Published in final edited form as:
J Mol Biol. 2012 August 24; 421(4-5): 466–490. doi:10.1016/j.jmb.2012.01.030.

Physical chemistry of polyglutamine: Intriguing tales of a monotonous sequence

Ronald Wetzel
Department of Structural Biology and Pittsburgh Institute for Neurodegenerative Disease, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA

Abstract

Polyglutamine sequences of unknown normal function are present in a significant number of proteins, and their repeat expansion is associated with a number of genetic neurodegenerative diseases. Polyglutamine solution structure and properties are not only important because of the normal and abnormal biology associated with these sequences, but also because they represent an interesting case of a biologically relevant homopolymer. As the common thread in the expanded polyglutamine repeat diseases, it is important to understand the structure and properties of simple polyglutamine sequences. At the same time, experience has shown that sequences attached to polyglutamine, whether in artificial constructs or in disease proteins, can influence structure and properties. The two major contenders for the molecular source of the neurotoxicity implicit in polyglutamine expansion within disease proteins are a populated toxic conformation in the monomer ensemble and a toxic aggregated species. This review summarizes experimental and computational studies on the solution structure and aggregation properties of both simple and complex polyglutamine sequences, and their repeat-length dependence. As a representative of complex polyglutamine proteins, the behavior of huntingtin N-terminal fragments, such as exon-1, receives special attention.

Although functionally obscure polyglutamine (polyQ) sequences are found in a variety of proteins [1–3], many of which are transcription factors [1,4], and although aggregation-prone plant storage proteins rich in Gln are also well-known [5, 6], it is unlikely that polyQ sequences would ever have become a major focus of research were it not for their pathogenic role in proteins associated with a series of inherited, devastating neurodegenerative diseases [7, 8]. At least 10 [8, 9] of these expanded CAG repeat diseases have been described, including the most prevalent and most well-known condition, Huntington's disease (HD) [10]. All of them are triggered by inheritance of a polyQ sequence with a repeat length longer than normal, where normal is in the 20–35 range for all but one of these diseases [8]. All of them feature the accumulation in particular brain regions of polyQ-containing neuronal aggregates, and all of them appear to be primarily gain-of-function diseases [8]. These and other data have suggested that expanded polyQ repeat proteins are defective in some aspect of protein folding, such as a lowered barrier to misfolding within the monomer and/or to aggregate formation.

Corresponding author: Ronald Wetzel, Dept. Structural Biology, University of Pittsburgh School of Medicine, Rm. 2046 Biomedical Sciences Tower 3, 3501 Fifth Avenue, Pittsburgh, PA 15260, USA. Phone: 412-383-5271; Fax: 412-648-9008; rwetzel@pitt.edu.

Because of the great interest in understanding the molecular mechanisms of these diseases and in developing therapies, the biophysical properties of polyQ sequences have received considerable scrutiny in the nearly 20 years since the pathogenic role of polyQ expansion was first realized [11]. While the exact folded and aggregated states of the toxic species in these diseases are yet to be worked out, a wealth of important biophysical data has been accumulated. In addition to any implications for mechanisms and therapeutic strategies of neurodegenerative diseases, some of these data have far-reaching implications for general protein solution behavior and aggregation. Experimental biophysical studies on polyQ solution structure and aggregation can also provide data to support computational analyses that have potential relevance to polyQ pathology as well as to fundamental polymer physics. PolyQ studies have provided important information on the solution structure of intrinsically disordered proteins, and on the all-important mechanisms by which amyloid growth is initiated.

This review focuses on *in vitro* experimental studies of polyQ solution and aggregate structure and on the mechanisms of spontaneous aggregation and its inhibition. Selected results from the wealth of *in silico* and *in vivo* studies are also discussed. The reader is referred to a recent review for more thorough coverage of biological aspects [10].

Polyglutamine Solution Structure

The polyQ monomer in water

The observation of ubiquitinylated, polyQ containing aggregates in neuronal nuclei of HD patients [12] and animal models [13], in addition to other factors, led directly to the hypothesis that expanded CAG repeat diseases should be included in the growing family of human pathologies involving protein misfolding and aggregation [14]. Since globular protein aggregation [15], including amyloid formation [16–18], is understood to be associated with breakdowns in protein conformational integrity, speculation on the nature and source of expanded polyQ gain-of-function has included not only a proposed leading role for protein aggregation but also the related, but distinct, hypothesis of an important, toxic misfolded state within the monomer ensemble [7] that might also, coincidentally, mediate aggregation. But what is the evidence for the existence of a distinct, highly populated folded state in either benign or pathogenic polyQ sequences? In this section I review what is known about populated folding states of polyQ monomers.

Initial evidence for a predominantly disordered structure for polyQ in solution came from CD analysis of solutions of short polyQ monomers [19]. Subsequent CD studies on pathological length polyQ sequences, either chemically synthesized [20] or biosynthetic [21], showed that even long polyQ sequences are largely disordered, and that there is no obvious difference in secondary structure between long and short polyQ. NMR studies [22] subsequently confirmed the disordered nature of polyQ monomers of all repeat lengths, and supported earlier CD studies [23] that showed the absence of detectible, conformationally distinct intermediates in the coil to sheet transition that occurs when polyQs grow into amyloid. X-ray diffraction structures of crystals of a sequence-modified huntingtin N-terminal fragment containing a short polyQ insert showed the intervening polyQ sequence to be in a variety of conformations, including α-helix [24]. CD analysis shows that a simple $K_2Q_{40}K_2$sequence has about 10% α-helix configurations at 35 °C and takes on an additional 10% on cooling to 5 °C [25]. The overall picture from solution phase experiments is that polyQ sequences, regardless of repeat length, are largely disordered chain with transient elements of regular secondary structure. This impression is supported by a number of computational analyses starting from unstructured chains that, while not agreeing completely with each other, suggest that both short or long polyQ sequences generally

exhibit mixtures in various proportions of disordered structure, α-helix, β-sheet, PP_{II}helix, and β-turns [26–32].

There is an additional aspect to the solution structure of the polyQ sequence that likely plays an important role in its behavior – the tendency of the sequence to collapse, in spite of its apparent hydrophilic nature, in aqueous solution. This behavior has been demonstrated in several biophysical studies. A fluorescence correlation spectroscopy (FCS) analysis of a collection of labeled polyQ molecules of differing repeat lengths [33] demonstrated two important points. First, that polyQ peptides behave in a simple aqueous buffer as a polymer in a poor solvent. That is, the polyQ chain in water is neither fully extended in a statistical coil conformation, as it would be in a good solvent, nor is it somewhat less extended, as occurs under "theta" conditions just short of the point where long-range interactions begin to contribute to polymer conformations. Rather, polyQ in water behaves as a polymer in a bad solvent, in which solvent water is excluded from the collapsed polymer chain. Second, that the compactness of the molecule does not change appreciably as polyQ repeat length increases from Q_{15} to Q_{54}, that is, through the typical disease threshold of ~35. The relatively compact structure of polyQ in solution was confirmed in two FRET studies [32, 34]. Atomic force stretching experiments on polyQ sequences embedded within artificial flanking sequences are also consistent with a collapsed structure in contact with water, suggesting a very high barrier to the unfolding or extension of hydrated polyQ sequences [35]. These results are also consistent with computer simulations that predict a collapsed coil structure for monomeric polyQ, likely held together by H-bonding between side chain and main chain amide groups, in preference to H-bonding to water [27, 29, 36, 37]. This internal H-bonded structure could well play a role in the expected highly inefficient formation of β-hairpin [30] or other regular structures [37] that might be implicated in aggregation nucleation (see below).

Is there a "misfolded" state of polyQ?

Inconsistent with the above data on a polyQ repeat-length independent disordered state are antibody binding data that have been interpreted to suggest the existence of populated alternative conformations of monomeric polyQ that become more highly populated at repeat lengths above the typical disease risk threshold of ~35–45. Trottier and colleagues showed that the anti-polyQ monoclonal antibody 1C2 binds considerably better, in Western blots, to longer polyQ sequences compared with short versions [38]. This led to the idea that 1C2 recognizes a specific polyQ conformation that is much more highly populated in long polyQ sequences than in short, and therefore to the idea that 1C2 recognizes a toxic conformation populated only in long polyQs. An alternative explanation emerged, however, when the Bjorkman group showed that a similar MAb, MW1, owes its preferential binding to long polyQ peptides to a "linear lattice" effect of the polyQ sequence [39]. According to their analysis, the affinity of MW1 for a polyQ sequence is enhanced according to the number of individual epitopes in the polyQ chain, even though each individual epitope might be short (as was, in fact, subsequently shown by an X-ray crystal structure of the complex [40]).

More recently, another antibody with a preference for long polyQ sequences, 3B5H10, has been described and its selective binding also interpreted as evidence for a toxic folded conformation in long polyQ sequences [41]. While this remains formally possible, antibody binding does not prove that a particular conformation is populated in the absence of the antibody. The thermodynamics of complex formation allows antibodies to recruit and enrich kinetically accessible conformations even if they are only minimally populated in the absence of the antibody [42]. Conformational recruitment is an especially attractive possibility when the target molecule is an intrinsically disordered protein like polyQ that is presumably capable of undergoing coupled folding and binding [43, 44] to more structurally organized binding partners.

Although there is no convincing evidence for a populated, pre-existing conformation in expanded polyQ sequences that might be considered a “toxic conformation”, it remains possible, nonetheless, that monomers may yet prove to be the toxic species. This could happen if the toxic event (a binding interaction, for example, to some cellular target) is only favorable with a polyQ sequence of a particular length. Models for such binding phenomena might be (a) the linear lattice [39] type of polydentate interaction, (b) a binding cavity (such as found in some molecular chaperones or proteasomes) satisfied by the compact coil state of a sufficiently large polyQ molecule, or (c) a templated, coupled folding and binding event [43] (such as aggregate elongation, for example) with a thermodynamic requirement for a certain amount of buried surface area that can only be supplied by expanded polyQs. Thus there are a number of mechanisms by which one would expect to observe a correlation between cell toxicity and cellular levels of monomeric expanded polyQs, as has been recently observed [41]. Such correlative studies do not prove the existence of a populated, toxic, misfolded conformation in the monomer ensemble, however, and in fact merely formally indicate a pool of molecules that are capable of undergoing a toxic binding event.

Solution structure of complex polyQ sequences

In all expanded CAG repeat proteins, the polyQ sequence is embedded within the protein and therefore has both N- and C-terminal flanking sequences. Sequence algorithms make somewhat contrasting predictions of secondary structure of both the polyQ insert and surrounding sequence, with one program predicting disorder [45] and another predicting the ability to assemble into coiled coils [46]. Besides the structural details of a particular peptide in solution, one would also like to know how the flanking sequences influence polyQ structure and properties, and how the polyQ sequence (and its expansion) influences surrounding structure. There are significant technical limitations on addressing such questions experimentally with intrinsically unstructured proteins, requiring specialized spectroscopic approaches [44, 47]. Perhaps because of this, a number of groups have looked at polyQ disease protein structure issues through computational approaches.

There has been a special emphasis on sequences representing the N-terminal 3% of the huntingtin sequence (Fig. 1), which includes the polyQ segment as well as an N-terminal sequence (htt^{NT}) of 17 amino acids and a C-terminal segment rich in proline residues. A hint that flanking sequence can influence polyQ conformation in the monomer ensemble comes from the ability a C-terminal P_{10} sequence to slow the nucleated growth of polyQ amyloid and to abrogate the ability of a temperature decrease to enhance α-helix formation in a polyQ sequence [25]. A series of crystal structures of a significantly altered htt N-terminal fragment containing a short polyQ repeat show the $htt^{NT}$to fold into an α-helix, the C-terminal proline rich sequence to explore PP_{II}structures, and the intervening polyQ to occupy a variety of structures, including α-helix [24]. The extent to which crystal packing forces dictate any of these results is not clear but certainly must be considered possible, especially for a peptide that is disordered in solution.

The solution structure of a protein includes its native oligomerization state, and oligomerization can influence secondary structure in profound ways. Recent evidence suggests that htt N-terminal fragments are capable of reversible tetramer formation, almost certainly mediated by the formation of an α-helical bundle centered in the htt^{NT} segment. The htt^{NT} sequence in isolation as a monomer is disordered [48], exhibiting a slight tendency toward α-helix [48, 49]. Recent work shows that the α-helix content of the htt^{NT} segment depends on concentration, and that at concentrations of 0.5 mM or higher, the peptide is essentially fully helical [50] (Fig. 2a). These results, coupled with an analytical ultracentrifugation (AUC) analysis suggesting that htt^{NT} and $htt^{NT}Q_{10}K_2$ molecules exist at modest concentrations as an equilibrium of monomer, tetramer, octomer, and perhaps higher oligomers [50] (Fig. 2b) suggests that htt N-terminal fragments are capable of facile

oligomerization through α-helix bundle formation *via* their htt^{NT} segments. Consistent with the AUC data, various fluorescence techniques previously indicated an early assembly in cells of GFP-tagged htt exon-1 proteins into small multimers [51, 52]. Small multimers were not observed in striatal neurons expressing htt exon-1, however [41]. The htt^{NT} sequence has been proposed to be an important regulatory appendage of htt, mediating – with or without post-translational modifications – a variety of targeting, trafficking, and clearance operations in the cell [49, 53–58]. In this context, it seems possible that reversible tetramer formation may be a means by which solution exposure of the promiscuous htt^{NT} sequence can be itself regulated. Through this or other means, it is possible that tetramer formation by expanded polyQ versions of htt N-terminal fragments may somehow play a direct role in toxicity. The implications of htt^{NT}-mediated oligomer formation for the aggregation nucleation mechanism of htt N-terminal fragments is discussed later.

Two groups have computationally studied the htt^{NT} segment in isolation. Kelley et al. found a high tendency for all or part of the htt^{NT}segment to form α-helix structure [59]. Rossetti et al. found a number of conformational families which, except for a small fraction of condensed coil, are dominated by extended, solvated structures with helical segments [60]. In general, these findings are in agreement with experimental data that show monomeric htt^{NT} to be disordered at low concentrations [48], and to form helical tetramers at higher concentrations [50].

Other groups have conducted simulations on various htt N-terminal fragments. Consistent with the polyproline effect on aggregation mentioned above, Lakhani et al. found that the polyproline region of htt exon-1 decreases the ability of polyQ to engage β-structure elements within the monomer [31]. This group also found that polyQ expansion leads to increased β-structure in the htt^{NT} segment [31]. In contrast, Williamson et al. found some α-helix tendency within the htt^{NT} segment of htt^{NT}-Q_N sequences (i.e., without the proline domain), and a tendency to increased disorder in htt^{NT} as polyQ repeat length increases [61]. Dlugosz and Trylska found an intriguing repeat length dependent conformational change within an htt N-terminal fragment [62]. For both long and short polyQ sequence versions, both the $htt^{NT}$and the polyQ were largely α-helical. In the short polyQ molecule, the PP_{II} helix of the proline-rich tail tends to dock to the htt^{NT} helix, while in the long polyQ version these elements do not interact and are therefore more solvated [62]. The variety of simulation results reflected by the above discussion presumably results from the use of different simulation methods [62].

A number of model studies have been conducted on polyQ fusion proteins, and attempts made to interpret CD or FTIR data to deduce structural consequences. Placing polyQ inserts between secondary structural elements within a folded protein to various extents can destabilize and/or alter the folding or local structure of the host protein [63–67]. In a roughly repeat-dependent fashion, polyQ inserts tend to explore disordered and/or β-structure [65, 67]. Interestingly, polyQ sequences fused to some intact folded domains tend to acquire additional α-helical character [67,68], although other protein hosts don't seem to detectibly affect polyQ structural preferences [21]. In general, polyQ adjacent to a folded domain does not appreciably affect native structure and/or folding stability of the host protein [66, 67].

What is the normal role of polyQ in protein structure and function?

In some settings, polyQ repeat length correlates inversely with protein activity [69, 70], suggesting the possibility that repeat expansion diseases, while predominantly gain-of-function [8], might also possess some loss-of-function character in some cases [71]. As a general rule, however, it is not clear that polyQ segments in proteins actually have important, specific functions. Rather, they might simply have developed over the course of evolution at protein sites where polyQ tracts are tolerated, undergoing relatively cost-free expansion at

these sites due to the replicational instability [72] of the CAG repeat sequence. An example is shown in Table 1, a line-up of the polyQ-proximal sequences of various TATA box binding proteins (TBPs) over a wide range of evolution. While the sequences surrounding the polyQ are well-conserved, the polyQ repeat length itself ranges from 4 to 38 (Table 1). Such sequence expansion in an otherwise highly conserved protein suggests a linker between, or adjacent to, structural and/or functional units. At such sites normal protein function might be largely retained even when the polyQ sequence expands to a pathological repeat length. Such a scenario is roughly compatible with the results from model system studies discussed in the preceding paragraph.

Aggregation of Simple Polyglutamine Sequences

Because the only apparent common element of the proteins associated with the 10 expanded CAG repeat diseases is the polyQ sequence, the solution and aggregation behavior of simple polyQ sequences has been seen by several groups as a reasonable starting point in studying this disease family [20, 23, 25, 33, 34, 45, 73–79]. One common non-Gln element of these sequences has been in inclusion of flanking charged residues [19, 73, 80, 81] which bestow on these peptides a degree of kinetic (i.e., short term) solubility that allows their solution properties, including nucleation of aggregation, to be studied. The presence of flanking charged residues appear to be quite important; a polyQ sequence lacking flanking charged residues is reasonably well-behaved during disaggregation, but precipitates immediately upon adjustment to PBS conditions [81]. While it might be argued that a polyQ peptide lacking flanking charged residues is a more appropriate setting in which to study the properties of polyQ sequences in a larger protein context, polyQ fused to a number of disease protein flanking sequences retains the good kinetic solubility of $K_2Q_NK_2$ molecules, in contrast to the poor solution properties of polyQ lacking flanking residues, [25, 45]. While charged flanking residues probably introduce some qualitative rate differences, most of the experiments described in this section employ flanking charged residues to make physical experiments and comparisons feasible.

Soon after the observation that expanded polyQ aggregates accumulate in neurons [12, 13], *in vitro* studies on recombinantly produced exon-1 fragments demonstrated robust aggregation into amyloid-like fragments that is both repeat-length dependent and concentration-dependent [82]. Subsequently, *in vitro* studies on chemically synthesized polyQ peptides in the $K_2Q_NK_2$ background showed similar repeat length dependent aggregation [20] into ordered, amyloid-like [81] aggregates. The repeat length dependence was observed in both a spontaneous aggregation reaction dependent on both nucleation and elongation properties and in a fibril growth assay requiring only elongation of existing seeds [20]. A nearly identical repeat length dependence was subsequently demonstrated in a *C. elegans* model [83]. The microtiter-based elongation assay also showed that amyloid-like aggregates of one polyQ repeat length are able to seed elongation of other polyQ repeat lengths [20], supporting recruitment-based mechanisms for aggregate toxicity [84].

Nucleation of simple polyQ aggregation

The concentration dependence of an aggregating system is often dictated by the nucleation mechanism [85]. Studying the concentration dependence not only reveals clues to the nucleation mechanism, it also provides insight into how the rates observed for relatively high, laboratory concentrations might translate into the lower, steady state concentrations in the cell. An initial investigation of the concentration dependence of spontaneous aggregation for relatively long polyQ peptides in the $K_2Q_NK_2$ series revealed the surprising result of a rather shallow concentration dependence consistent with a critical nucleus (n*) of one [23] (Fig. 3a). Subsequent studies on shorter polyQ sequences in the same series revealed that n* itself is repeat length dependent [45]. PolyQ peptides $K_2Q_{18}K_2$ and $K_2Q_{23}K_2$ follow

concentration dependence consistent with n* ≈ 4, while polyQ peptides of $K_2Q_{24}K_2$ and $K_2Q_{25}K_2$exhibit n* ≈ 2. PolyQ peptides of $K_2Q_{26}K_2$or longer have n* ≈ 1 [45]. This discovery put the n* ≈ 1 value from earlier work [23] into context, suggesting that all simple polyQ peptides aggregate according to a classical nucleated growth polymerization mechanism.

The nucleus for simple polyQ aggregation, regardless of the n* value, can be modeled as a small β-sheet assembly, as expected (although technically not required) for an amyloid growth nucleus. As repeat length increases, folded versions of monomers, themselves containing β-sheet elements, take on enhanced (but still very low) stabilities that allow them to simplify nucleus assembly by requiring the involvement of fewer molecules [45]. One attractive model for these transiently populated conformations is the β-turn [45, 74, 86]. Indeed, one study showed that polyQ peptides containing L-Pro-Gly pairs compatible with β-hairpins aggregate as rapidly as simple polyQ of the same repeat length, and a similar peptide containing D-pro-Gly pairs that encourage β-hairpin formation aggregate faster than simple polyQ [74]. Both peptides exhibit the same n* ≈ 1 as do unbroken polyQs in this repeat length range, consistent with models for nucleus structure invoking folding into a β-sheet assembly [45]. A recent study showed a similar rate enhancement for D-Pro-Gly in a short polyQ background [79]. The conformational details underlying nucleation and aggregation of polyQ sequences containing Pro-Gly interruptions have been examined by computation analyses [87, 88].

The model for nucleation of aggregation by a long polyQ peptide (Fig. 3a) depicts the monomer ensemble as a condensed coil structure, as suggested by FCS measurements [33], and the nucleus as a specific folded form possessing β-structure. While nucleus formation is therefore depicted as a folding reaction, it is a highly unfavorable folding reaction. This unfavorable folding reaction is a pre-equilibrium, and the nucleus, once it forms, can either collapse back to the condensed coil monomer ensemble or can undergo a productive elongation step by interacting with a condensed coil monomer, forming an elongation intermediate by (presumably) elaborating a β-stand already present in the nucleus. Subsequent folding steps consolidate β-sheet rich, amyloid-like structure in a now-stabilized dimer that is more likely to engage in a new elongation step, and less likely to disintegrate [30]. This multi-step elongation process can be thought of as a “dock-and-lock” type of elongation mechanism [89] in which newly added monomers organize at the growth point of the template fibril, rather than encounter the growth point in a pre-organized state.

The ideal behavior of the polyQ nucleation mechanism has allowed a number of studies to elucidate details of the mechanism. The mathematical treatment of the mechanism-based kinetic model yields not only n* but also a complex parameter containing K_n^*, the equilibrium constant for nucleus formation, and k_+, the second order rate constant for nucleus elongation [23, 85] (Fig. 3a). A value for k_+ for a $K_2Q_{47}K_2$ peptide of 1.14×10^4 $M^{-1}sec^{-1}$ was obtained by determining the pseudo-first order kinetics rate constant for a seeded elongation of this peptide, then converting that value into the second order rate constant using experimental estimates of the concentration of fibril growth points [75]. Using this second order rate constant, a value for K_n^* of 2.6×10^{-9} was obtained [75]. $\Delta G_{folding}$ of + 12.2 kcal/mol (compared to the typical negative $\Delta G_{folding}$ for a small protein of −5 to −15 kcal/mol [90]), illustrates the rarity of the critical nucleus and the low favorability of nucleus formation.

In the classical nucleated growth mechanism, the efficiency with which the ephemeral critical nucleus elongates is a critical component of the overall efficiency of nucleation. Along with K_n^*, the relative rates of nucleus disintegration and nucleus elongation dictate nucleation efficiency (Fig. 3a). Since the rate of nucleus elongation = k_+[nucleus]

[monomer], this rate can be stimulated, without affecting nucleus formation, by adding relatively high concentrations of short polyQ peptides [76]. The end result is an increase in nucleation efficiency by short polyQ sequences. This consequence of the promiscuity of polyQ amyloid elongation may have *in vivo* significance [76].

The ability to fit aggregation kinetics to a mechanism-based mathematical model also allows generation of simulated aggregation kinetics, using the parameters from the fit, under conditions where measurement is not possible. Such simulations predict that even very small differences in n* have huge consequences for nucleation efficiency at the low polyQ concentrations of the cell. Thus, projection of aggregation curves for various repeat lengths of $K_2Q_NK_2$ to 1 nM concentration shows onset of aggregation in the 10–100 year span for $K_2Q_{37}K_2$ and 10–100 billion years for $K_2Q_{23}K_2$ [45] (Fig. 4).

Deviations from the classical nucleated growth model

In the classical nucleated growth model, the overall aggregation reaction is depicted as a series of homologous equilibria, each of which involves the binding of a ground state monomer to the species A_n to generate the species A_{n+1} [85] (Fig. 3b). Initial binding equilibria (starting with the interaction of two monomers to make A_2) are thermodynamically unfavorable, and later equilibria (approaching a limiting, low monomer concentration) are favorable, and the dividing line between these two classes is the critical nucleus A_n^*, the least stable species on the aggregation pathway [85]. In the context of this formulation of the classical mechanism, it might be argued that a result of n* = 1 is impossible, since, according to that definition, all species on the assembly pathway are oligomeric. However, what if we slightly alter the definitation of the critical nucleus as a consolidation of structure (i.e., regardless of molecularity) with sufficient stability (and hence lifetime) to sometimes undergo a productive elongation encounter? This does not change the mathematical formalism, while easily accommodating the observation of n* = 1. In fact, it is worth noting that the surprising result of n* = 1 for long polyQ peptides is generated by a data treatment based on a classical model which never anticipated this most simple of solutions. In the next section, the classical nucleated growth model is also found to be too narrow to accommodate an aggregation mechanism for complex polyQ sequences, even though it too features the stochastic formation of a critical species required for amyloid growth.

Other mechanistic models

Based on computational modeling that suggests an ability of polyQ peptides to form dimers [37] and higher oligomers [91] through non-specific interactions, and the observation that such oligomers can form [92], including during fibril formation [77, 93], an alternative nucleation model has been proposed in which amyloid nuclei are generated from within these oligomeric intermediates [77, 94]. As further support for this model, an analysis showed that it is compatible with previously reported kinetics data, and could account for the non-integer values that sometimes are reported for n* [94]. Models involving non-amyloid intermediates do have a certain attraction, given the ubiquitous appearance of on- or off-pathway oligomeric intermediates in almost all other amyloid systems [95]. There are several issues with this alternative model, however, as it is applied to simple polyQ peptides. First, observations of non-amyloid oligomers from simple $K_2Q_NK_2$ peptides are quite possibly due to artifacts in sample preparation (see discussion in aggregate morphology section). Indeed, careful adherence to disaggregation protocols [45, 96] and a variety of tests focused on detecting low levels of oligomers failed to show any evidence for non-amyloid aggregates during amyloid formation by $K_2Q_NK_2$ peptides [45]. Furthermore, although fractional n* values are often reported [23, 25, 45, 74–76, 78], there are other theoretical rationalizations for non-integer n* values, and it has also proved technically difficult to determine n* values

with accuracies sufficient to show that non-integer values are significantly different than the nearest integer value [45]. Finally, direct conversion of amorphous aggregates into amyloid-like aggregates by computer simulations has not been reported, nor have observed non-amyloid oligomeric intermediates been successfully incorporated into kinetic models of aggregation. While it remains possible that very low levels of very small oligomers might be missed in experiments, it has not been necessary to invoke such species in order to account for observed nucleation kinetics [45].

Interestingly, it is possible that an oligomer-mediated mechanism applies to the aggregation of histidine-interruped polyQ peptides in the acid pH-range (see below). And there is a very clear case of oligomer-mediated amyloid nucleation in the case of htt N-terminal fragment aggregation, as discussed later.

PolyQ sequence interruptions

Between the world of simple, unbroken polyQ sequences, as discussed above, and complex polyQ proteins with extensive flanking sequences, as discussed below, are peptides that consist of mostly polyQ but containing sequence interruptions. The limited analysis of these sequences suggests that their aggregation is qualitatively similar to unbroken polyQ but with altered kinetics. Sequences containing Pro or Pro-Gly interruptions [74, 79, 97–99] are discussed elsewhere in this review.

Sequences containing His interruptions are of special interest, since in Spinocerebellar ataxia 1 (SCA1), one of the expanded CAG repeat diseases, repeat expansion is linked to loss of internal His residues. At the protein level, some individuals have expanded polyQ tracts but retain the His residues, and, perhaps because of that, are symptom-free [100]. Analysis showed that such His insertions do not appear to alter the disordered state of monomeric peptides [97], and in the neutral pH range do not alter the nucleation mechanism [78]. However, His interruptions do substantially reduce aggregation kinetics [78, 97]. A proposed model for aggregate structure [97] in which the His residues are placed in reverse turns was supported by a structural analysis of the aggregates [78]. Interestingly, His-containing polyQ peptides incubated at pH 6 experience a significantly reduced aggregation rate and exhibit an alternate aggregation reaction profile with clear evidence for a transient non-amyloid intermediate [78]. Whether this is on- or off-pathway, however, was not determined. These changes are presumably brought about by protonation of the His residues at the lower pH.

More extreme examples of mixed polyQ sequences should also be mentioned. In a dramatic example, the mixed Ala/Gln sequence K_2-$(QA)_{20}$-K_2, derived from the htt interacting protein CA150, forms amyloid fibrils with approximately the same kinetics as a $K_2Q_{40}K_2$ sequence [101]. Furthermore, amyloid-like aggregates of these two sequences can “cross-seed” the elongation of the other [101], which is typically not a very efficient process for amyloidogenic peptides of widely different sequences [102]. These results are especially interesting given the possible involvement of amyloid formation in some polyAla repeat diseases [103]. The other mixed sequences discussed in this section, containing His insertions or Pro-Gly insertions, are also very effective in cross-seeding reactions with simple polyQ sequences, suggesting a high degree of structural similarity [74, 78].

Complex Aggregation Mechanisms for Complex PolyQ Sequences

All known expanded CAG repeat disease proteins contain a polyQ sequence surrounded by mixed amino acid sequence peptides. Some of these flanking sequences do not alter the fundamental aggregation nucleation mechanism, although they may affect aggregation rates [25, 45]. In other cases, however, flanking sequences can play an enormous role, greatly

affecting rates, sometimes by fundamentally changing the aggregation mechanism. This multiplicity of effect is mirrored in studies on artificial polyQ fusion protein model systems. In some cases, producing the polyQ sequence as a fusion protein greatly suppresses amyloid formation, facilitating purification of the protein and allowing control over the start of aggregation via specific protease cleavage sites between fusion partner and polyQ domain [104]. Studies on intact fusion proteins give a range of results. Some fusion proteins produce only non-amyloid aggregates [67], some make amyloid via non-amyloid, oligomeric intermediates [105, 106] and some produce amyloid fibrils with little or no evidence for non-amyloid intermediates [66]. For those proteins that feature formation of non-amyloid intermediates, an important question is the extent to which polyQ expansion influences the first step, *versus* the second step, of the two-step mechanism of amyloid formation.

Ataxin-3

An excellent example of the ability of flanking sequence to dramatically alter the aggregation pathway is the role of the "Josephin" domain of the disease protein ataxin-3 (AT3). The Josephin domain by itself or with a short polyQ extension undergoes thermally induced formation of worm-like, SDS-soluble, ThT-positive fibrils [107, 108]. While short polyQ versions of AT3 aggregate about as efficiently as the isolated Josephin domain [107, 108], however, longer polyQ versions aggregate more rapidly and do so *via* a two-step mechanism [108]. In this mechanism, the initial product has similar structural features to the wormlike morphologies of the short polyQ AT3 molecules, while the final product is a more stable, amyloid-like fibril [108]. There are conflicting data on whether polyQ and its expansion influences the stability of the Josephin domain [107, 109–111], which, if true, could influence the first step of the aggregation mechanism. The lead role played by aggregation of the Josephin domain, and/or by the intervening flexible segment [112] is, however, well documented. Josephin point mutations [113, 114] or binding to chaperones [115] or normal ligands [114] have all been observed to suppress initial aggregation and/or amyloid formation.

Huntingtin N-terminal fragments

Huntingtin, the protein responsible for Huntington's disease (HD), is an intracellular protein of ~ 3,200 amino acids whose polyQ sequence resides very near the N-terminus (Fig. 1). A variety of studies have implicated one or more key proteolysis events in the disease mechanism, in which release of polyQ-containing N-terminal fragments is required for the emergence of toxicity. Animals expressing expanded polyQ versions of htt N-terminal fragments like exon-1 (Fig. 1) generally exhibit aggressive HD disease features such as neurological symptoms, cellular aggregates, cellular abnormalities, and death [116]. *In vitro* studies on recombinant exon-1 proteins exhibit a parallel polyQ repeat length dependent amyloid formation [82]. In contrast with simple polyQ aggregation, however, the aggregation pathway of htt N-terminal fragments *in vitro* features early formation of aggregates lacking typical amyloid morphologies [117].

Initial investigation of htt polyQ flanking sequence effects focused on the unusual Pro-rich segment (Fig. 1). In cell models, removal of the Pro-rich sequence from exon-1 significantly enhances aggregate formation and/or toxicity [118, 119]. *In vitro*, the previously discussed ability of a C-terminal polyproline segment to reduce the aggregation rate of a polyQ sequence [25] is also observed in the htt N-terminal fragment context [48] (Fig. 5a). The polyproline effect observed *in vitro* can be accounted for by its ability to alter the conformational mix in the polyQ component toward less aggregation-prone conformations [25, 120, 121]. This conformational induction depends on details of the connection of polyproline to polyQ: a P_{10} sequence at the polyQ C-terminus inhibits aggregation, but at the polyQ N-terminus does not [25]. The situation is more complicated in

yeast cells, where independently expressed polyproline sequences can interfere with polyQ aggregation *in trans* [119].

The mechanism of the htt^{NT} effect on polyQ aggregation

The addition of the 17 amino acid htt^{NT} sequence (Fig. 1) to the N-terminus of polyQ produces a dramatic increase in aggregation rate compared with the simple polyQ peptide [48] (Fig. 5a). In contrast with the polyproline effect, htt^{NT} has the same rate enhancing effect when placed either on the N- or C-terminus of polyQ [48] (Fig. 5a). In a series of htt$^{NT}Q_N$ peptides, there remains a very strong repeat length dependence [48] (Fig. 5b), consistent with previous observations on htt exon-1 sequences [82, 104]. In analogy to the AT3 system (see above), aggregation of htt$^{NT}Q_N$ proteins occurs by a two-step mechanism. Rough features of this two-step mechanism are now fairly well understood (Fig. 6).

After long incubation at relatively high concentration, htt^{NT} peptides without any attached polyQ form α-helix rich [50, 122] (Fig. 2a), pelletable [48] aggregates. This is now visualized as a continuation of the ability of htt^{NT}to reversibly form α-helix rich tetramers, octomers and dodecamers as part of a native state ensemble (Fig. 2). When polyQ sequences are attached to htt^{NT}, pelletable oligomers tend to form more quickly in a polyQ repeat length dependent fashion [48]. Once oligomers are formed, they have the ability to stochastically rearrange to form a polyQ-core amyloid nucleus, while retaining the α-helical structure in htt^{NT} [50] (Fig. 6). This enhanced polyQ amyloid formation is presumably related to the AT3 nucleation mechanism, being mediated by high local polyQ concentrations within the intermediate aggregates. This second, nucleation step appears to also be polyQ repeat length dependent [48, 50]. For example, an htt$^{NT}Q_7K_2$ peptide oligomer cannot undergo nucleation of polyQ amyloid structure, while an htt$^{NT}Q_8K_2$ oligomer can [50]. While nucleus formation in short polyQ htt N-terminal fragments appears to take place within large, pelletable aggregates [50], it is possible and even likely that long polyQ htt$^{NT}Q_N$ peptides undergo efficient nucleation within tetramers or octomers (Fig. 6). The critical role of α-helical bundle formation in htt^{NT}-mediated polyQ aggregation is further supported by inhibitor studies described below. The two-step mechanism involving an initial oligomer held in an α-helix bundle, followed by nucleation of amyloid structure in non-bundled segments, is similar in many respects to a recent proposed mechanism of amyloid nucleation in Aβ and IAPP [123].

There are many details of the nucleation process to work out, but typically this event is associated with a number of ensemble observations. After nucleation, aggregation rate increases dramatically [48], except with very low polyQ repeat lengths [50]. Aggregate morphology changes from oligomers and protofibrils to fibrils [48] (Fig. 7). Post-nucleation aggregates (a) respond to ThT, (b) become capable of seeding elongation, (c) exhibit a less solvent exposed htt^{NT} sequence, (d) begin to develop β-structure in FTIR spectra, and (e) have polyQ sequences that are no longer recognizable by an anti-polyQ antibody [48, 50]. While most of these details were worked out with relatively short htt N-terminal fragments of the general structure htt$^{NT}Q_NP_6K_2$ and htt$^{NT}Q_NP_{10}K_2$, recent studies with chemically synthesized htt exon-1 [124] show that these short polyproline analogs behave very similarly to htt exon-1 in aggregation rates, mechanism, and response to inhibitors (B. Sahoo, D. Singer, T. Zuchner and R. Wetzel, Ms. in preparation).

There are some alternative models for the aggregation mechanism of htt N-teminal fragments and how htt^{NT} contributes to it. Tam et al. posit that the role of α-helical htt^{NT} is to intramolecularly induce a more amyloidogenic conformation in the polyQ segment [125]. Others, based on experimental [126] or computational [61] studies, propose that initial oligomer formation by htt N-terminal fragments is mediated primarily by direct polyQ interactions.

Implications of the htt^{NT}-dependent aggregation pathway

The role played by the htt^{NT} segment in determining the physical properties of htt N-terminal fragments may greatly affect the behavior of such peptides in cells. In addition to the possible regulatory role of htt^{NT}-mediated, reversible oligomer formation (see above), the role of htt^{NT} in mediating amyloid formation suggests that post-translational modifications within this sequence might influence the rate or extent of pathological aggregation. A potential example is in the phosphorylation of Ser13 and Ser16. A double Ser to Asp phosphomimetic mutation within htt^{NT} in expanded polyQ htt in a tg mouse model was found to very effectively block aggregate formation and neurodegeneration [58]. In a peptide model system, the same double Ser to Asp mutation *in vitro* was found to both suppress aggregation rate and to alter final aggregate morphology [58] (Fig. 7l). These data suggest that protein phosphorylation might not always or only play a regulating role in the cell, but might sometimes lead to important cellular consequences simply due to changes in the biophysical properties of the protein substrate [58, 127].

The existence of two radically different aggregation mechanisms for polyQ sequences (i.e., in simple polyQ sequences and in complex sequences like htt fragments) suggests the possibility of mixed mode aggregation under specific conditions. In fact, incubation of an htt N-terminal fragment in the presence of an inhibitor of htt^{NT}-mediated aggregation (below) appears to open the door to the alternative aggregation mechanism engaged by simple polyQ sequences (M. Jayaraman, A. K. Thakur, and R. Wetzel, unpublished). This may be the source of temperature-dependent aggregate conformations reported by the Tanaka group [128]. In addition to HD-specific issues, the htt^{NT} story has more general implications for predictions of amyloidogenic sequences, providing an example of a sequence that greatly enhances amyloid formation rates without ever engaging an amyloid or β-sheet structure [50, 122].

The question of whether non-amyloid, oligomeric intermediates are on-pathway or off-pathway to amyloid formation has stimulated significant work and discussion in the field [95], and there is almost certainly no universal answer that will apply to all amyloid assembly reactions. The assembly mechanism suggested in Figure 6 presents an interesting case, in which α-helix rich oligomers play both an on-pathway role - as the substrate within which nucleation of amyloid structure occurs [129], and an off-pathway role – as a reservoir for the release of monomers [130] to feed amyloid elongation as solution phase monomer is depleted [50].

Inhibitors of Polyglutamine Aggregation

Characterization of aggregation inhibitors is important for several reasons. If accumulation of a particular type of aggregate in cells is responsible for at least some HD symptoms, then understanding inhibitor structure and mechanism of action can contribute directly to the discovery of lead compounds, and indirectly to providing a basis for inhibitor design. Even if the practical aspects of inhibitor delivery complicate an anti-aggregation approach to HD therapy, inhibitors should make excellent tools for determining aggregation mechanism and for understanding the cellular basis of aggregate toxicity. Inhibition of huntingtin aggregation is an especially interesting challenge, since, while the mechanism is quite different from that of simple polyQ amyloid formation, the polyQ core structure of the aggregates appears to be quite similar [122]. Although high-throughput screens conducted in cell or animal models are potentially powerful ways of cutting through bioavailability issues, they also have the downside of delivering hits whose mechanisms of action are initially obscure and possibly indirect. Given the focus of this review, cell screening based inhibitors are only described if subsequent characterization indicated a direct polyQ effect.

Inhibitors of simple polyQ aggregation

Directly targeting polyQ aggregation by focusing on simple polyQ is attractive because effective inhibitors might be useful in all 10 expanded CAG repeat diseases. Burke and co-workers identified peptide based inhibitors by using phage display to select short sequences with high affinities to immobilized polyQ sequences [131]. One peptide, the Trp-rich sequence QBP1, is particularly effective against polyQ aggregation *in vitro* [131] with an IC_{50} in the 5 μM range [132]. Subsequent studies characterized structure-activity relationships in this inhibitor [132, 133]. This peptide is effective in inhibiting aggregation and/or toxicity in cell [131, 134] and animal [135] models, and is widely used to test the role of polyQ amyloid formation in the aggregation of complex proteins (see, for example reverence [66]).

The β-sheet rich structure of the amyloid end products of polyQ aggregation also has provided the basis for several structure-based designs of inhibitors. By inserting periodic Pro-Gly pairs along a polyQ sequence, it is possible to control how that sequence engages an amyloid-like structure [74]. By then including additional Pro residues at mid-β-strand locations between the Pro-Gly pairs (i.e., $PGQ_9P^{1,2,3}$, Fig. 8b), inhibitors of polyQ amyloid elongation, with IC_{50} values in the 1–2 μM range, were obtained that are imagined to work by adding to fibril growing ends while being incapable of consolidating amyloid structure [87, 88, 98] (Fig. 8b). Prolines are anathema to β-sheet formation through a combination of constraining backbone torsion angles, steric clashes with β-sheet geometry, and the absence of an H-bonding N-H group [136]. The latter factor has also been used in polyQ aggregation inhibitor design by installing backbone N-methyl groups into oligoGln peptides [137].

Inhibitors of huntingtin N-terminal fragment aggregation

A number of library screening and structure-based design studies have been conducted with huntingtin N-terminal fragments like exon-1 as targets. Using a screening assay built around *in situ* proteolytic generation of htt exon-1 and a filter-trap aggregation assay [138], Wanker and co-workers have screened several hundred thousand compounds [139, 140]. Two classes of inhibitors thus identified are polyphenols [140, 141] and benzothiazoles [139]. Polyphenols like EGCG ((−)-epigallocatechin-gallate) appear to work by rapidly directing monomeric htt exon-1 down an alternative aggregation pathway leading to stable, non-amyloid aggregates [140]. Two benzothiazoles have been shown to be effective, with EC_{50} values of 40 μM, at reducing aggregate burden in organotypic brain slice cultures from the R6/2 htt exon-1 tg mouse [142]. Unfortunately, at the maximum tolerated dose, one of these compounds, riluzone, is ineffective in the R6/2 mouse, in spite of reaching 100 μM levels in the brain [142].

Another series of inhibitor can be considered to be structure and mechanism based. Peptides related to the isolated htt^{NT} sequence prove to be excellent inhibitors of htt N-terminal fragment amyloid formation [143]. The mechanism of action appears to be *via* their ability to make mixed oligomers with the exon-1-like molecules [143]. Co-assembly reduces the local concentration of polyQ chains in the oligomers and thereby decreases the frequency of amyloid-like nucleus formation (Fig. 8c). Thus, sequence variants that decrease the ability to fold into an amphipathic helix decrease inhibitory power [143]. In general, these inhibitors slow nucleation but do not totally eliminate it, presumably because, statistically, there are always a few oligomers formed that are predominantly composed of htt N-terminal fragments. In an extension of this design, the molecule htt^{NT}-$PGQ_9P^{1,2,3}$ – a combination of the htt^{NT} segment with the elongation inhibitor $PGQ_9P^{1,2,3}$ (see above) – is an even more effective inhibitor [143]. It presumably works by co-assembling with the htt N-terminal fragments, thereby diluting their solvent exposed polyQs with an active elongation inhibitor

(Fig. 8d). This mode of inhibition is similar to the proposed model for how Pro-containing variants of islet amyloid polypeptide (IAPP) inhibit WT IAPP amyloid formation [144].

A related mode of inhibition involving the role of htt^{NT} is to disallow homo-tetramerization by complexing the htt^{NT} moiety of htt N-terminal fragments with another protein. There are several examples. A recent X-ray crystal structure of a complex between htt^{NT} and an engineered antibody fragment with anti-aggregation and anti-toxicity activity in cells shows, in a different setting, the tendency of this peptide to adopt α-helix on binding [145]. Molecular chaperones have been reported to reduce htt exon-1 aggregation and toxicity in mammalian cells [146], and it now appears that at least some molecular chaperones do so by targeting the htt^{NT} sequence to suppress amyloid nucleation. TriC, for example, appears to inhibit htt exon-1 aggregation by binding and sequestering the htt^{NT} component [125]. It seems possible that the ability of Hsp70/Hsp40 to inhibit htt exon-1 aggregation [126] also relies on sequestration of the htt^{NT} segment. This seems especially likely given the reported ability of Hsp40/Hsp70 to reduce non-amyloid oligomer formation [126], as well as the observation that Hsp40/Hsp70 only blocks htt exon-1 aggregation when added during the lag phase [147]. The prospects for therapeutic exploitation of these observations are not clear. Although over-expression of Hsp70/Hsp40 suppresses neurodegeneration in a number of cell and animal studies [148], mixed results are obtained in the (very aggressive) R6/2 tg mouse model [148]. Especially given the MoRF-like character of the htt^{NT} sequence [48], there may be many other cellular proteins, besides chaperones and htt N-terminal fragments themselves, that bind to the htt^{NT} sequence during the normal cellular life of htt. These potentially very fluid interactions may play subtle modulating roles of htt aggregation.

Structural Features of Polyglutamine Aggregates

In vivo aggregates

As in other aggregation-associated neurodegenerative diseases, the disperse distribution and relatively light aggregate burden compared with total brain mass makes it difficult to obtain quality data on the underlying structural features of tissue-derived polyQ aggregates. PolyQ protein aggregates appear to be exclusively intracellular. The earliest observation of aggregates placed them in the neuronal nucleus [12, 13]. Subsequently it was recognized that aggregates are also often found outside the nucleus in the cell body [149, 150].

There is limited ultra-structural data on *in vivo* aggregates. High resolution electron micrographs of inclusions in brain slices sometimes reveal a substructure of well-defined fibrils[12]. In addition, aggregates isolated from the brains of a Huntington's disease (HD) mouse model exhibit a range of morphologies including a number of filamentous structures similar to those observed in the *in vitro* assembly of huntingtin fragment aggregates [151] (Fig. 7a,b). At least some of the end-stage aggregates in the brain appear to have amyloid-like structures, since they are capable of specifically and strongly binding biotinylated polyQ in a recruitment stain dependent on the normally highly specific amyloid elongation reaction [152]. The lack of specific stains for other types of (non-amyloid) aggregates, such as the oligomeric structures commonly observed in the early stages of the aggregation of huntingtin (htt) exon-1 and similar fragments (see below), makes it difficult to determine if these structures exist in tissue, and, if so, where they are localized and whether they persist.

Many cell models have been described in which polyQ proteins are fused to GFP. In these models, short polyQ sequences give diffuse fluorescence while long polyQs lead to one to several bright puncta either in the nucleus [153] or cytoplasm [118]. The limited number of inclusions per cell might be evidence for the rarity of an aggregation nucleation event [154] but could also reflect the ability of the cell to organize aggregates, such as in perinuclear aggresomes [155]. Since the assembly of some complex polyQ proteins such as htt exon-1

proceeds with intermediate formation of non-amyloid aggregates (see below), and since different aggregated states may have different toxicities [156], it is of considerable interest whether particular GFP puncta in different experiments represent accumulations of amyloid or other aggregates. AUC of cell contents [157], and fluorescence methods like FRET, FRAP, FLIP, FCS and split-GFP techniques [51, 52, 158–160] may ultimately clarify such issues.

In vivo aggregates are often found to be associated with other cellular proteins. In some cases this might reflect the operation of cellular mechanisms directed at protein misfolding or aggregation [161]. Due to the promiscuity of polyQ amyloid elongation [20], polyQ aggregates can also effectively recruit other polyQ-containing proteins [84]. Recruitment and sequestration [162], and/or recruitment-mediated elimination [163], of important cellular proteins such as transcription factors represent attractive mechanisms for how amyloid-like polyQ aggregates might be toxic. Such recruitment mechanisms may even sometimes contribute to identification of polyQ repeat-length dependent protein interaction partners [164, 165]. Although normally interpreted as evidence for specific binding partners for monomeric polyQ proteins, in at least some cases identification of binding partners may reflect the binding of proteins to already aggregated polyQ bait proteins[166].

Structures and properties of simple polyQ aggregates assembled in vitro

Mirroring the fine-structures of some polyQ aggregates formed *in vivo*, aggregates of simple polyQ peptides grown in aqueous buffers from disaggregated peptides are amyloid-like, exhibiting filamentous morphologies in EM images (Fig. 7c) [20, 81]. In spite of their filamentous morphologies, most aggregates grown *in vitro* from simple polyQ peptides do not exhibit the classical twisted fibril morphology seen in EM images, as typified, for example, by $A\beta_{40}$ fibrils (Fig. 7d). Other features of amyloid are well-replicated, however. Simple polyQ aggregates exhibit β-sheet features in circular dichroism [23], FTIR [48], X-ray diffraction [73, 167–169], and solid state NMR [122, 170]. Although the signal is unusually low, polyQ aggregates also bind thioflavin T (ThT) and exhibit the ThT fluorescence spectral shift characteristic amyloid fibrils [81]. Overall, there is no evidence that the polyQ component of simple or complex polyQ sequences in mature fibrillar aggregates exists in anything other than β-sheet rich, amyloid-like structure, in contrast to a recent proposal that α-helix-rich coiled coil might be the structural basis of such fibrils [46].

The absence of a 10 Å reflection in the X-ray powder diffraction of polyQ aggregates led Perutz to suggest a β-helical structure for these aggregates [167], and a number of computational analyses have been found to support the feasibility of such structures in polyQ sequences [26, 32, 91, 171–173]. Subsequent X-ray diffraction analyses, however, suggested a more conventional β-sheet structure for polyQ aggregates, although also finding, in contrast to many peptide-based amyloid fibrils [174], an anti-parallel β-sheet arrangement [168, 169] featuring β-hairpins [169]. Certain features of solid state NMR [122, 170] and CD (J. Kardos, personal communication) spectra are also consistent with an anti-parallel β-sheet structure in polyQ aggregates. The ability of polyQ sequences interspersed at regular intervals with Pro-Gly pairs to form aggregates with kinetics, mechanistic features, and morphologies similar to unbroken polyQ has also been interpreted as indicating an anti-parallel β-sheet arrangement [74]. Several anti-parallel four stranded sheet models have been studied *in silico* [30, 31, 36, 73, 175, 176]. While there is a lot of physical evidence (see above) that such structures are not measurably populated in the monomeric state, these would appear to be reasonable models for aggregation nuclei or for the fundamental repeat unit of polyQ amyloid fibrils.

Like other amyloid fibrils, simple polyQ aggregates exhibit a characteristic concentration of monomer when the fibril formation reaction reaches equilibrium [45, 81]. This monomer concentration can often be approached both in the growth direction and the dissociation

direction [45, 177] (Fig. 9), and is formally equivalent to the critical concentration, C_r, for aggregation (the concentration below which spontaneous fibril formation cannot be initiated). Even for less stable aggregates of short polyQ peptides, the dissociation of fibrils to the equilibrium position takes a very long time [45, 81] (Fig. 9). Since the C_r is related to the elongation equilibrium constant for fibril growth [177], the C_r is an indicator of fibril stability. Not surprisingly, the C_r decreases as polyQ chain length increases, with a value of ~ 3 μM for $K_2Q_{23}K_2$ [45] and 1 μM or less for longer sequences [81]. This trend suggests that aggregates of long polyQ peptides are more stable than those of short polyQs, and may have implications for the thermodynamic probability of aggregate formation in the cell.

There are many unresolved questions about amyloid-like polyQ aggregate structure. Although experiments with Pro-Gly interrupted polyQ sequences suggested that polyQ amyloid can be manufactured with β-sheet widths of 7–9 Gln residues [74], it is possible that unbroken polyQ might favor significantly wider sheets [170]. There are several chain connectivity patterns possible in an anti-parallel β-sheet [174], and it remains to be seen whether the pattern in polyQ amyloids involves β-hairpins or longer loops and whether it is uniform or variable. The structural role of the Gln sidechains has also not been established with certainty. Diffraction- or ssNMR-based models of fibrillar aggregates [122, 168–170, 178] or microcrystals [179, 180] of polyQ [168, 169] or Gln/Asn-rich [178–180] peptides, as well as AFM studies of Gln-rich peptides [86], exhibit different structural resolutions and degrees of Gln side-chain interdigitation.

Amyloid-like aggregates formed in frozen solutions

Simple polyQ peptides have the unusual ability to highly efficiently make amyloid-like aggregates in frozen solutions in PBS [20,81]. These aggregates share many properties of aggregates formed from the same peptides at 37 °C, but exhibit an atypical monofilament morphology (Fig. 7e). Perhaps because of this, compared with 37 °C aggregates, these assemblies are sometimes found to be much better seeds for elongating polyQ monomers [181] and are much more efficiently taken up by mammalian cells in a novel cell model for polyQ aggregate toxicity [182] and elongation [183]. Since aggregate formation depends on the amount of time incubated at −20 °C [81], this process seems to occur in the frozen solution and not as a consequence of the freezing and/or thawing processes. It is presumably brought about by the process of “freeze concentration” [184, 185], whereby solutes, at freezing temperatures above their eutectic points, accumulate in liquid fissures between ice lattices resulting in high solute concentrations. The atypical monofilamentous structures of these aggregates may reflect the confined environment in which aggregation takes place or a deviant β-sheet structure within the amyloid protofilaments. These aggregates have a number of potential uses for studies of polyQ phenomena.

Non-amyloid oligomers of simple polyQ peptides

There is a lively discussion in the literature as to whether simple polyQ peptides in the class $K_2Q_NK_2$ are capable of forming non-amyloid aggregates, and whether such aggregates play a role in the mechanism of spontaneous amyloid growth. Some computer simulations predict the ability of polyQ sequences to make such structures [37, 91], and aggregates described as non-amyloid oligomers have been observed in light scattering, EM and AFM measurements [77, 93]. These reports are inconsistent with other experiments, however, where oligomeric aggregates are looked for but not observed [23, 45]. These divergent results likely stem from differences in sample preparation. The required disaggregation protocol for synthetic polyQ peptides [45, 96, 186, 187] includes transient exposure to the volatile solvent hexamethylisopropanol (HFIP). This solvent is well known to dramatically modify assembly of some amyloidogenic peptides at 2–4% concentrations [188, 189], and thus must be rigorously removed under vacuum as part of the disaggregation protocol. In fact, it is

possible to drive the formation of homogeneous $K_2Q_NK_2$ oligomers in aqueous HFIP [92]. These data are consistent with the hypothesis that non-amyloid $K_2Q_NK_2$ oligomers arise due to incomplete HFIP removal. This question of non-amyloid oligomer formation is an important issue for several reasons as discussed elsewhere in this review.

In vitro aggregates of complex polyQ polypeptides

Limited available data suggests that the amyloid cores of mature aggregates of simple polyQ peptides and polyQ peptides with flanking sequences are essentially identical. In ss NMR analysis, the spectral features of this core in fibrils of an htt N-terminal fragment are identical to those of a $K_2Q_NK_2$ peptide of similar repeat length [122]. Analogous FTIR spectra are also very similar [48]. EM images of mature aggregates of htt N-terminal fragments [48, 104, 117, 151] show differences compared to aggregates of simple polyQ peptides, with filaments of greater thickness and rough edges that probably reflect the presence of coats of alternatively structured flanking sequences (Fig. 7f–h). For example, while it was expected that polyQ amyloid structure might propagate into the htt^{NT} flanking sequence at some point during amyloid growth, ss NMR showed the surprising result of a stable element of α-helix in a portion of the htt^{NT} sequence in mature fibrils [122]. This reflects the strong tendency of this sequence to form α-helix in various associated states, even though it is disordered in the monomeric state [48, 50].

In contrast with simple polyQ peptides, oligomer formation in the early stages of htt N-terminal fragment aggregation appears to be universally observed and has a fairly well-understood structural basis, as described in the section on complex polyQ peptide aggregation mechanisms. Similar small oligomers are formed from recombinant (Fig. 7i) or chemically synthesized (Fig. 7j) htt N-terminal fragments. *In vitro* produced protofibril-like structures (Fig. 7k) intermediate between spherical oligomers and mature fibrils resemble some aggregates isolated from tg HD mouse brains (Fig. 7b)

Conclusions

Expanded polyQ repeat diseases are often referred to as misfolding diseases, where misfolding can be meant to refer to the generation of a wide range of species, from alternative folded states of monomers to abnormal pathological self-assembled states ranging from non-amyloid oligomers to amyloid fibrils to the further self-associated inclusions seen in cell and animal models. The biophysical basis of these hypothesized states has been examined in this review.

How close are we to identifying the toxic species? Evidence for a stable or meta-stable specific folded state of expanded polyQ monomers, either in isolation or as part of a disease protein, remains elusive. At the same time, fundamental consideration of the intrinsically disordered nature of the polyQ sequence suggests that important complexes might be formed in the cell that do not depend on the existence of any populated, pre-existing, misfolded, “toxic” monomer. Thus, intrinsically disordered proteins (IDPs) are understood to be capable of highly specific coupled folding and binding [190] to cellular targets whose nature would never be apparent from only a knowledge of the especially rugged folding landscape of the IDP in isolation. In the context of this hypothetical cellular target, long polyQs might make stable complexes, and short polyQs unstable complexes, simply by virtue of the amount of buried surface area, H-bonds, etc., formed in the complex. Such a model positing that “long polyQs are toxic because they are long” may be scientifically unsatisfying, but is nonetheless consistent with the intrinsically disordered character of the polyQ sequence. The definitive test of the “toxic monomer hypothesis” will be to unequivocally demonstrate molecular events and structures associated with one of more early pathological cellular consequences. In contrast to misfolded polyQ monomers, there is no shortage of

demonstrated polyQ aggregates found both *in vitro* and *in vivo*. However, the same standard applies to the "toxic aggregate hypothesis" as to the toxic monomer hypothesis. Bad aggregates will have to be caught in the act; the circumstantial evidence of correlation does not provide enough evidence to lead to a conviction.

The physical chemistry of polyQ aggregation is perhaps more clear than its disease role, but is nonetheless not without controversy. PolyQ sequences are most thermodynamically stable in an aggregated state, and there are apparently a number of kinetically accessible paths, with varying efficiencies, to aggregate formation. Mechanisms for aggregation nucleation *in vitro* can be simple or ornate, but are ultimately driven by the high stability of the amyloid fold. Deciphering some of these mechanisms has allowed the design of structure and mechanism-based aggregation inhibitors which may allow tests of the aggregation hypothesis for expanded polyQ diseases, and ultimately perhaps an entrée into therapeutic design.

An important recent discovery is that at least one disease protein, the N-terminal fragment of huntingtin, is capable of a regular, reversible tetramerization through its htt^{NT} N-terminus. While this equilibrium might very well play a normal role in the cellular biology of huntingtin, it also offers a middle ground between monomers and higher-order aggregates as an expanded polyQ species which, in spite of its possible ubiquitous presence and useful functions, may take on an additional negative, and ultimately toxic, set of activities in longer polyQ molecules.

Belying the tedium of their amino acid sequences, the properties of polyglutamine segments are far from monotonous, offering continued surprises in the study of their biophysical behaviors and the search for molecular mechanisms of disease.

Acknowledgments

The author acknowledges NIH R01 AG019322 for financial support and Rakesh Mishra and Robert Fairman for helpful suggestions on the manuscript.

References

1. Karlin S, Burge C. Trinucleotide repeats and long homopeptides in genes and proteins associated with nervous system disease and development. Proc Natl Acad Sci U S A. 1996; 93:1560–5. [PubMed: 8643671]
2. Margolis RL, Abraham MR, Gatchell SB, Li SH, Kidwai AS, Breschel TS, Stine OC, Callahan C, McInnis MG, Ross CA. cDNAs with long CAG trinucleotide repeats from human brain. Hum Genet. 1997; 100:114–22. [PubMed: 9225980]
3. Butland SL, Devon RS, Huang Y, Mead CL, Meynert AM, Neal SJ, Lee SS, Wilkinson A, Yang GS, Yuen MMS, Hayden MR, Holt RA, Leavitt BR, Ouellette BFF. CAG-encoded polyglutamine length polymorphism in the human genome. Bmc Genomics. 2007; 8:18. [PubMed: 17229318]
4. Gerber HP, Seipel K, Georgiev O, Hofferer M, Hug M, Rusconi S, Schaffner W. Transcriptional activation modulated by homopolymeric glutamine and proline stretches. Science. 1994; 263:808–11. [PubMed: 8303297]
5. Wellner N, Belton PS, Tatham AS. Fourier transform IR spectroscopic study of hydration-induced structure changes in the solid state of omega-gliadins. Biochem J. 1996; 319(Pt 3):741–7. [PubMed: 8920975]
6. Ridgley DM, Ebanks KC, Barone JR. Peptide Mixtures Can Self-Assemble into Large Amyloid Fibers of Varying Size and Morphology. Biomacromolecules. 2011; 12:3770–3779. [PubMed: 21879764]
7. Zoghbi HY, Orr HT. Glutamine repeats and neurodegeneration. Annu Rev Neurosci. 2000; 23:217–47. [PubMed: 10845064]

8. Bates, GP.; Benn, C. The polyglutamine diseases. In: Bates, GP.; Harper, PS.; Jones, L., editors. Huntington's Disease. Oxford University Press; Oxford, U.K.: 2002. p. 429-472.
9. Wilburn B, Rudnicki DD, Zhao J, Weitz TM, Cheng Y, Gu XF, Greiner E, Park CS, Wang N, Sopher BL, La Spada AR, Osmand A, Margolis RL, Sun YE, Yang XW. An antisense CAG repeat transcript at JPH3 locus mediates expanded polyglutamine protein toxicity in Huntington's disease-like 2 mice. Neuron. 2011; 70:427–440. [PubMed: 21555070]
10. Zuccato C, Valenza M, Cattaneo E. Molecular mechanisms and potential therapeutical targets in Huntington's disease. Physiol Rev. 2010; 90:905–81. [PubMed: 20664076]
11. Biancalana V, Serville F, Pommier J, Julien J, Hanauer A, Mandel JL. Moderate instability of the trinucleotide repeat in spino bulbar muscular atrophy. Hum Mol Genet. 1992; 1:255–8. [PubMed: 1303195]
12. DiFiglia M, Sapp E, Chase KO, Davies SW, Bates GP, Vonsattel JP, Aronin N. Aggregation of huntingtin in neuronal intranuclear inclusions and dystrophic neurites in brain. Science. 1997; 277:1990–3. [PubMed: 9302293]
13. Davies SW, Turmaine M, Cozens BA, DiFiglia M, Sharp AH, Ross CA, Scherzinger E, Wanker EE, Mangiarini L, Bates GP. Formation of neuronal intranuclear inclusions underlies the neurological dysfunction in mice transgenic for the HD mutation. Cell. 1997; 90:537–48. [PubMed: 9267033]
14. Paulson HL. Protein fate in neurodegenerative proteinopathies: polyglutamine diseases join the (mis)fold. Am J Hum Genet. 1999; 64:339–45. [PubMed: 9973270]
15. Neurath H, Greenstein JP, Putnam FW, Erickson JO. The chemistry of protein denaturation. Chemical Reviews. 1944; 34:157–265.
16. Colon W, Kelly JW. Partial denaturation of transthyretin is sufficient for amyloid fibril formation in vitro. Biochemistry. 1992; 31:8654–60. [PubMed: 1390650]
17. Hurle MR, Helms LR, Li L, Chan W, Wetzel R. A role for destabilizing amino acid replacements in light chain amyloidosis. Proc Natl Acad Sci (USA). 1994; 91:5446–5450. [PubMed: 8202506]
18. Wetzel R. Mutations and off-pathway aggregation. Trends in Biotechnology. 1994; 12:193–198. [PubMed: 7764903]
19. Altschuler EL, Hud NV, Mazrimas JA, Rupp B. Random coil conformation for extended polyglutamine stretches in aqueous soluble monomeric peptides. J Pept Res. 1997; 50:73–5. [PubMed: 9273890]
20. Chen S, Berthelier V, Yang W, Wetzel R. Polyglutamine aggregation behavior *in vitro* supports a recruitment mechanism of cytotoxicity. J Mol Biol. 2001; 311:173–182. [PubMed: 11469866]
21. Masino L, Kelly G, Leonard K, Trottier Y, Pastore A. Solution structure of polyglutamine tracts in GST-polyglutamine fusion proteins. FEBS Lett. 2002; 513:267–72. [PubMed: 11904162]
22. Klein FA, Pastore A, Masino L, Zeder-Lutz G, Nierengarten H, Oulad-Abdelghani M, Altschuh D, Mandel JL, Trottier Y. Pathogenic and non-pathogenic polyglutamine tracts have similar structural properties: towards a length-dependent toxicity gradient. J Mol Biol. 2007; 371:235–44. [PubMed: 17560603]
23. Chen S, Ferrone F, Wetzel R. Huntington's Disease age-of-onset linked to polyglutamine aggregation nucleation. Proc Natl Acad Sci USA. 2002; 99:11884–11889. [PubMed: 12186976]
24. Kim MW, Chelliah Y, Kim SW, Otwinowski Z, Bezprozvanny I. Secondary structure of Huntingtin amino-terminal region. Structure. 2009; 17:1205–12. [PubMed: 19748341]
25. Bhattacharyya A, Thakur AK, Chellgren VM, Thiagarajan G, Williams AD, Chellgren BW, Creamer TP, Wetzel R. Oligoproline effects on polyglutamine conformation and aggregation. J Mol Biol. 2006; 355:524–35. [PubMed: 16321399]
26. Khare SD, Ding F, Gwanmesia KN, Dokholyan NV. Molecular origin of polyglutamine aggregation in neurodegenerative diseases. PLoS Comput Biol. 2005; 1:230–5. [PubMed: 16158094]
27. Wang X, Vitalis A, Wyczalkowski MA, Pappu RV. Characterizing the conformational ensemble of monomeric polyglutamine. Proteins. 2006; 63:297–311. [PubMed: 16299774]
28. Leitgeb B, Kerenyi A, Bogar F, Paragi G, Penke B, Rakhely G. Studying the structural properties of polyalanine and polyglutamine peptides. J Mol Model. 2007; 13:1141–50. [PubMed: 17805586]

29. Vitalis A, Wang X, Pappu RV. Quantitative characterization of intrinsic disorder in polyglutamine: insights from analysis based on polymer theories. Biophys J. 2007; 93:1923–37. [PubMed: 17526581]

30. Nakano M, Watanabe H, Rothstein SM, Tanaka S. Comparative characterization of short monomeric polyglutamine peptides by replica exchange molecular dynamics simulation. J Phys Chem B. 2010; 114:7056–61. [PubMed: 20441177]

31. Lakhani VV, Ding F, Dokholyan NV. Polyglutamine induced misfolding of huntingtin exon1 is modulated by the flanking sequences. PLoS Comput Biol. 2010; 6:e1000772. [PubMed: 20442863]

32. Laghaei R, Mousseau N. Spontaneous formation of polyglutamine nanotubes with molecular dynamics simulations. Journal of Chemical Physics. 2010; 132:8.

33. Crick SL, Jayaraman M, Frieden C, Wetzel R, Pappu RV. Fluorescence correlation spectroscopy shows that monomeric polyglutamine molecules form collapsed structures in aqueous solutions. Proc Natl Acad Sci U S A. 2006; 103:16764–9. [PubMed: 17075061]

34. Walters RH, Murphy RM. Examining polyglutamine peptide length: a connection between collapsed conformations and increased aggregation. J Mol Biol. 2009; 393:978–92. [PubMed: 19699209]

35. Dougan L, Li JY, Badilla CL, Berne BJ, Fernandez JM. Single homopolypeptide chains collapse into mechanically rigid conformations. Proc Natl Acad Sci U S A. 2009; 106:12605–12610. [PubMed: 19549822]

36. Starikov EB, Lehrach H, Wanker EE. Folding of oligoglutamines: a theoretical approach based upon thermodynamics and molecular mechanics. J Biomol Struct Dyn. 1999; 17:409–27. [PubMed: 10636078]

37. Vitalis A, Wang X, Pappu RV. Atomistic simulations of the effects of polyglutamine chain length and solvent quality on conformational equilibria and spontaneous homodimerization. J Mol Biol. 2008; 384:279–97. [PubMed: 18824003]

38. Trottier Y, Lutz Y, Stevanin G, Imbert G, Devys D, Cancel G, Saudou F, Weber C, David G, Tora L, et al. Polyglutamine expansion as a pathological epitope in Huntington's disease and four dominant cerebellar ataxias. Nature. 1995; 378:403–6. [PubMed: 7477379]

39. Bennett MJ, Huey-Tubman KE, Herr AB, West AP, Ross SA, Bjorkman PJ. A linear lattice model for poly-glutamine in CAG expansion diseases. Proc Natl Acad Sci U S A. 2002; 99:11634–11639. [PubMed: 12193654]

40. Li P, Huey-Tubman KE, Gao T, Li X, West AP Jr, Bennett MJ, Bjorkman PJ. The structure of a polyQ-anti-polyQ complex reveals binding according to a linear lattice model. Nat Struct Mol Biol. 2007; 14:381–7. [PubMed: 17450152]

41. Miller J, Arrasate M, Brooks E, Libeu CP, Legleiter J, Hatters D, Curtis J, Cheung K, Krishnan P, Mitra S, Widjaja K, Shaby BA, Lotz GP, Newhouse Y, Mitchell EJ, Osmand A, Gray M, Thulasiramin V, Saudou F, Segal M, Yang XW, Masliah E, Thompson LM, Muchowski PJ, Weisgraber KH, Finkbeiner S. Identifying polyglutamine protein species in situ that best predict neurodegeneration. Nat Chem Biol. 2011; 7:925–34. [PubMed: 22037470]

42. Pejchal R, Gach JS, Brunel FM, Cardoso RM, Stanfield RL, Dawson PE, Burton DR, Zwick MB, Wilson IA. A conformational switch in human immunodeficiency virus gp41 revealed by the structures of overlapping epitopes recognized by neutralizing antibodies. J Virol. 2009; 83:8451–62. [PubMed: 19515770]

43. Dyson HJ, Wright PE. Coupling of folding and binding for unstructured proteins. Curr Opin Struct Biol. 2002; 12:54–60. [PubMed: 11839490]

44. Dyson HJ. Expanding the proteome: disordered and alternatively folded proteins. Quarterly Reviews of Biophysics. 2011; 44:467–518. [PubMed: 21729349]

45. Kar K, Jayaraman M, Sahoo B, Kodali R, Wetzel R. Critical nucleus size for disease-related polyglutamine aggregation is repeat-length dependent. Nat Struct Mol Biol. 2011; 18:328–36. [PubMed: 21317897]

46. Fiumara F, Fioriti L, Kandel ER, Hendrickson WA. Essential role of coiled coils for aggregation and activity of Q/N-rich prions and PolyQ proteins. Cell. 2010; 143:1121–35. [PubMed: 21183075]

47. Eliezer D. Biophysical characterization of intrinsically disordered proteins. Current Opinion in Structural Biology. 2009; 19:23–30. [PubMed: 19162471]
48. Thakur AK, Jayaraman M, Mishra R, Thakur M, Chellgren VM, Byeon IJ, Anjum DH, Kodali R, Creamer TP, Conway JF, Gronenborn AM, Wetzel R. Polyglutamine disruption of the huntingtin exon 1 N terminus triggers a complex aggregation mechanism. Nat Struct Mol Biol. 2009; 16:380–9. [PubMed: 19270701]
49. Atwal RS, Desmond CR, Caron N, Maiuri T, Xia JR, Sipione S, Truant R. Kinase inhibitors modulate huntingtin cell localization and toxicity. Nature Chemical Biology. 2011; 7:453–460.
50. Jayaraman M, Kodali R, Sahoo B, Thakur AK, Mayasundari A, Mishra R, Peterson CB, Wetzel R. Slow amyloid nucleation via α-helix-rich oligomeric intermediates in short polyglutamine–containing Huntingtin fragments. J Mol Biol. 201110.1016/j.jmb.2011.12.011
51. Takahashi Y, Okamoto Y, Popiel HA, Fujikake N, Toda T, Kinjo M, Nagai Y. Detection of polyglutamine protein oligomers in cells by fluorescence correlation spectroscopy. J Biol Chem. 2007; 282:24039–48. [PubMed: 17573338]
52. Ossato G, Digman MA, Aiken C, Lukacsovich T, Marsh JL, Gratton E. A two-step path to inclusion formation of huntingtin peptides revealed by number and brightness analysis. Biophys J. 2010; 98:3078–85. [PubMed: 20550921]
53. Steffan JS, Agrawal N, Pallos J, Rockabrand E, Trotman LC, Slepko N, Illes K, Lukacsovich T, Zhu YZ, Cattaneo E, Pandolfi PP, Thompson LM, Marsh JL. SUMO modification of Huntingtin and Huntington's disease pathology. Science. 2004; 304:100–4. [PubMed: 15064418]
54. Atwal RS, Xia J, Pinchev D, Taylor J, Epand RM, Truant R. Huntingtin has a membrane association signal that can modulate huntingtin aggregation, nuclear entry and toxicity. Hum Mol Genet. 2007; 16:2600–15. [PubMed: 17704510]
55. Rockabrand E, Slepko N, Pantalone A, Nukala VN, Kazantsev A, Marsh JL, Sullivan PG, Steffan JS, Sensi SL, Thompson LM. The first 17 amino acids of Huntingtin modulate its sub-cellular localization, aggregation and effects on calcium homeostasis. Hum Mol Genet. 2007; 16:61–77. [PubMed: 17135277]
56. Angeli S, Shao J, Diamond MI. F-actin binding regions on the androgen receptor and huntingtin increase aggregation and alter aggregate characteristics. PLoS One. 2010; 5:e9053. [PubMed: 20140226]
57. Thompson LM, Aiken CT, Kaltenbach LS, Agrawal N, Illes K, Khoshnan A, Martinez-Vincente M, Arrasate M, JG OS-R, Khashwji H, Lukacsovich T, Zhu YZ, Lau AL, Massey A, Hayden MR, Zeitlin SO, Finkbeiner S, Green KN, Laferla FM, Bates G, Huang L, Patterson PH, Lo DC, Cuervo AM, Marsh JL, Steffan JS. IKK phosphorylates Huntingtin and targets it for degradation by the proteasome and lysosome. J Cell Biol. 2009; 187:1083–1099. [PubMed: 20026656]
58. Gu X, Greiner ER, Mishra R, Kodali R, Osmand A, Finkbeiner S, Steffan JS, Thompson LM, Wetzel R, Yang XW. Serines 13 and 16 are critical determinants of full-length human mutant huntingtin induced disease pathogenesis in HD mice. Neuron. 2009; 64:828–40. [PubMed: 20064390]
59. Kelley NW, Huang X, Tam S, Spiess C, Frydman J, Pande VS. The predicted structure of the headpiece of the Huntingtin protein and its implications on Huntingtin aggregation. J Mol Biol. 2009; 388:919–27. [PubMed: 19361448]
60. Rossetti G, Cossio P, Laio A, Carloni P. Conformations of the Huntingtin N-term in aqueous solution from atomistic simulations. Febs Letters. 2011; 585:3086–3089. [PubMed: 21889504]
61. Williamson TE, Vitalis A, Crick SL, Pappu RV. Modulation of polyglutamine conformations and dimer formation by the N-terminus of huntingtin. J Mol Biol. 2010; 396:1295–309. [PubMed: 20026071]
62. Dlugosz M, Trylska J. Secondary Structures of Native and Pathogenic Huntingtin N-Terminal Fragments. Journal of Physical Chemistry B. 2011; 115:11597–11608.
63. Stott K, Blackburn JM, Butler PJ, Perutz M. Incorporation of glutamine repeats makes protein oligomerize: implications for neurodegenerative diseases. Proc Natl Acad Sci U S A. 1995; 92:6509–13. [PubMed: 7604023]

64. Chen YW, Stott K, Perutz MF. Crystal structure of a dimeric chymotrypsin inhibitor 2 mutant containing an inserted glutamine repeat. Proc Natl Acad Sci U S A. 1999; 96:1257–61. [PubMed: 9990011]

65. Tanaka M, Morishima I, Akagi T, Hashikawa T, Nukina N. Intra- and intermolecular beta-pleated sheet formation in glutamine-repeat inserted myoglobin as a model for polyglutamine diseases. J Biol Chem. 2001; 276:45470–5. [PubMed: 11584007]

66. Robertson AL, Horne J, Ellisdon AM, Thomas B, Scanlon MJ, Bottomley SP. The structural impact of a polyglutamine tract is location-dependent. Biophys J. 2008; 95:5922–30. [PubMed: 18849414]

67. Tobelmann MD, Murphy RM. Location Trumps Length: Polyglutamine-Mediated Changes in Folding and Aggregation of a Host Protein. Biophysical Journal. 2011; 100:2773–2782. [PubMed: 21641323]

68. Nagai Y, Inui T, Popiel HA, Fujikake N, Hasegawa K, Urade Y, Goto Y, Naiki H, Toda T. A toxic monomeric conformer of the polyglutamine protein. Nat Struct Mol Biol. 2007; 14:332–40. [PubMed: 17369839]

69. Lamb DJ, Puxeddu E, Malik N, Stenoien DL, Nigam R, Saleh GY, Mancini M, Weigel NL, Marcelli M. Molecular analysis of the androgen receptor in ten prostate cancer specimens obtained before and after androgen ablation. J Androl. 2003; 24:215–25. [PubMed: 12634308]

70. Wang Q, Udayakumar TS, Vasaitis TS, Brodie AM, Fondell JD. Mechanistic relationship between androgen receptor polyglutamine tract truncation and androgen-dependent transcriptional hyperactivity in prostate cancer cells. J Biol Chem. 2004; 279:17319–28. [PubMed: 14966121]

71. Cattaneo E, Rigamonti D, Goffredo D, Zuccato C, Squitieri F, Sipione S. Loss of normal huntingtin function: new developments in Huntington's disease research. Trends Neurosci. 2001; 24:182–8. [PubMed: 11182459]

72. Pearson CE, Sinden RR. Trinucleotide repeat DNA structures: dynamic mutations from dynamic DNA. Curr Opin Struct Biol. 1998; 8:321–30. [PubMed: 9666328]

73. Perutz MF, Johnson T, Suzuki M, Finch JT. Glutamine repeats as polar zippers: their possible role in inherited neurodegenerative diseases. Proc Natl Acad Sci U S A. 1994; 91:5355–8. [PubMed: 8202492]

74. Thakur A, Wetzel R. Mutational analysis of the structural organization of polyglutamine aggregates. Proc Natl Acad Sci U S A. 2002; 99:17014–17019. [PubMed: 12444250]

75. Bhattacharyya AM, Thakur AK, Wetzel R. polyglutamine aggregation nucleation: thermodynamics of a highly unfavorable protein folding reaction. Proc Natl Acad Sci U S A. 2005; 102:15400–5. [PubMed: 16230628]

76. Slepko N, Bhattacharyya AM, Jackson GR, Steffan JS, Marsh JL, Thompson LM, Wetzel R. Normal-repeat-length polyglutamine peptides accelerate aggregation nucleation and cytotoxicity of expanded polyglutamine proteins. Proc Natl Acad Sci U S A. 2006; 103:14367–72. [PubMed: 16980414]

77. Lee CC, Walters RH, Murphy RM. Reconsidering the mechanism of polyglutamine Peptide aggregation. Biochemistry. 2007; 46:12810–20. [PubMed: 17929830]

78. Jayaraman M, Kodali R, Wetzel R. The impact of ataxin-1-like histidine insertions on polyglutamine aggregation. Protein Eng Des Sel. 2009; 22:469–78. [PubMed: 19541676]

79. Walters RH, Murphy RM. Aggregation kinetics of interrupted polyglutamine peptides. J Mol Biol. 2011; 412:505–19. [PubMed: 21821045]

80. Sharma D, Sharma S, Pasha S, Brahmachari SK. Peptide models for inherited neurodegenerative disorders: conformation and aggregation properties of long polyglutamine peptides with and without interruptions. FEBS Lett. 1999; 456:181–5. [PubMed: 10452554]

81. Chen S, Berthelier V, Hamilton JB, O'Nuallain B, Wetzel R. Amyloid-like features of polyglutamine aggregates and their assembly kinetics. Biochemistry. 2002; 41:7391–9. [PubMed: 12044172]

82. Scherzinger E, Sittler A, Schweiger K, Heiser V, Lurz R, Hasenbank R, Bates GP, Lehrach H, Wanker EE. Self-assembly of polyglutamine-containing huntingtin fragments into amyloid-like fibrils: implications for Huntington's disease pathology. Proc Natl Acad Sci U S A. 1999; 96:4604–9. [PubMed: 10200309]

83. Morley JF, Brignull HR, Weyers JJ, Morimoto RI. The threshold for polyglutamine-expansion protein aggregation and cellular toxicity is dynamic and influenced by aging in Caenorhabditis elegans. Proc Natl Acad Sci U S A. 2002; 99:10417–22. [PubMed: 12122205]
84. McCampbell A, Fischbeck KH. Polyglutamine and CBP: fatal attraction? Nat Med. 2001; 7:528–30. [PubMed: 11329046]
85. Ferrone F. Analysis of protein aggregation kinetics. Meths Enzymol. 1999; 309:256–274.
86. Smith MH, Miles TF, Sheehan M, Alfieri KN, Kokona B, Fairman R. Polyglutamine fibrils are formed using a simple designed beta-hairpin model. Proteins. 2010; 78:1971–9. [PubMed: 20408173]
87. VanSchouwen BM, Oblinsky DG, Gordon HL, Rothstein SM. Structure propensities in mutated polyglutamine peptides. Interdiscip Sci. 2011; 3:1–16. [PubMed: 21369882]
88. VanSchouwen BMB, Nakano M, Watanabe H, Tanaka S, Gordon HL, Rothstein SM. Molecular mechanics and all-electron fragment molecular orbital calculations on mutated polyglutamine peptides. Journal of Molecular Structure-Theochem. 2010; 944:12–20.
89. Esler WP, Stimson ER, Jennings JM, Vinters HV, Ghilardi JR, Lee JP, Mantyh PW, Maggio JE. Alzheimer's disease amyloid propagation by a template-dependent dock-lock mechanism. Biochemistry. 2000; 39:6288–95. [PubMed: 10828941]
90. Pfeil W. The problem of the stability globular proteins. Mol Cell Biochem. 1981; 40:3–28. [PubMed: 7031463]
91. Marchut AJ, Hall CK. Effects of chain length on the aggregation of model polyglutamine peptides: molecular dynamics simulations. Proteins. 2007; 66:96–109. [PubMed: 17068817]
92. Demuro A, Mina E, Kayed R, Milton SC, Parker I, Glabe CG. Calcium dysregulation and membrane disruption as a ubiquitous neurotoxic mechanism of soluble amyloid oligomers. J Biol Chem. 2005; 280:17294–300. [PubMed: 15722360]
93. Legleiter J, Mitchell E, Lotz GP, Sapp E, Ng C, DiFiglia M, Thompson LM, Muchowski PJ. Mutant huntingtin fragments form oligomers in a polyglutamine length-dependent manner in vitro and in vivo. J Biol Chem. 2010; 285:14777–90. [PubMed: 20220138]
94. Vitalis A, Pappu RV. Assessing the contribution of heterogeneous distributions of oligomers to aggregation mechanisms of polyglutamine peptides. Biophysical Chemistry. 2011; 159:14–23. [PubMed: 21530061]
95. Kodali R, Wetzel R. Polymorphism in the intermediates and products of amyloid assembly. Curr Opin Struct Biol. 2007; 17:48–57. [PubMed: 17251001]
96. O'Nuallain B, Thakur AK, Williams AD, Bhattacharyya AM, Chen S, Thiagarajan G, Wetzel R. Kinetics and thermodynamics of amyloid assembly using a high-performance liquid chromatography-based sedimentation assay. Methods Enzymol. 2006; 413:34–74. [PubMed: 17046390]
97. Sen S, Dash D, Pasha S, Brahmachari SK. Role of histidine interruption in mitigating the pathological effects of long polyglutamine stretches in SCA1: A molecular approach. Protein Sci. 2003; 12:953–62. [PubMed: 12717018]
98. Thakur AK, Yang W, Wetzel R. Inhibition of polyglutamine aggregate cytotoxicity by a structure-based elongation inhibitor. FASEB J. 2004; 18:923–5. [PubMed: 15001566]
99. Popiel HA, Nagai Y, Onodera O, Inui T, Fujikake N, Urade Y, Strittmatter WJ, Burke JR, Ichikawa A, Toda T. Disruption of the toxic conformation of the expanded polyglutamine stretch leads to suppression of aggregate formation and cytotoxicity. Biochem Biophys Res Commun. 2004; 317:1200–6. [PubMed: 15094397]
100. Quan F, Janas J, Popovich BW. A novel CAG repeat configuration in the SCA1 gene: implications for the molecular diagnostics of spinocerebellar ataxia type 1. Hum Mol Genet. 1995; 4:2411–3. [PubMed: 8634720]
101. Arango M, Holbert S, Zala D, Brouillet E, Pearson J, Regulier E, Thakur AK, Aebischer P, Wetzel R, Deglon N, Neri C. CA150 expression delays striatal cell death in overexpression and knock-in conditions for mutant huntingtin neurotoxicity. J Neurosci. 2006; 26:4649–59. [PubMed: 16641246]
102. O'Nuallain B, Williams AD, Westermark P, Wetzel R. Seeding specificity in amyloid growth induced by heterologous fibrils. J Biol Chem. 2004; 279:17490–17499. [PubMed: 14752113]

103. Scheuermann T, Schulz B, Blume A, Wahle E, Rudolph R, Schwarz E. Trinucleotide expansions leading to an extended poly-L-alanine segment in the poly (A) binding protein PABPN1 cause fibril formation. Protein Sci. 2003; 12:2685–92. [PubMed: 14627730]
104. Scherzinger E, Lurz R, Turmaine M, Mangiarini L, Hollenbach B, Hasenbank R, Bates GP, Davies SW, Lehrach H, Wanker EE. Huntingtin-encoded polyglutamine expansions form amyloid-like protein aggregates in vitro and in vivo. Cell. 1997; 90:549–58. [PubMed: 9267034]
105. Ignatova Z, Gierasch LM. Extended polyglutamine tracts cause aggregation and structural perturbation of an adjacent beta barrel protein. J Biol Chem. 2006; 281:12959–67. [PubMed: 16524881]
106. Bulone D, Masino L, Thomas DJ, San Biagio PL, Pastore A. The Interplay between PolyQ and Protein Context Delays Aggregation by Forming a Reservoir of Protofibrils. PLoS ONE. 2006; 1:e111. [PubMed: 17205115]
107. Masino L, Nicastro G, Menon RP, Dal Piaz F, Calder L, Pastore A. Characterization of the structure and the amyloidogenic properties of the Josephin domain of the polyglutamine-containing protein ataxin-3. J Mol Biol. 2004; 344:1021–35. [PubMed: 15544810]
108. Ellisdon AM, Thomas B, Bottomley SP. The two-stage pathway of ataxin-3 fibrillogenesis involves a polyglutamine-independent step. J Biol Chem. 2006; 281:16888–96. [PubMed: 16624810]
109. Bevivino AE, Loll PJ. An expanded glutamine repeat destabilizes native ataxin-3 structure and mediates formation of parallel beta -fibrils. Proc Natl Acad Sci U S A. 2001; 98:11955–60. [PubMed: 11572942]
110. Chow MK, Ellisdon AM, Cabrita LD, Bottomley SP. Polyglutamine expansion in ataxin-3 does not affect protein stability: implications for misfolding and disease. J Biol Chem. 2004; 279:47643–51. [PubMed: 15345714]
111. Chow MK, Paulson HL, Bottomley SP. Destabilization of a non-pathological variant of ataxin-3 results in fibrillogenesis via a partially folded intermediate: a model for misfolding in polyglutamine disease. J Mol Biol. 2004; 335:333–41. [PubMed: 14659761]
112. Santambrogio C, Frana AM, Natalello A, Papaleo E, Regonesi ME, Doglia SM, Tortora P, Invernizzi G, Grandori R. The role of the central flexible region on the aggregation and conformational properties of human ataxin-3. FEBS J. 201110.1111/j.1742–4658.2011.08438.x
113. Saunders HM, Gilis D, Rooman M, Dehouck Y, Robertson AL, Bottomley SP. Flanking domain stability modulates the aggregation kinetics of a polyglutamine disease protein. Protein Science. 2011; 20:1675–1681. [PubMed: 21780213]
114. Masino L, Nicastro G, Calder L, Vendruscolo M, Pastore A. Functional interactions as a survival strategy against abnormal aggregation. FASEB J. 2011; 25:45–54. [PubMed: 20810784]
115. Robertson AL, Headey SJ, Saunders HM, Ecroyd H, Scanlon MJ, Carver JA, Bottomley SP. Small heat-shock proteins interact with a flanking domain to suppress polyglutamine aggregation. Proceedings of the National Academy of Sciences of the United States of America. 2010; 107:10424–10429. [PubMed: 20484674]
116. Walker FO. Huntington's disease. Lancet. 2007; 369:218–28. [PubMed: 17240289]
117. Poirier MA, Li H, Macosko J, Cai S, Amzel M, Ross CA. Huntingtin spheroids and protofibrils as precursors in polyglutamine fibrilization. J Biol Chem. 2002; 277:41032–7. [PubMed: 12171927]
118. Apostol BL, Kazantsev A, Raffioni S, Illes K, Pallos J, Bodai L, Slepko N, Bear JE, Gertler FB, Hersch S, Housman DE, Marsh JL, Thompson LM. A cell-based assay for aggregation inhibitors as therapeutics of polyglutamine-repeat disease and validation in Drosophila. Proc Natl Acad Sci U S A. 2003; 100:5950–5. [PubMed: 12730384]
119. Duennwald ML, Jagadish S, Muchowski PJ, Lindquist S. Flanking sequences profoundly alter polyglutamine toxicity in yeast. Proc Natl Acad Sci U S A. 2006; 103:11045–50. [PubMed: 16832050]
120. Darnell G, Orgel JP, Pahl R, Meredith SC. Flanking Polyproline Sequences Inhibit beta-Sheet Structure in Polyglutamine Segments by Inducing PPII-like Helix Structure. J Mol Biol. 2007; 374:688–704. [PubMed: 17945257]

121. Darnell GD, Derryberry J, Kurutz JW, Meredith SC. Mechanism of Cis-Inhibition of PolyQ Fibrillation by PolyP: PPII Oligomers and the Hydrophobic Effect. Biophysical Journal. 2009; 97:2295–2305. [PubMed: 19843462]
122. Sivanandam VN, Jayaraman M, Hoop CL, Kodali R, Wetzel R, van der Wel PC. The aggregation-enhancing huntingtin N-terminus is helical in amyloid fibrils. J Am Chem Soc. 2011; 133:4558–66. [PubMed: 21381744]
123. Abedini A, Raleigh DP. A role for helical intermediates in amyloid formation by natively unfolded polypeptides? Phys Biol. 2009; 6:015005. [PubMed: 19208933]
124. Singer D, Zauner T, Genz M, Hoffmann R, Zuchner T. Synthesis of pathological and nonpathological human exon 1 huntingtin. Journal of Peptide Science. 2010; 16:358–363. [PubMed: 20552561]
125. Tam S, Spiess C, Auyeung W, Joachimiak L, Chen B, Poirier MA, Frydman J. The chaperonin TRiC blocks a huntingtin sequence element that promotes the conformational switch to aggregation. Nat Struct Mol Biol. 2009; 16:1279–85. [PubMed: 19915590]
126. Wacker JL, Zareie MH, Fong H, Sarikaya M, Muchowski PJ. Hsp70 and Hsp40 attenuate formation of spherical and annular polyglutamine oligomers by partitioning monomer. Nat Struct Mol Biol. 2004; 11:1215–22. [PubMed: 15543156]
127. Broncel M, Falenski JA, Wagner SC, Hackenberger CP, Koksch B. How post-translational modifications influence amyloid formation: a systematic study of phosphorylation and glycosylation in model peptides. Chemistry. 2010; 16:7881–8. [PubMed: 20491120]
128. Nekooki-Machida Y, Kurosawa M, Nukina N, Ito K, Oda T, Tanaka M. Distinct conformations of in vitro and in vivo amyloids of huntingtin-exon1 show different cytotoxicity. Proc Natl Acad Sci U S A. 2009; 106:9679–84. [PubMed: 19487684]
129. Serio TR, Cashikar AG, Kowal AS, Sawicki GJ, Moslehi JJ, Serpell L, Arnsdorf MF, Lindquist SL. Nucleated conformational conversion and the replication of conformational information by a prion determinant. Science. 2000; 289:1317–21. [PubMed: 10958771]
130. Powers, ET.; Ferrone, F. Kinetic models for protein misfolding and association. In: Dobson, CM.; Kelly, JW.; Ramirez-Alvarado, M., editors. Protein Misfolding Diseases: Current and Emerging Principles and Therapies. Wiley; New York: 2009. p. 73-92.
131. Nagai Y, Tucker T, Ren H, Kenan DJ, Henderson BS, Keene JD, Strittmatter WJ, Burke JR. Inhibition of polyglutamine protein aggregation and cell death by novel peptides identified by phage display screening. J Biol Chem. 2000; 275:10437–42. [PubMed: 10744733]
132. Tomita K, Popiel HA, Nagai Y, Toda T, Yoshimitsu Y, Ohno H, Oishi S, Fujii N. Structure-activity relationship study on polyglutamine binding peptide QBP1. Bioorg Med Chem. 2009; 17:1259–63. [PubMed: 19121945]
133. Hamuro L, Zhang G, Tucker TJ, Self C, Strittmatter WJ, Burke JR. Optimization of a polyglutamine aggregation inhibitor peptide (QBP1) using a thioflavin T fluorescence assay. Assay Drug Dev Technol. 2007; 5:629–36. [PubMed: 17939755]
134. Popiel HA, Nagai Y, Fujikake N, Toda T. Protein transduction domain-mediated delivery of QBP1 suppresses polyglutamine-induced neurodegeneration in vivo. Mol Ther. 2007; 15:303–9. [PubMed: 17235308]
135. Nagai Y, Fujikake N, Ohno K, Higashiyama H, Popiel HA, Rahadian J, Yamaguchi M, Strittmatter WJ, Burke JR, Toda T. Prevention of polyglutamine oligomerization and neurodegeneration by the peptide inhibitor QBP1 in Drosophila. Hum Mol Genet. 2003; 12:1253–9. [PubMed: 12761040]
136. Richardson, JS.; Richardson, DC. Principles and patterns of protein conformation. In: Fasman, GD., editor. Prediction of Protein Structure and the Principles of Protein Conformation. Plenum; New York: 1989. p. 1-98.
137. Lanning JD, Hawk AJ, Derryberry J, Meredith SC. Chaperone-like N-methyl peptide inhibitors of polyglutamine aggregation. Biochemistry. 2010; 49:7108–18. [PubMed: 20583779]
138. Heiser V, Scherzinger E, Boeddrich A, Nordhoff E, Lurz R, Schugardt N, Lehrach H, Wanker EE. Inhibition of huntingtin fibrillogenesis by specific antibodies and small molecules: implications for Huntington's disease therapy. Proc Natl Acad Sci U S A. 2000; 97:6739–44. [PubMed: 10829068]

139. Heiser V, Engemann S, Brocker W, Dunkel I, Boeddrich A, Waelter S, Nordhoff E, Lurz R, Schugardt N, Rautenberg S, Herhaus C, Barnickel G, Bottcher H, Lehrach H, Wanker EE. Identification of benzothiazoles as potential polyglutamine aggregation inhibitors of Huntington's disease by using an automated filter retardation assay. Proc Natl Acad Sci U S A. 2002; 99(Suppl 4):16400–6. [PubMed: 12200548]

140. Ehrnhoefer DE, Duennwald M, Markovic P, Wacker JL, Engemann S, Roark M, Legleiter J, Marsh JL, Thompson LM, Lindquist S, Muchowski PJ, Wanker EE. Green tea (−)-epigallocatechin-gallate modulates early events in huntingtin misfolding and reduces toxicity in Huntington's disease models. Hum Mol Genet. 2006; 15:2743–51. [PubMed: 16893904]

141. Ehrnhoefer DE, Bieschke J, Boeddrich A, Herbst M, Masino L, Lurz R, Engemann S, Pastore A, Wanker EE. EGCG redirects amyloidogenic polypeptides into unstructured, off-pathway oligomers. Nat Struct Mol Biol. 2008; 15:558–66. [PubMed: 18511942]

142. Hockly E, Tse J, Barker AL, Moolman DL, Beunard JL, Revington AP, Holt K, Sunshine S, Moffitt H, Sathasivam K, Woodman B, Wanker EE, Lowden PA, Bates GP. Evaluation of the benzothiazole aggregation inhibitors riluzole and PGL-135 as therapeutics for Huntington's disease. Neurobiol Dis. 2006; 21:228–36. [PubMed: 16111888]

143. Mishra R, Jayaraman M, Roland BP, Landrum E, Fullam R, Kodali R, Thakur AK, Arduini I, Wetzel R. Inhibiting nucleation of amyloid structure in a huntingtin fragment by targeting α-helix rich oligomeric intermediates. J Mol Biol. 201110.1016/j.jmb.2011.12.010

144. Cao P, Meng F, Abedini A, Raleigh DP. The ability of rodent islet amyloid polypeptide to inhibit amyloid formation by human islet amyloid polypeptide has important implications for the mechanism of amyloid formation and the design of inhibitors. Biochemistry. 2010; 49:872–81. [PubMed: 20028124]

145. Schiefner A, Chatwell L, Korner J, Neumaier I, Colby DW, Volkmer R, Wittrup KD, Skerra A. A Disulfide-Free Single-Domain V(L) Intrabody with Blocking Activity towards Huntingtin Reveals a Novel Mode of Epitope Recognition. J Mol Biol. 2011; 414:337–55. [PubMed: 21968397]

146. Carmichael J, Chatellier J, Woolfson A, Milstein C, Fersht AR, Rubinsztein DC. Bacterial and yeast chaperones reduce both aggregate formation and cell death in mammalian cell models of Huntington's disease. Proc Natl Acad Sci U S A. 2000; 97:9701–5. [PubMed: 10920207]

147. Muchowski PJ, Schaffar G, Sittler A, Wanker EE, Hayer-Hartl MK, Hartl FU. Hsp70 and hsp40 chaperones can inhibit self-assembly of polyglutamine proteins into amyloid-like fibrils. Proc Natl Acad Sci U S A. 2000; 97:7841–6. [PubMed: 10859365]

148. Hay DG, Sathasivam K, Tobaben S, Stahl B, Marber M, Mestril R, Mahal A, Smith DL, Woodman B, Bates GP. Progressive decrease in chaperone protein levels in a mouse model of Huntington's disease and induction of stress proteins as a therapeutic approach. Hum Mol Genet. 2004; 13:1389–405. [PubMed: 15115766]

149. Gutekunst CA, Li SH, Yi H, Mulroy JS, Kuemmerle S, Jones R, Rye D, Ferrante RJ, Hersch SM, Li XJ. Nuclear and neuropil aggregates in Huntington's disease: relationship to neuropathology. J Neurosci. 1999; 19:2522–34. [PubMed: 10087066]

150. Li H, Li SH, Cheng AL, Mangiarini L, Bates GP, Li XJ. Ultrastructural localization and progressive formation of neuropil aggregates in Huntington's disease transgenic mice. Hum Mol Genet. 1999; 8:1227–36. [PubMed: 10369868]

151. Sathasivam K, Lane A, Legleiter J, Warley A, Woodman B, Finkbeiner S, Paganetti P, Muchowski PJ, Wilson S, Bates GP. Identical oligomeric and fibrillar structures captured from the brains of R6/2 and knock-in mouse models of Huntington's disease. Hum Mol Genet. 2010; 19:65–78. [PubMed: 19825844]

152. Osmand AP, Berthelier V, Wetzel R. Imaging polyglutamine deposits in brain tissue. Methods Enzymol. 2006; 412:106–22. [PubMed: 17046655]

153. Moulder KL, Onodera O, Burke JR, Strittmatter WJ, Johnson EM Jr. Generation of neuronal intranuclear inclusions by polyglutamine-GFP: analysis of inclusion clearance and toxicity as a function of polyglutamine length. J Neurosci. 1999; 19:705–15. [PubMed: 9880591]

154. Colby DW, Cassady JP, Lin GC, Ingram VM, Wittrup KD. Stochastic kinetics of intracellular huntingtin aggregate formation. Nat Chem Biol. 2006; 2:319–23. [PubMed: 16699519]

155. Waelter S, Boeddrich A, Lurz R, Scherzinger E, Lueder G, Lehrach H, Wanker EE. Accumulation of mutant huntingtin fragments in aggresome-like inclusion bodies as a result of insufficient protein degradation. Mol Biol Cell. 2001; 12:1393–407. [PubMed: 11359930]

156. Ross CA, Poirier MA. Opinion: What is the role of protein aggregation in neurodegeneration? Nat Rev Mol Cell Biol. 2005; 6:891–8. [PubMed: 16167052]

157. Olshina MA, Angley LM, Ramdzan YM, Tang J, Bailey MF, Hill AF, Hatters DM. Tracking mutant huntingtin aggregation kinetics in cells reveals three major populations that include an invariant oligomer pool. J Biol Chem. 2010; 285:21807–16. [PubMed: 20444706]

158. Chai Y, Shao J, Miller VM, Williams A, Paulson HL. Live-cell imaging reveals divergent intracellular dynamics of polyglutamine disease proteins and supports a sequestration model of pathogenesis. Proc Natl Acad Sci U S A. 2002; 99:9310–5. [PubMed: 12084819]

159. Stenoien DL, Mielke M, Mancini MA. Intranuclear ataxin1 inclusions contain both fast- and slow-exchanging components. Nat Cell Biol. 2002; 4:806–10. [PubMed: 12360291]

160. Lajoie P, Snapp EL. Formation and Toxicity of Soluble Polyglutamine Oligomers in Living Cells. PLoS ONE. 2010; 5:e15245. [PubMed: 21209946]

161. Cummings CJ, Mancini MA, Antalffy B, DeFranco DB, Orr HT, Zoghbi HY. Chaperone suppression of aggregation and altered subcellular proteasome localization imply protein misfolding in SCA1. Nat Genet. 1998; 19:148–54. [PubMed: 9620770]

162. Nucifora FC Jr, Sasaki M, Peters MF, Huang H, Cooper JK, Yamada M, Takahashi H, Tsuji S, Troncoso J, Dawson VL, Dawson TM, Ross CA. Interference by huntingtin and atrophin-1 with cbp-mediated transcription leading to cellular toxicity. Science. 2001; 291:2423–8. [PubMed: 11264541]

163. Jiang H, Nucifora FC Jr, Ross CA, DeFranco DB. Cell death triggered by polyglutamine-expanded huntingtin in a neuronal cell line is associated with degradation of CREB-binding protein. Hum Mol Genet. 2003; 12:1–12. [PubMed: 12490527]

164. Engelender S, Sharp AH, Colomer V, Tokito MK, Lanahan A, Worley P, Holzbaur ELF, Ross CA. Huntingtin-associated protein 1 (HAP1) interacts with the p150Glued subunit of dynactin. Hum Mol Genet. 1997; 6:2205–12. [PubMed: 9361024]

165. Boutell JM, Thomas P, Neal JW, Weston VJ, Duce J, Harper PS, Jones AL. Aberrant interactions of transcriptional repressor proteins with the Huntington's disease gene product, huntingtin. Hum Mol Genet. 1999; 8:1647–55. [PubMed: 10441327]

166. Davranche A, Aviolat H, Zeder-Lutz G, Busso D, Altschuh D, Trottier Y, Klein FAC. Huntingtin affinity for partners is not changed by polyglutamine length: aggregation itself triggers aberrant interactions. Human Molecular Genetics. 2011; 20:2795–2806. [PubMed: 21518730]

167. Perutz MF, Finch JT, Berriman J, Lesk A. Amyloid fibers are water-filled nanotubes. Proc Natl Acad Sci U S A. 2002; 99:5591–5. [PubMed: 11960014]

168. Sharma D, Shinchuk LM, Inouye H, Wetzel R, Kirschner DA. Polyglutamine homopolymers having 8–45 residues form slablike beta-crystallite assemblies. Proteins. 2005; 61:398–411. [PubMed: 16114051]

169. Sikorski P, Atkins E. New model for crystalline polyglutamine assemblies and their connection with amyloid fibrils. Biomacromolecules. 2005; 6:425–32. [PubMed: 15638548]

170. Schneider R, Schumacher MC, Mueller H, Nand D, Klaukien V, Heise H, Riedel D, Wolf G, Behrmann E, Raunser S, Seidel R, Engelhard M, Baldus M. Structural characterization of polyglutamine fibrils by solid-state NMR spectroscopy. J Mol Biol. 2011; 412:121–36. [PubMed: 21763317]

171. Rossetti G, Magistrato A, Pastore A, Persichetti F, Carloni P. Structural Properties of Polyglutamine Aggregates Investigated via Molecular Dynamics Simulations. Journal of Physical Chemistry B. 2008; 112:16843–16850.

172. Ogawa H, Nakano M, Watanabe H, Starikov EB, Rothstein SM, Tanaka S. Molecular dynamics simulation study on the structural stabilities of polyglutamine peptides. Comput Biol Chem. 2008; 32:102–10. [PubMed: 18243803]

173. Zhou ZL, Zhao JH, Liu HL, Wu JW, Liu KT, Chuang CK, Tsai WB, Ho Y. The Possible Structural Models for Polyglutamine Aggregation: A Molecular Dynamics Simulations Study. Journal of Biomolecular Structure & Dynamics. 2011; 28:743–758. [PubMed: 21294586]

174. Kajava AV, Baxa U, Steven AC. Beta arcades: recurring motifs in naturally occurring and disease-related amyloid fibrils. FASEB J. 2010; 24:1311–9. [PubMed: 20032312]
175. Esposito L, Paladino A, Pedone C, Vitagliano L. Insights into structure, stability, and toxicity of monomeric and aggregated polyglutamine models from molecular dynamics simulations. Biophys J. 2008; 94:4031–40. [PubMed: 18234827]
176. Zhang QC, Yeh TL, Leyva A, Frank LG, Miller J, Kim YE, Langen R, Finkbeiner S, Amzel ML, Ross CA, Poirier MA. A compact beta model of huntingtin toxicity. J Biol Chem. 2011; 286:8188–96. [PubMed: 21209075]
177. O'Nuallain B, Shivaprasad S, Kheterpal I, Wetzel R. Thermodynamics of abeta(1–40) amyloid fibril elongation. Biochemistry. 2005; 44:12709–18. [PubMed: 16171385]
178. van der Wel PC, Lewandowski JR, Griffin RG. Solid-state NMR study of amyloid nanocrystals and fibrils formed by the peptide GNNQQNY from yeast prion protein Sup35p. J Am Chem Soc. 2007; 129:5117–30. [PubMed: 17397156]
179. Nelson R, Sawaya MR, Balbirnie M, Madsen AO, Riekel C, Grothe R, Eisenberg D. Structure of the cross-beta spine of amyloid-like fibrils. Nature. 2005; 435:773–8. [PubMed: 15944695]
180. Sawaya MR, Sambashivan S, Nelson R, Ivanova MI, Sievers SA, Apostol MI, Thompson MJ, Balbirnie M, Wiltzius JJ, McFarlane HT, Madsen AO, Riekel C, Eisenberg D. Atomic structures of amyloid cross-beta spines reveal varied steric zippers. Nature. 2007; 447:453–7. [PubMed: 17468747]
181. Berthelier V, Hamilton JB, Chen S, Wetzel R. A microtiter plate assay for polyglutamine aggregate extension. Anal Biochem. 2001; 295:227–36. [PubMed: 11488626]
182. Yang W, Dunlap JR, Andrews RB, Wetzel R. Aggregated polyglutamine peptides delivered to nuclei are toxic to mammalian cells. Hum Mol Genet. 2002; 11:2905–2917. [PubMed: 12393802]
183. Ren PH, Lauckner JE, Kachirskaia I, Heuser JE, Melki R, Kopito RR. Cytoplasmic penetration and persistent infection of mammalian cells by polyglutamine aggregates. Nature Cell Biology. 2009; 11:219–U232.
184. Franks, F. Biophysics and biochemistry at low temperatures. Cambridge University Press; Cambridge: 1985.
185. Franks, F. Storage stabilization of proteins. In: Franks, F., editor. Protein Biotechnology: Isolation, Characterization and Stabilization. Humana Press; New York: 1993. p. 489-531.
186. Chen S, Wetzel R. Solubilization and disaggregation of polyglutamine peptides. Protein Science. 2001; 10:887–891. [PubMed: 11274480]
187. Jayaraman M, Thakur AK, Kar K, Kodali R, Wetzel R. Assays for studying nucleated aggregation of polyglutamine proteins. Methods. 2011; 53:246–54. [PubMed: 21232603]
188. Padrick SB, Miranker AD. Islet amyloid: phase partitioning and secondary nucleation are central to the mechanism of fibrillogenesis. Biochemistry. 2002; 41:4694–703. [PubMed: 11926832]
189. Nichols MR, Moss MA, Reed DK, Cratic-McDaniel S, Hoh JH, Rosenberry TL. Amyloid-beta protofibrils differ from amyloid-beta aggregates induced in dilute hexafluoroisopropanol in stability and morphology. J Biol Chem. 2005; 280:2471–80. [PubMed: 15528204]
190. Sugase K, Dyson HJ, Wright PE. Mechanism of coupled folding and binding of an intrinsically disordered protein. Nature. 2007; 447:1021–5. [PubMed: 17522630]
191. Palidwor GA, Shcherbinin S, Huska MR, Rasko T, Stelzl U, Arumughan A, Foulle R, Porras P, Sanchez-Pulido L, Wanker EE, Andrade-Navarro MA. Detection of alpha-rod protein repeats using a neural network and application to huntingtin. PLoS Comput Biol. 2009; 5:e1000304. [PubMed: 19282972]
192. Williams AD, Portelius E, Kheterpal I, Guo JT, Cook KD, Xu Y, Wetzel R. Mapping abeta amyloid fibril secondary structure using scanning proline mutagenesis. J Mol Biol. 2004; 335:833–42. [PubMed: 14687578]

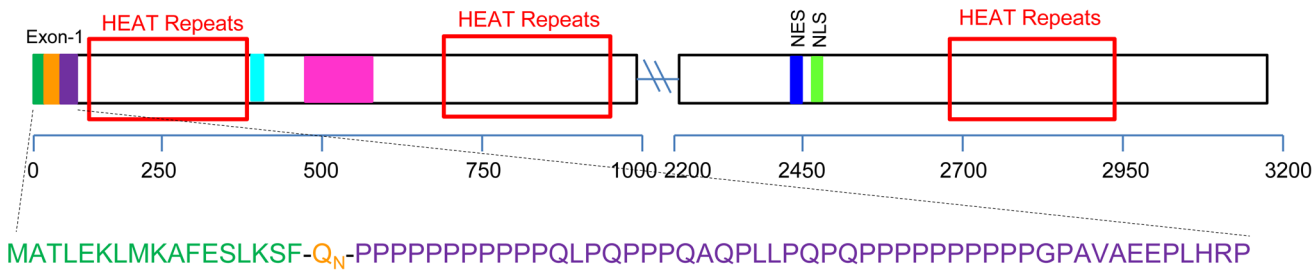

Figure 1.

Structure of huntingtin. The 3144 amino acid long human huntingtin protein (numbering based on a Q_{23} repeat length) showing the N-terminal exon-1 segment and other regions of interest. Boxed in red are three domains rich in ~40 amino acid long HEAT repeats implicated in protein-protein interactions [10], as identified by a neural network analysis [191]. Dark green, htt^{NT} segment; orange, polyQ; purple, proline-rich; cyan, phosphorylation and acetylation sites; magenta, *in vivo* caspase and calpain cleavage sites suggesting a disordered region; blue, nuclear export sequence; light green, nuclear import sequence. Based on information reviewed in reference [10].

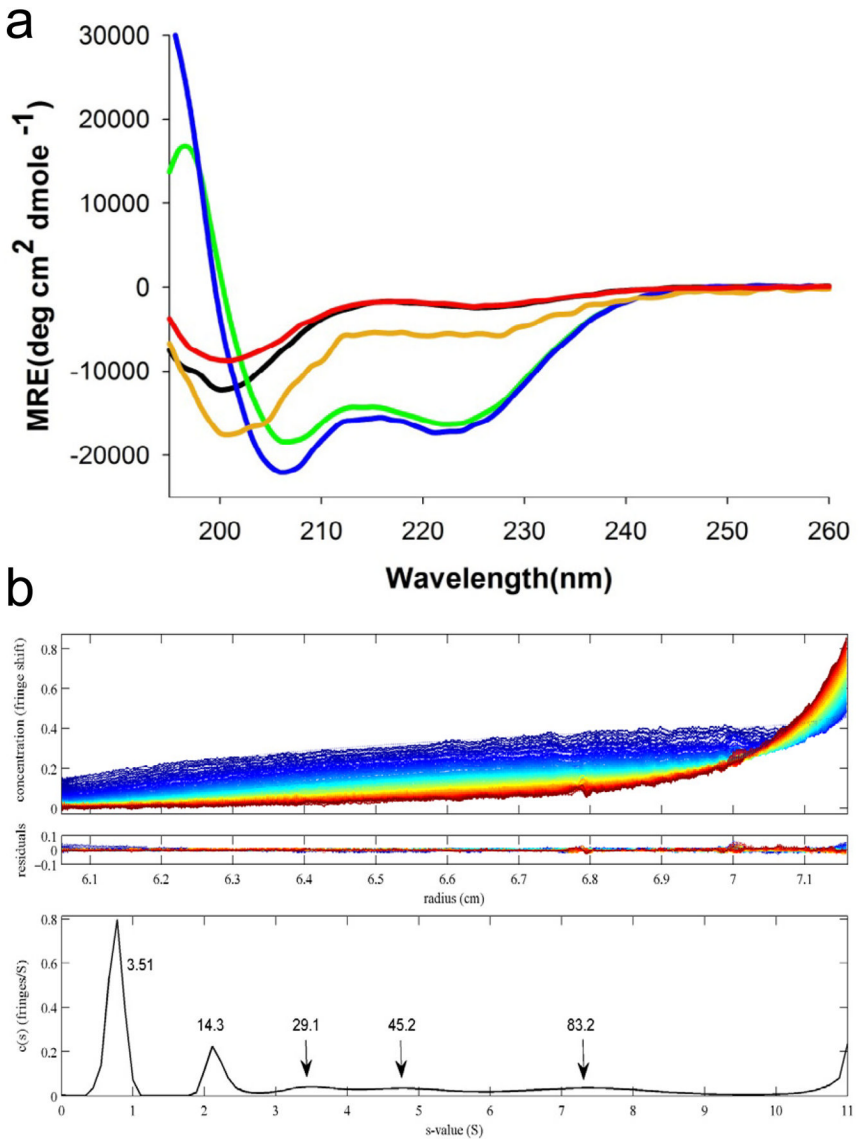

Figure 2.

Htt^{NT}-mediated oligomerization. (a) Concentration dependence of the CD spectra of $htt^{NT}Q$: green, 1.3 mM; blue, 0.67 mM; orange, 0.33 mM; black, 0.15 mM; red, 0.10 mM obtained by dilution from a 0.76 mM solution. (b) Sedimentation velocity analysis of 50 μM $htt^{NT}Q_{10}K_2$ at 20 °C. Top panel, sedimentation data; middle panel, residuals from the global fit; bottom panel, c(s) profile with peaks labeled (in kDa). Data from reference [50].

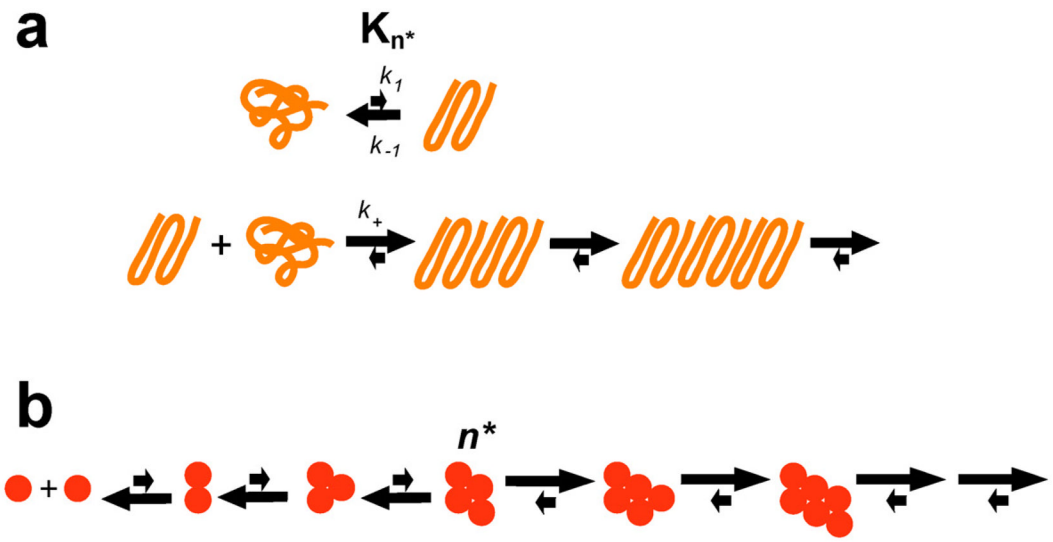

Figure 3.

Nucleation mechanisms. (a) Mechanism for simple polyQ aggregation nucleation by an energetically unfavorable conformation of the condensed coil polyQ monomer, where k_1 and k_{-1} are the forward and reverse rates of nucleus formation, K_{n^*} is the equilibrium constant for nucleus formation, and k_+ is the second order rate constant for elongation of the nucleus [23, 45, 75, 76]; (b) classical thermodynamic model for nucleated growth polymerization (from reference [45]).

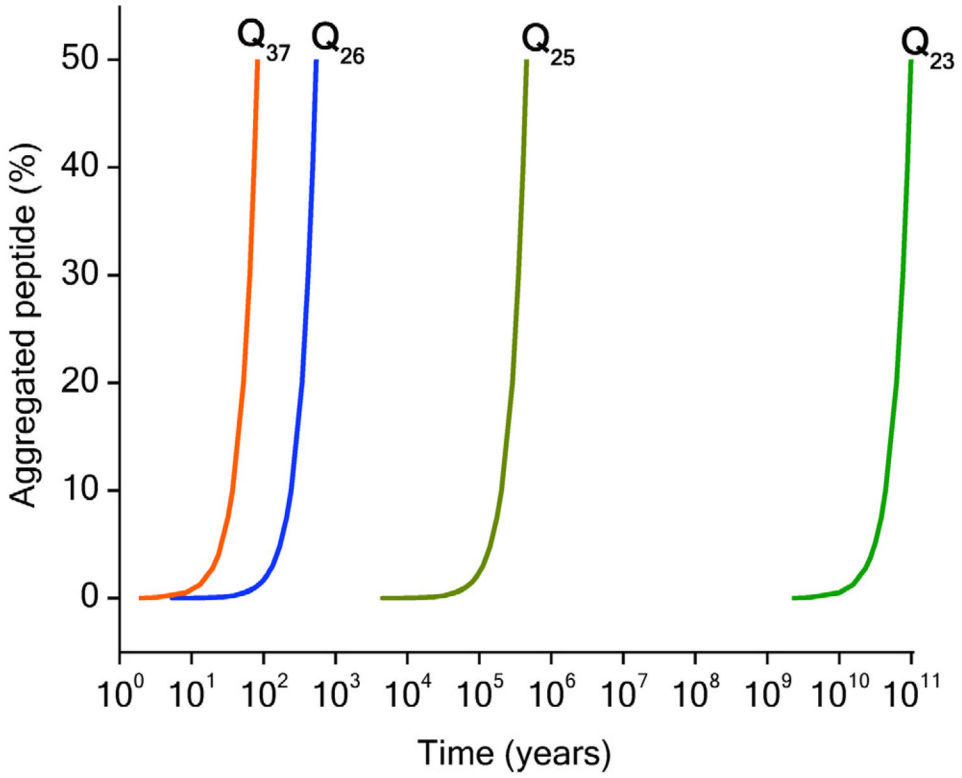

Figure 4.
Calculated aggregation kinetics curves for 1 nM of various polyQ peptides using parameters derived from nucleation kinetics analyses. From reference [45].

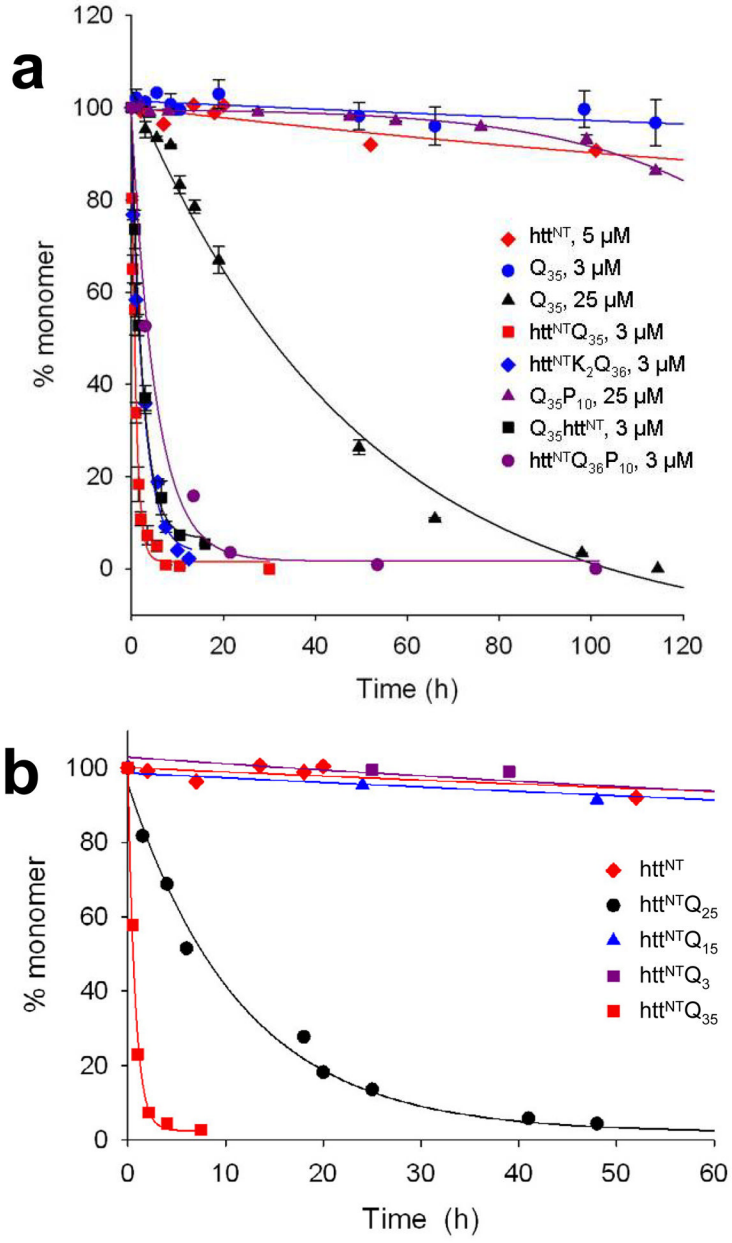

Figure 5.
Kinetics of spontaneous aggregation of polyQ peptides linked to htt^{NT}. (a) Effect of presence and location of htt^{NT}; (b) effect of polyQ length on $htt^{NT}Q_NK_2$ aggregation. From reference [48].

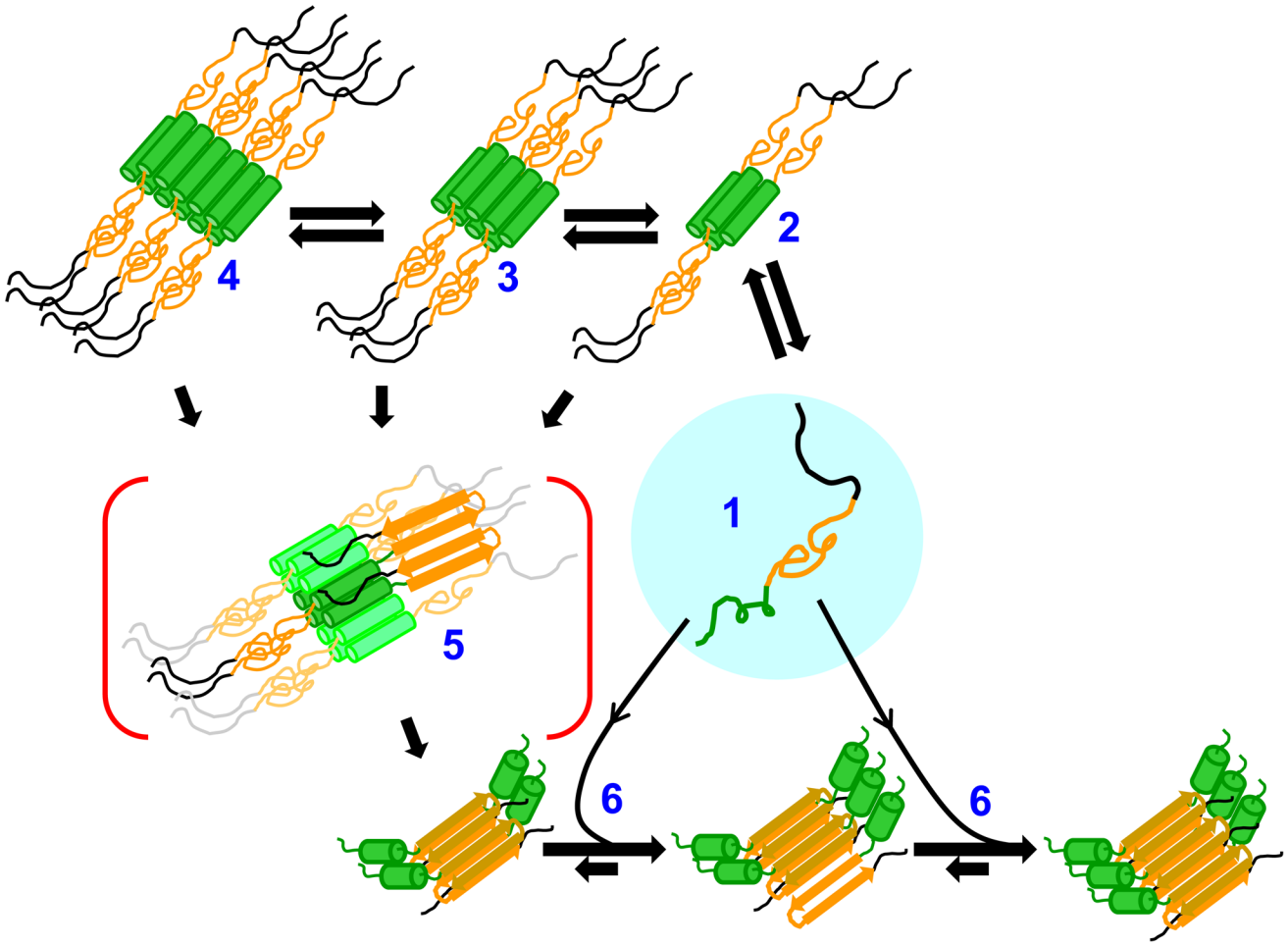

Figure 6.

Mechanism of htt^{NT}-mediated amyloid formation of htt N-terminal fragments. Disordered htt N-terminal fragment monomers (1; htt^{NT} = green; polyQ = orange; polyproline = black) reversibly form α-helix rich tetramers (2) via the htt^{NT} domain, and these tetramers can further reversibly assemble into octamers (3), dodecamers (4), and higher order oligomers. Any oligomer has a certain potential to undergo nucleation of amyloid structure (5), supported by the highly concentrated polyQ chains on its periphery; the propensity of any oligomer to undergo nucleation is probably linked to polyQ repeat length and the presence or absence of polyproline. Once polyQ-core amyloid nuclei are formed, amyloid elongation proceeds by monomer addition (6). As monomer concentration decreases due to fibril elongation, oligomers (2,3,4) that have not undergone nucleation dissociate to replenish the monomer pool and continue to support fibril elongation. Adapted from references [50, 143].

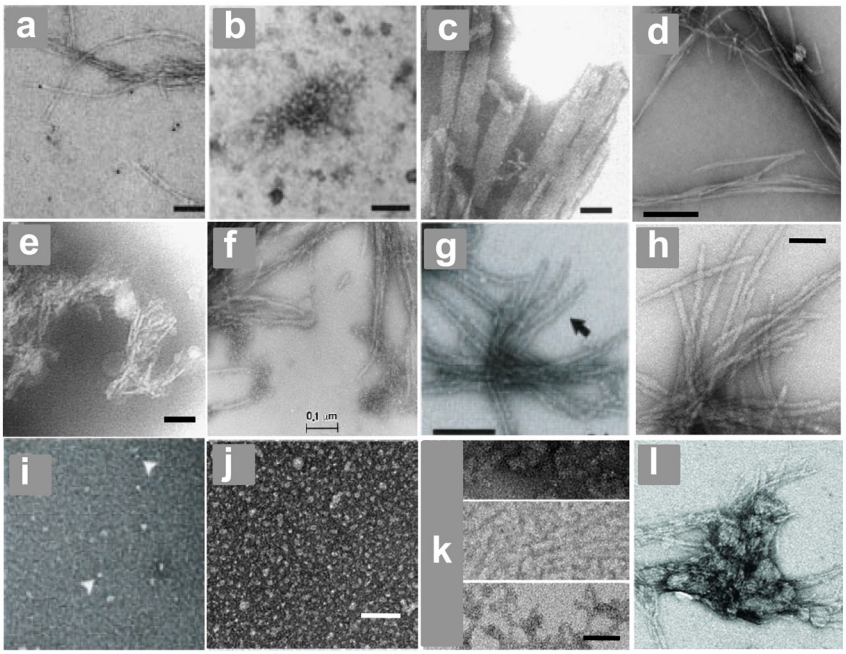

Figure 7.

Electron micrographs. (a, b) Htt N-terminal fragment aggregates isolated from tg HD Q_{150} mouse brains, scale bar = 100 nm [151]; (c) PolyQ mature aggregates produced in vitro at 37 °C from $K_2Q_{37}K_2$, scale bar = 50 nm [20]; (d) $A\beta_{40}$ amyloid fibrils generated at 37 °C in PBS without agitation, scale bar = 200 nm [192]; (e) aggregates of $K_2Q_{37}K_2$ generated in PBS solution incubated frozen at −20 °C, scale bar = 50 nm [20]; (f - l) aggregates of htt N-terminal fragments generated at 37 °C, as follows: (f) mature fibrils of Q_{51} exon-1, scale bar = 100 nm [104], (g) mature fibrils of Q_{44} exon-1, scale bar = 200 nm [117] (h) mature fibrils of $htt^{NT}Q_{30}P_6K_2$, scale bar = 50 nm [48]; (i) initial oligomers of Q_{44} exon-1 [117]; (j) initial oligomers of $htt^{NT}Q_{30}P_6K_2$, scale bar = 50 nm [48]; intermediate protofibrillar aggregates of $htt^{NT}Q_{30}P_6K_2$, scale bar = 50 nm [48]; final aggregates of $htt^{NT}Q_{37}P_{10}K_2$ (S13D/S16D) [58].

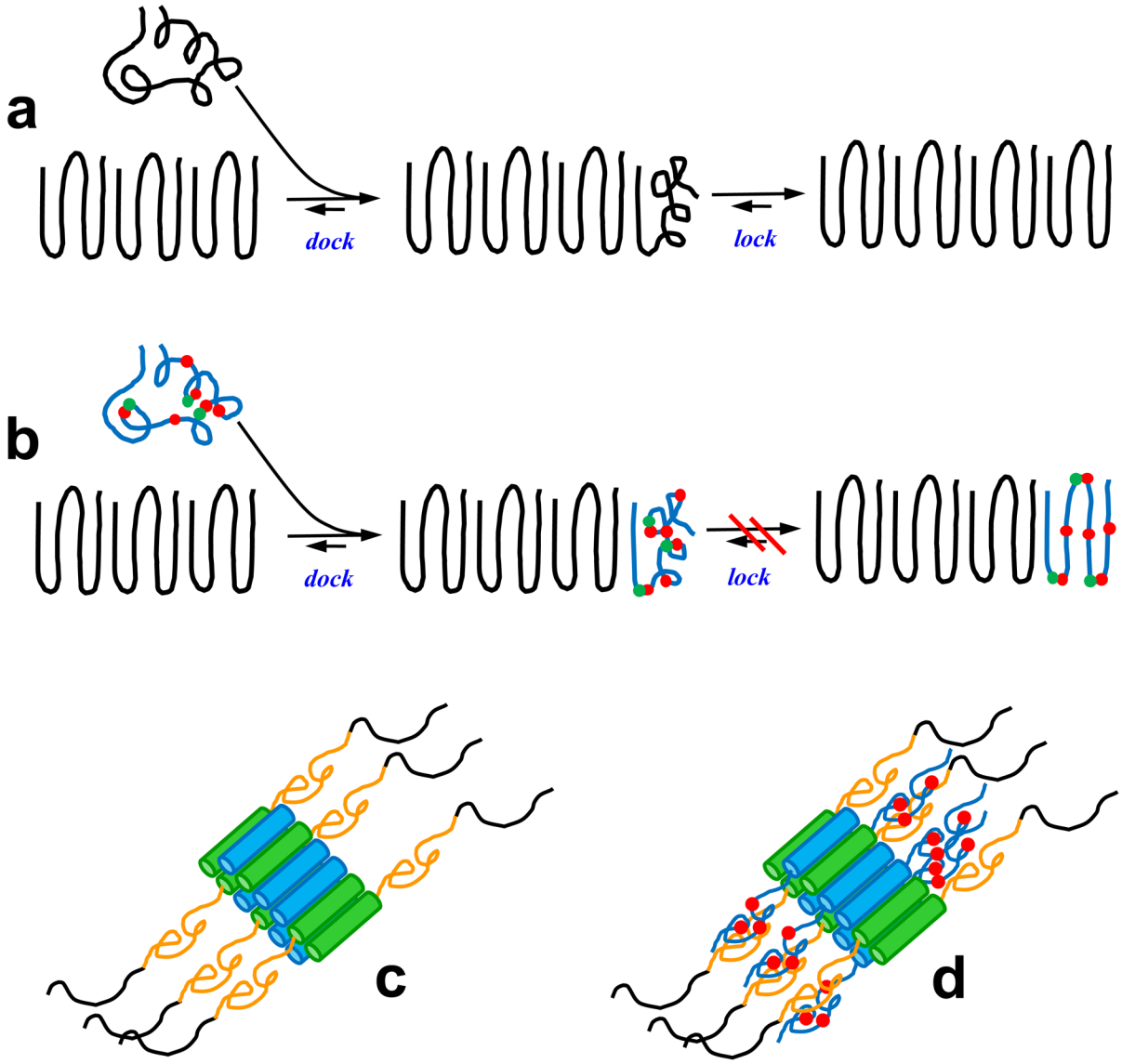

Figure 8.

Structure and mechanism-based inhibitors of polyglutamine amyloid formation. (a) dock and lock mechanism of amyloid fibril elongation from an encounter of the growth end of a fibril with a condensed coil polyQ monomer; (b) encounter of the growth end of the fibril with disordered $PGQ_9P^{1,2,3}$ monomer (K_2Q_9-PG-Q_4PQ_4-PG-Q_4PQ_4-PG-$Q_4PQ_4K_2$; proline, red; glycine, green) showing docking via the single unbroken Q_9 repeat, but an inability to rearrange into stable β-sheet in the locking step due to the mid-strand proline residues; (c) hypothetical inhibition complex when htt^{NT} (blue) is incubated with an htt N-terminal fragment (color scheme, see Figure 6); hypothetical inhibition complex when htt^{NT}-$PGQ_9P^{1,2,3}K_2$ (blue htt^{NT} and polyQ chain showing only the mid-strand Pro residues in red) is incubated with an htt N-terminal fragment.

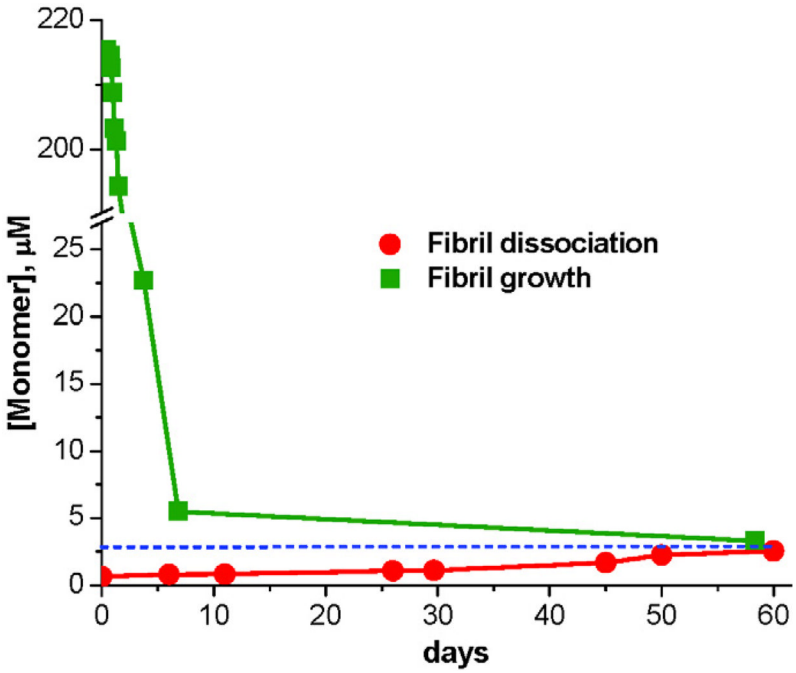

Figure 9.

Confirmation, for $K_2Q_{23}K_2$, of a dynamic equilibrium for amyloid fibril elongation with a characteristic end-point concentration of monomer equivalent to the C_r. From reference [45].

Table 1

PolyQ Repeat Lengths in TATA Box Binding Proteins (TBPs) of Various Species[a]

Source - GenBank	N-terminal to polyQ	Q_N	C-terminal to polyQ
Xenopus NP_001089038.1	-TYGTGLTPQPVQTTNSLSILEEQQR	4	--------TQQSTLQQGNQG-SGQTPQ
Zebrafish AAH55549.1	-PYGTGLTPQPVQNSNSLSLLEEQQR	6	---------AASQQQGGMVGGSGQTPQ-
Viper Q92146.1	-PYGTGLTPQPAQSTNSLSILEEQQR	6	-------AAAQQSTSQPTQAPSGQTPQ-
Chicken NP_990434.1	-PYGTGLTPQPVQSTNSLSILEEQQR	6	-------AAQSSTSQQATQGTSGQTPQ-
Rat AAH81939.1	-PYGTGLTPQPVQNTNSLSILEEQQR	15	AVATAAASVQQSTSQQPTQGASGQTPQ-
Cow AAI13309.1	-PYGTGLTPQPIQNTNSLSILEEQQR	19	---AAVAAVQQSTSQQATQGPSGQTPQ-
Rabbit XP_002723543.1	-PYGTGLTPQPIQNTNSLSILEEQQR	20	AVAAAAAAVQQSTSQQAAQGVSGQTPQ-
Chimpanzee NP_001098077.1	-PYGTGLTPQPIQNTNSLSILEEQQR	35	--AVAAAAVQQSTSQQATQGTSGQAPQ-
Human AAI10342.1	-PYGTGLTPQPIQNTNSLSILEEQQR	38	--AVAAAAVQQSTSQQATQGTSGQAPQ-

[a]Alignments fron Kalign (EMBL).