smaller aggregates during amyloid assembly: spherical oligomers with monomer contents in the range of 100–600 and short, linear, protofibril-like structures with monomer contents in the range of 500–2600. These are in addition to much longer amyloid fibrils that develop later in aggregation reactions. While our data do not prove that such small aggregates exist or accumulate in the cell, it is clear that these aggregates have sufficient kinetics of formation and stabilities \textit{in vitro} that their formation in cells is energetically feasible.

It is interesting to put these findings in the context of results of other studies. EM-characterized aggregates of similar sizes, much smaller than inclusions but much larger than monomer, have been generated either \textit{in vitro} incubation of recombinantly produced HTT exon1 \cite{1} or by isolation from brain homogenates of a HD mouse model. A variety of aggregate sizes has been revealed by analysis of lysates of cell or tissue from HD models or HD patients, which, in order of increasing estimated number of peptide monomers per aggregate, are 4-mers (B. Sahoo et al., msc submitted), 5-mers to 15-mers, \cite{2} 10-mers (assuming a composition of HTT exon1-sized proteolytic fragments), \cite{3} 200-mers, \cite{4} and 50–10000-mers (B. Sahoo et al., msc submitted).

Recombinant DNA methods of synthesizing useful amounts of peptides and proteins \cite{5} have revolutionized protein biochemistry and biophysics research, oftentimes making studies possible that otherwise simply could not be undertaken. The HTT exon1 peptide lies at an intermediate size range, which, ironically, provides different but equally significant challenges for both recombinant and chemical synthesis, perhaps in particular in conducting aggregation studies. While this peptide is somewhat longer than the typical upper limit to solid phase synthesis, it is too small and aggregation prone for successful cell production unless it is fused with a partner that limits its aggregation. This fusion partner must then be removed chemically or proteolytically \textit{in vitro}.

For peptides with lengths within the synthesis limits, chemical synthesis can deliver precisely designed molecules. However, in addition to the size limit, chemical synthesis can generate difficult-to-remove peptide side products and other impurities, including \textalpha{}-amino acids from epimerization during synthesis that can be very difficult to detect but that, even in small amounts, can compromise amyloid formation. Recombinant synthesis, in contrast, can generate peptides and proteins of much greater length with relative ease and initially high sequence fidelity. However, it appears that the properties of the same sequence produced in different cell types can sometimes differ considerably, presumably due to variations in post-translational modifications. Furthermore, exigencies of cloning or proteolytic release of desired peptides from fusion partners can sometimes lead to small but potentially significant modifications in the recombinant product, such as in extensions \cite{6,7} or truncations \cite{8} of the HTT \cite{9} sequence. If the modifications occur in a portion of the molecule that is critical to its solution properties, the seemingly small compromises made to produce the recombinant peptide could have unintended consequences. For example, the use of trypsin as one option for cleaving HTT exon1 from an N-terminal fusion partner \cite{10} almost certainly leads to cleavage at one or more of the lysine residues within HTT \cite{11}. This appears to generate an amyloid fibril morphology \cite{12} that more resembles the ribbon-like aggregates produced from simple polyQ peptides \cite{13} (Figure 4C) than the isolated filaments produced from HTT exon1-like peptides with intact N-termini (Figure 4F). In situations such as that presented by HTT exon1, it will continue to be important to investigate the possible impact of the effects of these compromises on experimental results.

Chemically accessible, C-terminally truncated versions of HTT exon1 and related molecules have been utilized to deduce the mechanism of the aggregation enhancing effect of the HTT \cite{14} N-terminal segment of HTT, \cite{15,16,17,18} assess the impact of serine phosphorylation within HTT \cite{19} on that mechanism, \cite{20,21} characterize the magnitude and nature of the polyproline effect on aggregation, \cite{22,23} determine aspects of HTT exon1 amyloid structure, \cite{24} characterize aggregation inhibition by HTT \cite{25} peptides, \cite{26} and monitor HTT exon1–membrane interactions. The recent availability of chemically synthesized, full-length HTT exon1 peptides \cite{27,28} has offered the possibility that, given sufficient yields of pure material, \textit{systematic \textit{in vitro} studies} can now be conducted on molecules more accurately resembling the HTT exon1 fragments generated in the cell. In this paper, we describe in detail characterization of the aggregation properties of such chemically synthesized full-length HTT exon1 containing both benign and pathological polyQ repeat lengths. Our data show that the PRD of HTT exon1 beyond the first 10 prolines contributes only modest quantitative differences in some biophysical properties, while other properties seem to be essentially unchanged. The data suggest that these truncated molecules can be valuable substitutes for full-length HTT exon1 when chemically precise versions of the latter are not available.

### References

1. Bates, G., Harper, P. S., and Jones, L., Eds. (2002) \textit{Huntington’s Disease}, 3rd ed., Oxford University Press, Oxford, U.K.
2. Bates, G., Tabrizi, S. J., and Jones, L., Eds. (2014) \textit{Huntington’s Disease}, 4th ed., Oxford University Press, Oxford, U.K.
Intriguing tales of a monotonous sequence. Huntington pp 429
Masliah, E., Thompson, L. M., Muchowski, P. J., Weisgraber, K. H.,
Gray, M., Thulasiramin, V., Saudou, F., Segal, M., Yang, X. W.,
Hatters, D., Curtis, J., Cheung, K., Krishnan, P., Hatters, D., Brooks, E., Widjaja, K.,
Taran, T., Mitra, S., Arrasate, M., Mosquera, L. A., Taylor, D., Weisgraber, K. H.,
and Finkbeiner, S. (2012) Disease-associated polyglutamine stretches in monomeric huntingtin adopt a compact structure. J. Mol. Biol. 421, 887–860.

Bates, G. P., Miller, J., Rutenber, E., Newhouse, Y.,
Krishnan, P., Cheung, K., Hatters, D., Brooks, E., Widjaja, K., Tran, T., Mitra, S., Arrasate, M., Mosquera, L. A., Taylor, D., Weisgraber, K. H.,
and Finkbeiner, S. (2012) Disease-associated polyglutamine stretches in monomeric huntingtin adopt a compact structure. J. Mol. Biol. 421, 887–860.

Bates, G. P., Miller, J., Rutenber, E., Newhouse, Y.,
Krishnan, P., Cheung, K., Hatters, D., Brooks, E., Widjaja, K., Tran, T., Mitra, S., Arrasate, M., Mosquera, L. A., Taylor, D., Weisgraber, K. H.,
and Finkbeiner, S. (2012) Disease-associated polyglutamine stretches in monomeric huntingtin adopt a compact structure. J. Mol. Biol. 421, 887–860.

Bates, G. P., Miller, J., Rutenber, E., Newhouse, Y.,
Krishnan, P., Cheung, K., Hatters, D., Brooks, E., Widjaja, K., Tran, T., Mitra, S., Arrasate, M., Mosquera, L. A., Taylor, D., Weisgraber, K. H.,
and Finkbeiner, S. (2012) Disease-associated polyglutamine stretches in monomeric huntingtin adopt a compact structure. J. Mol. Biol. 421, 887–860.
formation by the N-terminus of huntingtin. *J. Mol. Biol.* 396, 1295–1309.
(53) Singer, D., Zauner, T., Genz, M., Hoffmann, R., and Zucher, T. (2010) Synthesis of pathological and nonpathological human exon 1 huntingtin. *J. Pept. Sci.* 16, 358–363.
(54) O’Nouallain, B., Thakur, A. K., Williams, A. D., Bhattacharrya, A. M., Chen, S., Thiagarajan, G., and Wetzel, R. (2006) Kinetics and thermodynamics of amyloid assembly using a high-performance liquid chromatography-based sedimentation assay. *Methods Enzymol.* 413, 34–74.
(55) Chemur, S., Kodali, R., and Wetzel, R. (2014) Improved chemical synthesis of hydrophobic Abeta peptides using addition of C-terminal lysines later removed by carbapenemase B. *Biopolymers* 102, 206–221.
(56) Kar, K., Jayaraman, M., Sahoo, B., Kodali, R., and Wetzel, R. (2011) Critical nucleus size for disease-related polyglutamine aggregation is repeat-length dependent. *Nat. Struct. Mol. Biol.* 18, 328–336.
(57) Quillin, M. L., and Matthews, B. W. (2000) Accurate calculation of the density of proteins. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 56, 791–794.
(58) Haris, P. I., and Chapman, D. (1995) The conformational analysis of peptides using Fourier transform IR spectroscopy. *Biopolymers* 37, 251–263.
(59) O’Nouallain, B., Williams, A. D., Westermark, P., and Wetzel, R. (2004) Seeding specificity in amyloid growth induced by heterologous fibrils. *J. Biol. Chem.* 279, 17490–17499.
(60) O’Nouallain, B., Shivaprasad, S., Kheterpal, I., and Wetzel, R. (2005) Thermodynamics of Abeta(1–40) amyloid fibril elongation. *Biochemistry* 44, 12709–12718.
(61) Williams, A. D., Shivaprasad, S., and Wetzel, R. (2006) Alamine scanning mutagenesis of Abeta(1–40) amyloid fibril stability. *J. Mol. Biol.* 357, 1283–1294.
(62) Kodali, R., Williams, A. D., Chemur, S., and Wetzel, R. (2010) Abeta(1–40) forms five distinct amyloid structures whose beta-sheet contents and fibril stabilities are correlated. *J. Mol. Biol.* 401, 503–517.
(63) Kar, K., Hoop, C. L., Drombosky, K. W., Baker, M. A., Kodali, R., Arduini, I., van der Wel, P. C., Horne, W. S., and Wetzel, R. (2013) beta-hairpin-mediated nucleation of polyglutamine amyloid formation. *J. Mol. Biol.* 425, 1183–1197.
(64) Bugg, C. W., Isas, J. M., Fischer, T., Patterson, P. H., and Langen, R. (2012) Structural features and domain organization of huntingtin fibrils. *J. Biol. Chem.* 287, 31739–31746.
(65) Bhattacharyya, A., Thakur, A. K., Chellgren, V. M., Thiagarajan, G., Williams, A. D., Chellgren, B. W., Creamer, T. P., and Wetzel, R. (2006) Oligopropene effects on polyglutamine conformation and aggregation. *J. Mol. Biol.* 355, 524–535.
(66) Martin-Aparicio, E., Yamamoto, A., Hernandez, F., Hen, R., Avila, J., and Lucas, J. J. (2001) Proteasomal-dependent aggregate reversal and absence of cell death in a conditional mouse model of Huntington’s disease. *J. Neurosci.* 21, 8772–8781.
(67) Orr, H. T. (2004) Neurodegenerative disease: Neuron protection agency. *Nature* 431, 747–748.
(68) Sathasivam, K., Lane, A., Legleiter, J., Warley, A., Woodman, B., Finkbeiner, S., Paganetti, P., Muchowski, P. J., Wilson, S., and Bates, G. P. (2010) Identical oligomeric and fibrillar structures captured from the brains of R6/2 and knock-in models of Huntington’s disease. *Hum. Mol. Genet.* 19, 65–78.
(69) Wetzel, R., and Goeddel, D. V. (1983) Synthesis of polypeptides by recombinant DNA methods, in *The Peptides: Analysis, Synthesis, Biology* (Meienhofer, J., and Gross, E., Eds.), pp 1–64, Academic Press, New York.
(70) Finder, V. H., Vodopivec, I., Nitsch, R. M., and Glickshuber, R. (2010) The recombinant amyloid-beta peptide Abeta1–42 aggregates faster and is more neurotoxic than synthetic Abeta1–42. *J. Mol. Biol.* 396, 9–18.
(71) Scherzer, E., Lurz, R., Turkmen, M., Mangiarini, L., Hollenbach, B., Hasenbank, R., Bates, G. P., Davies, S. W., Lehrach, H., and Wanker, E. E. (1997) Huntington-encoded polyglutamine
expansions form amyloid-like protein aggregates in vitro and in vivo. 

(72) Burke, K. A., Hensal, K. M., Umbaugh, C. S., Chaibva, M., and Legleiter, J. (2013) Huntingtin disrupts lipid bilayers in a polyQ-length dependent manner. *Biochim. Biophys. Acta: Biomembr.* 1828, 1953–1961.

(73) Ansaloni, A., Wang, Z. M., Jeong, J. S., Ruggeri, F. S., Dietler, G., and Lashuel, H. A. (2014) One-Pot Semisynthesis of Exon 1 of the Huntingtin Protein: New Tools for Elucidating the Role of Posttranslational Modifications in the Pathogenesis of Huntington’s Disease. *Angew. Chem., Int. Ed.* 53, 1928–1933.

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