

Self-assembly and sequence length dependence on nanofibrils of polyglutamine peptides



Mohammed Inayathullah^{a,f,k,1}, Aaron Tan^{a,b,c,**,1}, Rebecca Jeyaraj^b, James Lam^b, Nam-Joon Cho^{d,g}, Corey W. Liu^e, Martin A.C. Manoukian^h, Keyoumars Ashkanⁱ, Morteza Mahmoudi^{a,j,k}, Jayakumar Rajadas^{a,k,*}

^a Biomaterials & Advanced Drug Delivery Laboratory (BioADD), Stanford University School of Medicine, Stanford University, Palo Alto, CA, USA

^b UCL Medical School, University College London (UCL), London, UK

^c University College London Hospitals NHS Foundation Trust, London, UK

^d Division of Gastroenterology and Hepatology, Stanford University School of Medicine, Stanford University, Palo Alto, CA, USA

^e Stanford Magnetic Resonance Laboratory, Stanford University, Palo Alto, CA, USA

^f Bioorganic and Neurochemistry Laboratory, Central Leather Research Institute, Adyar, Chennai, Tamilnadu, India

^g School of Materials Science and Engineering, Nanyang Technological University, Singapore

^h Department of Dermatology, Stanford University School of Medicine, Stanford University, Palo Alto, CA, USA

ⁱ Department of Neurosurgery, King's College Hospital NHS Foundation Trust, King's College London, London, UK

^j Nanotechnology Research Center, Faculty of Pharmacy, Tehran University of Medical Sciences, Tehran, Iran

^k Cardiovascular Pharmacology Division, Cardiovascular Institute, Stanford University School of Medicine, Stanford University, Palo Alto, CA, USA

ARTICLE INFO

Article history:

Received 11 November 2015

Received in revised form 11 January 2016

Accepted 31 January 2016

Available online 3 February 2016

Keywords:

Huntington's disease

Polyglutamine disease

Misfolded polyglutamine

Nanofibrils

PolyQ peptides

Neurodegenerative disease

ABSTRACT

Huntington's disease (HD) is recognized as a currently incurable, inherited neurodegenerative disorder caused by the accumulation of misfolded polyglutamine (polyQ) peptide aggregates in neuronal cells. Yet, the mechanism by which newly formed polyQ chains interact and assemble into toxic oligomeric structures remains a critical, unresolved issue. In order to shed further light on the matter, our group elected to investigate the folding of polyQ peptides – examining glutamine repeat lengths ranging from 3 to 44 residues. To characterize these aggregates we employed a diverse array of technologies, including: nuclear magnetic resonance; circular dichroism; Fourier transform infrared spectroscopy; fluorescence resonance energy transfer (FRET), and atomic force microscopy. The data we obtained suggest that an increase in the number of glutamine repeats above 14 residues results in disordered loop structures, with different repeat lengths demonstrating unique folding characteristics. This differential folding manifests in the formation of distinct nano-sized fibrils, and on this basis, we postulate the idea of 14 polyQ repeats representing a critical loop length for neurotoxicity – a property that we hope may prove amenable to future therapeutic intervention. Furthermore, FRET measurements on aged assemblages indicate an increase in the end-to-end distance of the peptide with time, most probably due to the intermixing of individual peptide strands within the nanofibril. Further insight into this apparent time-dependent reorganization of aggregated polyQ peptides may influence future disease modeling of polyQ-related proteinopathies, in addition to directing novel clinical innovations.

© 2016 Elsevier Ltd. All rights reserved.

Abbreviations: HD, Huntington's disease; PolyQ, polyglutamine; FRET, fluorescence resonance energy transfer; Htt, huntingtin protein; CD, circular dichroism; FTIR, Fourier transform infrared spectroscopy; XRD, X-ray diffraction; NMR, nuclear magnetic resonance; AFM, atomic force microscopy; TOCSY, total correlation spectroscopy; HSQC, heteronuclear single-quantum coherence; NOESY, nuclear overhauser effect spectroscopy; BPTI, bovine pancreatic trypsin inhibitor.

* Correspondence to: J. Rajadas, Biomaterials & Advanced Drug Delivery Laboratory (BioADD), Stanford University, 1050 Arastradero Road, Building A, Room A163, Palo Alto, CA 94304.

** Correspondence to: A. Tan, UCL Medical School, University College London, Gower Street, London WC1E 6BT.

E-mail addresses: aaron.tan@ucl.ac.uk (A. Tan), jayraja@stanford.edu (J. Rajadas).

¹ Equal first authors.

1. Introduction

1.1. Molecular pathophysiology of HD

HD is a neurodegenerative disease characterized by fragmentation and intraneuronal aggregation of the mutant huntingtin protein (Htt) – an aberrant polypeptide consisting of misfolded polyQ segments (Katsuno et al., 2008). The mechanistic hypothesis that links CAG (the three-letter genetic code for the amino acid glutamine) repeat expansion to neurotoxicity involves the accumulation of polyQ sequences in the Htt protein. Expansion of polyQ repeats above a certain critical length induces proteolytic cleavage of the mutant polypeptide – with the corollary N-terminal fragments forming insoluble aggregates (Bates, 2003; Orr, 2012; Wetzel, 2012). These self-assembled aggregates are associated

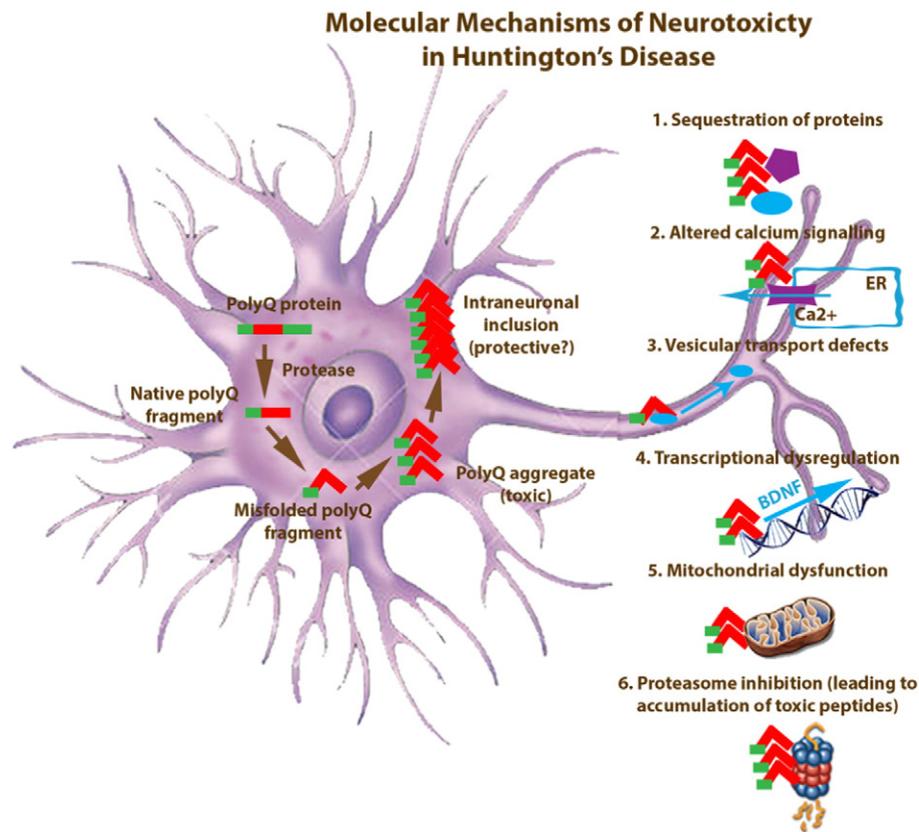


Fig. 1. Mutant Htt is cleaved enzymatically into peptide fragments containing *expanded* polyQ tracts. It is thought that over a critical length, these fragments assume a misfolded configuration, which promotes self-aggregation into toxic oligomers that disrupt various cellular functions.

with cellular dysfunction and consequent neurotoxicity (Fig. 1) (Blum et al., 2013; Cisbani and Cicchetti, 2012; Zheng and Diamond, 2012) and ultimately, are thought to give rise to the neurological sequelae of HD.

Oligomerization of polyQ and other amyloid-forming proteins has long been thought of as a crucial determinant in the development of cytotoxicity (Hands and Wyttenbach, 2010; Hatters, 2012; Janciauskiene et al., 1999; Legleiter et al., 2010; Ono et al., 2009; Shao and Diamond, 2007). Yet, perhaps less well-recognized is that this conversion of polyQ peptides into insoluble aggregates appears to be driven by more than a simple propensity for self-association – with current evidence indicating that a more complex, highly regulated process may be responsible (Bates, 2003; La Spada et al., 2011). In spite of accumulating data linking polyQ length to toxicity, the structural basis for the influence of polyQ repeat length on the folding and self-assembly mechanism has yet to be fully elucidated. In fact, it has been demonstrated that Htt fragments with expanded polyQ segments can sequester transcription factors with polyQ domains as small as 18 residues, suggesting that even smaller polyQ repeat lengths can adopt aberrant structures (Kar et al., 2011; Robertson et al., 2010) and highlighting key gaps in our understanding of the misfolding process.

1.2. Uncertainties in current understanding

One of the major bottlenecks in our understanding of polyQ toxicity is the lack of structural details on these assemblages. The disordered nature of polyQ peptides and their tendency to aggregate make spectroscopic interpretations of structure difficult (Napolitano et al., 2011). Nonetheless, various biophysical methods have met with some success in investigating the gross structure of polyQ peptides; prominent examples include: (1) circular dichroism (CD), which has linked β -sheet formation to aggregation; (2) Fourier-transform infrared (FTIR) spectroscopy, which has linked hydrogen bond formation to aggregation;

and (3) X-ray diffraction (XRD) measurements, which have yielded inter-strand distances (Davranche et al., 2011; Ortega et al., 2010).

At present, there is little consensus on the precise conformation adopted by polyQ peptides. Atom energy minimization studies of polyQ with implicit solvation showed that CHARMM parameters produce a β -hairpin structure, whereas AMBER parameters produce a random coil structure (Finke et al., 2004; Tsukamoto et al., 2006; Vanschouwen et al., 2011; Vitalis et al., 2007; Zanuy et al., 2006). In contrast, homology modeling predicted the polyQ region of ataxin-3 to be an α -helix while in vitro coherent anti-stokes Raman microscopy revealed that polyQ fibers assemble into highly rigid β -sheet structures (Perney et al., 2012).

Current literature is also equivocal about whether the length-dependence of toxicity is related to a conformational change in the monomeric state of expanded polyQ peptides. For instance, the Wetzel group (Thakur and Wetzel, 2002; Wetzel, 2012) reported that polyQ peptides with repeat lengths of 5, 15, 28, and 44 residues all adopt random coil structures in solution, whereas a Flory-Huggins mean-field lattice model postulated that polyQ increasingly prefers a β -hairpin state in a length-dependent manner (Crick et al., 2006; Masino et al., 2004). Notably, however, host-guest studies where increasing lengths of a polyQ “guest” were inserted into “host” chymotrypsin inhibitor 2 (CI 2) mutants demonstrated increasing destabilization of the host CI 2 protein. This length-dependent destabilization of monomeric polyQ is consistent with the observed increase in kinetic and thermodynamic stability of ordered, amyloid-like aggregates for polyQ peptides with more than 37 residues (Bhattacharyya et al., 2006; Chen et al., 2001; Chen et al., 2002b; Crick et al., 2006; Thakur and Wetzel, 2002). Overall, these studies suggest that repeat length is likely to be a crucial factor in the oligomerization – and hence neurotoxicity – of destabilized polyQ peptides, although the structural mechanism of this has yet to be clarified.

It has also been shown that polyQ peptides with repeat lengths in the 15–20 residue range may attach to pre-formed aggregates (analogous to the elongation stage in the formation of amyloidic proteins (Ghavami et al., 2012; Mahmoudi et al., 2013a; Mahmoudi et al., 2013b)), despite their limited ability to support spontaneous self-nucleated aggregation. This suggests that even shorter polyQ peptides can be induced to misfold – a phenomenon which might have important implications in the disruption of cellular function.

With the lack of clarity on polyQ structure, several conflicting models have been proposed to describe the polyQ aggregation process. For instance, Perutz et al. (Perutz et al., 2002; Perutz and Windle, 2001) proposed a model in which shorter polyQ repeats have a random coil conformation while longer repeats have β -hairpin or β -strand structures that promote aggregation through the formation of intermolecular polar zippers. An alternative hypothesis, presented by Landrum and Wetzel, (2014), suggests that polyQ aggregation is initiated by a monomer functioning as a critical nucleus for random coil to β -sheet transition – with subsequent linear additions of single polyQ molecules resulting in fibril formation (Crick et al., 2006). Interestingly, the reaction kinetics and fibrillar stability in this model appear to have a non-linear dependence on polyQ repeat length – mirroring the non-linear dependence of age-of-onset and severity of polyQ diseases on repeat length. In other words, it is possible that age-of-onset and disease severity may be directly related to the rate of polyQ aggregation, and the stability of formed aggregates (Landrum and Wetzel, 2014).

Clearly, given the large discrepancies between polyQ structures derived from several lines of study, as well as the potential implications for disease modeling and therapeutics, further experimental evidence is needed to characterize polyQ peptide folding in aqueous solution. Here, using diverse spectroscopic and microscopic evaluations, we aim to better elucidate the folding mechanism of polyQ peptides. In comparison to previously published studies, we believe that the wide range of biophysical techniques employed here allows us to reliably integrate our findings, and present a more consistent and convincing model of how aggregation may operate at the molecular level.

2. Experimental procedure

2.1. Design and synthesis of PolyQ peptides

Peptides were designed and synthesized to study the intramolecular folding interaction in polyQ peptide chains. The peptides were synthesized by solid phase methods (Supplementary Methods), and some of the peptides were obtained from the PAN facility of Stanford University and from the Keck facility of Yale University. In order to overcome the issue of the low solubility of the polyQ peptides, lysine residues were added to the N and C-terminals to increase the solubility of the peptides. Dansyl-Trp and pyrene-Trp FRET pairs were used in the present studies. The peptides Q20 and Q44 containing Trp were labeled with a pyrene fluorescent probe for the FRET experiments. Pyrene labeling for the peptides was done coupling pyrene maleimide to the SH group of the side chain of cysteine present on the C-terminus. All the peptides were purified by RP-HPLC and characterized by MALDI-TOF mass. Before each experiment, the lyophilized peptides were freshly disaggregated. The list of fluorescently labeled and unlabeled polyQ peptides used in the present study and the abbreviations used are given in Supplementary Table S1.

2.2. Sample preparation

The peptides were solubilized and disaggregated using TFA/HFIP procedures (Chen and Wetzel, 2001). The samples in aqueous medium for CD and IR were prepared from the disaggregated stock solution.

2.3. CD Spectroscopy

CD spectra were recorded using a JASCO model J-715 (Japan) spectropolarimeter. Spectra were recorded with 1 nm bandwidth and 0.2 nm step resolution at a scan speed of 100 nm/min over the range of 190–260 nm. The CD data are representative spectra of three experiments and the averages from five recordings each, and processed for routine baseline correction and noise reduction. The peptide solutions were prepared in 10 mM phosphate buffer, pH 7.4, incubated at 25 °C and the spectra were recorded at various time points. Quantitative estimation of secondary structure was obtained by using computer programs such as Selcon3 and Contin-LL of the CDPro software package (Sreerama and Woody, 2000). The input file containing CD data ranges from 190 to 260 nm.

2.4. FTIR spectroscopy

FTIR spectra were recorded on a Thermo Nicolet (AVATAR 320 Model) at 25 °C. The spectra were all processed by OMNIC (version 6.0) software supplied by Thermo Nicolet. The samples were dissolved in D₂O with 200 μ l of the peptide solutions loaded into the cell. Peptide solutions were prepared for infrared measurement in a CaF₂ cell with a 0.015 μ m spacer. For each spectrum, 500-scan interferograms were collected in single-beam mode with a 4 cm⁻¹ resolution. The spectrum of D₂O was subtracted from the observed spectra of peptide according to previously established criteria (Dong et al., 1990). The final spectrum was processed for baseline correction and noise reduction routines.

2.5. Nuclear magnetic resonance (NMR) spectroscopy

NMR experiments were performed on 600 and 800 MHz Varian Inova spectrometers. Sample concentrations were 1 mM in Q20 or Q44, in 35 mM phosphate buffer (pH 7.1) of 90% H₂O and 10% (v/v) D₂O at 10 °C. 1D ¹H NMR spectra of Q20 and Q44 were acquired immediately following solvation. Two-dimensional ¹H-¹H NOESY was acquired on Q20 and Q44 acquired at 600 MHz at 10 °C with 200 ms mixing times; F2: 8000 Hz spectral width, 4096 points, 1 s prescan delay; F1: 8000 Hz spectral width, 128 points.

2.6. FRET measurement

Fluorescence measurements were made on a Perkin Elmer LS45 model spectrofluorimeter. Fluorescence emission spectra of fluorescently labeled polyQ were measured in the presence of 1:10 M excess of unlabeled polyQ peptide in order to minimize the intermolecular FRET. Trp was used as the FRET donor, and pyrene (py) or dansyl (dan) was used acceptor. Excitation and emission slit widths were fixed at 10 nm. All experiments were recorded at 25 °C. The Trp was excited at 280 nm and the emission spectra were recorded for dansyl or pyrene group along with Trp from 300 to 540 nm. The peptide solutions were incubated at 25 °C and the spectra were recorded at various time points.

2.7. Atomic force microscopy (AFM)

For AFM measurements, a freshly cleaved mica surface was used to examine the growth of polyQ peptides. Prior to analysis, a fixed concentration of Q20 and Q44 was deposited onto mica and allowed to adsorb for time courses. The samples were washed with milli-Q water, dried with nitrogen gas, and imaged in air. The AFM experiments were carried out on an XE-100 (PSIA Inc. Santa Clara, CA, USA) in contact mode and non-contact mode. Rectangular-shaped SiO₂ cantilevers were used (NCHR, Nanosensors, Switzerland). The cantilevers had a force constant of $k = 42$ N/m. All measurements were obtained in air for both the “contact” with an imaging force of less than 100 pN and “non-contact” modes. However, all images presented in this manuscript were obtained in the “non-contact” mode in air. The scan line speed was optimized

between 0.5 Hz to 2 Hz with a pixel number of 256×256 , depending on the scan size. Images were recorded in height, amplitude, phase, and error modes. All measurements were done on the height images. Whole illustrated images were subjected to a first order plane-fitting procedure to compensate for sample tilt. The cross-sectional analysis was carried out on images subjected only to a first order plane-fitting procedure. Topographical and grain analyses of polypeptides were performed using the software XEI 1.5 supplied by PSIA (Sungnam, Korea).

3. Results

3.1. Elucidation of PolyQ conformation and assembly

PolyQ peptides of different lengths (Q3, Q6, Q10, Q14, Q20 and Q44) were synthesized (Supplementary Table S1). Q20 and Q40 correspond to the polyQ repeat length of wild-type Htt (6–37 residues) and mutant Htt (>37 residues) respectively (Chen and Wetzel, 2001; Hoffner and Djian, 2014; Waelter et al., 2001). Shorter polyQ repeat lengths (Q3, Q6, Q10, Q14) were used to monitor the change in peptide conformation with a sequential increase in length, and potentially determine the threshold length for peptide misfolding.

The conformation and self-assembly of these polyQ peptides were studied using CD, FTIR and nuclear magnetic resonance spectroscopies. The folding, aggregation and morphology of fibrils were studied by FRET and AFM techniques.

3.1.1. CD Spectroscopy

CD spectroscopy exploits the unequal absorption of left-handed and right-handed circularly polarized light to determine the secondary structure and folding of a protein (Greenfield, 2006). Far UV-CD spectra were measured on Q20 and Q44 to analyze their conformational propensities (Fig. 2). Secondary-structure content for both peptides in the freshly solubilized form was estimated by Selcon to be predominantly disordered. Acquiring CD spectra of the Q20 peptide solution over three days shows that the peptide backbone adopts an increasingly β -sheet-like conformation with time. The π - π^* excitation component at 200 nm was observed to change from $-10,000$ deg. cm^2/dmol to 9000 deg. cm^2/dmol . Similar CD spectral changes were observed for Q44 in the 200 nm region, resulting in changes in molar ellipticity values from $-10,000$ deg. cm^2/dmol to 6000 deg. cm^2/dmol upon aging. However, interesting differences are noted in the CD spectra at 222 nm. For Q20, only moderate change was observed over three days, indicating that the structural transition in the aggregation/self-assembly process involves only the random coil region. In contrast, considerable change was observed for Q44 over three days, revealing that the total sheet content in Q20 and Q44 may be different. The time-course measurements

also demonstrate that Q44 adopts a β -sheet conformation with increased propensity compared to Q20. The mid-point of the conformational transition of Q44 is 10 h compared to 45 h for Q20, indicating a faster rate of Q44 self-assembly.

3.1.2. FTIR spectroscopy

FTIR is a sensitive and commonly used technique for monitoring disordered β -sheet transitions (Haris, 2010). The sensitivity of vibrational spectroscopy to protein structure is well documented, particularly with regard to the use of the amide I carbonyl stretches to probe hydrogen bonding and secondary structure formation (Manning, 2005). Earlier studies have used IR spectroscopy to examine secondary structure in β -sheet-rich proteins and monitor intermolecular β -sheet formation, peptide aggregation, and amyloid formation (Altschuler et al., 2000; Natalello et al., 2011; Poirier et al., 2005). The infrared spectra of Q20 and Q44 in D_2O are displayed in Supplementary Figure S1. Upon aggregation, the secondary structure-sensitive amide I vibrational band shifts from 1640 cm^{-1} , which is recognized as the typical wave number for disordered proteins, to 1617 cm^{-1} . This region is assigned to the β -sheet, which is the predominant secondary structure in amyloid assemblages. The Q20 peptide displayed peaks indicating a predominately disordered conformation while the Q44 peptide showed significant β -sheet content as evidenced by its 1617 cm^{-1} peak immediately after solubilization. This structural transition may be attributable to the very high concentration of peptide employed for IR measurements (2 mM). Chen et al. (Chan et al., 2011; Chen et al., 2002a; Chen et al., 2002b; Chen and Wetzel, 2001) have reported such a concentration-dependent enhancement of β -sheet formation in polyQ peptides.

The (C = O) vibration of Gln side chains that appears near 1680 cm^{-1} is a relatively strong infrared absorber. The downshift of 30 cm^{-1} , due to hydrogen bonding, is considerably larger than that of the amide I bands of the protein backbone and helps distinguish the amide side chain absorption in Gln from the amide backbone absorption. Furthermore, glutamine (C = O) bands are sensitive to hydrogen bonding, with band position lowering as hydrogen bond strength increases (Natalello et al., 2011). Weakly hydrogen bonded side chain amide groups are highly blue-shifted and appear around 1704 cm^{-1} . A closer examination of the 1680 cm^{-1} peak indicates it can be resolved into two components with peaks at 1675 cm^{-1} and 1690 cm^{-1} . The lower wavelength component is presumably the hydrogen bonded side chain that occurs in the interior of the folded domain, while the higher peak at the longer wavelength may correspond to the solvent-exposed outer form. It is interesting that the non-hydrogen bonded side chains are not presented in the polyQ fold. This suggests that the main driving force for peptide folding might be achieved due to side chain interactions in the polypeptide fold. Overall, the FTIR and CD

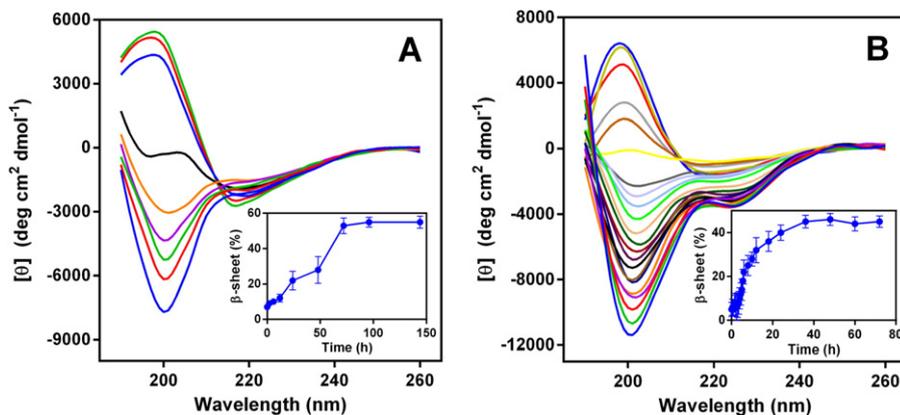


Fig. 2. Far-UV CD spectra of 0.09 mg/ml Q20 & Q44 during its aggregation in 10 mM phosphate buffer, pH 7.4 at 25°C . (A) Q20 at different time points 0 to 144 h, [0 h, 2, 6, 12, 24, 48, 72, 96 and 144 h], transition from the negative peak at ~ 200 nm at 0 h to a positive peak at 144 h. (B) Q44 at different time points, 0 to 72 h, [0 h, 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 5.5, 6, 8, 10, 12, 18, 24, 36, 48, 60, 72 h], transition from the negative peak at ~ 200 nm at 0 h to a positive peak at 72 h. The CD spectra represent the average of ten spectral scans each of three experiments. Inset: Plot showing a time dependent increase of β -sheet conformation.

data suggest that both the Q20 and Q44 peptides initially adopt disordered conformations upon fresh dissolution and subsequently develop a β -sheet form on aging.

3.1.3. NMR spectroscopy

NMR is a powerful method that allows the study of the solution-state interaction of peptides and protein aggregates (Billeter et al., 2008). In this study, we used NMR to study the backbone and side chain interactions among polyQ peptides in the aggregates. One-dimensional ^1H spectra of freshly solvated Q20 and Q44 were very similar, indicating comparable starting solution conformations between the two peptides (Supplementary Figure S2). However, acquiring 1D ^1H spectra over time reveals that both peptides exhibit reductions in signal intensities, with Q44 showing a more rapid drop-off compared to Q20 (Supplementary Figure S3). Visual inspection of the samples revealed the solutions to remain clear and free of precipitate, suggesting that the signal loss is due to the formation of large solution aggregates which are very broad in standard solution-state NMR experiments.

Two-dimensional ^1H - ^1H total correlation spectroscopy (TOCSY) and ^{13}C - ^1H heteronuclear single-quantum coherence (HSQC) experiments were acquired on Q20 and Q44 to verify resonance assignments and to assess solution-state conformations. The bulk Gln and terminal Lys amide resonances can be identified in the TOCSY spectra (Supplementary Figure S4), as can the aliphatic side chain resonances in the HSQC spectra (Supplementary Figure S5). The bulk Gln signals are consistent with random coil chemical shifts for Gln. The degeneracy of the Gln resonances precludes a true Chemical Shift Index analysis for secondary structure elements.

Two-dimensional ^1H - ^1H nuclear overhauser effect spectroscopy (NOESY) experiments acquired on Q20 and Q44 were unable to discern any proton-proton contacts other than what appear to be intra-residue

contacts (Supplementary Figure S6). Though it is certainly possible that the inter-residue nuclear overhauser effects are being shielded by the highly degenerate glutamine resonances, no specific structure or folding can be ascertained from the NOESY data.

The bulk Gln side chain $\text{H}\epsilon$ resonances are narrowly dispersed with chemical shifts very close to random coil for Gln. The aforementioned polar zipper model proposed by Perutz et al. (1994) should result in highly downfield shifted side chain resonances due to strong hydrogen bonding. The chemical shifts observed here do not support such a model for the NMR observable solution-state.

Instead, the NMR data are consistent with a model where the polyQ peptides have highly flexible and poorly defined structures in the NMR observable solution-state. This indicates that the conformations observed with CD and IR spectroscopy are present in large aggregates, which are invisible to standard solution-state of NMR experiments.

3.2. Elucidation of PolyQ folding by FRET

Though NMR can be used to investigate the tertiary structure of proteins, the self-assembling nature of polyQ peptides precludes such analysis. FRET is another method for probing longer-range structure – it has long been used as a spectroscopic ruler to give accurate measurements of distances within proteins and other macromolecules (Lakowicz, 2006). These measurements are based on the efficiency of energy transfer between two intramolecular fluorophores, an energy donor and an energy acceptor. In our experiments, tryptophan was used as a donor while dansyl or pyrene was employed as an acceptor. Both pairs adhere to the Forster theory with Forster distances of 22 Å (distance determination range of 11–32 Å) for dansyl-Trp and 28 Å (distance range 14–42 Å) for pyrene-Trp (Lakowicz, 2006). Fig. 3 shows the FRET results of pyrene-Q20 incubated for different time intervals. For the pyrene-Trp

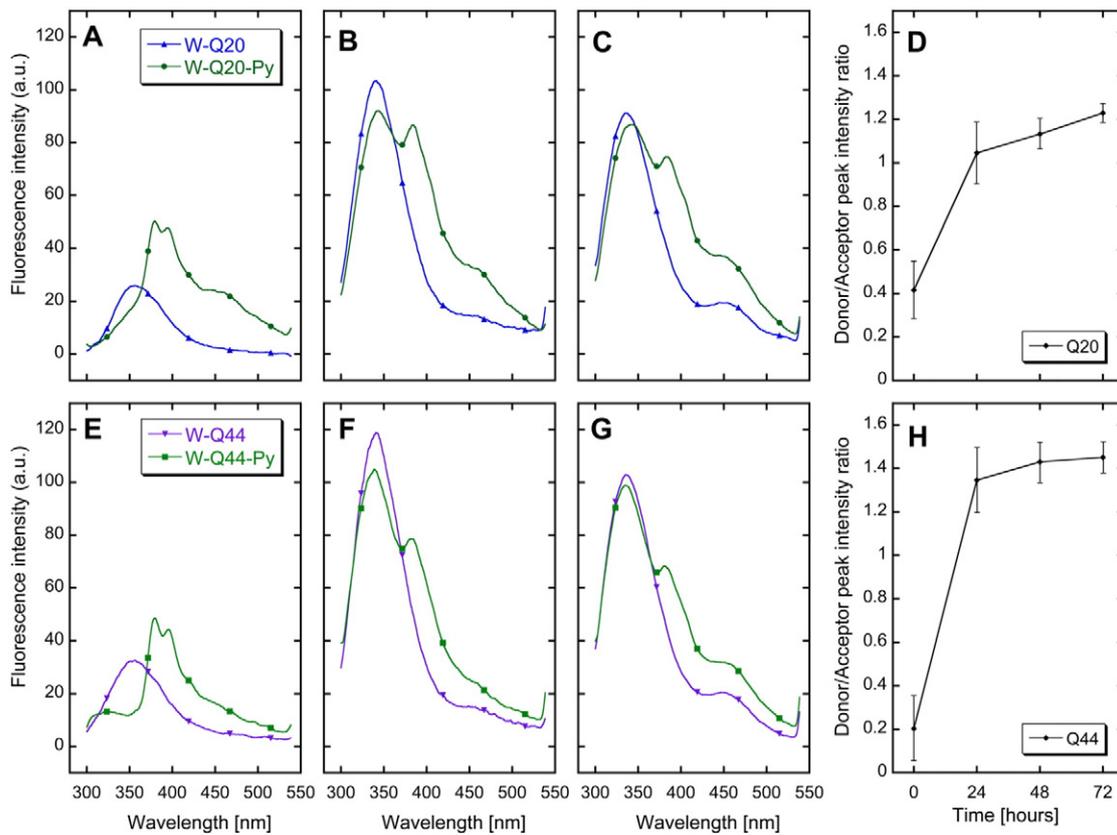


Fig. 3. Fluorescence emission spectra of fluorescently labeled polyQ (Q20 and Q44) in presence of 1:10 M excess of unlabeled polyQ peptide (Q20 and Q44 respectively). FRET donor and acceptor used were Trp (W) and pyrene (py) respectively. Concentration of each labeled peptide was 2 μM and the total concentration of peptide (labeled + unlabeled) was 20 μM in 10 mM phosphate buffer, pH 7.4. The peptide solutions (Q20 & Q44) were incubated at 25 $^{\circ}\text{C}$ and the spectra were recorded at various time points, (A) & (E) 0 h, (B) & (F) 24 h, (C) & (G) 48 h. (D) & (H) Plot of donor/acceptor intensity ratio with respect to time.

FRET pair, 50% transfer occurs at 28 Å (Lakowicz, 2006). Complete transfer takes place in the case that the distance falls below 14 Å. Assuming that the polyQ peptide acts as a flexible polymer, the average end-to-end distance of random coil chains of polyQ can be calculated. Using a sequential distance between adjacent α -carbon atoms of the peptide chains ($C\alpha_i-C\alpha_j$) as ~ 3.8 Å, with N number of residues connecting the donor and acceptor chain, the following equation yields the average root mean square (or end-to-end) distance, of a random coil population: $(C\alpha_i-C\alpha_j) \times \sqrt{N}$. For a length of 21 amino acid residues (i.e. $N = 21$), this equation yields an end-to-end distance of 17.4 Å if the peptide were to adopt a random coil conformation.

Upon dissolution of the polyQ peptides, we observed a high efficiency of resonance energy transfer. Our data demonstrates that the $C\alpha_3-C\alpha_{24}$ distance in Q20 is below 14 Å, which is representative of the spatial proximity between residues in the Q20 fold. This observed distance is smaller than the previously calculated distance of 17.4 Å for a predicted random coil conformation. Instead, such a distance of less than 14 Å might be expected if the peptide were to fold into a β -hairpin-like structure. However, on aging, we observed a reduction in energy transfer, and after 24 h it reached 50%. This is indicative of the N and C termini opening up and moving back to a distance above ~ 28 Å, probably due to conformational changes or intermixing of individual peptide monomers in an aggregate structure. The intermixing of polyQ peptides results in the formation of oligomeric aggregates with greater stability, and this stability increases with an increasing number of glutamine residues in the peptide. It should be noted that β -sheet formation as supported by the CD spectra correlates with $C\alpha_3-C\alpha_{24}$ opening up, thus indicating that intermixing of polyQ peptide monomers is the probable explanation for the increase in distance between $C\alpha_3-C\alpha_{24}$ atoms.

The pyrene-Trp fluorescence was monitored in the folding process of K(biot)K WQ44 C(Py)KK. Unexpectedly, the similar FRET profiles for both Q20 and Q44 suggest that their $C\alpha_3-C\alpha_{45}$ separation in the fold is the same (Fig. 3). In the 46 residue case, the calculated $C\alpha_3-C\alpha_{45}$ length for the peptide with random coil conformation was ~ 26 Å; however, the observed $C\alpha_3-C\alpha_{45}$ distance is below 14 Å, which is similar to the Q20 and indicates their spatial proximity in the Q44 fold. This interpretation is possible if the peptide folds are punctuated by three β -hairpin-like structures. Nonetheless, on aging, we observed a reduction in energy transfer efficiency which reached 50% at 24 h, indicating the N and C termini moved back to a distance above

~ 28 Å – probably due to the conformational change or intermixing of strands in the self-assembled peptides. Intermixing may be the result of domain swelling due to polyQ aggregation, which increases with the number of glutamine residues.

Ittah and Haas (1995) have measured intramolecular segmental end-to-end distance distributions for bovine pancreatic trypsin inhibitor (BPTI) protein with an unfolded peptide stretch of 15, 26, 41, and 46 residues. The measured distance between the labeled N-terminal and C-terminal ends were indicative of partial folding conditions of low GuHCl concentration, with reduced BPTI giving rise to either an open or compact state (Navon et al., 2001). The dominant subpopulation was compact, with native-like intramolecular segmental end-to-end distance distribution, while the other subpopulation was unfolded. The pairs of sites, residues 3 and 24 and residues 3 and 46, showed close proximity in the compact subpopulation. These contacts form two loops, probably collapsed (Chintapalli et al., 2014): one consists of the first 26 residues, and the second comprises the full length of the chain from the N- to the C-terminal segments, which is in fact made up to two shorter loops (1–26 and 27–46). Hence, it is quite unlikely that the intramolecular FRET observed in the freshly soluble and aged solutions is due to statically collapsed coil conformation. Rather, it may be caused by a fold with an initial disordered polypeptide chain that upon aging develops into a sheet conformation, with more of the population moving towards opening or intermixing of the strands that spatially keep the N to C termini away from each other – thus displaying a lower FRET efficiency compared to that of the freshly soluble form.

3.3. Elucidation of PolyQ aggregation and fibrillary morphology by AFM

While the UV-CD spectra provide information about the secondary structure of Q20 and Q44 and demonstrate disorder within both freshly solvated peptides, AFM offers valuable morphological information on polyQ peptide aggregation and fibrillation as a function of time. AFM confirms and directly monitors the growth of the disordered peptide structure. We first examined the bare mica surface as a control and calculated an average root mean roughness (Rq) of ~ 2 Å (Supplementary Figure S7). A grain analysis was applied in order to identify and count the spherical structures of polyQ peptide. We first investigated the early stage of polyQ peptide assembly at a low concentration (0.009 mg/ml) as depicted in Fig. 4. With a small scan size (500 nm²),

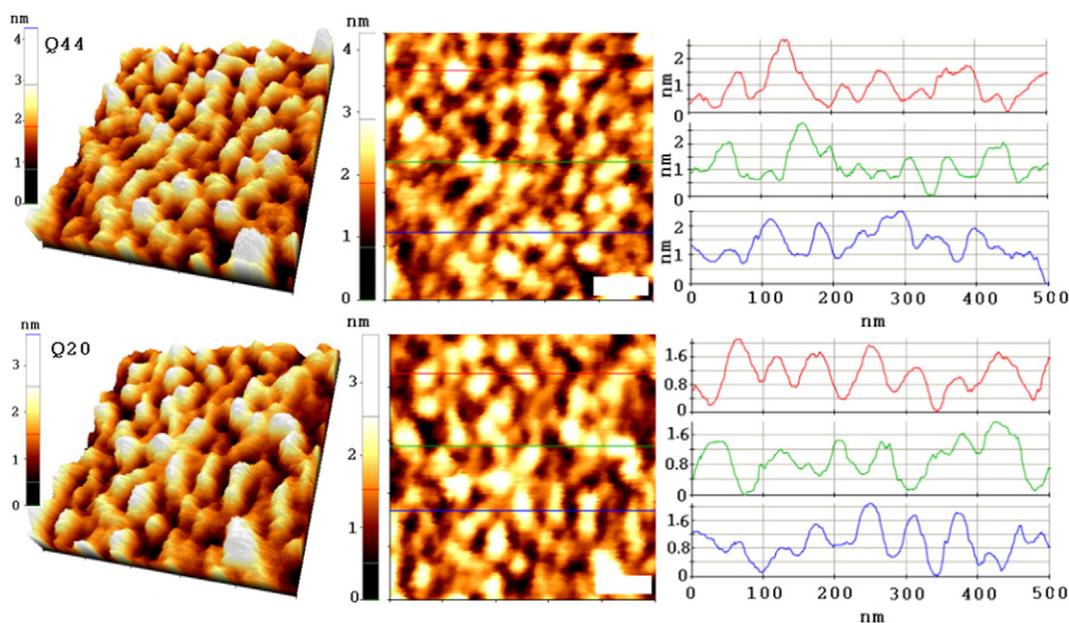


Fig. 4. The left panels show the height mode images of 3D and 2D for Q44 and Q20. The right panels show three different line profiles of the height mode. The scan size is 500 X 500 nm, and the scale bar is 100 nm.

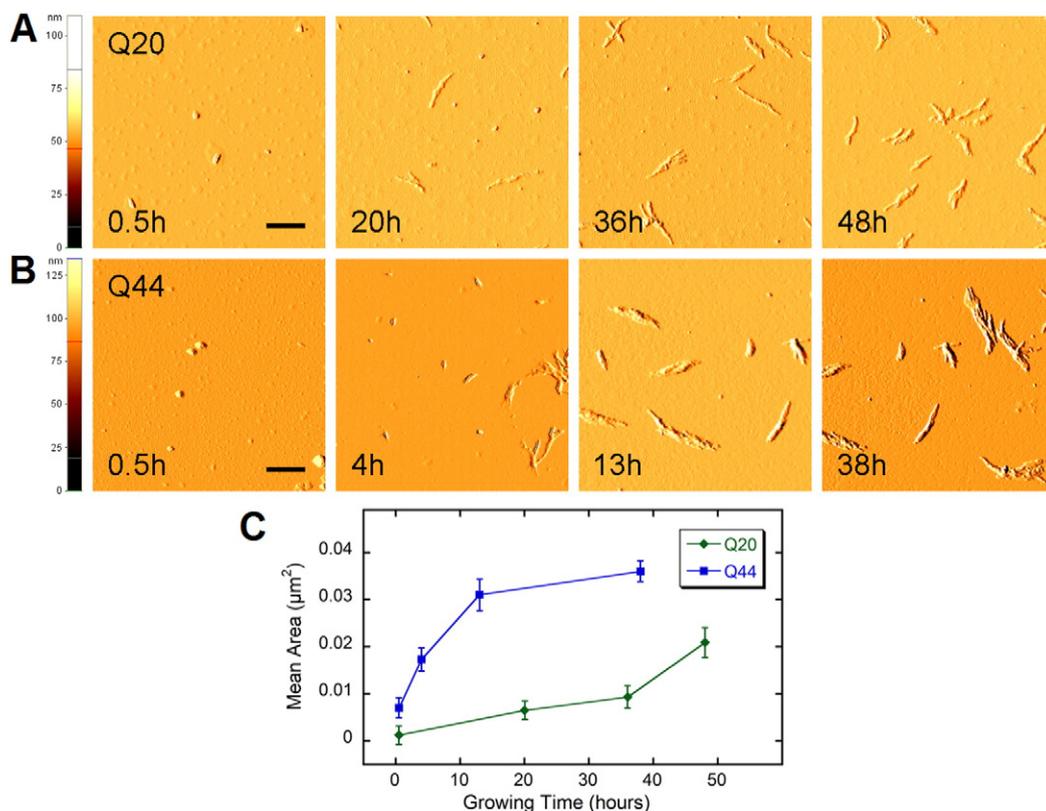


Fig. 5. AFM amplitude images of (A) Q20 and (B) Q44 at different time points during the aggregation process. From left to right are the representative images as a function of time; 0.5 to 48 h for Q20 and 0.5 to 38 h for Q44. (C) The plot values the calculated from the grain analysis by XE 1.5 software, indicating the rate of growth of the fibril structure of the peptides. Scale bar represent 500 nm.

both Q20 and Q44 form distinct spherical and annular oligomers with intermediate aggregates of globular structures (arrows). Putative prefibrillar polyQ structures have been reported using electron microscopes and resemble the spherical oligomers which were in agreement with our observations (Altschuler et al., 2000; Poirier et al., 2005; Tsukamoto et al., 2006; Vanschouwen et al., 2011). The detected spherical assemblies were relatively homogeneous with a diameter of 10–75 nm for both Q20 and Q44. Average height profiles for Q20 and Q44 are calculated to be 1.582 ± 0.06 nm and 2.065 ± 0.134 nm respectively (Fig. 4), indicating that fundamental self-assembly involves such aggregates. Incidentally, the size of the fundamental monomer of N-terminal Htt fragments containing 91 glutamines fused to different

affinity tags were seen to be 4.7 nm, which is slightly higher than twice the value for the structure obtained for Q44.

To further investigate the polyQ-mediated fibril assembly pathway and to directly compare it to the UV-CD data, we employed the same concentration (0.09 mg/ml) as used in the UV-CD experiments. The kinetics of AFM amplitude data are illustrated in Fig. 5. AFM analysis displays morphological features evolving over time intervals, which was in good agreement with the UV-CD results. Spherical assemblies start aggregating to form amyloid-like fiber structures in the early stages of deposition. Consistently, the number of spherical features decreased with a coincident increase in fibers as reported by other groups (Wong and Heremans, 1988). Though both Q20 and Q44 residues show similar

Table 1
Summary of key findings about the structure, folding and aggregation of Q20 and Q44.

Technique	Findings	
	Q20	Q44
<i>Structure</i>		
CD	Transition from disordered to β -sheet structure on aging	Transition from disordered to β -sheet structure on aging – 4× faster than Q20
FTIR	Disordered on solubilization Increased β -sheet structure on aging	β -sheet on solubilization, probably due to high concentration used
NMR	Hydrogen bonding between side chains drives polyQ folding Random coil structure is flexible and poorly defined Formation of large soluble aggregates on aging Inconsistent with polar zipper model	
<i>Folding and Aggregation</i>		
FRET	β -hairpin-like structure initially Intermixing of strands on aging	PolyQ fold punctuated by three β -hairpin-like structures initially Intermixing of strands on aging
AFM	Distinct spherical and annular oligomeric assemblies Homogenous intermediate aggregates of globular structures (prefibrillar polyQ) Nucleation of protofibrils occurs in x,y and z directions Q44 aggregation rate 4× faster than Q20	

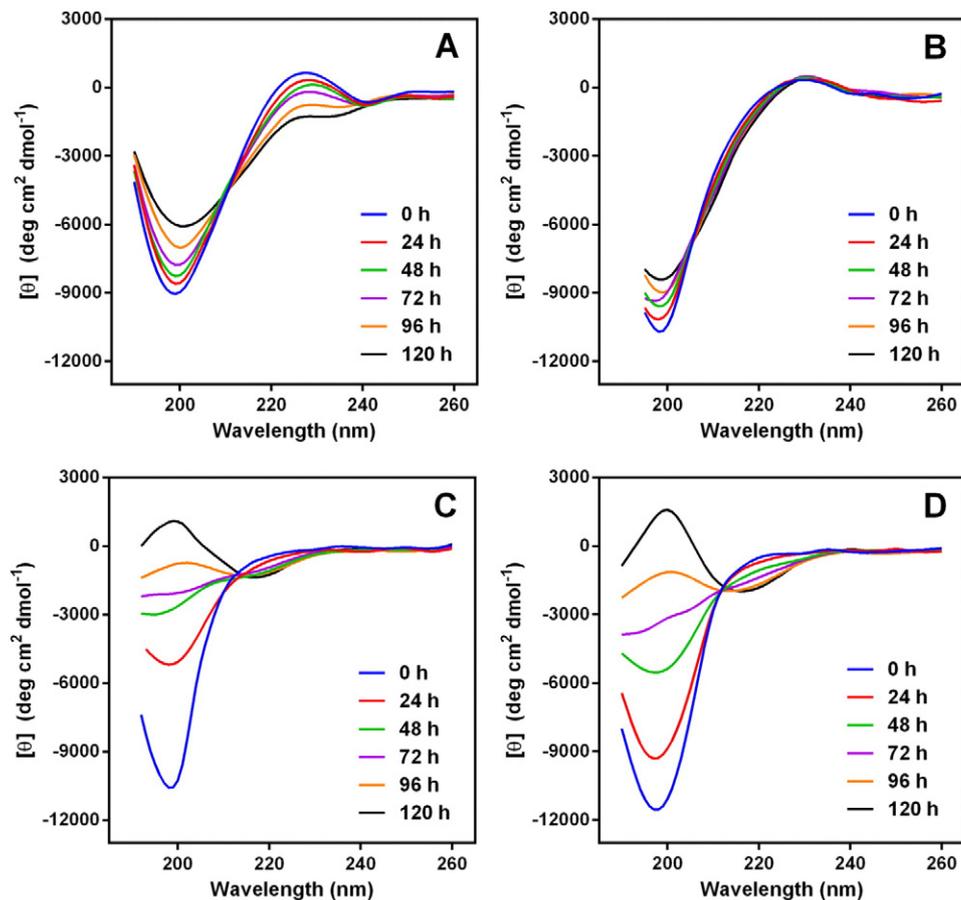


Fig. 6. Far-UV CD spectra of 0.3 mg/mL (A) Q3, (B) Q6, (C) Q10 and (D) Q14 showing the time-dependent conformational changes due to aggregation at different time points in 10 mM phosphate buffer, pH 7.4, at 25 °C. The CD spectra represent the average of ten spectral scans each of three experiments.

structural changes over 3 days of experimental schemes, the rate of Q44 peptide aggregation is different from that of the Q20 peptide. By grain analysis, the rate of growth of fibrillar structures of Q44 ($0.035 \mu\text{m}^2$) is $4 \times$ faster than that of Q20 ($0.0087 \mu\text{m}^2$), consistent with data derived from other spectroscopic techniques (Fig. 5C). It is noteworthy that amyloid growth occurred in all three directions as demonstrated in the series panels in Fig. 4A and B. The fibrillar structure tends to grow in a parallel fashion with a height of 200 nm, and it also grows in the z direction as shown in Fig. 5. This suggests that nucleation for the protofibrils is not restricted to the “head or tail” of the protofibrils.

3.4. Summary of findings for longer Q20 and Q44

Table 1 consolidates and compares the key findings about Q20 and Q44 from each of the aforementioned biophysical analyses.

3.5. Elucidation of conformation and folding of shorter PolyQ repeat lengths

CD spectra of Q3 and Q6 (Fig. 6A, B) indicate a random coil conformation initially with a characteristic negative peak at 198 nm. A weak positive band around 225–230 nm was also observed which may be due to the mixture of other conformations such as β -turn or polyproline-like structures (Rucker and Creamer, 2002). Time-dependent CD spectra of Q3 show disappearance of the positive band at 225 nm along with a decrease in the intensity of the negative peak at 198 nm, indicating a structural transition to neat random coil conformation. Q6 showed minimum change in the intensity of the negative peak at 198 nm and the weak positive band at 230 nm, which indicates that there is no significant conformational change on aging.

The CD spectra of Q10 and Q14 showed an initial random coil conformation (Fig. 6C, D). Incubation at 25 °C until 120 h showed a decrease in the intensity of the band at 198 nm and the appearance of a positive peak at 200 nm along with a negative band at 218 nm, characteristic of a β -sheet conformation. The isodichroic point at ~ 213 nm indicates a two-state transition process from a random coil to β -sheet structure on aggregation. These results are summarized in Table 2.

Further details about the intermolecular proximity of C and N terminal ends of polyQ were obtained in FRET experiments using dansyl as the N-terminal and Trp at the C-terminal of the chain. In the results depicted in the Fig. 7, Q3 revealed the 50% FRET. On increasing the length of the peptides viz., Q6, Q10, and Q14, the FRET decreases to 25%, 19%, and 14% respectively. The achieved declines are due to an increase in the distance between fluorophores in the N and C termini. In the case of Q20, 20% FRET was observed. The 6% increase from the Q14 FRET of 14% may be due to the formation of twisted or disordered loop structures in Q20, which increases the efficiency of energy transfer.

Table 2
Summary of the conformations of Q3, Q6, Q10, Q14 and Q20, initially and on aging, as elucidated from CD data.

Peptide	Conformation based on CD analysis	
	Initial	On aging
Q3	Random coil, mixed with other (e.g. β -turn, polyproline-like) structures	Neat random coil
Q6	Random coil, mixed with other (e.g. β -turn, polyproline-like) structures	Neat random coil
Q10	Random coil	Transition to β -sheet
Q14	Random coil	Transition to β -sheet
Q20	Disordered loop	Transition to β -sheet

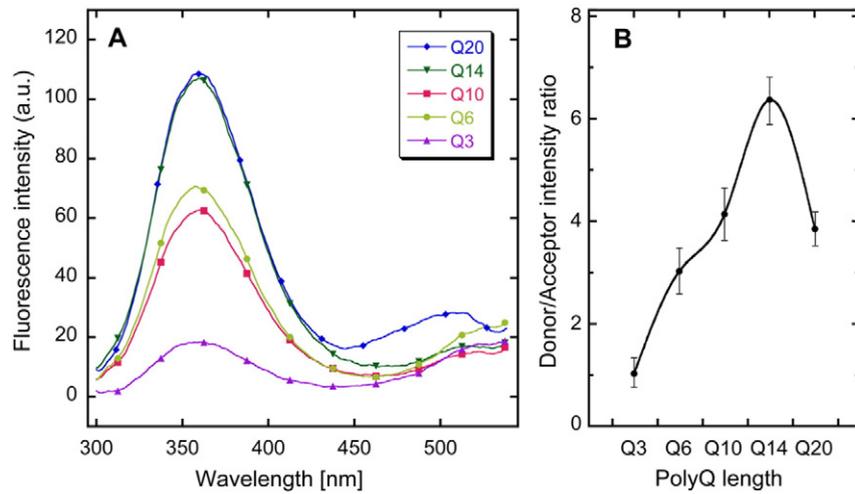


Fig. 7. (A) Fluorescence emission spectra of DanQ3WK, DanQ6WK, DanQ10WK, DanQ14WK and DanQ20WK at 0 h in the presence of molar excess of unlabelled Q3, Q6, Q10, Q14 and Q20 respectively (1:10). The spectra (dansyl–tryptophan FRET) were recorded in 10 mM phosphate buffer, pH 7.4 at 25 °C. Concentration of labeled peptide is 2 μ M and the total peptide concentration is 20 μ M. (B) Ratio of emission maximum (E_{max}) of donor/acceptor with respect to polyQ length.

Fluorescence emission spectra of dan-Q20 in the presence of a molar excess of unlabelled Q20 (1:10) on aging are depicted in Supplementary Figure S8. The intensity of emission increases to its maximum, and FRET efficiency is altered significantly, indicating structural transitions in the peptide fold. The Trp peak shows a blue shift on aging. Quantum mechanical calculations of excited state indole systems have characterized the energy transitions which give rise to wavelength shifts in fluorescence due to environmental effects. An instantaneous response

(relative to the fluorescence lifetime) by the environment, such as in a polar solvent, will stabilize the excited state, and result in fluorescence emission that is red-shifted as compared with a nonpolar environment (Eftink and Chiron, 1981; Valeur, 2001a). In principle, this result can be applied to the polyQ system, where the local environment of a Trp residue would be massively influenced by the nature of the proposed hydrophilic cleft as a result of the polyQ fold. Hence, the magnitude of the fluorescence transition and the resultant red shift in Trp fluorescence

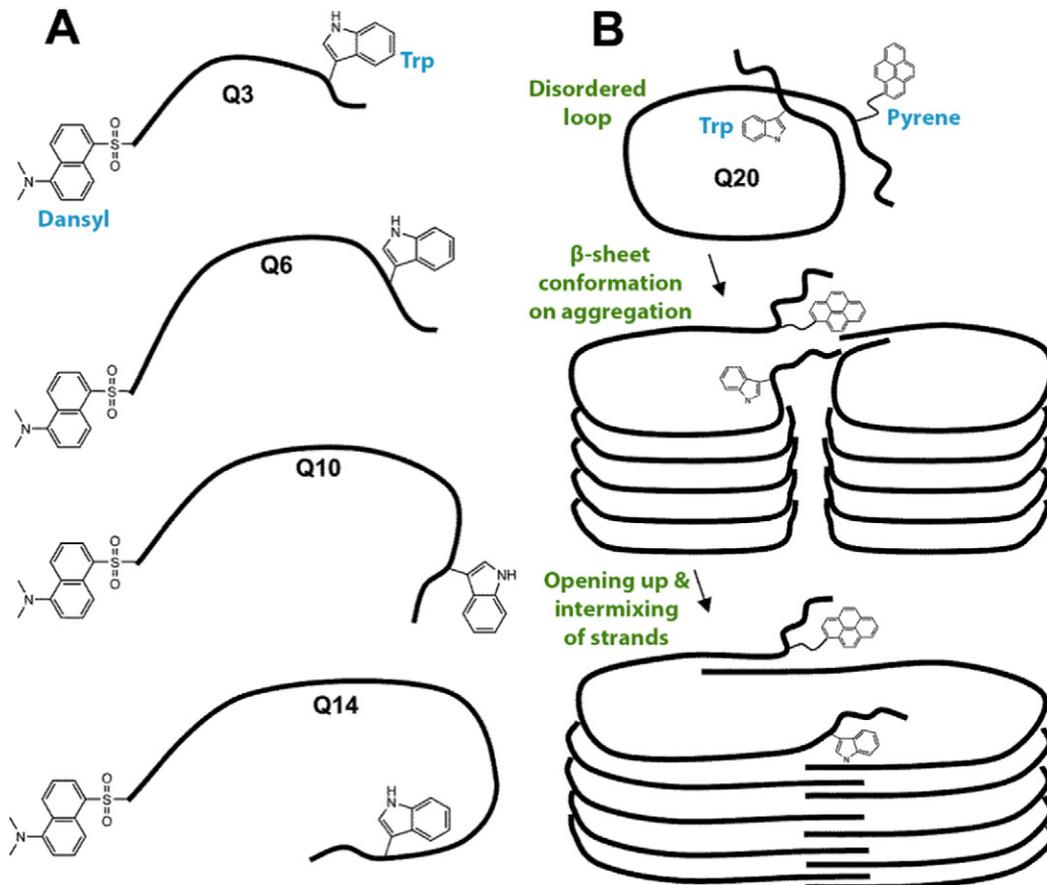


Fig. 8. Model for polyQ folding. (A) Schematic representation of the length dependence of polyQ peptides with fluorophores tryptophan and dansyl groups. (B) Schematic representation of the aggregation of Q20 in solutions along with fluorophores tryptophan and pyrene groups.

relative to free peptide can be interpreted based on the stabilization of the excited state. In contrast, we observed an increase in the fluorescence. This is due to protection of Trp by free water molecules in a bi-molecular quenching process. This phenomenon results in the relaxation of the excited state of the fluorophore by a non-radiative process, affecting the quantum yield and, thus, the intensity of the achieved fluorescence (Eftink and Ghiron, 1981; Semrouni et al., 2013; Valeur, 2001b). The rise in the intensity is predicted to be a consequence of a folded polyQ form. The dansyl peak also shows the distinct blue shift when compared to the other smaller homologs of the polyQ studied.

Based on the obtained results of length- and time-dependent aggregation studies by FRET, a schematic model for polyQ folding and aggregation is proposed as depicted in Fig. 8. The folding of polyQ peptides of different lengths with dansyl and Trp fluorophores in N and C-terminal positions respectively are shown with a molar excess of unlabelled peptide in Fig. 8A.

4. Discussion

4.1. Loop formation

Our work supports the model of instantaneous domain folding due to loop formation in the unstructured regions of both Q20 and Q44. This is in agreement with work by Ittah and Haas (1995), which proposed that non-local interactions stabilize long range loops in the initial folding intermediates of reduced BPTI – possibly protecting the side chains of glutamine from the solvent. The model can be visualized most simply by contact between two residues of an unfolded polypeptide chain. For generic polypeptides, using $n^{-3/2}$ scaling, the contact formation resulting from loop folding occurred in $\sim 3 \mu\text{s}$ for peptides of 10 residues, compared to $\sim 40 \mu\text{s}$ for peptides of 50 residues (Eaton et al., 2000). Folding of longer chains is associated with larger decreases in entropy while inter-residue contact formation is avoided in chains shorter than 10 residues due to chain stiffness – thereby predicting maximum probability of loop formation in polyQ sequences of 10 residues. This model, however, is based on a BPTI disulfide formation rate (Camacho and Thirumalai, 1995). Considering the higher persistent length of glutamine (and thus increased chain stiffness of polyQ), the maximum probability of loop formation could occur at a length higher than the predicted length of 10 residues for other polypeptides (Singh and Lapidus, 2008). Our results suggest that above 14 residues, polyQ peptides adopt a disordered loop structure instead of a random coil. Future studies could more precisely explore the idea of a threshold length for loop formation by comparing the conformational transitions in peptides ranging from 10 to 20 residues in length.

To interpret the results of this study, it is critical to understand that polyQ folding may alter solvent accessibilities to Trp and dansyl, and subsequently affect Trp fluorescence (Munishkina and Fink, 2007). This may account for the fact that total quenching of Trp fluorescence by dansyl was not observed in Q20 as in the case of pyrene. The FRET data shows that the distance between the Trp-dansyl FRET pair is in the range of 32 Å. By subtracting the non-specific spacer extension estimate of 5.8 Å, the distance between the two centers becomes 26.2 Å. This distance is well above that expected for the proposed disordered loop structure. However, it should be noted that the dansyl extension might be a non-participant in loop formation – a non-participating three-residue amino acid will place the Trp away from the C-terminal group more than 11 Å, which may explain the discrepancy between the two FRET pair experiments.

4.2. Fibril formation

Understanding of the polyQ fibrillation process is complicated by the hydrophilic amino acid content of polyQ peptides. Though polyQ peptides are composed totally of hydrophilic Gln sequences, they share common gross structural features with other amyloid fibrils

(Hao et al., 2010; Nelson et al., 2005; Sambasivam et al., 2011). For instance, polyQ peptides form unbranched fibrils in a cross- β core structure, with a diameter of 4–12 nm, in which continuous β -sheets are formed with β -strands running perpendicular to the long axis of fibrils (Altschuler et al., 2000; Masino et al., 2002; Masino et al., 2004; Navon et al., 2001; Poirier et al., 2005; Vanschouwen et al., 2011). Most of these pathogenic peptides appear to be unstructured in their globular form, but undergo structural conversion during fibril formation to adopt a substantial amount of β -sheet character.

Previous NMR and CD studies of polyQ appended with soluble globular proteins revealed the presence of unstructured polyQ domains (Bulone et al., 2006; Kim et al., 2009; Robertson et al., 2008); this is compatible with the disordered loop structure of Q20 and Q44 observed in our study. However, based on XRD data, Perutz proposed a β -helical model in which polypeptides contain a unique “single-layer” β -sheet segment (Iengar et al., 2006; Lodish et al., 2000). The single-layer β -sheet segment consists of β -strands (20 residue per turn), and contains side chains of Gln on both faces of the sheet (Cheng et al., 2013). This amino acid sequence is distinct from the amphipathic nature of most anti-parallel β -sheets in natural proteins. Most studies on fibril formation to date have been performed on proteins and peptides that are either amphipathic or highly hydrophobic (Bulone et al., 2006; Poirier et al., 2002; Tanaka et al., 2003; Zanuy et al., 2006). Despite the lack of a hydrophobic core in this region, the single-layer β -sheet segment is highly stable, and stability can be attributed to side chain hydrogen bonding between Gln residues (Koide et al., 2000). We previously found that the single-layer β -sheet segment buries its polar side chain with an energetic cost similar to that found in structures triggered by cavitation forces (Satheeshkumar and Jayakumar, 2003). This is accomplished by the formation of “hydrophobic layers” consisting of aliphatic moieties of amino acid residues normally classified as hydrophilic, such as Gln, Glu, and Lys (Moses et al., 2003).

Moreover, other studies show that as soon as a polyQ peptide is cleaved, it adopts a folded conformation which appears to be toxic in the soluble condition (Altschuler et al., 1997). These observations suggest that the ability to form toxic structures is nucleation-dependent, involving a specific conformer as the amyloidogenic species that induces template-dependent fibrillation (Bhak et al., 2009) – in agreement with our AFM results that demonstrated fibrillar growth in all three directions.

4.3. Summary

It is known that the backbone of an unfolded polypeptide has some propensity for adopting the C-N terminals in close proximity if the persistence length falls within a few amino acids (Krishna and Englander, 2005), as well as moderate propensity for adopting extended polyproline-II helical structures, which could bend to form a circular polypeptide fold as proposed in this work (Rucker and Creamer, 2002). Using a short alanine peptide model, Pappu and Rose (2002) have shown that conformations corresponding to the general polyproline-II structure are well-represented, even when just steric interactions are taken into account. The ensemble of conformations is likely to be far smaller than that of previously reported data (Vitalis et al., 2007) on polyQ peptide folds and is also likely to possess persistent structural elements with insipient β -helical folding modules.

The proposed model for the folding of polyQ of different lengths and the aggregation of Q20 is depicted in the Fig. 8B, based on the obtained results for the length and time dependent studies. Extended loop structure first form β -sheet structures; domain swapping between different peptides in the assemblages next leads to the formation of mixed aggregates. The Wetzel group (Chen et al., 2002a; Chen et al., 2002b; Chen and Wetzel, 2001) has proposed polyQ aggregation could be associated with femto molar concentrations of intermediates in the presence of micromolar concentrations of polyQ. The proposed disordered loop structure may not favor formation of such intermediates for the aggregation process.

However, longer polyQ repeats may have a higher percentage of ordered structures that would facilitate intermediate formation.

In summary, the flexible disordered region of polyQ peptides have the ability to adopt an ordered structure in which the C and N termini occur in close proximity, above the critical length of 14 polyQ repeats. The FRET data indicate that peptides with a higher number of Q repeats (Q44) fold into self-catenated forms with a more β -sheet structure compared to peptides with a shorter repeat length (Q20). Aging of such loops resulted in more β -sheet formation and the intermixing of different peptide strands within the assembling monomers.

5. Conclusion

5.1. Therapeutic possibilities

Using a range of protein-characterizing technologies we have been able to further elucidate the structure of polyQ aggregates, providing us with a platform to suggest a novel mechanism of polyQ folding and aggregation. This mechanistic model offers a new insight into the molecular pathophysiology of diseases such as HD, spinocerebellar ataxia, spinobulbar muscular dystrophy and dentatorubral-pallidoluysian atrophy (Fan et al., 2014) – with potential ramifications for future therapeutic efforts aimed at preventing the pathological misfolding and oligomerisation of polyQ peptides underlying these diseases (Fig. 9). For instance, TRiC is an endogenous molecular chaperone that has been found to inhibit polyQ misfolding and aggregation in HD by binding to the N17 terminal of mutant Htt peptide fragments (Arrasate and Finkbeiner, 2012). Drugs that induce expression of these chaperones or prevent their degradation via the VRK2 pathway can thus prevent the conformational changes that lead to oligomerisation of polyQ peptide fragments (Kim and Kim, 2014). One agent that has displayed pharmacological promise in cellular and fly studies is the benzyl pyrazole derivative HSF1A, which is thought to interact with TRiC to enhance its anti-aggregative effects (Calamini and Morimoto, 2012; Neef et al., 2011; Schipper-Krom et al., 2012).

5.2. Disease modeling

Greater polyQ repeat length in Huntington's Disease is associated with both an earlier age-of-presentation, and worse clinical severity – due to the accelerated aggregation of fragmented mutant Htt peptides.

As such, the effects of polyQ repeat length have important implications for screening and prognosis. It has long been established that the majority of people with HD have more than 36 CAG repeats on the htt gene, compared with the normal 6–36 repeats in unaffected individuals – and given the autosomal dominant mode of inheritance, only one allele needs to have a trinucleotide repeat expansion beyond this threshold for a person to develop HD. While it is rare for a patient to be homozygous for mutant alleles for the Htt gene, there is evidence to suggest that homozygotes have a more severe form of disease (but not necessarily a decreased age of onset) (Squitieri et al., 2003). Our study provides evidence for significant molecular-level aggregation events beyond as few as 14 polyQ residues, which could suggest that heterozygous individuals who have one expanded mutant allele, as well as another “normal” allele with more than 14 polyQ repeats, may actually develop a more severe form of disease. Further research is warranted into this possibility, as well as towards linking the critical polypeptide misfolding threshold, identified here as 14 glutamine residues, with the clinically observed genetic threshold of 36 CAG trinucleotide repeats.

Further elucidation of the folding and aggregation process of polyQ peptides paves the way for better disease modeling. This in turn promises to enhance our understanding of how polyQ aggregates interact to disrupt cellular processes, and ultimately, supply us with the link between neurotoxicity and the clinical manifestations of these conditions. Better models of disease may also shed further insight into the neuroprotective mechanism whereby polyQ aggregates form intraneuronal inclusions, which was alluded to in Fig. 1. It is our hope that future studies could form a platform for examining other proteinopathies with a view of altering the disease course. Protein-misfolding disorders are notoriously refractory to current treatment efforts; yet, if there is a therapeutic answer to be found, it undoubtedly lies in understanding the molecular basis of the pathology. Efforts such as this may prove critical in establishing the foundations that perhaps, one day, will lead us to that goal.

Acknowledgments

We thank Prof. Ron Kopito for his support and insightful discussion during the initiation of this work. We thank Dr. Eric Bennett for his help in editing this manuscript, Prof. R. L Baldwin for proposing NMR experiments, and Prof. Mike Fayer and his colleagues in the Department of Chemistry for their help with infrared measurements.

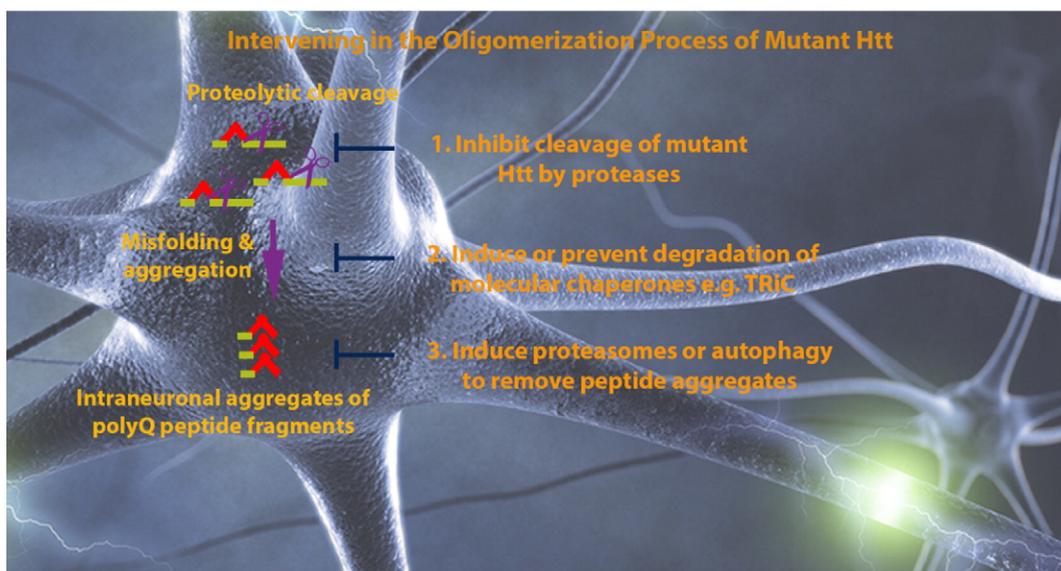


Fig. 9. Pharmacological agents that 1) inhibit protease-mediated cleavage of mutant Htt, 2) induce molecular chaperones such as TRiC and 3) induce proteasome-mediated removal of polyQ aggregates may be useful in preventing the pathological misfolding and oligomerisation that occur in HD.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.npep.2016.01.011>.

References

- Altschuler, E.L., Hud, N.V., Mazrimas, J.A., Rupp, B., 1997. Random coil conformation for extended polyglutamine stretches in aqueous soluble monomeric peptides. *J. Pept. Res.* 50, 73–75.
- Altschuler, E.L., Hud, N.V., Mazrimas, J.A., Rupp, B., 2000. Structure of polyglutamine. *FEBS Lett.* 472, 166–168.
- Arrasate, M., Finkbeiner, S., 2012. Protein aggregates in Huntington's disease. *Exp. Neurol.* 238, 1–11.
- Bates, G., 2003. Huntingtin aggregation and toxicity in Huntington's disease. *Lancet* 361, 1642–1644.
- Bhak, G., Choe, Y.J., Paik, S.R., 2009. Mechanism of amyloidogenesis: nucleation-dependent fibrillation versus double-concerted fibrillation. *BMB Rep.* 42, 541–551.
- Bhattacharyya, A., Thakur, A.K., Chellgren, V.M., Thiagarajan, G., Williams, A.D., Chellgren, B.W., Creamer, T.P., Wetzel, R., 2006. Oligoproline effects on polyglutamine conformation and aggregation. *J. Mol. Biol.* 355, 524–535.
- Billeter, M., Wagner, G., Wuthrich, K., 2008. Solution NMR structure determination of proteins revisited. *J. Biomol. NMR* 42, 155–158.
- Blum, E.S., Schwendenam, A.R., Shaham, S., 2013. PolyQ disease: misfiring of a developmental cell death program? *Trends Cell Biol.* 23, 168–174.
- Bulone, D., Masino, L., Thomas, D.J., Biagio, P.L.S., Pastore, A., 2006. The interplay between PolyQ and protein context delays aggregation by forming a reservoir of protofibrils. *PLoS One* 1, e111.
- Calamini, B., Morimoto, R.I., 2012. Protein homeostasis as a therapeutic target for diseases of protein conformation. *Curr. Top. Med. Chem.* 12, 2623–2640.
- Camacho, C.J., Thirumalai, D., 1995. Theoretical predictions of folding pathways by using the proximity rule, with applications to bovine pancreatic trypsin inhibitor. *Proc. Natl. Acad. Sci. U. S. A.* 92, 1277–1281.
- Chan, W.M., Tsoi, H., Wu, C.C., Wong, C.H., Cheng, T.C., Li, H.Y., Lau, K.F., Shaw, P.C., Perrimon, N., Chan, H.Y., 2011. Expanded polyglutamine domain possesses nuclear export activity which modulates subcellular localization and toxicity of polyQ disease protein via exportin-1. *Hum. Mol. Genet.* 20, 1738–1750.
- Chen, S., Wetzel, R., 2001. Solubilization and disaggregation of polyglutamine peptides. *Protein Sci.* 10, 887–891.
- Chen, S., Berthelie, V., Yang, W., Wetzel, R., 2001. Polyglutamine aggregation behavior in vitro supports a recruitment mechanism of cytotoxicity. *J. Mol. Biol.* 311, 173–182.
- Chen, S., Berthelie, V., Hamilton, J.B., O'Nuallain, B., Wetzel, R., 2002a. Amyloid-like features of polyglutamine aggregates and their assembly kinetics. *Biochemistry* 41, 7391–7399.
- Chen, S., Ferrone, F.A., Wetzel, R., 2002b. Huntington's disease age-of-onset linked to polyglutamine aggregation nucleation. *Proc. Natl. Acad. Sci. U. S. A.* 99, 11884–11889.
- Cheng, P.N., Pham, J.D., Nowick, J.S., 2013. The supramolecular chemistry of beta-sheets. *J. Am. Chem. Soc.* 135, 5477–5492.
- Chintapalli, S.V., Illingworth, C.J., Upton, G.J., Sacquin-Mora, S., Reeves, P.J., Mohammedali, H.S., Reynolds, C.A., 2014. Assessing the effect of dynamics on the closed-loop protein-folding hypothesis. *J. R. Soc. Interface* 11, 20130935.
- Cisbani, G., Cicchetti, F., 2012. An in vitro perspective on the molecular mechanisms underlying mutant huntingtin protein toxicity. *Cell Death Dis.* 3, e382.
- Crick, S.L., Jayaraman, M., Frieden, C., Wetzel, R., Pappu, R.V., 2006. Fluorescence correlation spectroscopy shows that monomeric polyglutamine molecules form collapsed structures in aqueous solutions. *Proc. Natl. Acad. Sci. U. S. A.* 103, 16764–16769.
- Davranche, A., Aviat, H., Zeder-Lutz, G., Busso, D., Altschuh, D., Trotter, Y., Klein, F.A., 2011. Huntingtin affinity for partners is not changed by polyglutamine length: aggregation itself triggers aberrant interactions. *Hum. Mol. Genet.* 20, 2795–2806.
- Dong, A., Huang, P., Caughey, W.S., 1990. Protein secondary structures in water from 2nd-derivative amide-I infrared-spectra. *Biochemistry* 29, 3303–3308.
- Eaton, W.A., Munoz, V., Hagen, S.J., Jas, G.S., Lapidus, L.J., Henry, E.R., Hofrichter, J., 2000. Fast kinetics and mechanisms in protein folding. *Annu. Rev. Biophys. Biomol. Struct.* 29, 327–359.
- Eftink, M.R., Ghiron, C.A., 1981. Fluorescence quenching studies with proteins. *Anal. Biochem.* 114, 199–227.
- Fan, H.C., Ho, L.I., Chi, C.S., Chen, S.J., Peng, G.S., Chan, T.M., Lin, S.Z., Harn, H.J., 2014. Polyglutamine (PolyQ) diseases: genetics to treatments. *Cell Transplant.* 23, 441–458.
- Finke, J.M., Cheung, M.S., Onuchic, J.N., 2004. A structural model of polyglutamine determined from a host-guest method combining experiments and landscape theory. *Biophys. J.* 87, 1900–1918.
- Ghavami, M., Rezaei, M., Ejtehadi, R., Lotfi, M., Shokrgozar, M.A., Abd Emamy, B., Raush, J., Mahmoudi, M., 2012. Physiological temperature has a crucial role in amyloid beta in the absence and presence of hydrophobic and hydrophilic nanoparticles. *ACS Chem. Neurosci.* 4, 375–378.
- Greenfield, N.J., 2006. Using circular dichroism spectra to estimate protein secondary structure. *Nat. Protoc.* 1, 2876–2890.
- Hands, S.L., Wyttenbach, A., 2010. Neurotoxic protein oligomerisation associated with polyglutamine diseases. *Acta Neuropathol.* 120, 419–437.
- Hao, J., Zhang, W., Zhang, P., Liu, R., Liu, L., Lei, G., Su, C., Miao, J., Li, Z., 2010. Abeta20–29 peptide blocking apoE/abeta interaction reduces full-length Abeta42/40 fibril formation and cytotoxicity in vitro. *Neuropeptides* 44, 305–313.
- Haris, P.I., 2010. Can infrared spectroscopy provide information on protein–protein interactions? *Biochem. Soc. Trans.* 38, 940–946.
- Hatters, D.M., 2012. Putting huntingtin “aggregation” in view with windows into the cellular milieu. *Curr. Top. Med. Chem.* 12, 2611–2622.
- Hoffner, G., Djan, P., 2014. Monomeric, oligomeric and polymeric proteins in Huntington disease and other diseases of polyglutamine expansion. *Brain Sci.* 4, 91–122.
- Iengar, P., Joshi, N.V., Balam, P., 2006. Conformational and sequence signatures in beta helix proteins. *Structures* 14, 529–542.
- Ittah, V., Haas, E., 1995. Nonlocal interactions stabilize long range loops in the initial folding intermediates of reduced bovine pancreatic trypsin inhibitor. *Biochemistry* 34, 4493–4506.
- Janciauskiene, S., Wright, H.T., Lindgren, S., 1999. Fibrillar Alzheimer's amyloid peptide abeta(1–42) stimulates low density lipoprotein binding and cell association, free radical production and cell cytotoxicity in PC12 cells. *Neuropeptides* 33, 510–516.
- Kar, K., Jayaraman, M., Sahoo, B., Kodali, R., Wetzel, R., 2011. Critical nucleus size for disease-related polyglutamine aggregation is repeat-length dependent. *Nat. Struct. Mol. Biol.* 18, 328–336.
- Katsuno, M., Banno, H., Suzuki, K., Takeuchi, Y., Kawashima, M., Tanaka, F., Adachi, H., Sobue, G., 2008. Molecular genetics and biomarkers of polyglutamine diseases. *Curr. Mol. Med.* 8, 221–234.
- Kim, S., Kim, K.T., 2014. Therapeutic approaches for inhibition of protein aggregation in Huntington's disease. *Exp. Neurobiol.* 23, 36–44.
- Kim, M.W., Chelliah, Y., Kim, S.W., Otwinowski, Z., Bezprozvanny, I., 2009. Secondary structure of huntingtin amino-terminal region. *Structures* 17, 1205–1212.
- Koide, S., Huang, X., Link, K., Koide, A., Bu, Z., Engelman, D.M., 2000. Design of single-layer beta-sheets without a hydrophobic core. *Nature* 403, 456–460.
- Krishna, M.M., Englander, S.W., 2005. The N-terminal to C-terminal motif in protein folding and function. *Proc. Natl. Acad. Sci. U. S. A.* 102, 1053–1058.
- La Spada, A.R., Weydt, P., Pineda, V.V., 2011. Huntington's disease pathogenesis: mechanisms and pathways. In: Lo, D.C., Hughes, R.E. (Eds.), *Neurobiology of Huntington's Disease: Applications to Drug Discovery*. CRC Press, Boca Raton (FL).
- Lakowicz, J.R., 2006. *Principles of Fluorescence Spectroscopy*. 3 ed. Springer, London (Limited).
- Landrum, E., Wetzel, R., 2014. Biophysical underpinnings of the repeat length dependence of polyglutamine amyloid formation. *J. Biol. Chem.* 289, 10254–10260.
- Legleiter, J., Mitchell, E., Lotz, G.P., Sapp, E., Ng, C., DiFiglia, M., Thompson, L.M., Muchowski, P.J., 2010. Mutant huntingtin fragments form oligomers in a polyglutamine length-dependent manner in vitro and in vivo. *J. Biol. Chem.* 285, 14777–14790.
- Lodish, H., BA, Zipursky, S.L., et al., 2000. *Hierarchical Structure of Proteins*. In: Freeman, W.H. (Ed.), *Molecular Cell Biology*, fourth ed. (New York).
- Mahmoudi, M., Monopoli, M.P., Rezaei, M., Lynch, I., Bertoli, F., McManus, J.J., Dawson, K.A., 2013a. The protein Corona Mediates the impact of nanomaterials and slows amyloid Beta Fibrillation. *ChemBiochem* 14, 568–572.
- Mahmoudi, M., Quinlan-Pluck, F., Monopoli, M.P., Sheibani, S., Vali, H., Dawson, K.A., Lynch, I., 2013b. Influence of the physicochemical properties of superparamagnetic iron oxide nanoparticles on amyloid β protein fibrillation in solution. *ACS Chem. Neurosci.* 4, 475–485.
- Manning, M.C., 2005. Use of infrared spectroscopy to monitor protein structure and stability. *Expert Rev. Proteomics* 2, 731–743.
- Masino, L., Kelly, G., Leonard, K., Trotter, Y., Pastore, A., 2002. Solution structure of polyglutamine tracts in GST-polyglutamine fusion proteins. *FEBS Lett.* 513, 267–272.
- Masino, L., Nicastro, G., Menon, R.P., Dal Piaz, F., Calder, L., Pastore, A., 2004. Characterization of the structure and the amyloidogenic properties of the josephin domain of the polyglutamine-containing protein ataxin-3. *J. Mol. Biol.* 344, 1021–1035.
- Moses, J.P., Satheeshkumar, K.S., Murali, J., Alli, D., Jayakumar, R., 2003. Self-assembly of the synthetic polymer (leu-glu)(n): An amyloid-like structure formation. *Langmuir* 19, 3413–3418.
- Munishkina, L.A., Fink, A.L., 2007. Fluorescence as a method to reveal structures and membrane-interactions of amyloidogenic proteins. *Biochim. Biophys. Acta* 1768, 1862–1885.
- Napoleitano, F., Occhi, S., Calamita, P., Volpi, V., Blanc, E., Charroux, B., Royet, J., Fanto, M., 2011. Polyglutamine atrophin provokes neurodegeneration in drosophila by repressing fat. *EMBO J.* 30, 945–958.
- Natalello, A., Frana, A.M., Relini, A., Apicella, A., Invernizzi, G., Casari, C., Gliozzi, A., Doglia, S.M., Tortora, P., Regonesi, M.E., 2011. A major role for side-chain polyglutamine hydrogen bonding in irreversible ataxin-3 aggregation. *PLoS One* 6, e18789.
- Navon, A., Ittah, V., Landsman, P., Scheraga, H.A., Haas, E., 2001. Distributions of intramolecular distances in the reduced and denatured states of bovine pancreatic ribonuclease A. Folding initiation structures in the C-terminal portions of the reduced protein. *Biochemistry* 40, 105–118.
- Neef, D.W., Jaeger, A.M., Thiele, D.J., 2011. Heat shock transcription factor 1 as a therapeutic target in neurodegenerative diseases. *Nat. Rev. Drug Discov.* 10, 930–944.
- Nelson, R., Sawaya, M.R., Balbirnie, M., Madsen, A.O., Riekel, C., Grothe, R., Eisenberg, D., 2005. Structure of the cross-beta spine of amyloid-like fibrils. *Nature* 435, 773–778.
- Ono, K., Condrion, M.M., Teplow, D.B., 2009. Structure-neurotoxicity relationships of amyloid beta-protein oligomers. *Proc. Natl. Acad. Sci. U. S. A.* 106, 14745–14750.
- Orr, H.T., 2012. Polyglutamine neurodegeneration: expanded glutamines enhance native functions. *Curr. Opin. Genet. Dev.* 22, 251–255.
- Ortega, Z., Diaz-Hernandez, M., Maynard, C.J., Hernandez, F., Dantuma, N.P., Lucas, J.J., 2010. Acute polyglutamine expression in inducible mouse model unravels ubiquitin/proteasome system impairment and permanent recovery attributable to aggregate formation. *J. Neurosci.* 30, 3675–3688.
- Pappu, R.V., Rose, G.D., 2002. A simple model for polyproline II structure in unfolded states of alanine-based peptides. *Protein Sci.* 11, 2437–2455.
- Perney, N.M., Braddick, L., Jurna, M., Garbacik, E.T., Offerhaus, H.L., Serpell, L.C., Blanch, E., Holden-Dye, L., Brocklesby, W.S., Melvin, T., 2012. Polyglutamine aggregate structure

- in vitro and in vivo; new avenues for coherent anti-stokes Raman scattering microscopy. *PLoS One* 7, e40536.
- Perutz, M.F., Windle, A.H., 2001. Cause of neural death in neurodegenerative diseases attributable to expansion of glutamine repeats. *Nature* 412, 143–144.
- Perutz, M.F., Johnson, T., Suzuki, M., Finch, J.T., 1994. Glutamine repeats as polar zippers: their possible role in inherited neurodegenerative diseases. *Proc. Natl. Acad. Sci. U. S. A.* 91, 5355–5358.
- Perutz, M.F., Pope, B.J., Owen, D., Wanker, E.E., Scherzinger, E., 2002. Aggregation of proteins with expanded glutamine and alanine repeats of the glutamine-rich and asparagine-rich domains of Sup35 and of the amyloid beta-peptide of amyloid plaques. *Proc. Natl. Acad. Sci. U. S. A.* 99, 5596–5600.
- Poirier, M.A., Li, H., Macosko, J., Cai, S., Amzel, M., Ross, C.A., 2002. Huntingtin spheroids and protofibrils as precursors in polyglutamine fibrillization. *J. Biol. Chem.* 277, 41032–41037.
- Poirier, M.A., Jiang, H., Ross, C.A., 2005. A structure-based analysis of huntingtin mutant polyglutamine aggregation and toxicity: evidence for a compact beta-sheet structure. *Hum. Mol. Genet.* 14, 765–774.
- Robertson, A.L., Horne, J., Ellisdon, A.M., Thomas, B., Scanlon, M.J., Bottomley, S.P., 2008. The structural impact of a polyglutamine tract is location-dependent. *Biophys. J.* 95, 5922–5930.
- Robertson, A.L., Headey, S.J., Saunders, H.M., Ecroyd, H., Scanlon, M.J., Carver, J.A., Bottomley, S.P., 2010. Small heat-shock proteins interact with a flanking domain to suppress polyglutamine aggregation. *Proc. Natl. Acad. Sci. U. S. A.* 107, 10424–10429.
- Rucker, A.L., Creamer, T.P., 2002. Polyproline II helical structure in protein unfolded states: lysine peptides revisited. *Protein Sci.* 11, 980–985.
- Sambasivam, D., Sivanesan, S., Ashok, B.S., Rajadas, J., 2011. Structural preferences of abeta fragments in different micellar environments. *Neuropeptides* 45, 369–376.
- Satheeshkumar, K.S., Jayakumar, R., 2003. Conformational polymorphism of the amyloidogenic peptide homologous to residues 113–127 of the prion protein. *Biophys. J.* 85, 473–483.
- Schipper-Krom, S., Juenemann, K., Reits, E.A., 2012. The ubiquitin-proteasome system in Huntington's disease: are proteasomes impaired, initiators of disease, or coming to the rescue? *Biochem. Res. Int.* 2012, 837015.
- Semrouni, D., Clavaguera, C., Ohanessian, G., Parks, J.H., 2013. Relationship between conformational dynamics and electron transfer in a desolvated peptide. Part I. Structures. *J. Phys. Chem. B* 117, 1746–1755.
- Shao, J., Diamond, M.I., 2007. Polyglutamine diseases: emerging concepts in pathogenesis and therapy. *Hum. Mol. Genet.* 16, R115–R123 (Spec No. 2).
- Singh, V.R., Lapidus, L.J., 2008. The intrinsic stiffness of polyglutamine peptides. *J. Phys. Chem. B* 112, 13172–13176.
- Squitieri, F., Gellera, C., Cannella, M., Mariotti, C., Cislighi, G., Rubinsztein, D.C., Almqvist, E.W., Turner, D., Bachoud-Levi, A.C., Simpson, S.A., Delatycki, M., Maglione, V., Hayden, M.R., Donato, S.D., 2003. Homozygosity for CAG mutation in Huntington disease is associated with a more severe clinical course. *Brain* 126, 946–955.
- Sreerama, N., Woody, R.W., 2000. Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal. Biochem.* 287, 252–260.
- Tanaka, M., Machida, Y., Nishikawa, Y., Akagi, T., Hashikawa, T., Fujisawa, T., Nukina, N., 2003. Expansion of polyglutamine induces the formation of quasi-aggregate in the early stage of protein fibrillization. *J. Biol. Chem.* 278, 34717–34724.
- Thakur, A.K., Wetzel, R., 2002. Mutational analysis of the structural organization of polyglutamine aggregates. *Proc. Natl. Acad. Sci. U. S. A.* 99, 17014–17019.
- Tsukamoto, K., Shimizu, H., Ishida, T., Akiyama, Y., Nukina, N., 2006. Aggregation mechanism of polyglutamine diseases revealed using Quantum Chemical calculations, fragment molecular orbital calculations, molecular dynamics simulations, and binding free energy calculations. *J. Mol. Struct. THEOCHEM* 778, 85–95.
- Valeur, B., 2001a. Effect of Polarity on Fluorescence Emission. *Polarity Probes, Molecular Fluorescence*. Wiley-VCH Verlag GmbH, pp. 200–225.
- Valeur, B., 2001b. Resonance energy transfer and its applications, molecular fluorescence. *Wiley-VCH Verlag GmbH*, pp. 247–272.
- Vanschouwen, B.M., Oblinsky, D.G., Gordon, H.L., Rothstein, S.M., 2011. Structure propensities in mutated polyglutamine peptides. *Interdiscip. Sci.* 3, 1–16.
- Vitalis, A., Wang, X., Pappu, R.V., 2007. Quantitative characterization of intrinsic disorder in polyglutamine: insights from analysis based on polymer theories. *Biophys. J.* 93, 1923–1937.
- Waelter, S., Boeddrich, A., Lurz, R., Scherzinger, E., Lueder, G., Lehrach, H., Wanker, E.E., 2001. Accumulation of mutant huntingtin fragments in aggresome-like inclusion bodies as a result of insufficient protein degradation. *Mol. Biol. Cell* 12, 1393–1407.
- Wetzel, R., 2012. Physical chemistry of polyglutamine: intriguing tales of a monotonous sequence. *J. Mol. Biol.* 421, 466–490.
- Wong, P.T., Heremans, K., 1988. Pressure effects on protein secondary structure and hydrogen deuterium exchange in chymotrypsinogen: a Fourier transform infrared spectroscopic study. *Biochim. Biophys. Acta* 956, 1–9.
- Zanuy, D., Gunasekaran, K., Lesk, A.M., Nussinov, R., 2006. Computational study of the fibril organization of polyglutamine repeats reveals a common motif identified in beta-helices. *J. Mol. Biol.* 358, 330–345.
- Zheng, Z., Diamond, M.I., 2012. Huntington disease and the huntingtin protein. *Prog. Mol. Biol. Transl. Sci.* 107, 189–214.