Putting huntingtin “aggregation” in view with windows into the cellular milieu.†

Danny M. Hatters¹,*

¹ Department of Biochemistry and Molecular Biology
Bio21 Molecular Science and Biotechnology Institute
30 Flemington Road
The University of Melbourne
Melbourne VIC 3010 Australia
Ph: +61 3 8344 2530
Fax: +61 3 9348 1421
Email: dhatters@unimelb.edu.au

Corresponding author:
* Corresponding author.

† This article is published as part of a themed issue on Protein Misfolding in Conformational Disorders, Guest Edited by Cláudio M. Gomes.
Abstract
Huntington’s disease arises from CAG codon-repeat expansions in the Htt gene, which leads to a Htt gene product with an expanded polyglutamine (polyQ) sequence. The length of the polyQ expansion correlates with an increased tendency to form aggregates and clustering into micrometer-plus sized inclusion bodies in neurons and other cell types. Yet after nearly 20 years since the genetic basis for HD was identified, our knowledge of how polyQ-expanded Htt fragment aggregation relates to disease mechanisms remains fragmentary and controversial. Challenges remain in defining the aggregation process at the molecular level and how this process is influenced by, or influences cellular activities. Insight is further confounded by the term “aggregation” being used to describe a composite of distinct processes that may have opposing consequences to cell health and survival. This review discusses these issues in light of a historic summary of Htt aggregation in the cellular milieu and the intrinsic attributes of polyQ-expanded Htt that lead to aggregation. Finally, discussion centers on strategies forward to improve our knowledge for how aggregation relates to cellular dysfunction.

Keywords: Huntington’s disease (HD), Huntingtin, polyglutamine (polyQ), misfolding, aggregation.
Introduction

Huntington’s disease (HD) is an untreatable and fatal condition that commonly strikes between the ages of 35 and 45 years and in 7% cases before the age of 21 [1-2]. HD arises from a CAG codon repeat expansion in the Huntingtin IT15 gene, which leads to a huntingtin (Htt) gene product containing an expanded polyglutamine (polyQ) sequence beyond a threshold of about 36 glutamines [3]. The mutant Htt clusters into micrometer-plus sized inclusion bodies in neurons and other cell types [4-6]. Purified polyQ peptides also form fibrils with a classic β-sheet amyloid-like motif [7-8], which is a characteristic of a large category of diseases associated with protein aggregation (reviewed in [9] and [10]). These data compelling suggest pathogenesis arises as a direct consequence from polyQ-induced aggregation – yet after nearly 20 years since the genetic basis for HD was identified the role of aggregation in pathogenesis is still unclear and controversial. A major reason is that cellular dysfunction appears to occur prior to the formation of microscopically visible aggregates or is non-correlative to their presence in human HD pathology and animal models [11-12]. In addition, RNA based mechanisms of toxicity at the level of CAG mRNA transcripts have emerged to challenge Htt protein-based mechanisms altogether (which have been reviewed recently elsewhere [13-15]). The focus of this review is to revisit “aggregation” of Htt in light of the historical research into this process and discuss the challenges deciphering a complete understanding of the relationship between Htt aggregation and cellular reactions to this process. Also included is discussion of common misconceptions about what aggregation means, which are contributing to our difficulties in resolving how and whether aggregation is involved in pathogenesis.

What is an aggregate?

The concept of protein folding, misfolding and aggregation as an integrated process in neurodegenerative disease has come to light in the 1990s and is hypothesized to be involved in over 35 diseases [16-17]. The central gamut of this hypothesis is that when protein conformations are destabilized, such as through mutations, they can fold into non-native conformations (ie misfolding), which then triggers assembly into β-sheet rich amyloid fibrils.
that accumulate into large deposits during disease progression [10, 18]. Traditionally amyloid was defined as an extracellular substance, but the key principles underlying amyloid-fibrillization also occur within cells in several diseases, which includes HD [16]. In Htt, polyQ expansions promote an inherent tendency to aggregate into fibrils [19].

The most obvious indication of “protein aggregation” in HD is the appearance of inclusion bodies in striatal and cortical neurons, which are punctate clusters of Htt and proteolytic fragments corresponding to the exon 1 polyQ repeat region (Httex1) [4]. Inclusion bodies can form inside the nucleus as well as peri-nuclearly and in dendrites [4-5]. While inclusion bodies look like aggregates under a microscope the molecular process of aggregation is not technically equivalent to the condensation of a protein’s cellular localization [20]. Yet, a distinction between molecular aggregation of Htt and the formation of inclusion bodies is often not made in the field, which confounds insight to how aggregation impacts on cellular function. For example, it is conceivable that early aggregation steps, or even the tendency of a single Htt molecule to aggregate, influences a cell profoundly irrespective to the location of Htt molecules within a cell. In addition, various studies on a range of different aggregation-prone proteins suggest a general phenomenon for small sized aggregates to be more toxic to cells than large aggregates or for different aggregate structures to have very different degrees of toxicity [21-25]. Knowledge of how these parameters relate to Htt or its proteolytic fragments remains by and large fragmentary and is discussed below.

**Inclusion bodies – a cellular response to aggregation?**

Historically inclusion bodies were originally proposed to cause toxicity by co-aggregating other cellular components into the inclusion body [20, 26-30]. However, other evidence suggested the converse view that inclusion bodies are not toxic, but the result of quality control systems that sequester smaller, more toxic mutant Htt forms from the cytosol [31-32]. One example demonstrating that inclusion bodies can be protective is that pharmacological treatment of cells with molecules such as B2 (5-[4-(4-chlorobenzoyl)-1-piperazinyl]-8-nitroquinoline), promoted inclusion body formation while also protecting cells from toxicity [33]. Cultured rat neurons
transfected with Htt\textsubscript{ex1}-GFP were more likely to survive longer in neurons that had formed an inclusion body than neurons retaining dispersed Htt throughout the cytosol [32]. Genetic and biochemical evidence also suggests inclusion bodies form via dynein-mediated retrograde transport of diffuse precursor forms to the microtubule-organizing center, analogous to the aggresome model developed to describe the organized aggregation of the cystic fibrosis transmembrane conductance regulator (CFTR) protein at the microtubule organizing center [34-39].

But are all inclusion bodies protective or formed through the same mechanisms? Inclusion bodies in human HD and model systems are found in multiple cellular locations: intranuclearly as well as cytoplasmically, which is at odds with a unique location at the microtubule organizing center as characterized by the classic aggresome model [35, 37]. In yeast, inclusion bodies formed by Htt\textsubscript{ex1} can be toggled from a single aggresome-like structure into scattered multiple puncta by deletion of Cdc48 (the VCP/p97 yeast homologue), which is an AAA chaperone involved in ubiquitination of certain substrates [39]. Different misfolding proteins can also form mutually exclusive puncta when expressed in the same cell which indicates aggresome-like structures are more generally polymorphic and arise by different cellular processes [40-43]. Hence it remains a distinct possibility that Htt inclusion bodies \textit{in vivo} are also polymorphic in structure or mode of formation, depending on cell, intracellular location or time point along the disease process.

The complexity and polymorphic nature of inclusion bodies was recently partially explained by the identification of the two novel aggresome-like compartments of the IPOD and JUNQ [41-43]. The IPOD contains immobile protein aggregates, to which mutant Htt and yeast prion proteins are terminally sorted in yeast and mammalian cell culture systems [41]. The JUNQ by contrast contains primarily mobile proteins, which may be salvageable and refolded back into the properly functional forms or degraded [41]. Both JUNQ and IPOD form as a result of microtubule-dependent active processes, however unlike classic aggresomes, neither JUNQ or IPOD is localized to the microtubule organizing centre [41]. Whether these or similar structures
have relevance to HD is not clear, but it does indicate the possibility for inclusion bodies to play different beneficial roles that could be engaged under different conditions of stress, development or cell type.

Increasing evidence also points to heat shock protein chaperones as being involved in Htt-relevant inclusion body formation as part of a tight integration with overall protein quality control mechanisms against unfoldable/misfolded proteins. Hsp70 and its co-chaperone Hsp40 are key elements of this network that have canonical functions in assisting proteins to fold correctly [44]. Hsp70 can also direct misfolded/unfoldable proteins to degradation via the proteasome or lysosome (reviewed in [45] and [46]). Hsp40 and 70 potently inhibit toxicity of Httex1 in model systems [47-49]. They also can alter the Htt aggregation process in vitro, remodel the structure of the Htt aggregates that form, and are enriched in inclusion bodies [48, 50-52]. A recent study has unmasked a new triage role for Hsp70 in directing misfolded proteins to aggresomes via interactions with Bcl2-associated athanogene 3 [53]. Support for this mechanism as being relevant for Htt is that Hsp70 can also facilitate inclusion body formation of Httex1 in some circumstances and/or protect cells from toxicity without reducing inclusion body load [49, 54-59]. Hence, chaperone networks govern an intimate interplay between refolding (which is not possible with an unfoldable expanded polyQ sequence) and removal options of degradation and sequestration into the inclusion body. Such a mechanism implies that unfoldable nature of the pathogenic-length polyQ-sequences may be a more direct trigger for cellular responses than aggregates, and hence could be more pertinent to mechanisms relating to toxicity. The movement of polyQ-expanded Htt to inclusions may be a last resort when other means such as degradation are exhausted. While this hypothesis needs to be experimentally investigated, recent evidence supports a role for expanded polyQ in full-length Htt as suppressing the capacity of the heat shock protein system to respond to stresses before any evidence of aggregation [60]. Furthermore, the presence of polyQ-related aggregation has been proposed to correlate with a global deficiency in the protein quality control network by overdrawing on the available chaperone network, which in itself could be a major non-specific driver of toxicity by making cells more vulnerable to stresses [61-62].
Antibodies as probes for Htt conformation, aggregation and as tools to regulate Htt toxicity

What do we know about the conformation of Htt in an aggregated state? While in the test tube Htt forms fibrils, it is less clear that these structures are populated in vivo or if all fibrils have the same structure. Tools to assess this are limited as the complex cellular environment impedes our capacity to use traditional approaches for high resolution insight. A major approach has been the use of antibodies, several of which bind the Httex1 sequence in a manner sensitive to conformation or aggregation state, and these collectively have provided much utility to probe how Htt conformations and aggregation states are involved in pathogenesis (Fig 1). One set of eight monoclonal antibodies, MW1 to MW8, target different regions of the Httex1 sequence [63]. Of these, MW1 binds to an extended linear polyQ conformation in monomers [63-64]. Antibody 2B7, which binds first 17 amino acids of Htt, also binds Htt monomers and with MW1 has been used quantitate soluble Htt load in cells and tissues [65-66]. In contrast, antibodies 4C9 and MW8 preferentially bind aggregated Htt, with utility to quantitate aggregate load in mouse models of HD and complement the 2B7/MW1 assays for monomers [67].

The adaptation of these antibodies to be expressed in cells as “nanobodies” also has provided insight to how different forms relate to Htt toxicity. Antibodies targeting the N-terminal 20 amino acids (C4 and V_L12.3) or the proline rich region of Httex1 (mEM48, Happ1 & Happ3) in general prevented inclusion body formation and/or promoted Htt degradation in cell culture or HD mouse models [68-73]. One exception was with the V_L12.3 antibody which slightly worsened pathology in the YAC128 mouse model, despite decreasing aggregate load [72]. The C4 could rapidly degrade Htt and alleviate pathology in HD mice when it was fused to a degradation-promoting PEST sequence, suggesting that reducing monomeric Htt levels could be a useful therapeutic strategy and method to mitigate aggregation [74]. Conversely, an antibody that targets Htt aggregates but not monomers, 6E, enhanced Htt inclusion body formation and toxicity [75]. Furthermore, 6E could not stimulate Htt degradation when fused to a PEST sequence, suggesting that aggregates are not amenable to degradation via the proteasome [76].
What is the conformation of Htt in inclusion bodies?

Fibrillar material has also been observed in electron micrographs of inclusion bodies, suggesting Htt adopts an amyloid fibril conformation in inclusion bodies [4, 77-78]. This is supported by the observation that Htt becomes SDS insoluble in cells when inclusion bodies form, which is a property of Htt fibrils formed in vitro [79-80]. An agarose electrophoresis based method (AGERA) to resolve SDS-insoluble material into different sized fragments has also indicated aggregates form before detectable changes in pathology and behaviour in cell culture and in mice (within 2 weeks for the R6/2 HD mouse model and 6 months for the HdhQ150 knock-in mouse model) and became larger in size over time [80]. The AGERA technique also showed that cytoplasmic aggregates from mouse brain had a smaller SDS-distribution than nuclear aggregates in brains of 6 or 8 week-old R6/2 mice, which could reflect a difference in architecture between inclusion bodies in the nucleus relative to the cytosol [80].

Because SDS is a harsh denaturant it also remains possible that some aggregates or inclusion body substructures formed in cells could be dissociated if the cellular machinery molds the architecture and aggregation pathways. This is supported by one study where inclusion bodies were purified and washed in SDS before being imaged by electron microscopy finding inclusion bodies were left as a shell of fibrillar material with the core dissolved [81]. Httex1 is also capable of forming fibrils with distinct structural architectures [82], a process with precedence in other amyloid systems including yeast prions [83], Aβ [84] and other proteins [85-86]. Httex1 fibrils formed in vitro at 4 ºC had different structural properties than fibrils formed at 37 ºC and were more toxic to cells when applied exogenously [82]. These 4 ºC-fibrils also displayed greater binding to the 3B5H10 antibody, which binds to a compact hairpin monomer conformation of polyQ and has independently been shown to correlate highly with toxicity in neuronal cultures transfected with Httex1 [82, 87].

Inclusion bodies also contain Htt-immunoreactive amorphous structure to varying extents, consistent with non-fibrillar forms of Htt as a key element in inclusion bodies [4-5, 77, 81, 88]. One study used a seprion ligand assay to isolate Htt-immunoreactive aggregate structures from
HD mouse brain that spanned a range of different morphologies – some more amorphous and globular and others more fibrillar [89]. Seprion is a heterogeneous class of high molecular weight polymeric molecules of polyionic charge and hydrophobic that was originally described to bind selectively to the aggregated scrapie form of prion protein but not the unaggregated counterpart [90]. The mechanism of binding is not well described, but thought to arise from avidity of clustered polar and hydrophobic regions interacting with aggregated proteins [89-90]. Of note was that the aggregates could be divided into classes based on their morphology and reactivity to different Htt monoclonal antibodies that are sensitive to epitope conformation (MW1, MW8 and 3B5H10; Fig 1) [89]. Some of the more amorphous classes of aggregates bound to 3B5H10 [87], which binds a compact hairpin polyQ monomer whereas others bound MW1 which binds an extended polyQ monomer sequence, while others bound to antibodies more specific to aggregates over monomers, suggesting these different classes of structures comprise Htt with distinct conformations (Fig 2; [91]).

**Htt fragment length dictates toxicity, aggregation and inclusion body heterogeneity**

Different fragment lengths of Htt have quite distinct effects on inclusion body formation and apparent toxicity. Transgenic mice expressing polyQ-expanded Htt_{ex1} (N90 1) or a longer N171 fragment generate a severe degenerative pathology [92-93]. Strikingly in contrast, a transgenic YAC mouse expressing exon 1 & 2 of Htt (N117; also called the Short-stop Htt fragment) with 120Q under the human Htt promoter formed many inclusion bodies in the brain but had no neurodegeneration [94]. Different truncation products of Htt produced very different effects on cytotoxicity in cell culture in a manner uncorrelated with inclusion body formation: N117 was not toxic whereas N90, N167 and N171 were, even though the N117 formed more inclusion bodies than the other fragments [95]. These different fragments also formed inclusion bodies

---

1 Nomenclature hereon defines the number of amino acids from the N-terminus in wild-type Htt with a Q23 length. Eg, Htt_{ex1} = N90, regardless of the Q-length.
with different morphological and biochemical properties, which further suggests heterogeneity in inclusion body structure and their possible mechanisms of formation [95]. N117 expressed and purified from cell culture and then assayed for aggregation kinetics in vitro formed a smaller size pool of spherically shaped oligomers than N90 suggesting that the early steps of the aggregation process could be differentially effected by 27 amino acids that differ between these fragments [96]. In a further twist, an independently constructed N117 transgenic mouse line expressed by the prion promoter had a moderately severe phenotype, which indicates that expression levels and patterns likely are also important in dictating toxicity [97].

Cell to cell variability in expression levels could indeed be important in dictating the type of inclusion bodies that form, and whether they are protective or toxic. In a study mapping 14 N-terminal proteolytic fragments of Htt in the HdhQ150 full length HT knock-in mouse, it was found that N90-like fragments decreased upon age and seemed to be restricted to the cytoplasm when soluble [98]. When N90 was SDS-insoluble it was present in the nucleus. The nuclear Htt seemed to be in both diffuse and inclusion body localizations, consistent with molecular aggregation preceding inclusion body formation [98]. Other work has found that Htt fragments shorter than N115 assemble into nuclear aggregates whereas fragments longer than this (between about 146 and 214 amino acids) are mostly cytoplasmic in cell culture and human HD brain [99]. One study proposed that N586 generated by caspase-6 cleavage is toxic in a YAC 128Q mouse, whereas slightly shorter fragments (N552 and N513) generated by putative caspase 2 and 3 cleavage seemed to not be involved in pathogenesis [100]. When N586 was expressed in mice it formed extensive cytoplasmic inclusion bodies comprising just N117 [101, 102], which raises the possibility protein quality control mechanisms cleave toxic fragments into less toxic “Short-stop”-like forms during or after initial aggregation processes.

Flanking sequence also can play a role in toxicity by altering the aggregation process or aggregate structure in manners that differentially exposes epitopes that interact with other cellular ligands. For example, studies in yeast found certain amino acid sequences to have the capacity to unmask toxicity or alter inclusion body properties of otherwise non-toxic polyQ
sequences or vice versa when fused to the polyQ sequence [39, 103]. The charged FLAG tag (DYKDDDDK) made long polyQ sequences lethal no matter where in the protein sequence it was appended, which suggests that this tag makes the protein acquire new features that are important for toxicity (eg new interaction partners) [103]. A particularly intriguing feature is the ability of these modifying sequences to abrogate or unmask toxicity of toxic polyQ contexts in trans – ie when they are attached to a short inert polyQ sequence, they become co-aggregated with the non-toxic expanded polyQ sequence and mediate the toxic status [104]. A similar in trans effect was observed for the poly-pro rich sequences modulating the ability of aggregates to have “aggresome”-like features or not [39]. These data suggest that the exogenous sequences such as the FLAG tag and the proline-rich region recruit third party proteins to the aggregate, a process by which cellular health and viability are affected. While the positively charged FLAG tag or effects of other artificial sequences is not relevant to HD pathology, the experiments demonstrate the capability for different fragments of Htt – or alternate aggregate structures – to display different interaction motifs. Htt aggregates co-aggregate other polyQ-rich sequences [26-27] and possibly other proteins via the other domains (eg pro-rich sequence, which is a known ligand of the SH3 domain [105]). PolyQ expanded Htt binds an SH3 domain from SH3GL3 specifically when in an aggregated state, indicating exposure of the pro-rich region in the Htt aggregates can inappropriately recruit proteins [106-107]. Further support for such a mechanism comes from co-expression of Httex1 with antibody MW7, which targets the pro-rich region leading to colocalization in inclusion bodies and suppression of toxicity, suggesting MW7 shields the pro-rich regions in aggregates from aberrant interactions [108].

**What is the conformation and aggregation state of Htt outside inclusion bodies?**

A number of studies have pointed to smallest molecular aggregates of Htt in cells, collectively called oligomers, as being nanometer-scale and spherically-shaped [51, 78, 109-113]. Attention has focused on these structures as possible toxins based on evidence similar-sized oligomers are in general toxic in amyloid-related aggregation phenomena, and because inclusion bodies in the
general sense are protective to neurons, which could act to sequester the oligomers from the cytosol [21, 32]. Electron microscopy has indicated globular oligomer-like structures ~30-100 nm in size accumulate at the periphery of an inclusion bodies in human HD brain and in transfected cell culture, suggestive of their transportation into the inclusion bodies from the surrounding cellular environment (Figure 3; [78]).

Atomic force microscopy on purified Htt\textsubscript{ex1} indicates it inherently forms small oligomers less than 10 nm in diameter regardless of Q-length – but pathogenic Q lengths formed a distinct larger set of oligomers (25-130 nm in diameter) that were transient to forming fibrils [78]. Oligomers formed by non pathogenic Htt\textsubscript{ex1} (23Q) have also been detected in cells using a split GFP system, which involves the complementation of two structural fragments of GFP to form a fluorophore upon binary complexation [113-114]. Interestingly, pathogenic Q-lengths led to a reduction in oligomers by this method, which could indicate that the “abnormal” larger sized oligomers form structures non-permissive to GFP complementation [113]. It is also noteworthy that fusion of full length GFP to Htt\textsubscript{ex1} abrogates “wild-type” oligomerization: ie Htt\textsubscript{ex1}-GFP (25Q) forms only monomers whereas the 46Q counterpart still formed a pool of oligomers in the range of 25nm-100 nm [51]. The fusion of GFP to Htt\textsubscript{ex1} does not abolish the pathogenicity of Htt\textsubscript{ex1} in a transgenic mouse model [115] or indeed other animal models of HD [103, 116-117].

The larger spherical shaped oligomers (up to ~ 50 nm) associated with pathogenic polyQ expanded Htt\textsubscript{ex1} have been purified from cells transfected with Htt\textsubscript{ex1} with no evidence of fibrils, which suggests that cells restrict fibrillization in the cytosol or that fibrils only form when Htt has moved into the inclusion body [109]. Further evidence for cells actively regulating cytosolic Htt aggregations into an organized class of oligomers comes from our study revealing the total pool of oligomers in cell lysates expressing Htt\textsubscript{ex1}-GFP (46Q) to be invariant in size distribution and proportion of the total Htt\textsubscript{ex1}-GFP molecules over a 3 day transfection [51].

Oligomers have been suggested to progressively assemble from monomers throughout the cytosol prior to inclusion body formation in individual cells. One fluorescence correlation
spectroscopy (FCS) study indicated the oligomers to comprise 3-6 monomers based on fluorescent intensity per particle and diffusion rates [110]. Fluorescence recovery after photobleaching (FRAP) also revealed the diffuse pools of polyQ-expanded Htt^{ex1} to become less mobile over time suggesting the concentration or size of oligomers increases [113]. An independent “numbers and brightness” time-lapse confocal imaging method indicated monomers assembled into oligomers of 5-15 proteins in a two-step process once a critical threshold concentration of Htt^{ex1} is reached [112]. In this study, once an inclusion body formed oligomers were rapidly depleted leaving only a low concentration of monomers in the cytosol. A similar effect was also observed using fluorescent resonance energy transfer (FRET) reporters of polyQ self-interactions [111]. Longitudinal analysis in neuronal-like SY5Y cells indicated cells with higher diffuse FRET had a greater risk of death, consistent with the accumulation of oligomers leading to heightened toxicity [111].

We have also developed tetracysteine-based Htt^{ex1}-CFP biosensors that can differentially distinguish monomers from oligomers in individual cells [118]. In this biosensor, the tetracysteine tag is occluded from binding biarsenical dyes FlAsH or ReAsH when Htt oligomerizes [118]. We found that the diffuse Htt^{ex1} (46Q) in cells tended to fall into two classes – those enriched with monomers and those where monomers were almost entirely depleted at the expense of submicroscopic oligomers. Often this coincided with a small, apparently developing inclusion body, which is consistent with individual cells rapidly accumulating dispersed oligomers preceding an inclusion body forming [118]. We extended this method to quantitate whole cell populations further by developing new flow cytometry methods to independently track inclusion body formation from oligomerization [119]. With this system we found a subset of cells that were enriched with oligomers but without an inclusion body [119]. Interestingly, these cells were more abundant at lower levels of Htt^{ex1} expression and increased in proportion of the cell population over time [119]. Hence while higher expression levels may trigger a rapid condensation of oligomers into the inclusion body, at lower expression levels oligomers, possibly of a different composition or architecture, may instead accrue as part of a
parallel quality control machinery with the net effect of oligomers remaining relatively constant in proportion to all Htt molecules over time [51].

**Toxic compact monomers?**

There is also emerging data pointing directly to an abnormal monomer conformation that is toxic in addition to the roles described earlier for chaperones targeting unfoldable proteins. Theoretical and biophysical studies have suggested polyQ monomers are mechanically rigid, compact and heterogeneous as a result of extensive intra-chain hydrogen bonding via glutamine — even though they otherwise have disordered or low β-sheet secondary structure [91, 120-127]. Compact polyQ monomers have been proposed to be difficult to “unfold” and hence obstruct the proteasome during degradation, leading to toxicity [125, 128-129].

At least one conformation or a subset of similar conformations relevant to toxicity and aggregation may have a hairpin β-structure. Purified thioredoxin-polyQ fusions converted from an α-helix rich monomer to a β-sheet monomer that preceded assembly into fibrils and the β-rich form was toxic when injected into cells [130]. Another study showed Httex1(53Q) to have a conformational rearrangement into a compact form when cleaved from a GST fusion by an engineered protease site [131]. This compact form could interact with and deactivate transcription factors TBP and CBP, which suggests a novel mode of toxicity for the monomer [131].

Recently an antibody specific to a compact polyQ hairpin structure (3B5H10) strongly predicted neuronal death in situ over other antibodies to polyQ or Httex1 monomer (eg MW1), concordant with a compact hairpin monomer conformation precipitating cell death [87, 132]. 3B5H10 also was strongly immunoreactive against short polyQ sequences (9Q) or tandem repeats of 6Q/12Q linked together with pro-gly β-turn linkers relative to short polyQ sequences linked with non-β-turn permissive linkers [133]. β-turn linked short polyQ sequences can aggregate in vitro, form inclusion bodies in cells and confer cellular toxicity to a similar manner as long contiguous polyQ sequences (eg 76Q), suggesting that the β-turn conformation stabilizes a template for aggregation [134-135]. However, when 6Q and 12Q alternating repeats were used instead of
just Q9 repeats, which was predicted to offset a β-sheet register in the aggregate fibril, this impaired inclusion body formation in cells yet the protein retained high 3B5H10 reactivity and was still highly toxic, supporting the hypothesis that a soluble compact β-turn monomer was the main toxic agent independent to aggregation state [133].

In a recent genetic screen for inhibitors of polyQ aggregation in C. elegans Moag-4 was revealed as a potent inhibitor of aggregation and toxicity indicating it could provide a critical link in the assembly of inclusion bodies and compact monomers [136]. Using a native agarose-gel electrophoresis (NAGE) assay to follow polyQ proteins, they found MOAG-4 promoted a compact misfolded faster migrating form of polyQ, which was SDS-soluble. This was suggested as MOAG-4 being be part of a quality control pathway that drives toxic, aggregation-prone proteins into inclusion bodies for sequestration.

**Interplay of post-translational modifications and membrane interfaces with aggregation**

Post-translational modifications of the extreme N-terminus of Htt is another parameter with effects on aggregation and toxicity and interactions with protein quality control machinery. The N-terminal 17 amino acids of Htt can be projected as an amphipathic α-helix and has putative roles in binding to phospholipid membranes of the mitochondria and ER [137-138]. Mutations in Httex1 that disrupt the amphipathicity have large effects on cellular localization and inclusion body formation rates [137]. Recent studies have suggested that the α-helical properties of this sequence is important for driving early steps in Httex1 aggregation and can be adopted in the fibril form [139-140]. Several other amyloid-fibril forming proteins also share an ability to switch from a relatively disordered conformation into a amphipathic α-helix upon lipid-binding, which has huge influences on amyloid-fibril formation rates and aggregate structure [85, 141-142]. This suggests events that influence the interaction of Htt with lipids could profoundly alter the intrinsic tendency to aggregate in a highly localized manner within the cell. The N-terminal 17 residue sequence of Htt is also subjected to a number of post translational events, which alter the aggregation propensity or lipid binding properties. Lysines at positions 6, 9 and/or 15 have
been reported to be sumoylated and ubiquitinated, which reduces aggregation [143]. In addition, Htt\textsubscript{ex1} is N-terminally acetylated, which could alter the properties of the Htt such as helical propensity and lipid interactions [144-145]. Indeed, N-terminal acetylation of the amphipathic N-terminal sequence of α-synuclein increases its helical propensity and affects its fibrillization [146]. Threonine 3 has been reported phosphorylated and experiments with phospho-mimetic (T3D)and nonphosphorylatable (T3A) mutants suggest threonine 3 phosphorylation enhances inclusion body formation rates [144]. Interestingly both mutations reduced toxicity of Htt in a Drosophila model of disease. Serines 13 and 16 are also reported to be phosphorylated by IκB kinase, which alters the nuclear localization of Htt and initiates other post-translational modifications of Htt\textsubscript{ex1} including lysine-16 acetylation and ubiquitination [147]. These findings suggest post-translational modifications can toggle multiple interactions with cellular machinery, the aggregation tendency and pathways to degradation and inclusion body formation. Phospho-mimetic mutations also suggest those that decrease toxicity can increase solubility, and it remains to be determined whether protein quality control pathways that are linked to degradation might link also with inclusion body formation, such as through a Hsp70-mediated mechanism.

**Outlook and remaining challenges**

The field is rapidly making advances towards a more complete understanding of how aggregation proceeds *in situ* by using a range of model systems that recapitulate the biochemical aggregation properties of Htt\textsubscript{ex1}. Nevertheless we are still somewhat short in a capacity to develop therapeutic strategies to target pathological consequences arising from aggregation without further knowledge of how aggregation fundamentally impacts on cellular activities. One strategy to circumvent this requirement is to block the inherent tendency of expanded polyQ to aggregate altogether, which is a strategy recently been proven as a fruitful (and so far the only successful) road to therapeutics against any misfolding-related disease [148]. The drug tafamidis meglumine is a pharmacological chaperone, which stabilizes the native conformation of a conformationally-destabilized transthyretin mutant and prevents aggregation. The drug reverses
an amyloid disease and is now approved for clinical use in Europe [148]. This example exemplifies how preventing all steps in an aggregation process can ameliorate pathogenesis and offers much optimism for a parallel strategy in HD.

Other areas of fundamental biology need further examination to untangle aggregation from other biological properties of the Htt protein. Since Htt is such a large protein with so many interactions aggregation will invariably interfere with many normal functions of Htt and could do so in quite non-specific manners, which is a major challenge to further insight. Particular consideration is needed to the context of the HD models that often require extreme manipulations to generate an observable phenotype (e.g., high expression, very long polyQ lengths, short Htt fragments), parameters which could be expected to be particularly sensitive to non-specific mechanisms of toxicity.

Another continuing challenge remains defining heterogeneity in inclusion bodies and the mechanisms by which they form. If inclusion bodies are protective, then defining elements of the cell that govern their formation would be an attractive avenue for therapeutics. Yet, if inclusion bodies are heterogeneous in how they form, then they may not necessarily always be protective. Questions also remain as to how we consider heterogeneity of the aggregation process – i.e., misfolded monomers, oligomers of various sizes which is a broader issue of relevance to protein misfolding and aggregation [149]. Furthermore whether “aggregates” in cells are biochemically equivalent to “aggregates” in the test tube in structure and/or composition needs further investigation since aggregates as a toxic interface to cellular ligands could have different interactions if in a different structural configuration.

Other clues to how Htt-relevant inclusion bodies might be formed come from genome-wide RNAi genetic screens in *Drosophila* cells for suppressors of Htt*-ex1* inclusion body formation [150-151]. Interestingly at least 16 genes were identified that have (putative) functions in cytoskeletal- and protein trafficking, which seem particularly attractive targets for these mechanisms. A similar strategy was applied to find all proteins that bind to Htt*-ex1* and which also modify toxicity in a *Drosophila* model of HD [152]. A total of 27 genes were characterized
and shown to modify Htt\textsubscript{ex1} toxicity in \textit{Drosophila} [152], many of which also have putative functions in transport and cytoskeleton-related activities.

Advances in new technologies, such as our recent flow cytometry assay to track different aggregation states of Htt in cells, now enable a new “aggreomics” view of HD to track the changes in cells enriched with specific aggregation states of expanded polyQ [119]. This will enable gene expression or metabolite changes in cells to be followed more explicitly with aggregation state, which provides new avenues towards understanding the molecular mechanisms influenced by the aggregation process. Another recent approach with similar capabilities is a luciferase activity reporter that folds in a manner sensitive to Htt\textsubscript{ex1} aggregation in the cell [153]. Screens using this reporter identified drugs leflunomide and teriflunomide as inhibitors of aggregation and which has potential to be used in assays linking the efficiency of the protein quality control network to Htt aggregation state [153]. A further feature of an aggreomics platform is that it enables new tools (eg biosensors) to be utilized to probe the human HD condition and relate it back to aggregation state. This is necessary to validate the physiological relevance of the findings of the model systems but also will be useful to gauge how these processes operate as the human pathology evolves presymptomatically to postsymptomatically in, for example, biopsy samples.

\textbf{Acknowledgements}

This work was funded by grants to DMH (Australian National Health and Medical Research Council Project grant 566640, Australian Research Council Discovery grant DP120102763 and the Hereditary Disease Foundation). DMH is a Grimwade Fellow, supported by the Miegunyah Trust. We also thank Gillian Bates, Paul Muchowski and Marian DiFiglia for providing the data in Figures 2 and 3.
References


Figure legends

Figure 1. Repertoire of antibodies for probing Httex1 structure and aggregation. Antibodies that have been used widely and/or with features useful to probe conformation of the Httex1 sequence (shown). References: C4 [68], V1.12.3 [69], 2B7 [66], 3B5H10 [87, 132], MW1 [63], 1C2 [154], MW7 [63], Happ1 [70], Happ3 [70], 4C9 [67], mEM48 [155], and MW8 [63].

Figure 2. Heterogeneity of Httex1 aggregate structure and conformation in vivo (adapted from [89]). Aggregated protein was extracted from the R6/2 mouse brain at 12 weeks of age using the Seprion ligand assay, and assessed by transmission electron microscopy. Different morphological classes of aggregates immunoreactive to Htt were observed, which also display different epitopes as indicated by the immunoreactivity to the indicated antibodies. Scale bar = 100 nm.

Figure 3. Complex structural composition of an inclusion body (adapted from [78]). A N969 Htt fragment (100Q) tagged with a FLAG sequence was expressed in MCF7 cells and detected using an immunoperoxidase method and FLAG antibody M5 (dark electron dense material). A. One cytoplasmic inclusion body is shown (open arrow) with nucleus (nuc) indicated. B. Shows a higher magnification: the core of the INCLUSION BODY contains fibrils that do not label with M5 (arrows) and the periphery comprises granular structures immunoreactive to M5 (arrowheads). Left scale bar = 1 μm and right scale bar = 100 nm.