Conformational landscapes of Huntingtin's polyQ domain and variations in its aggregate morphology

Neha Nanajkar[1◕], Abhilash Sahoo[2◕¤], Silvina Matysiak[3*],

1 Department of Biology, University of Maryland, College Park, Maryland, USA
2 Center for Computational Biology, Flatiron Institute, New York, NY, USA
3 Center for Computational Mathematics, Flatiron Institute, New York, NY, USA
4 Fischell Department of Bioengineering, University of Maryland, College Park, Maryland, USA

◕These authors contributed equally to this work.
¤Current Address: Flatiron Institute, New York, USA
* matysiak@umd.edu

Abstract

Huntington's disease is a fatal neurodegenerative disease caused by an aberrant polyQ expansion in the Huntingtin protein (htt). In cases with polyQ$\geq$35, the expanded polyQ tract misfolds and forms large intra-cellular aggregates. Despite significant attention, no clear consensus has been reached on the aggregate structures of the Huntingtin exon 1 protein fragments. Here, we apply coarse grained molecular dynamics to examine the mechanisms driving the aggregation process and resulting aggregate structures of 5 different N17 + polyQ systems, with polyQ lengths of 7, 15, 35, 40 and 45, using the ProMPT forcefield. Elongation of the polyQ domain results in an increase in β-sheet rich structures. At longer polyQ lengths, intra-molecular β-sheets are preferred over inter-molecular sheets, due to the collapse of the elongated polyQ domain into hairpin rich conformations. Crucially, we have discovered that changes in the polyQ length affect the resulting aggregated structure. With short polyQ domains, the N17 domains of htt bundle together, forming a hydrophobic core. As the polyQ length increases, competition between the hydrophobic interactions of N17 and the polar interactions of polyQ arises, leading to a tightly packed polyQ core, with an external distribution of the N17 domains. These findings can help explain the toxic gain-of function that mutant htt acquires.

Author summary

In this work, we consider 5 huntingtin peptide systems- N17 with polyQ lengths of 7, 15 35, 40 and 45, and examine structural characteristics of both the monomeric and aggregated peptide systems. We observe the presence of helices and β-sheet in the polyQ domain. Most importantly, we discovered that the polyQ length affects the resultant aggregate structure. With short polyQ lengths, there is a competition between the hydrophobic interactions of N17 and the polar interactions of the polyQ domain, leading to a bundling of the N17 domains. At longer lengths of polyQ, the polar interactions dominate leading to a loss of N17 bundling and external distribution of the N17 domains. We suspect these variations in the aggregate morphology could explain possible downstream effects in Huntington's disease pathology.

Introduction

The Huntingtin protein (htt) is a large multidomain protein that is involved in a plethora of homeostatic processes including axonal transport, signalling pathways, and interactions with many other proteins [1–3]. Most notably, elongations in the glutamine repeat region (polyQ) of htt result in Huntington's Disease (HD), a progressive neurodegenerative disorder that leads to physical and cognitive decline in affected individuals. A healthy form of the protein typically has between 5 to 35 glutamine repeats, while mutant htt has a polyQ length over 35 residues [4]. These extended polyQ regions misfold and form intra-cellular aggregates, ultimately leading to HD. The length of the extended polyQ domain has been linked with an earlier onset and increased severity of the disease. It has been observed that only the first exon of the Huntington protein (httex1) is necessary to recreate disease conditions in yeast [5].

Although Huntington's disease is one of 9 known polyQ based diseases, a proper understanding of the aggregation process and associated mechanisms is still lacking [6]. PolyQ is especially difficult to study experimentally due to its rapid and length dependent aggregation kinetics, conformational flexibility, and its insolubility [7, 8]. Nevertheless, a number of polyQ's structural characteristics have been uncovered. By itself, polyQ forms polar zippers, where in the beta sheets are stabilised by the glutamine sidechains [9]. Evidence for coiled coil, α-helices and β- helical structures in the polyQ domain also exists [10–12], highlighting the flexible nature of the segment. polyQ is known to be insoluble in solution and prefers self-interaction, leading to collapsed and more globular structures with increase in polyQ length [13, 14]. The structure and dynamics of polyQ are greatly affected by its flanking sequences, namely the N-terminal N17 domain and the C-terminal proline rich domain (PRD) [15–17]. Reports have suggested that N17 can exists in a coiled-coil state or as an amphipathic helix and is responsible for in a rise in htt aggregation [18, 19]. The helicity of N17 can propagate into the polyQ domain leading to an initial region of helicity, after which polyQ assumes a more extended structure [15].

Various aggregation models have been put forth, but a consensus on the aggregate structure is lacking. In one study, the authors performed tryptophan fluorescence experiments on htt, reporting that the aggregates consist of an amyloid core of both the polyQ and N17 domains [20]. These results were echoed by others who suggest that N17 and polyQ domains engage in ternary and quaternary interactions at the fibril core, and the PRD protrudes from the core [21, 22] . In contrast, using ssNMR on the htt fragment $N17Q_{30}P_{10}K_2$, Sivandandam *et al.* reported that the amyloid core consists only of the polyQ domain, and the N17 domain is located along the periphery of the aggregate and is more solvent exposed [23]. The implicit solvent htt model presented by Chen and Wolynes indicates the N17 domains forming bundles, with the polyQ region arranging into long, trailing inter-peptide β-sheets [24].

A clear understanding of the aggregation process and obtained structures is crucial for the development of protein targeting therapeutics. Molecular dynamics can fill these gaps, and provide a molecular mechanism behind the observed phenomenon. Biological processes such as protein folding and protein aggregation occur at timescales inaccessible with unbiased all-atom molecular dynamics simulations. Instead, a reduced resolution approach, called coarse-graining, can be adopted. Here, local atoms are grouped together into coarse-grained beads resulting in smoother free energy landscapes that allow the simulation of larger systems and for longer, biologically relevant timescales. In this work, we have used our in-house developed ProMPT coarse grained forcefield [25] that features free-energy based interaction parameterization, and explicit dipoles to represent polar regions of the protein. The addition of dipoles allows the forcefield to generate protein structural transitions in response to the changing local environment.

In this study, we have used coarse-grained molecular dynamics to examine the

oligomeric states of both wild-type and mutant lengths of htt. 5 N17 + polyQ systems, of which 3 systems are considered wild type, or healthy versions of the huntingtin protein(7Q, 15Q and 35Q) and 2 are characterized as mutant systems (40Q and 45Q) were created. Simulations of these peptide systems were performed at temperatures ranging from 380K to 500K, and reweighted them to 380K. It is important to note that these coarse-grain temperatures do not correspond real temperatures. High temperature simulations are commonly utilized by the molecular dynamics community as a way to better sample the conformational landscape and accelerate the aggregation process [26–28].

We examine the structural characteristics of each peptide system and reach good agreement with existing studies. Given its flexible nature, the polyQ domain exists in a combination of helical, unstructured and sheet like conformations. As the polyQ length increases, the β-sheets tend towards intra-molecular sheets. We also observed a bundling of the N-terminal N17 domain in the aggregate systems, which reduces upon increase in polyQ length.

Materials and methods

All simulations were carried out on GROMACS 2019.4 simulator [29], visualised on VMD 1.9.4 [30], and analysed using in-house scripts using MDAnalysis version 1.1.1 [31].

Forcefield

We used the Protein Model with Polarizability and Transferability (ProMPT) for all simulations [25]. ProMPT is a coarse-grained protein model that is capable of capturing secondary and tertiary structural changes. This is made possible by the addition of dummy charges to all polar beads, which enables the capture of structural transitions in response to changing local environments. ProMPT has been validated on small proteins with varying structural characteristics [32]. Additional details about parameterization and validation can be found in the ProMPT paper. Fig. S1 presents the coarse-grained representation of the N17 domain. ProMPT was used in conjunction with the MARTINI polarizable water model. [33].

Simulation Parameters

For the monomeric N17+polyQ, a single peptide was placed in a cubic water box 8 nm in length. To characterise the aggregate, 6 peptides were used in a 12 nm box. Peptides were initially in an unfolded state. In this model, a helical dihedral potential with a competing minima has been applied to the N17 backbone beads with a force constant of 10 kcal/mol (Fig. S2). Experiments have suggested that N17 observes an unstructured/coil like structure at low concentrations [20], but at high concentrations exists as an amphipathic helix [34]. No dihedral potential was applied to the polyQ domain. Restraints were applied to hold the peptides in place while the system underwent steepest descent energy minimization for 5000 steps, followed by a 5000 step NPT equilibration. Sodium and chloride salt ions, taken from the MARTINI model, were added to the simulation box for an effective salt concentration of 125 mM, as in experimental studies of the huntingtin protein. Upon addition of salt ions, a second round of energy minimization and equilibration was carried out, for 5000 steps each. Upon release of the constraints, a canonical production run, with a timestep of 10 femtoseconds was carried out. Monomeric simulations were run for 200 nanoseconds and aggregate simulations were run for 350 nanoseconds.

Peptide systems were run at a range of temperatures starting from 380K to 500K, in 20° increments, amounting to 7 temperatures for each peptide system. Here we would like to note that the coarse-grained temperatures used in this work are not exactly comparable to the real or experimental temperatures. Variation of the temperature allows for a better exploration of the free energy landscape, and is a common technique in molecular dynamics [26]. As these are NVT simulations, the solvent density remains consistent for all temperatures. Temperatures were fixed using the Nose-Hoover thermostat with a time constant of 1 picosecond. A relative dielectric constant of 2.5 was used. Electrostatics were computed using Particle Mesh Ewald and the LINCS algorithm was used to evaluate constraints. Neighborlist was updated every 10 steps.

Analysis

For the monomeric simulations, analysis was carried out on the last 150 nanoseconds, and for the aggregated systems, the last 260 nanoseconds. Inter-peptide sheet contacts are characterised as the number of contiguous contacts occurring between the polyQ domains of the aggregate. PolyQ interactions within a single peptide are not counted. For intra-peptide sheet contacts, non-helical contiguous contacts within the polyQ domain of a single peptide are considered. In both inter and intra-peptide sheet contacts, a minimum of 4 contiguous contacts is required to qualify as a sheet. The inter-peptide and intra-peptide contacts were normalized by the maximum possible contacts in each system. To create protein backbone contact maps, positions of the protein backbone beads were collected, and a distance matrix was constructed between all backbone beads. Contact is assumed if residues were found within 7 Å of each other. To calculate bundling of N17 domains, a distance matrix was constructed using the backbone and sidechain beads of the N17 domain. Residues that were within 7 Å of each other were defined as in contact. To filter out conformations in which the N17 domains undergo transient interactions, we defined N17 domains as interacting if there at least 4 or more interactions between the domains. The number of Q-Q interactions was obtained by calculating the average number of glutamine residues that a single Q can interact with. Analysis with water molecules was performed by counting the number of distinct water molecules in contact with each glutamine residue. Contact was assumed if within 7 Å of the water bead. All analysis carried out was reweighted to 380K using the Multiscale Bennet Acceptance Ratio, implemented in the pyMBAR package [35].

Results

0.1 Aggregation Pathway

To assess the secondary structure characteristics of the aggregate, we calculated the fraction of inter-peptide and intra-peptide β-sheet contacts for the all polyQ lengths (7Q, 15Q, 35Q, 40Q, 45Q), as seen in Fig. 1. At longer polyQ lengths, the glutamine residues display a preference for intra-peptide β-sheet structures as compared to inter-peptide β-sheets. Fig. 2 presents representative snapshots of each peptide system, where it is evident that larger polyQ domains collapse, leading to more intra-peptide β-sheets and hairpin-like structures. Similar findings have been proposed by Yushchenko *et al.* using FTIR spectroscopy on htt fragments of varying polyQ lengths [36]. The monomeric structures follow the same trend, where in peptides with longer lengths of polyQ assume more collapsed and globular structures, as seen by asphericity measurements in Fig. S3 and the representative images in Fig. S4. A number of studies have shown that longer polyQ domains tend to assume more collapsed, globular structures [14, 37]. Most recently, Kang *et al.* demonstrated this with varying lengths of polyQ in htt [14].

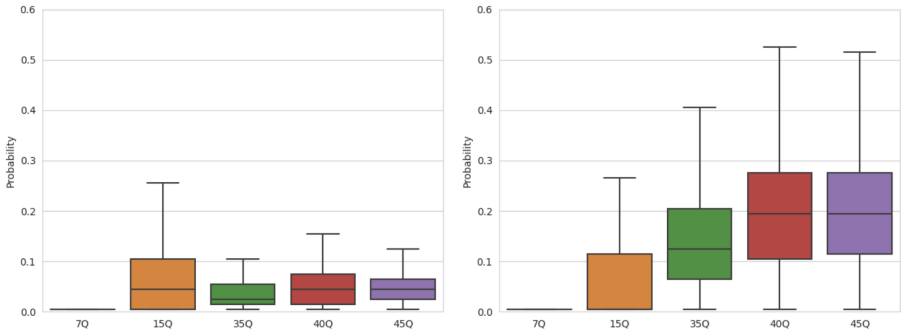

Fig 1. Beta sheet preferences of the polyQ domain, reweighted to 380K. A) Normalized inter-peptide β-sheet contacts, B) Normalized intra-peptide β-sheet contacts

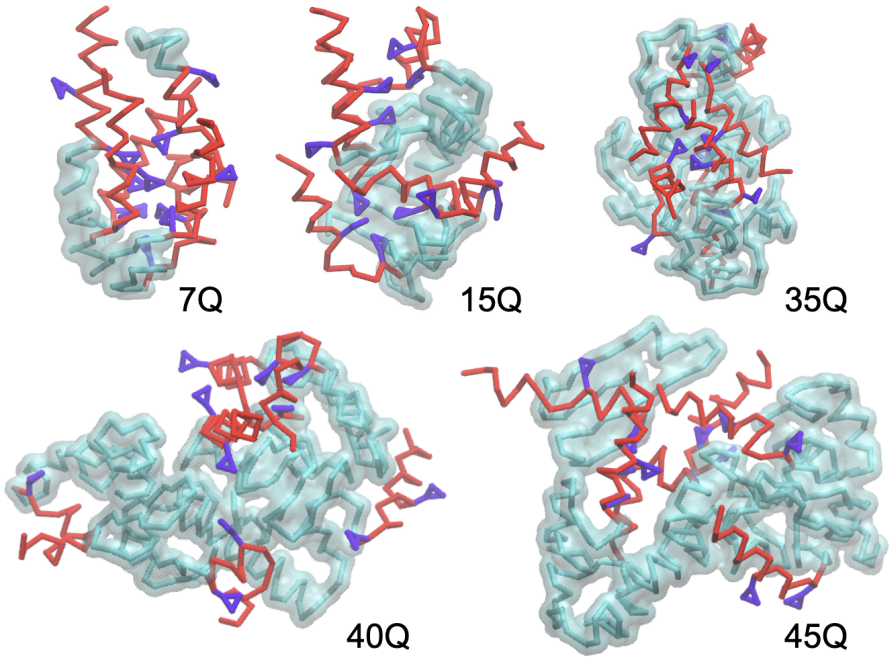

Fig 2. Representative snapshots of aggregated N17 + polyQ at 380K. Red, cyan and violet are used to denote the N17, polyQ and phenylalanine (PHE) residues respectively. Images were chosen by selecting snapshots where the probability of bundled N17 domains is highest, ie. snapshot of the 7Q system has all 6 N17 domains interacting while the 45Q system only has 2.

Fig. 3 presents backbone contact maps for the aggregated peptide systems. Here, we notice i:i+4 contacts in both the N17 and polyQ domains, indicating the presence of helices. Fig. S5 presents the helical fraction of each peptide system. The N17 domain is highly helical, while the polyQ domain has a helical fraction of 0.2. The monomeric systems behave similarly, with helicity in N17 and polyQ domains, as seen in Fig. S6. Various biophysical studies have highlighted the conformational flexibility of the polyQ domain, noting its ability to form helices, coiled coils and β-sheet rich structures [10–12, 38]. In our simulations, we do not observe the presence of coiled coils. The increase in contiguous off-diagonal contacts in the contact maps of the monomeric systems (Fig. S6) indicates an enhancement of β-sheet contacts in the mutant systems (polyQ $>$ 35). Fig. 2 reflects the same trend in the aggregated peptide systems.

Anti-parallel β-sheets appear to dominate, consistent with findings by Bugg *et al.* [21], pictured in Fig. S7. This is in contrast to the results of Chen and Wolynes, who observed extremely long sheets in simulations with $N17Q_{20}$ and observe coiled-coil structures in the N17 domain [24].

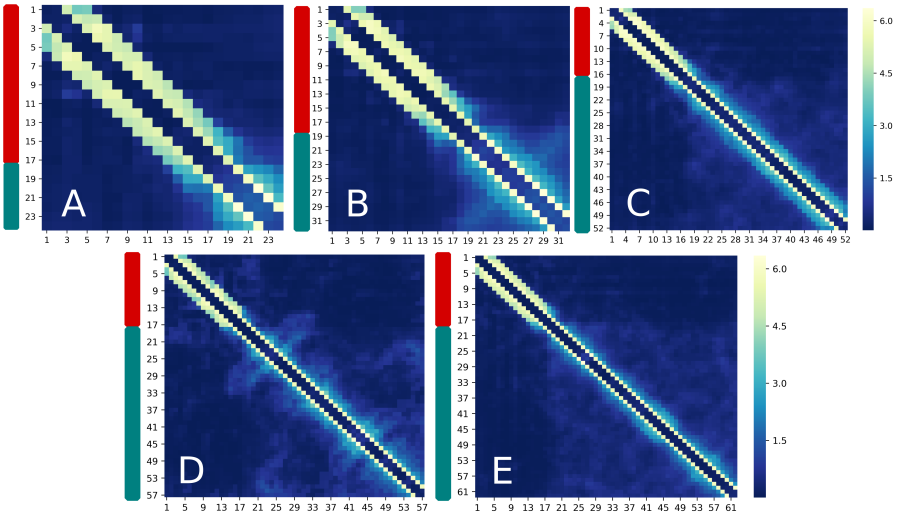

Fig 3. Backbone contacts maps for aggregated N17+ polyQ A) N17 + 7Q, B) N17 + 15Q, C) N17 + 35Q, D) N17 + 40Q, E) N17 + 45Q reweighted to T = 380K. N17 and polyQ domains are marked in red and cyan respectively

To get a broader overview on the structural attributes, domain-specific interactions were considered. Fig. 4A presents a probability of the having a certain number of interacting N17 domains. At lower lengths of polyQ, there is a large probability of having all 6 N17 domains interacting, indicating grouping or bundling together of the N17 domains. In the 7Q system, the probability of having all 6 N17 domains interacting with each other is approximately 0.55. As the polyQ length increases, the probability of observing the bundling between N17 domains drops, with the 45Q system having highest probability of 2 interacting N17 domains. This trend is observable from Fig. 2, where the bundling of N17 domains (denoted in red) declines with increase in polyQ length. Given its amphipathic nature, N17 attempts to bundle with neighboring N17 domains, but this is hindered by increasing polyQ length. Concurrently, an increase in the average number of Q-Q interactions assumed by each glutamine residue is observed, as seen in Fig. 4B. In polyQ systems of 35Q, 40Q and 45Q, each glutamine residue interacts, on average, with 4 or more other glutamine residues. This is indicative of a vast polar, dipole bonding network between the glutamine residues. These results are consistent with reports on the ability of the polyQ domain to form hydrogen bonds (which are a type of dipole interaction) through its sidechains, and its ability to self-associate [9]. This behaviour is also consistent with the results of the monomeric simulations. Fig S8 presents a breakdown of interactions between polyQ and water in the monomeric htt systems. With an increase in polyQ length, the average number of Q-Q interactions per glutamine residue increases, while the average number of water beads interacting with each glutamine residue reduces. These results are in agreement with atomistic simulations of polyQ performed by Wen *et al.* and a growing body of work that posit water is a poor solvent for polyQ [39–41].

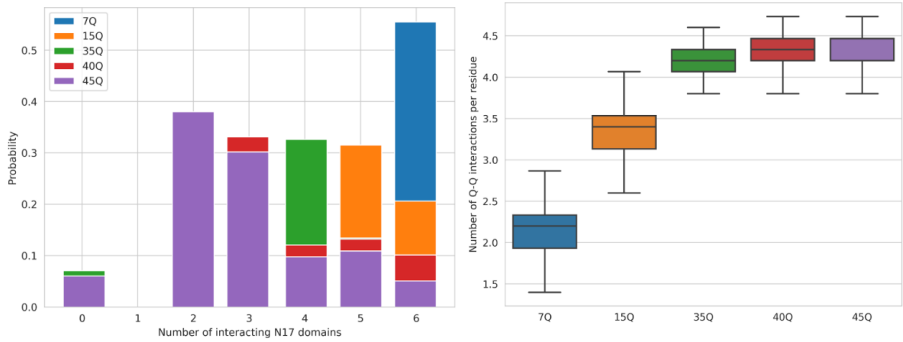

Fig 4. Domain specific interactions of N17 and polyQ domains. A) Probabilities of the number of interacting N17 domains in each protein system. B) Number of Q-Q interactions per residue for each peptide system.

0.2 Competition between hydrophobic and polar interactions

Upon further inspection of the bundling of N17, we notice the formation of a hydrophobic core, in systems with shorter lengths of polyQ. In Fig. 2 , this can been seen by the close grouping of phenylalanine (PHE) groups (pictured in violet) in polyQ = 7, 15. As the polyQ length increases, the bundling of N17 reduces, leading to a loss of the hydrophobic core. To quantify this trend, the average distance of the PHE beads from the center of mass of the aggregate was plotted (Fig. 5A). As expected, the distance of PHE residues from the aggregate center of mass increases with increase in polyQ length, indicating the PHE residues and N17 domains have been pushed to the periphery of the aggregate and the core of the aggregate is instead comprised of the polyQ bonding network. Fig. 5B presents the radius of gyration of the PHE sidechains which increases with polyQ length. At lower lengths of polyQ the PHE residues point inwards due to the bundling of N17, resulting in lower values of radius of gyration. However, in systems with longer polyQ domains and subsequently less N17 bundling, the PHE residues become more solvent exposed, leading to higher radius of gyration. and present short movies of the final aggregate structures for polyQ lengths of 7 and 45 respectively,

The external distribution of N17 has been noted in prior studies on the $N17Q_{30}P_{10}K_2$ fragment of htt using ssNMR techniques [23]. This variation in morphology can be attributed to a competition between the hydrophobic effects of the N17 domain and the polar Q-Q interactions. At mutant lengths of polyQ, the polar Q-Q interactions begin to dominate and outperform the hydrophobic interactions, leading to a more solvated N17 domain.

Discussion

In this study, a coarse-grained molecular dynamics approach was employed to examine the oligomeric forms of wild-type and mutant htt. Here, we present a detailed description of the aggregation process and the resultant structures for 3 wild type systems (N17+7Q, N17+15Q, N17+35Q) and 2 mutant systems (N17+40Q, N17+45Q). The application of coarse grained molecular dynamics to this problem has allowed us to adequately sample the conformational landscapes of these structures, a non-trivial task when using unbiased all-atom molecular dynamics methods.

Fig. 6 presents a schematic of the aggregation process. The conformations shown in the figure are not meant to indicate the structural features of the peptide but rather explain the aggregation process. From the initial disordered conformation, the peptide

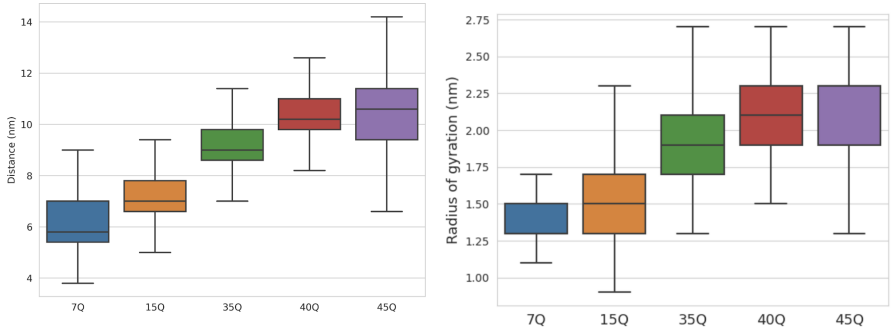

Fig 5. Gradual loss of the hydrophobic core with increasing polyQ length. A) Distance between the center of of mass of PHE sidechain beads from the center of mass of the whole aggregate. B) Radius of gyration of the PHE sidechain beads.

can either bind to neighboring peptides while still in a disordered or partially ordered state and slowly start developing its ordered structure, as seen in pathway A. Alternatively, the peptide can develop some of its secondary structure while still in monomeric form and later join other semi-ordered aggregates, as depicted by pathway B. Both intermediates converge towards a final ordered state that depends on the length of the polyQ domain. At short polyQ lengths, the N17 domains bundle together, forming a hydrophobic core. As the polyQ length increases, the probability of encountering bundled N17 domains reduces as they are pushed to the periphery of the aggregate, and become more solvent exposed.

The resulting aggregate structures are in good agreement with existing experimental studies. Our results indicate that the length of the polyQ sequence determines whether inter-molecular sheet contacts are preferred over intra-molecular sheet contacts at increasing lengths of polyQ. These results are in conjunction with spectroscopic data from Yushchenko *et al.* [36]. Figures 3 and 2 indicate that β-sheet propensity is more likely at increasing lengths of polyQ. We have uncovered the dependence of the aggregate morphology on the length of the polyQ segment. At shorter lengths of polyQ, the aggregate possesses a hydrophobic core, characterised by the bundling of N17 domains such that the hydrophobic side faces the interior of the aggregate. Upon elongation of the polyQ stretch, the degree to which N17 is able to bundle together drops. In such cases, the core of the aggregate is comprised of inter-linking glutamine regions, while the N17 gets re-positioned to the surface of the aggregate. This is in agreement with work by [21] *et al.* , who indicate that the N17 domains engage in tertiary and quaternary interactions between themselves and [23] who demonstrate that N17 positions itself on the periphery of the amyloid core. This occurs due to competition between the hydrophobic effects of the N17 and the polar interactions of the polyQ tract. At shorter lengths, the hydrophobic interactions dominate. We hypothesize that these results could suggest an explanation for the increased toxicity of mutant htt aggregates. A number of studies have indicated that in Huntington's disease, the protein undergoes a toxic gain of function by interfering with normal function. One such gain-of-function can occur through the isolation of transcription factors such as CREB binding protein [42]. Htt has been shown to co-aggregate with CREB, enhance cell toxicity, and cause depletion of CREB binding protein [43, 44]. As shown in Figs. 2 and 5, in the mutant forms of the aggregate, N17 is significantly more solvent exposed. In particular, the outward distribution of PHE residues in the N17 domain could make for an attractive binding site. Thus, we hypothesize that these aggregates may bind with key enzymes and proteins, and ultimately deplete cellular resources required for homeostasis.

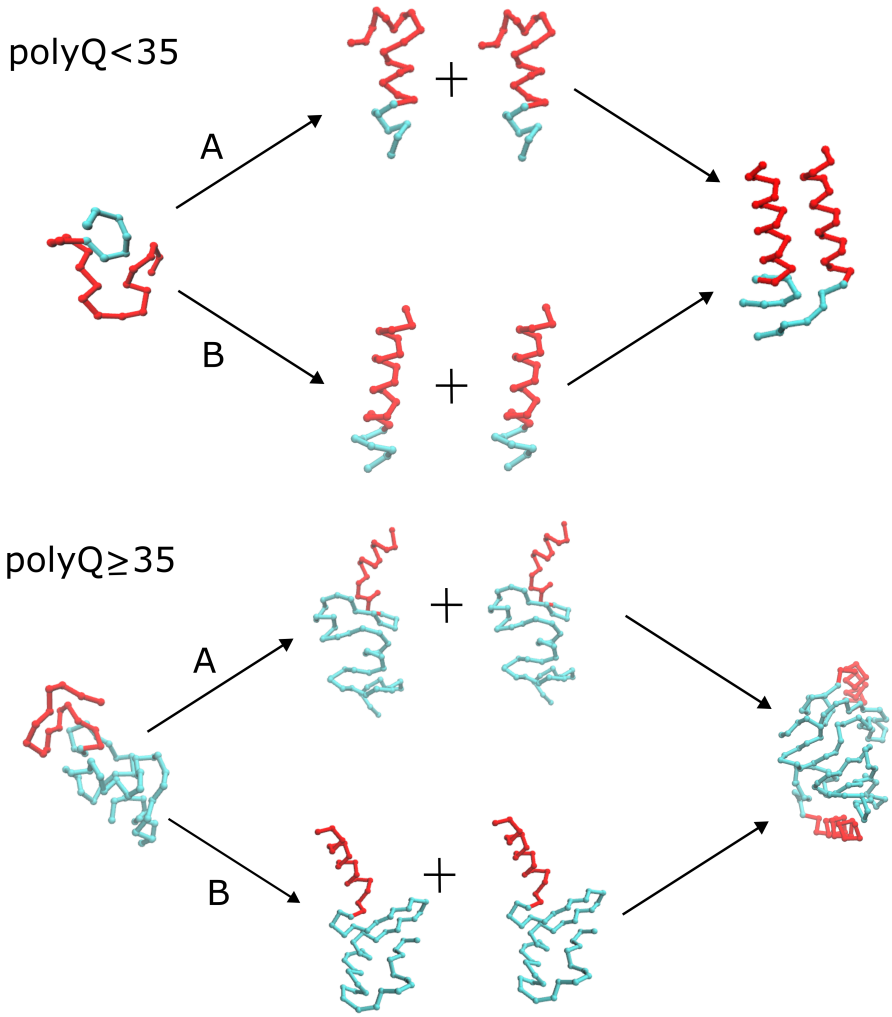

Fig 6. Proposed aggregation mechanism of N17+polyQ, with N17 and polyQ domains depicted in red and cyan respectively. A) Depicts semi-ordered/disordered monomers interacting to form an intermediate state, B) Ordered monomers interact to form the final aggregate. Both intermediate steps ultimately converge to similar final forms.

N17 has also been implicated as one of the membrane binding domains of htt, responsible for the subcellular localization to various organelles in the cell [45]. It carries a nuclear export sequence that directs its movement into the cytoplasm [46]. Chaibva *et al.* uncovered N17's ability to sense membrane curvature [47]. Thus, the external distribution of N17 in mutant htt aggregates could explain one possible gain-of-function that occurs in diseased states, particularly how the aggregate interacts with membranes. Our future works involves the characterization of wild type and mutant aggregates with lipid membranes of varying curvatures and compositions. In addition, we will be assessing the effects of the C-terminal proline segment of htt on the aggregate morphology.

Supporting information

Fig. S1 Coarse grained representation of the N17 domain. The forcefield was implemented as described in the ProMPT paper [25], with the exception of cation pi interactions.

Fig. S2 Dihedral potential applied to the backbone beads of the N17 domain.

Fig. S3 Potential of mean force of monomeric N17+ polyQ reweighted to T=380K, with asphericity as the reaction coordinate. An asphericity value of one indicates a more extended, planar structure, while a value of zero would indicate a a compact, spherical structure.

Fig. S4 Representative snapshots of monomeric N17+ polyQ at 380K. A) N17 + 7Q, B) N17 + 15Q, C) N17 + 35Q, D) N17 + 40Q, E) N17 + 40Q. Snapshots selected are the corresponding asphericity values of the global minima, for each peptide system.

Fig. S5 Helical fraction for aggregated peptide systems. A) 7Q, B) 15Q, C) 35Q, D) 40Q, E) 45Q.

Fig. S6 Backbone contacts maps for monomeric peptide systems A) N17 + 7Q, B) N17 + 15Q, C) N17 + 35Q, D) N17 + 40Q, E) N17 + 45Q, reweighted to T= 380K. N17 and polyQ domains are marked in red and cyan respectively.

Fig. S7 Snapshots of anti-parallel sheet formation in N17+45Q system, at T=460K. N17 is denoted in red, and the polyQ regions of each peptide is coloured in lavender, olive green, cyan, magenta and blue respectively.

Fig. S8 Domain specific interactions of the polyQ domain in monomeric htt, reweighted to 380K. A) Number of Q-Q interactions per residue for each peptide system. B) Number of waters interacting with each glutamine residue

Video S1 Bundling of N17 domains in systems with short polyQ domains. N17, polyQ and PHE are denoted in red, cyan and violet respectively.

Video S2 External distribution of N17 in systems with long polyQ domains (polyQ≥35). N17, polyQ and PHE are denoted in red, cyan and violet respectively.

Acknowledgments

This research was supported by the National Science Foundation under grants CHE-1454948 and CHE-2202281. The authors acknowledge the University of Maryland supercomputing resources (http://hpcc.umd.edu) made available for conducting the research reported in this paper.

References

1. DiFiglia M, Sapp E, Chase K, Schwarz C, Meloni A, Young C, et al. Huntingtin is a cytoplasmic protein associated with vesicles in human and rat brain neurons. Neuron. 1995;14(5):1075–1081.

2. Zuccato C, Tartari M, Crotti A, Goffredo D, Valenza M, Conti L, et al. Huntingtin interacts with REST/NRSF to modulate the transcription of NRSE-controlled neuronal genes. Nature genetics. 2003;35(1):76–83.

3. Schulte J, Littleton JT. The biological function of the Huntingtin protein and its relevance to Huntington's Disease pathology. Current trends in neurology. 2011;5:65.

4. Donaldson J, Powell S, Rickards N, Holmans P, Jones L. What is the pathogenic CAG expansion length in Huntington's disease? Journal of Huntington's Disease. 2021;10(1):175–202.

5. Meriin AB, Zhang X, He X, Newnam GP, Chernoff YO, Sherman MY. Huntingtin toxicity in yeast model depends on polyglutamine aggregation mediated by a prion-like protein Rnq1. The Journal of cell biology. 2002;157(6):997–1004.

6. Cohen-Carmon D, Meshorer E. Polyglutamine (polyQ) disorders: the chromatin connection. Nucleus. 2012;3(5):433–441.

7. Temussi PA, Masino L, Pastore A. From Alzheimer to Huntington: why is a structural understanding so difficult? The EMBO journal. 2003;22(3):355–361.

8. Kim MW, Chelliah Y, Kim SW, Otwinowski Z, Bezprozvanny I. Secondary structure of Huntingtin amino-terminal region. Structure. 2009;17(9):1205–1212.

9. Perutz MF, Johnson T, Suzuki M, Finch JT. Glutamine repeats as polar zippers: their possible role in inherited neurodegenerative diseases. Proceedings of the National Academy of Sciences. 1994;91(12):5355–5358.

10. Chellgren BW, Miller AF, Creamer TP. Evidence for polyproline II helical structure in short polyglutamine tracts. Journal of molecular biology. 2006;361(2):362–371.

11. Masino L, Kelly G, Leonard K, Trottier Y, Pastore A. Solution structure of polyglutamine tracts in GST-polyglutamine fusion proteins. Febs Letters. 2002;513(2-3):267–272.

12. Khare SD, Ding F, Gwanmesia KN, Dokholyan NV. Molecular origin of polyglutamine aggregation in neurodegenerative diseases. PLoS computational biology. 2005;1(3):e30.

13. Warner IV JB, Ruff KM, Tan PS, Lemke EA, Pappu RV, Lashuel HA. Monomeric huntingtin exon 1 has similar overall structural features for wild-type and pathological polyglutamine lengths. Journal of the American Chemical Society. 2017;139(41):14456–14469.

14. Kang H, Vázquez FX, Zhang L, Das P, Toledo-Sherman L, Luan B, et al. Emerging β-sheet rich conformations in supercompact huntingtin exon-1 mutant structures. Journal of the American Chemical Society. 2017;139(26):8820–8827.

15. Urbanek A, Popovic M, Morató A, Estaña A, Elena-Real CA, Mier P, et al. Flanking regions determine the structure of the poly-glutamine in huntingtin through mechanisms common among glutamine-rich human proteins. Structure. 2020;28(7):733–746.

16. Lakhani VV, Ding F, Dokholyan NV. Polyglutamine induced misfolding of huntingtin exon1 is modulated by the flanking sequences. PLoS computational biology. 2010;6(4):e1000772.

17. Crick SL, Ruff KM, Garai K, Frieden C, Pappu RV. Unmasking the roles of N-and C-terminal flanking sequences from exon 1 of huntingtin as modulators of polyglutamine aggregation. Proceedings of the National Academy of Sciences. 2013;110(50):20075–20080.

18. Kelley NW, Huang X, Tam S, Spiess C, Frydman J, Pande VS. The predicted structure of the headpiece of the Huntingtin protein and its implications on Huntingtin aggregation. Journal of molecular biology. 2009;388(5):919–927.

19. Michalek M, Salnikov ES, Werten S, Bechinger B. Membrane interactions of the amphipathic amino terminus of huntingtin. Biochemistry. 2013;52(5):847–858.

20. Thakur AK, Jayaraman M, Mishra R, Thakur M, Chellgren VM, L Byeon IJ, et al. Polyglutamine disruption of the huntingtin exon 1 N terminus triggers a complex aggregation mechanism. Nature structural & molecular biology. 2009;16(4):380–389.

21. Bugg CW, Isas JM, Fischer T, Patterson PH, Langen R. Structural features and domain organization of huntingtin fibrils. Journal of Biological Chemistry. 2012;287(38):31739–31746.

22. Hoop CL, Lin HK, Kar K, Magyarfalvi G, Lamley JM, Boatz JC, et al. Huntingtin exon 1 fibrils feature an interdigitated β-hairpin–based polyglutamine core. Proceedings of the National Academy of Sciences. 2016;113(6):1546–1551.

23. Sivanandam V, Jayaraman M, Hoop CL, Kodali R, Wetzel R, van der Wel PC. The aggregation-enhancing huntingtin N-terminus is helical in amyloid fibrils. Journal of the American Chemical Society. 2011;133(12):4558–4566.

24. Chen M, Wolynes PG. Aggregation landscapes of Huntingtin exon 1 protein fragments and the critical repeat length for the onset of Huntington's disease. Proceedings of the National Academy of Sciences. 2017;114(17):4406–4411.

25. Sahoo A, Lee PY, Matysiak S. Transferable and Polarizable Coarse Grained Model for Proteins- ProMPT. Journal of Chemical Theory and Computation. 2022;18(8):5046–5055.

26. Toofanny RD, Daggett V. Understanding protein unfolding from molecular simulations. Wiley Interdisciplinary Reviews: Computational Molecular Science. 2012;2(3):405–423.

27. Khurshid B, Rehman AU, Luo R, Khan A, Wadood A, Anwar J. Heparin-Assisted Amyloidogenesis Uncovered through Molecular Dynamics Simulations. ACS omega. 2022;7(17):15132–15144.

28. Cino EA, Choy WY, Karttunen M. Comparison of secondary structure formation using 10 different force fields in microsecond molecular dynamics simulations. Journal of chemical theory and computation. 2012;8(8):2725–2740.

29. Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX. 2015;1:19–25.

30. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. Journal of molecular graphics. 1996;14(1):33–38.

31. Gowers RJ, Linke M, Barnoud J, Reddy TJ, Melo MN, Seyler SL, et al. MDAnalysis: a Python package for the rapid analysis of molecular dynamics simulations. In: Proceedings of the 15th python in science conference. vol. 98. SciPy Austin, TX; 2016. p. 105.

32. Lee PY, Sahoo A, Matysiak S. Folding and modulation of helical conformation of Glycophorin A by point mutations. Physical Chemistry Chemical Physics. 2023;.

33. Yesylevskyy SO, Schäfer LV, Sengupta D, Marrink SJ. Polarizable water model for the coarse-grained MARTINI force field. PLoS computational biology. 2010;6(6):e1000810.

34. Jayaraman M, Kodali R, Sahoo B, Thakur AK, Mayasundari A, Mishra R, et al. Slow amyloid nucleation via α-helix-rich oligomeric intermediates in short polyglutamine-containing huntingtin fragments. Journal of molecular biology. 2012;415(5):881–899.

35. Shirts MR, Chodera JD. Statistically optimal analysis of samples from multiple equilibrium states. The Journal of chemical physics. 2008;129(12):124105.

36. Yushchenko T, Deuerling E, Hauser K. Insights into the aggregation mechanism of PolyQ proteins with different glutamine repeat lengths. Biophysical journal. 2018;114(8):1847–1857.

37. Dougan L, Li J, Badilla CL, Berne B, Fernandez JM. Single homopolypeptide chains collapse into mechanically rigid conformations. Proceedings of the National Academy of Sciences. 2009;106(31):12605–12610.

38. Fiumara F, Fioriti L, Kandel ER, Hendrickson WA. Essential role of coiled coils for aggregation and activity of Q/N-rich prions and PolyQ proteins. Cell. 2010;143(7):1121–1135.

39. Wen J, Scoles DR, Facelli JC. Molecular dynamics analysis of the aggregation propensity of polyglutamine segments. PloS one. 2017;12(5):e0178333.

40. Crick SL, Jayaraman M, Frieden C, Wetzel R, Pappu RV. Fluorescence correlation spectroscopy shows that monomeric polyglutamine molecules form collapsed structures in aqueous solutions. Proceedings of the National Academy of Sciences. 2006;103(45):16764–16769.

41. Vitalis A, Wang X, Pappu RV. Atomistic simulations of the effects of polyglutamine chain length and solvent quality on conformational equilibria and spontaneous homodimerization. Journal of molecular biology. 2008;384(1):279–297.

42. Landles C, Bates GP. Huntingtin and the molecular pathogenesis of Huntington's disease: Fourth in Molecular Medicine Review Series. EMBO reports. 2004;5(10):958–963.

43. Steffan JS, Kazantsev A, Spasic-Boskovic O, Greenwald M, Zhu YZ, Gohler H, et al. The Huntington's disease protein interacts with p53 and CREB-binding protein and represses transcription. Proceedings of the National Academy of Sciences. 2000;97(12):6763–6768.

44. Jiang H, Poirier MA, Liang Y, Pei Z, Weiskittel CE, Smith WW, et al. Depletion of CBP is directly linked with cellular toxicity caused by mutant huntingtin. Neurobiology of disease. 2006;23(3):543–551.

45. Rockabrand E, Slepko N, Pantalone A, Nukala VN, Kazantsev A, Marsh JL, et al. The first 17 amino acids of Huntingtin modulate its sub-cellular localization, aggregation and effects on calcium homeostasis. Human molecular genetics. 2007;16(1):61–77.

46. Zheng Z, Li A, Holmes BB, Marasa JC, Diamond MI. An N-terminal nuclear export signal regulates trafficking and aggregation of Huntingtin (Htt) protein exon 1. Journal of Biological Chemistry. 2013;288(9):6063–6071.

47. Chaibva M, Burke KA, Legleiter J. Curvature enhances binding and aggregation of huntingtin at lipid membranes. Biochemistry. 2014;53(14):2355–2365.