Article

A Coarse-Grained Model for Polyglutamine Aggregation Modulated by Amphipathic Flanking Sequences

Kiersten M. Ruff,[1,2] Siddique J. Khan,[1] and Rohit V. Pappu[1,*]
[1]Department of Biomedical Engineering and Center for Biological Systems Engineering and [2]Division of Biology and Biomedical Sciences, Computational and Systems Biology Program, Washington University in St. Louis, St. Louis, Missouri

ABSTRACT The aggregation of proteins with expanded polyglutamine (polyQ) tracts is directly relevant to the formation of neuronal intranuclear inclusions in Huntington's disease. In vitro studies have uncovered the effects of flanking sequences as modulators of the driving forces and mechanisms of polyQ aggregation in sequence segments associated with HD. Specifically, a seventeen-residue amphipathic stretch (N17) that is directly N-terminal to the polyQ tract in huntingtin decreases the overall solubility, destabilizes nonfibrillar aggregates, and accelerates fibril formation. Published results from atomistic simulations showed that the N17 module reduces the frequency of intermolecular association. Our reanalysis of these simulation results demonstrates that the N17 module also reduces interchain entanglements between polyQ domains. These two effects, which are observed on the smallest lengthscales, are incorporated into phenomenological pair potentials and used in coarse-grained Brownian dynamics simulations to investigate their impact on large-scale aggregation. We analyze the results from Brownian dynamics simulations using the framework of diffusion-limited cluster aggregation. When entanglements prevail, which is true in the absence of N17, small spherical clusters and large linear aggregates form on distinct timescales, in accord with in vitro experiments. Conversely, when entanglements are quenched and a barrier to intermolecular associations is introduced, both of which are attributable to N17, the timescales for forming small species and large linear aggregates become similar. Therefore, the combination of a reduction of interchain entanglements through homopolymeric polyQ and barriers to intermolecular associations appears to be sufficient for providing a minimalist phenomenological rationalization of in vitro observations regarding the effects of N17 on polyQ aggregation.

INTRODUCTION

Aggregation and phase separation of polyQ tracts contribute directly to pathological processes that underlie neurodegeneration in Huntington's disease (1). Insoluble inclusions that accumulate within neuronal nuclei encompass the exon 1-encoded portion of the huntingtin protein (2). This block copolymeric sequence includes a 17-residue amphipathic segment (N17), a homopolymeric polyQ tract (Q_n), and a 38-residue proline-rich segment (C38). The amino acid sequences of N17 and C38 are MATLEKLMKAFESLKSF and P_{11}-QLPQPPPQAQPLLPQPQ-P_{10}, respectively. Exon 1 spanning sequences can be generated either by proteolysis of mutant huntingtin (3) or by translation of aberrantly spliced mutant mRNA transcripts with expanded CAG repeats (4).

Systematic studies have investigated the synergy between polyQ length and sequence context on various properties. These include the conformational properties of monomeric forms of exon 1 spanning fragments (5–11), temperature dependent solubilities (12,13), rates of loss of monomers into growing aggregates (14–23), rates of growth of fibrillar aggregates (24,25), and morphologies of aggregates in insoluble inclusions (26,27). Of particular relevance are the recent in vitro studies of Thakur et al. (18) and Crick et al. (12). The latter used a combination of methods to uncover the effects of N17 and C38 as modulators of driving forces for and mechanisms of polyQ aggregation. Although N17 has high intrinsic solubility, it lowers the overall solubility of N17-Q_n-$(Lys)_2$ constructs vis-à-vis Q_n-$(Lys)_2$ (Q_n-K2). Conversely, the high intrinsic solubility of C38 causes an enhancement of the solubility of Q_n-C38 vis-à-vis Q_n-K2. Experiments to probe the intrinsic aggregation mechanisms of polyQ tracts and the modulation of these mechanisms by N17 and C38 were performed under conditions where the magnitudes of driving forces for aggregation are equivalent for all constructs, i.e., the solutions were equivalently supersaturated.

Aggregation mechanisms were probed using two assays, viz., the rate of loss of tetramethyl rhodamine (TMR) fluorescence and the rate of gain of thioflavin T (ThT) fluorescence (see Fig. 1). In the TMR assay (28), disaggregated molecules with covalently attached TMR labels are incorporated into growing aggregates. Their fluorescence is quenched only when TMR dimers form within the aggregates. The loss of fluorescence depends on the formation of aggregates, and it appears to be agnostic regarding aggregate morphologies or specific structures formed within the aggregates (28).

Submitted May 22, 2014, and accepted for publication July 9, 2014.

*Correspondence: pappu@wustl.edu

Kiersten Ruff and Siddique Khan contributed equally to this work.

Editor: Michael Feig.

0006-3495/14/09/1226/10 $2.00

http://dx.doi.org/10.1016/j.bpj.2014.07.019

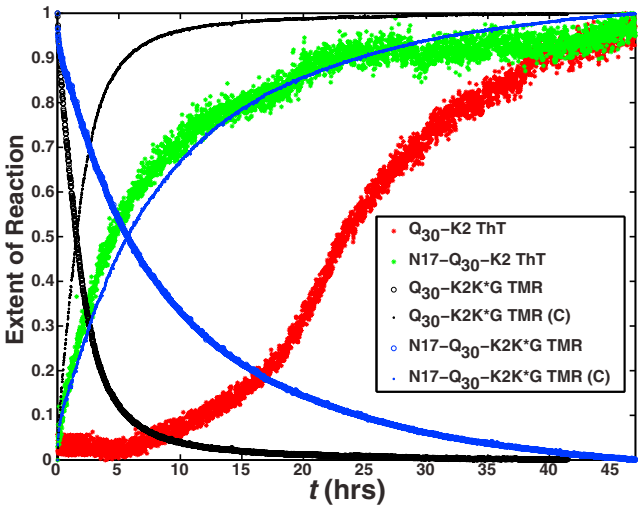

FIGURE 1 Summary of the overall rates of aggregation for polyQ constructs with and without N17. The data are from the work of Crick et al. (12). The rate of loss of TMR fluorescence for Q_{30} peptides with and without the N17 module is quantified by the blue and black open circles, respectively. The blue and black stars plot the complement of these data, which provide an assessment of the rates of growth of aggregates as monitored by the TMR assay. These data are to be compared to the green and red stars that quantify the rate of increase in ThT fluorescence for Q_{30} peptides with and without the N17 module. The rate of loss of TMR fluorescence and the rate of gain in ThT fluorescence are plotted in normalized units and depict the rates of the change of reactions that are being monitored by the distinct probes. Here, K* refers to the lysine residue that is covalently modified by attachment of the TMR dye.

The TMR assay affords enhanced temporal and spatial resolution (12,28). In contrast, the gain in ThT fluorescence measures the rate of growth of fibrillar species.

Based on the data in Fig. 1, we make the following observations. The rate of loss of TMR fluorescence is slowed by the presence of N17, as shown by the comparison of data for Q_{30}-K2K*G to those for N17-Q_{30}-K2K*G. Aggregates detected by the TMR assay, which were shown to be spherical clusters using electron microscopy for Q_{30}-K2K*G, form early and on distinct timescales when compared to the slow fibril formation observed from the gain in ThT fluorescence for Q_{30}-K2. In the presence of N17, aggregates monitored in the TMR assay form on timescales equivalent to that for the formation of fibrillar aggregates as monitored by the ThT assay. It appears that N17 destabilizes nonfibrillar states and accelerates the formation of fibrils. The effect of N17 as an accelerator of fibril formation is also observed in exon 1-mimicking fragments such as N17-Q_n-C38, yielding equivalent fibril growth rates as compared to N17-Q_n-K2 for similar degrees of supersaturation (12).

The effects of N17 on the conformational properties of monomeric forms of polyQ and the frequencies of intermolecular associations have been investigated in atomistic simulations by Williamson et al. (29). These simulations were based on the ABSINTH implicit solvation model and forcefield paradigm (30). The physical underpinnings of the ABSINTH paradigm have been detailed in previous publications (30–32). It has been used successfully—as judged by comparisons to experimental data—in simulations of intrinsically disordered proteins (33–39), proline-rich systems (31), protein unfolded states (40), and polyQ-containing systems (7,8,29,41). Williamson et al. compared the conformational properties and intermolecular associations of N17-Q_n peptides to that of Q_n. The helical propensity of N17 increases for short polyQ tracts ($n < 10$). It decreases sharply for $10 < n < 20$, and becomes negligibly small as the polyQ length increases ($n > 20$). N17 unfolds and adsorbs by forming an intramolecular, interdomain interface with the surface of the globular polyQ. The latter mimics a high local concentration of urea, thus denaturing N17. A similar proposal has been put forth for the effect of polyQ on cytosolic proteins (42). N17-Q_n peptides form patchy-colloid-like architectures (43,44) for $n > 20$ (see Fig. S1 in the Supporting Material). This weakens the effective surface tension, as well as the effective favorable energy that is associated with interactions between glutamine residues. Williamson et al. (29) also quantified the comparative frequency of forming dimers for Q_n versus N17-Q_n for $n =$ 15, 25, and 35. N17 diminishes the frequency of dimer formation by up to a factor of 3 in the temperature range 298–315 K. By imparting directionality and specificity to intermolecular associations, N17 helps reduce the random associations characteristic of the uniformly sticky, intrinsically disordered homopolymeric polyQ globules.

Here, we develop a phenomenological model that provides a conceptual framework to bridge the observations regarding N17-modulated polyQ aggregation as quantified in vitro and in simulations. We seek to understand how N17 diminishes the frequency of random intermolecular associations while simultaneously destabilizing small, nonfibrillar aggregates and accelerating fibril formation. The physics of flexible homopolymers in poor solvents provides the conceptual lynchpin. Intermolecular associations of homopolymers in poor solvents, such as polyQ in aqueous milieus, promote entanglements between pairs of molecules (45,46). Within dimers or higher-order clusters, the polymers are likely to have liquidlike organization, and chains can become entangled with each other because the intrachain and interchain interactions are equivalent for homopolymers (47,48). Our reanalysis of the simulation results of Williamson et al. (29) shows clear evidence for entanglements between polyQ domains. We also find that N17 reduces interchain entanglements between polyQ domains. Therefore, N17 diminishes intermolecular associations between globular polyQ while reducing interchain entanglements through these domains.

Coarse-grained models provide access to large spatial scales and morphologies (49–51). Here, we introduce an ultra-coarse-graining approach (52) that leverages the information gleaned from analysis of atomistic simulation results to develop a phenomenological (as opposed to transferrable) coarse-grained model that resembles efforts often deployed

in studies of colloidal particles. We model globule-forming polyQ molecules as spherical colloidal particles. The effects of N17 that include a barrier to interparticle interactions and the quenching of interparticle entanglements are modeled using suitable pair potentials. The remainder of the narrative is organized as follows: First, we present the methodological details of our Brownian dynamics (BD) simulations. This is followed by a presentation of the results obtained from BD simulations designed to study the aggregation of colloidal particles that interact via distinct classes of pair potentials. We show that the phenomenological pair potentials provide a minimalist framework for understanding the effects of the N17 module on polyQ aggregation kinetics and morphologies.

MATERIALS AND METHODS

We performed three-dimensional BD simulations composed of 2×10^4 spherical particles in a cubic box of length $L = 256R_g$ with periodic boundary conditions. Here, R_g corresponds to the particle radius. We set $R_g = 1$ unit, and all distances are measured in units of R_g. The volume fraction is $f_v = 0.005$, which is in accord with a supersaturated system of polyQ molecules (12). The equation of motion for each particle i is given by

$$\ddot{\vec{r}}_i = -\vec{\nabla}U_i - \Gamma\dot{\vec{r}}_i + \vec{W}_i(t). \quad (1)$$

Here, Γ is the monomer friction coefficient and W_i is the random force on each particle i, defined as a Gaussian white noise that satisfies the fluctuation-dissipation relation. In our simulations, $\Gamma = 0.5$ and $\Delta t = 0.005$ in reduced time units of $R_g(m/kT)^{1/2}$, where $m = 1$. Accordingly, $t = 1$ is the time it takes to diffuse a distance of R_g.

Developing a suitable set of phenomenological models

In the ultra-coarse-grained model, each molecule (Q_n or N17-Q_n) is modeled as a spherical particle. This is justified by the observation that monomeric polyQ forms compact globules in aqueous solvents, and this property is preserved for N17-Q_n molecules providing $n > 20$ (29) (see Fig. S2). Therefore, the early stages of aggregation in supersaturated solutions will most likely involve the diffusion and collision of globules that can coalesce through their polyQ domains. We use reduced units and normalize all lengthscales by the radius of gyration (R_g) of the molecule in question. Accordingly, an interparticle distance, R, quantified as the distance between the centers of particles, is normalized and denoted by X, where $X = R/R_g$. $X = 1$ implies that $R = R_g$ and $X = 2$ implies that $R = 2R_g$. Therefore, X is a normalized interparticle separation, i.e., it is R normalized by R_g. We stipulate that $X = 1$ corresponds to the fully entangled state, because values of $X < 1$ are rarely observed in atomistic simulations of monomer-dimer equilibria, indicating a significant excluded volume barrier below $X = 1$; $X = 2$ corresponds to the docking state, and $X > 2$ corresponds to dissociated particles.

The results presented in the main text are based on four different interparticle potentials, $U(X)$ (Fig. 2, *A–D*). We use two potentials to model different archetypes of docking plus entanglement (DPE) and these are designated DPE-1 and DPE-2. A potential with a single well and a featureless entangled state is shown in Fig. 2 *A* for $1 < X < 2$. For this potential (DPE-1), different degrees of entanglement are energetically equivalent, and the realization of entangled states arises purely due to entropic considerations. Alternatively, the fully entangled state can be of equivalent stability to the docked state, because interchain contacts can be maximized in fully entangled states. Accordingly, DPE-2 (Fig. 2 *B*) shows a bistable double-well potential with two minima, one at $X = 1$ and one at $X = 2$, implying that a pair of particles can either dock ($X = 2$) or entangle ($X = 1$).

Fig. 2 *C* shows a single-well potential with no docking barrier (docking with no barrier (DNB)). This pair potential is designed to capture one of the effects of the N17 module, quenching of interchain entanglements. Finally, Fig. 2 *D* shows a single-well potential with a barrier to docking (docking with a barrier (DWB)). This pair potential is designed to capture the two effects attributed to N17, namely, minimization of interchain entanglements and creation of a barrier that limits intermolecular associations. Note that the DNB and DWB potentials are designed to accommodate some degree of interchain entanglement, as evidenced by the location of the steep excluded-volume wall at values of $X < 2$ rather than at $X = 2$.

Functional forms for DPE-1, DPE-2, DNB, and DWB

The interparticle interaction potentials for DPE-1, DPE-2, DNB, and DWB are linear combinations of three terms, U_M, U_D, and U_B, that are pair potentials written in terms of the normalized interparticle distance, X. The term U_M is a $2n - n$ Mie potential and is written as

$$U_M(X) = \varepsilon_M\left[\left(\frac{X_m}{X}\right)^{2n} - 2\left(\frac{X_m}{X}\right)^{n}\right]. \quad (2)$$

Here, $n = 18$, ε_M is the well depth of the potential, and X_m corresponds to the location of the minimum of the Mie potential. We use the Mie potential instead of a Lennard Jones potential because we wish to have precise control over the location of minima, and because this potential is used in simulations of particles that are agglomerations of atoms, i.e., colloidal particles. The Mie potential provides a useful way to model a steeply rising excluded-volume wall and short-range attractions that allow us to

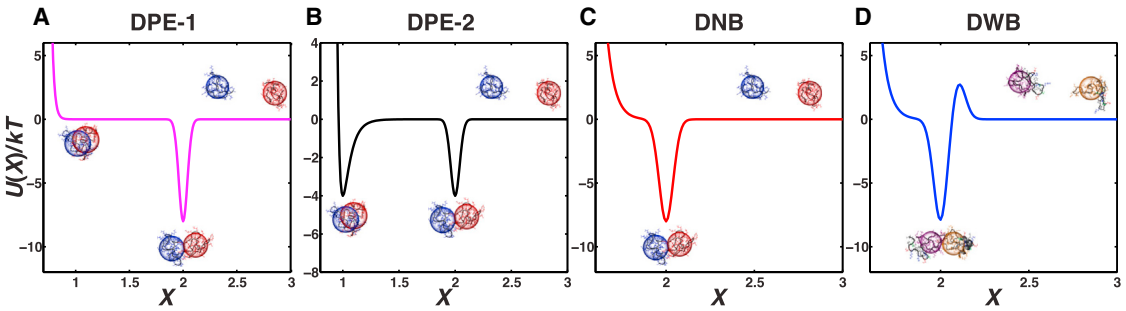

FIGURE 2 Phenomenological pair potentials designed to model the effective interactions in coarse-grained simulations of Q_n and N17-Q_n globules. To see this figure in color, go online.

incorporate additional terms to capture specific effects such as entanglements or docking. The term U_D is an inverted Gaussian written as

$$U_D(X) = -\varepsilon_D \exp\left[-\frac{(X-2)^2}{2\delta_D^2}\right]. \quad (3)$$

Here, ε_D is the depth of the potential well and δ_D controls the width of the well. In all cases, $\delta_D = 0.04$. The term U_B is a Gaussian written as

$$U_B(X) = e_B \exp\left[-\frac{(X-2.1)^2}{2\delta_B^2}\right]. \quad (4)$$

Here, ε_B corresponds to the height of the barrier and δ_B controls the width of the barrier. In all cases, $\delta_B = 0.04$.

The DPE-1, DPE-2, DNB, and DWB models take the forms shown in Eqs. 5–8 below, with all well depths and barrier heights written in units of kT, where k is Boltzmann's constant and T is the simulation temperature.

$$\begin{gathered} U_{\text{DPE-1}} = U_M(X) + U_D(X) \\ \varepsilon_M = 0.001kT, X_m = 1, \varepsilon_D = 8kT \end{gathered}; \quad (5)$$

$$\begin{gathered} U_{\text{DPE-2}} = U_M(X) + U_D(X) \\ \varepsilon_M = 4kT, X_m = 1, \varepsilon_D = 4kT \end{gathered}; \quad (6)$$

$$\begin{gathered} U_{\text{DNB}} = U_M(X) + U_D(X) \\ \varepsilon_M = 0.01kT;\ X_m = 2,\ \text{and}\ \varepsilon_D = 8kT \end{gathered}; \quad (7)$$

$$\begin{gathered} U_{\text{DWB}} = U_M(X) + U_D(X) + U_B(X) \\ \varepsilon_M = 0.01kT;\ X_m = 2,\ \varepsilon_D = 8kT,\ \text{and}\ \varepsilon_B = 3kT \end{gathered}. \quad (8)$$

The cutoff for the range of interparticle interactions was set to $2.2R_g$ for the DPE and DNB models and to $2.3R_g$ for the DWB model. The slightly larger cutoff for the latter is to accommodate the inclusion of a barrier beyond $X = 2$. All contributions to the interparticle forces beyond the respective cutoffs are neglected. Each simulation starts from a random configuration and is run for 4×10^6 time steps, corresponding to a simulation time of $t = 2 \times 10^4$. This is sufficient for the achievement of steady-state behavior. All simulation results are averaged over five runs for each model.

Parameterization of the pair potentials

We used a single-well potential based on Eq. 2 and varied the depth of the well in this potential until we achieved the steady-state behavior of spherical clusters coexisting with large, aspherical fractal aggregates (see Figs. S3–S5). In addition, the crossover lengthscale between spheres and clusters was required to reach a saturation value. Based on these considerations, we converged on an energy scale of $8kT$ for the system of interacting particles. Accordingly, for the DPE-1, DNB, and DWB systems, the well depth was set to $8kT$, and for the DPE-2 system, the two wells were each set to $4kT$, such that the overall energy scale for interparticle interactions remains $8kT$, except that it is now split between the docked and entangled states. The overall energy scale $8kT$ is set to control the extent of aggregation. This energy scale does not distinguish aggregation that yields small spherical aggregates versus aggregation that yields large linear species. Instead, setting the overall energy scale is equivalent to fixing the degree of supersaturation, and this should be the same for all four models. If the stability of the entangled state is to be increased, then there has to be a concordant weakening of the stability of the docking state. This feature allows us to mimic the experimental approach of assessing aggregation mechanisms at equivalent degrees of supersaturation as opposed to equivalent concentrations. Finally, the barrier for interparticle interactions in the DWB system was parameterized to yield a slower convergence toward smaller spherical aggregates than is seen for the DPE systems. This choice reflects the mapping of the DWB potential to systems with the N17 module, whereas the DPE potentials map to the polyQ system.

RESULTS

Reanalysis of atomistic simulation results shows evidence for interchain entanglements between polyQ domains and modulation of these entanglements by N17. We analyzed the simulation results of Williamson et al. (29) for dimers of Q_{35} and N17-Q_{35}. The chain-length-normalized radii of gyration are equivalent for monomeric forms of Q_n and N17-Q_n ($n = 30$ and 35, respectively) (see Fig. S2). This implies that their densities are equivalent, and therefore, we use the simulation results for Q_{35}-containing systems as proxies for Q_{30}-containing systems. For each of the associated molecules in the simulated ensembles, we calculated theintersected volume fraction as $V_F = SSI/(V_{m1} + V_{m2})$. Here, SSI corresponds to the sphere-sphere intersection and is defined as

$$SSI = \frac{\pi\left(R_{g,m1} + R_{g,m2} - d_{com}\right)^2\left(d_{com}^2 + 2d_{com}R_{g,m2} - 3R_{g,m2}^2 + 2d_{com}R_{g,m1} + 6R_{g,m1}R_{g,m2} - 3R_{g,m1}^2\right)}{12d_{com}},$$

where d_{com} is the distance between the centers-of-mass of two interacting Q_{35} domains and $R_{g,m1}$ and $R_{g,m2}$ correspond to the radii of gyration of the Q_{35} domains in molecules 1 and 2, respectively. In addition, V_{m1} and V_{m2} are the volumes of the spheres with radii $R_{g,m1}$ and $R_{g,m2}$, respectively. Unlike other measures of size, the radius of gyration is a formal order parameter in polymer physics theories and it provides a direct assessment of chain density, which is altered by aggregation via docking and/or entanglement.

If dimers form strictly through the docking of undeformed surfaces, then $V_F = 0$. Conversely, V_F increases as the degree of entanglement between polyQ domains increases. Fig. 3, *A* and *B*, shows comparative assessments of the probability densities of realizing conformations for dimers with specific degrees of entanglements between polyQ domains, as observed in simulations of Q_{35} (Fig. 3 *A*) versus N17-Q_{35} (Fig. 3 *B*) molecules. Fig. 3 *C* provides a quantitative summary of these comparisons. Here, we

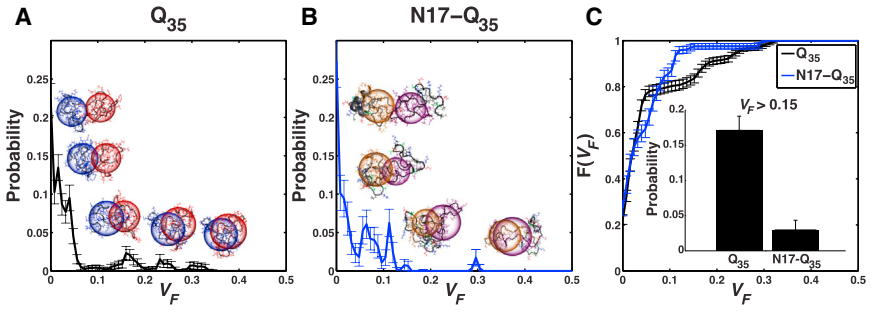

FIGURE 3 Quantifying the degree of entanglement between polyQ domains (T = 315 K) as observed for Q_{35} peptides in atomistic simulations for Q_{35} and N17-Q_{35}. (*A* and *B*) Probability density distributions using a bin width of 0.008. This quantifies the likelihood of observing conformations of dimers that have intersection volume fractions, V_F, that lie within specific intervals. These intersected volumes are computed over the polyQ domains alone, as described in the main text. Shown are two sets of representative conformations corresponding to distinct values of V_F. The radius of each translucent sphere corresponds to the R_g of each Q_{35} domain for the corresponding conformation. (*C*) Cumulative distribution functions for Q_{35} (*black*) and N17-Q_{35} (*blue*). The bumpiness evident in the probability distribution functions is the result of finite-sized artifacts, which are intrinsic to quantifying entanglements from simulations of monomer-dimer equilibria as opposed to monomer-oligomer equilibria. The latter simulations were not performed using atomistic descriptions, because these are computationally prohibitive at this juncture. $F(V_F)$ quantifies the probability of finding dimer conformations where the intersected volume fraction is less than or equal to a particular value for V_F. (*Inset*) Probability of finding entangled conformations for Q_{35} versus N17-Q_{35} at volume fractions of $V_F > 0.15$. To see this figure in color, go online.

calculated the cumulative distribution function, which allows us to compute the probability of realizing dimer conformations that have values of V_F greater than a specific threshold. It can be seen that the N17 module diminishes the degree of entanglement through polyQ domains, such that the probability of realizing conformations with $V_F > 0.15$ is an order of magnitude higher for Q_{35} than for N17-Q_{35} (Fig. 3 *C*, *inset*).

How to analyze steady-state and kinetics profiles to distinguish aggregation mechanisms

Our goal is to achieve a comprehensive analysis of the contributions of all of the particles in the system and analyze the aggregation process using theories that yield precise predictions in terms of the overall structure factor. Further, we avoid analyses that are ad hoc in nature, because the results can depend on nuances such as the definitions of aggregate sizes and criteria for deciding on aggregation. The experimental data (Fig. 1) suggest that a class of models based on diffusion-limited aggregation (53–57) and diffusion-limited cluster aggregation (58–61) have relevance to the early phases of polyQ aggregation. Guided by theories for the kinetics of gelation and aggregation, we analyzed the temporal evolution and steady-state properties of the structure factor $S(q)$, for a system of spherical particles (62–64). $S(q)$ provides information regarding the structures that form across distinct lengthscales. Here, q has units of inverse wavelength. High-q values provide information regarding structures at small lengthscales and low-q values provide structural information regarding those at larger lengthscales. In the early stages of aggregation, spherical clusters can form, and according to Porod's law (65), the presence of spherical clusters is detectable, because $S(q)$ should scale as q^{-4}. This would imply a straight line with a slope of –4 in a log-log plot of $S(q)$ versus q. Spherical clusters form during the early phases, but these spheres do not grow into macroscopic aggregates. If the arrival time of new clusters is faster than the rearrangement timescales within the aggregates, then the spherical structures that form on small lengthscales give way to large, aspherical, and extended structures (66) such that $S(q)$ scales as $q^{-1.8}$. In this work, these large, open aspherical structures will serve as proxies for linear, fibrillar morphologies that form in the postelongation phase of protein aggregation.

At steady state, the log-log plots of $S(q)$ versus q will have two linear regimes, one with a slope of ~–4 and the other with a slope of ~–1.8. The former corresponds to spherical aggregates that form across small lengthscales (high-q) and the latter to those that form across large lengthscales (low-q). The q-regime at which the slopes cross over reflects the joint contributions of the average aggregate sizes and shapes. If large aggregates form in solution, then the crossover lengthscale, q_c shifts to lower values. In a similar way, if aspherical aggregates form in the high-q regime, then q_c also shifts to lower values. Section S1 in the Supporting Material summarizes our formal approach for detecting the two linear regimes for steady-state structure factors. Once we identify the corresponding linear regimes, we monitor the temporal evolution of the slopes within each region, and this helps us analyze aggregation kinetics across distinct lengthscales.

Analysis of structure factors

The preceding paragraphs set the expectations for our analysis of the steady-state structure factors and their temporal evolution. We first present a summary of the observations extracted from the steady-state structure factors and their temporal evolutions for each of the four models. For a system of N particles, the structure factor at time t is written in terms of the q-vectors as

$$S(q;t) = \sum_{i=1}^{N}\sum_{j=1}^{N}\exp\left[i\vec{q}\times\left(\vec{x}_i-\vec{x}_j\right)_t\right]. \qquad (9)$$

Here, $\vec{x}_i$ and $\vec{x}_j$ are the position vectors of particles i and j, respectively. The relevant results are shown in Figs. 4–6.

Fig. 4, *A–D*, shows the temporal evolution of $S(q)$ for particles interacting via the DPE-1, DPE-2 DNB, and DWB potentials, respectively. These are shown as log-log plots. For each model, we analyzed the long-time steady-state structure factors to uncover the model-specific high-q and low-q regions that yield slopes of ~–4 and ~–1.8, respectively (Fig. 4). The steady-state structure factors were also analyzed to extract the crossover lengthscale between the two regimes. The resultant values for q_c are 0.71 ± 0.01, 0.90 ± 0.01, 0.67 ± 0.01, and 0.59 ± 0.02 for the DPE-1, DPE-2, DNB, and DWB systems, respectively. These values show a shift toward higher q values for DPE-1 and DPE-2 vis-à-vis DNB and DWB. Therefore, the steady-state aggregates are smaller in the presence of entanglements and are largest when entanglements are quenched and there is a barrier to interparticle association.

Having identified the high-q and low-q regimes that yield distinct slopes for each steady-state structure factor, we quantify the evolution of the slopes in these regimes as a function of time for each of the models (Fig. 5, *A–D*). The regimes within which these slope evolutions are calculated are different for each model because of the different q_c values (see summary in Table S1). As noted in Materials and Methods, we use reduced units for time by calibrating the simulation time against the average time it takes for a single particle to traverse a distance R_g via translational diffusion. For the DPE-1 model, the energetic equivalence between different entangled states leads to slightly aspherical aggregates as steady-state species in the high-q regime—note the convergence toward a slope of –3.73 as opposed to –4. The timescales for converging on steady-state slopes corresponding to the high-q and low-q regimes are an order of magnitude different from each other for DPE-1, and this is true for DPE-2 as well. Inasmuch as either model mimics the behavior of polyQ sans the N17 module, the results show qualitative agreement with expectations based on in vitro experiments, because distinct types of aggregates form on distinct timescales. Clearly, interchain entanglements engender a separation of timescales, and this separation is more pronounced for the DPE-2 model.

The results for DPE-1 and DPE-2 are to be contrasted with results from equivalent analyses for the DNB and DWB models (Figs. 4 and 5, *C* and *D*, respectively). The DNB and DWB systems mimic the ability of N17 to reduce interchain entanglements. Fig. 5 *C* shows that the temporal gap between the formation of early spherical clusters and the cluster-cluster associations that lead to structures that yield the $q^{-1.8}$ scaling behavior is narrowed considerably in the absence of entanglements. Thus, the formation of spherical clusters and large aspherical/linear aggregates is temporally synchronized in congruence with the observations of Fig. 1. In the DWB system, entanglements are quenched, this leads to overall synchrony between the aggregation timescales across the two distinct lengthscales. However, the presence of a barrier to interparticle interactions slows down the overall timescales for aggregation compared to that of the DNB system. Further, the smaller aggregates are relatively aspherical, as evidenced by the slope of –3.5 as opposed to –4 for the steady-state value of the slope in the high-q regime. Therefore, the DWB system that incorporates both effects of N17 from atomistic simulations leads to overall aggregation kinetics that agrees with in vitro observations.

Based on the preceding analysis, we take the DPE-2 and DWB models as the optimal approximations for the aggregation mechanisms of Q_n versus N17-Q_n molecules, respectively. In Fig. 6, we show comparative kinetics for the evolution of the slopes corresponding to the high-q and low-q regimes extracted from the steady-state structure factors for DPE-2 and DWB, respectively. The presence of N17 should reduce the frequency of intermolecular associations and hence slow down the kinetics of early aggregation. This behavior was observed when comparing the rates of loss of TMR fluorescence with and without the N17 module

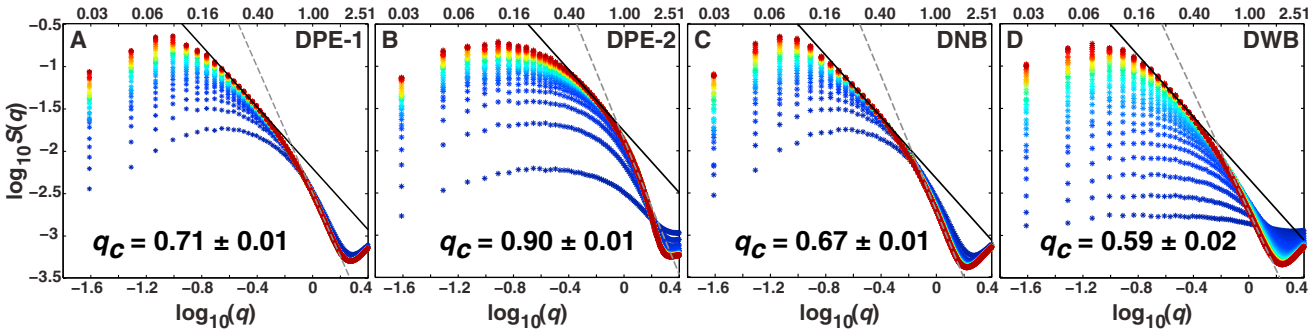

FIGURE 4 Log-log plots of the temporal evolution of the structure factor for aggregation modeled using the DPE-1, DPE-2, DNB, and DWB pair potentials. The structure factors plotted with cooler-colored symbols correspond to early time points and those plotted with hotter-colored symbols correspond to the later times and include steady-state behavior. The structure factors were calculated between $t = 500$ and $t = 20{,}000$ in increments of 500. Solid black lines correspond to the line of best fit to the low-q regime of the steady-state structure factor and dashed lines to that of the high-q regime. The crossover lengthscale, q_c, is calculated from the steady-state structure factor for each potential as the point where the two lines intersect.

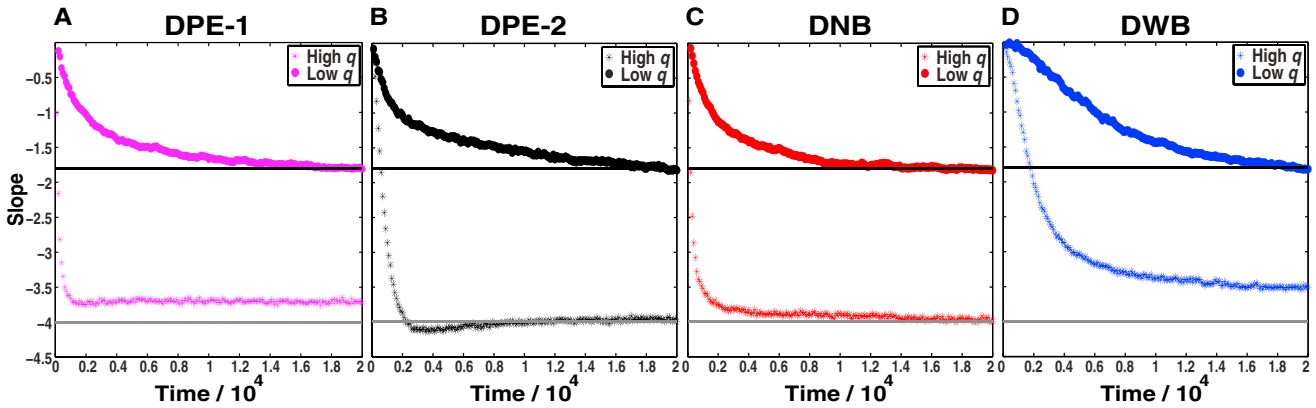

FIGURE 5 Temporal evolution of the slopes in the high- and low-q regimes for each of the four models. (*A*) The DPE-1 model, the formation of small, nearly spherical aggregates is detected by monitoring the evolution of the slope of the straight line in the high-q region ($1.18 \leq q \leq 1.40$) and the formation of linear aggregates is detected by monitoring the temporal evolution of the slope of the straight line in the low-q region ($0.22 \leq q \leq 0.44$). (*B*) The corresponding analysis for the DPE-2 model analyzes the evolution of slopes in the regimes $1.08 \leq q \leq 1.30$ and $0.44 \leq q \leq 0.66$, respectively. (*C*) For the DNB model, the evolution of slopes is followed in the regimes $0.96 \leq q \leq 1.18$ and $0.20 \leq q \leq 0.42$. (*D*) The same analysis for the DWB model by monitoring the slopes in the regimes $0.98 \leq q \leq 1.20$ and $0.17 \leq q \leq 0.39$, respectively. To see this figure in color, go online.

(Fig. 1, *A* and *B*, respectively), and it is reproduced in the coarse-grained model (Fig. 6 *A*). This plot compares the rate of convergence to the steady-state value for the slope in the high-q regime for DPE-2 and DWB. In accordance with the experimental data, Fig. 6 *A* shows that the presence of a barrier slows the early stages of aggregation such that the lack of entanglement is sufficient to synchronize the timescales for aggregation on small and large scales, whereas the presence of entanglements leads to significant slow-down in the growth of larger, linear aggregates (see Fig. 6 *B* and Fig. 5, *A–D*). Fig. S6 shows a plot of the rate of loss of monomers into aggregates for each of the four models, and Fig. S7 shows plots equivalent to those in Fig. 6 for all four models. These plots support the overall conclusions summarized above.

Fig. 7 shows representative snapshots for the DPE-2 and DWB models drawn from the steady-state ensembles. Fig. S8 shows an equivalent plot for all four models. The aggregates are colored based on the number of particles within each aggregate. This provides a visual depiction of the prevalence of numerous small, mostly spherical aggregates for the DPE-2 system. This highlights how interchain entanglements suppress growth into large aspherical aggregates. Conversely, for the DWB system, the snapshots reveal a variety of large aspherical aggregates and few small spherical aggregates. Indeed, for the DWB system, even the small aggregates are mildly aspherical and this shows that

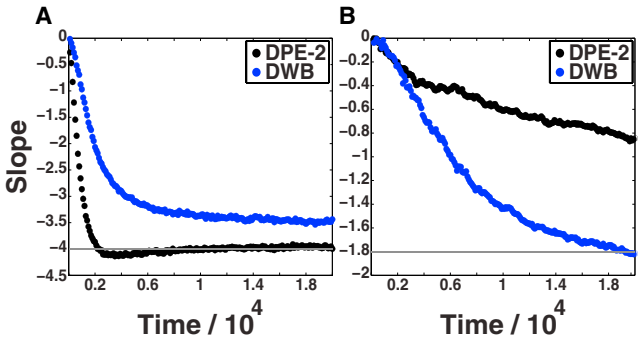

FIGURE 6 Comparative assessments of the rates of aggregation in the high-q (*A*) and low-q (*B*) regimes for DPE-2 and DWB models. We choose the high-q regime ($1.08 \leq q \leq 1.30$) (*A*) and the low-q regime ($0.17 \leq q \leq 0.39$) (*B*) from the steady-state structure factors for the DPE-2 and DWB models, respectively. These choices are made to quantify comparative kinetics on equivalent lengthscales. Accordingly, we note that stable entanglements arrest the ability to form large linear aggregates, as evidenced by the inability of the DPE-2 system to achieve a slope of –1.8 on lengthscales corresponding to the steady-state low-q regime for the DWB system (*B*). To see this figure in color, go online.

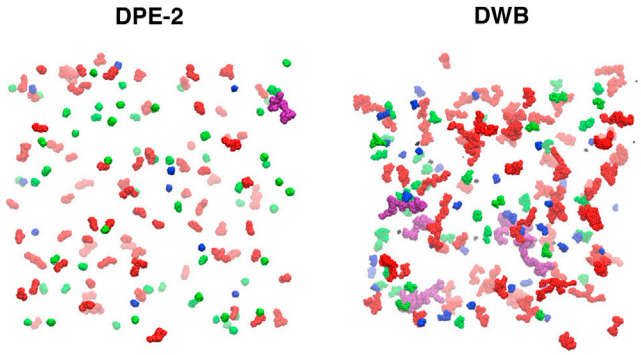

FIGURE 7 Representative snapshots of the steady-state aggregates that result from the simulations based on DPE-2 and DWB (*left* and *right*, respectively). Molecules in the snapshots are colored according to the size of the corresponding aggregates, i.e., the number of molecules, n, within the aggregate to which they belong. The scheme is as follows: *black*, $n \leq 10$; *blue*, $10 < n \leq 50$; *green*, $50 < n \leq 100$; *red*, $100 < n \leq 400$; *purple*, $n > 400$. The snapshots were generated using the Visual Molecular Dynamics package (67).

the reduction of entanglements promotes the formation of large aspherical aggregates through enhanced cluster-cluster associations.

DISCUSSION

We have introduced a phenomenological coarse-grained model designed to connect the observations from the small lengthscales accessible to atomistic simulations to those from considerably larger lengthscales in in vitro experiments. The structure factors provide a concise and complete analysis of the structures that form on all lengthscales and have the advantage of enabling direct connections to multiangle dynamic light-scattering experiments (5,9,14,68–70).

Results from the coarse-grained BD simulations demonstrate that small spherical clusters and large aspherical aggregates form on distinct timescales for particles interacting via the DPE potentials. Further, the overall aggregate sizes, characterized by the crossover lengthscale are smallest for the DPE-2 system. In contrast, eliminating entanglements (DNB) synchronizes the timescales for forming small/spherical aggregates and large aspherical/linear aggregates, whereas the introduction of a docking barrier in the DWB system slows the overall aggregation kinetics while further destabilizing small spherical aggregates. It appears that the observations regarding the incipient interactions between pairs of molecules can be used to extract insights regarding the coarse-grained morphologies and dynamics that develop on larger lengthscales.

Explanations for the coarse-grained simulation results

In diffusion-limited cluster aggregation, which is captured by the DNB system, the spherical clusters do not grow isotropically, because the arrival times of the monomers and clusters at the growing front are faster than the rearrangement times required to achieve a macroscopic spherical aggregate. Conversely, thermal fluctuations set up deformations on the surfaces of spherical clusters, and the size scale of these deformations can be commensurate with the sizes of the spherical clusters (56). The deformations are akin to instabilities, and these grow through diffusion-limited cluster aggregation, which gives rise to open, aspherical structures that generate a distribution of voids around them, thus making these aggregates fundamentally different from densely packed, spherical macroscopic aggregates.

We find that entanglements destabilize the growth into large, aspherical structures. The result is a distribution of spherical aggregates that coexist with shorter aspherical aggregates and this seems to be consistent with the recent electron microscopy data of Crick et al. for peptides that approximate homopolymeric polyQ. In addition, the quantum yield for ThT fluorescence is an order of magnitude smaller for Q_{30}-K2 vis-à-vis N17-Q_{30}-K2 and is suggestive of smaller, less fibrillar species forming for homopolymeric polyQ.

Entanglements are the result of the homopolymeric nature of polyQ and reflect the intrinsic conformational heterogeneity and degeneracy of interchain interaction possibilities through polyQ domains. The presence of entanglements offsets the deformation of spherical clusters, because the particles have an additional lengthscale to explore. Therefore, unlike the DNB and DWB systems, where rearrangements within spherical clusters are quenched by design, the presence of entanglements allows for rearrangements, thus arresting the molecules within long-lived spherical aggregates. In the DPE-2 case, these cannot grow into macroscopic sizes because the entangled and docked states are equivalent in energy for a pair of particles. This leads to energetic frustration and local intracluster rearrangement, which can interfere with cluster-cluster associations.

Connecting the results to polyQ biology

Experiments highlight the importance of heterogeneous distributions of small, spherical, and diffuse aggregates for polyQ toxicity in living cells and model organisms (71,72). In direct contrast, large, well-ordered aggregates are thought to protect cells against polyQ toxicity. Within the framework of our model, it would appear that enhanced interchain entanglements that derive from the absence of flanking sequence modules or increased polyQ length could give rise to a long-lived distribution of smaller, spherical aggregates that contribute to cellular toxicity. Inasmuch as this is a prediction, it should be possible to design sequences that modulate the degree of entanglement and assess the effects of this modulation on in vitro aggregation as well as cellular toxicities.

Making the models transferrable and development of multiresolution models

The models introduced here are purely phenomenological as opposed to being transferrable. It should be possible to design transferrable potentials based on quantitative statistics extracted from atomistic simulations of varying polyQ lengths and different sequence contexts, such as amphipathic-flanking sequences from other polyQ-disease-related proteins. We expect that the overall energy scale for aggregation will be governed by the decrease in saturation concentrations with increased polyQ length (12). Further, we expect that the precise nature of the entangled state, i.e., whether it is featureless or whether it allows for one or more minima, will be governed by the nature of the interface between the flanking sequence and the polyQ globule. It is worth emphasizing that the framework developed here applies only for flanking sequences and polyQ lengths where the overall density is akin to that of polyQ globules. Accordingly, the modeling of molecules as spherical particles will

not apply for constructs with the C38 module, because this semiflexible, proline-rich region behaves differently from N17 in atomistic simulations (K. Ruff and R. Pappu, unpublished).

The ultra-coarse-graining method has limitations in that it lacks the necessary internal structure to enable the distinction between fractal-like structures and fibrillar, β-sheeted aggregates or other amorphous higher-order aggregates. Systematic coarse-graining methods that interpolate between the atomistic simulations and the ultra-coarse-grained description used here might provide an attractive multiscale, renormalizable description of aggregation and phase-ordering processes.

SUPPORTING MATERIAL

Eight figures and one table are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(14)00737-1.

We are grateful to Dr. Scott Crick for helpful discussions and for the use of his data for Fig. 1. We thank the anonymous reviewers for their constructive criticisms and comments.

This work was supported by grant 5R01NS056114 from the National Institutes of Health.

REFERENCES

1. Walker, F. O. 2007. Huntington's disease. *Semin. Neurol.* 27:143–150.
2. Becher, M. W., J. A. Kotzuk, …, C. A. Ross. 1998. Intranuclear neuronal inclusions in Huntington's disease and dentatorubral and pallidoluysian atrophy: correlation between the density of inclusions and IT15 CAG triplet repeat length. *Neurobiol. Dis.* 4:387–397.
3. Landles, C., K. Sathasivam, …, G. P. Bates. 2010. Proteolysis of mutant huntingtin produces an exon 1 fragment that accumulates as an aggregated protein in neuronal nuclei in Huntington disease. *J. Biol. Chem.* 285:8808–8823.
4. Sathasivam, K., A. Neueder, …, G. P. Bates. 2013. Aberrant splicing of HTT generates the pathogenic exon 1 protein in Huntington disease. *Proc. Natl. Acad. Sci. USA.* 110:2366–2370.
5. Chen, S., V. Berthelier, …, R. Wetzel. 2001. Polyglutamine aggregation behavior in vitro supports a recruitment mechanism of cytotoxicity. *J. Mol. Biol.* 311:173–182.
6. Crick, S. L., M. Jayaraman, …, R. V. Pappu. 2006. Fluorescence correlation spectroscopy shows that monomeric polyglutamine molecules form collapsed structures in aqueous solutions. *Proc. Natl. Acad. Sci. USA.* 103:16764–16769.
7. Vitalis, A., X. Wang, and R. V. Pappu. 2008. Atomistic simulations of the effects of polyglutamine chain length and solvent quality on conformational equilibria and spontaneous homodimerization. *J. Mol. Biol.* 384:279–297.
8. Vitalis, A., N. Lyle, and R. V. Pappu. 2009. Thermodynamics of β-sheet formation in polyglutamine. *Biophys. J.* 97:303–311.
9. Walters, R. H., and R. M. Murphy. 2009. Examining polyglutamine peptide length: a connection between collapsed conformations and increased aggregation. *J. Mol. Biol.* 393:978–992.
10. Thakur, A. K., and R. Wetzel. 2002. Mutational analysis of the structural organization of polyglutamine aggregates. *Proc. Natl. Acad. Sci. USA.* 99:17014–17019.
11. Caron, N. S., C. R. Desmond, …, R. Truant. 2013. Polyglutamine domain flexibility mediates the proximity between flanking sequences in huntingtin. *Proc. Natl. Acad. Sci. USA.* 110:14610–14615.
12. Crick, S. L., K. M. Ruff, …, R. V. Pappu. 2013. Unmasking the roles of N- and C-terminal flanking sequences from exon 1 of huntingtin as modulators of polyglutamine aggregation. *Proc. Natl. Acad. Sci. USA.* 110:20075–20080.
13. Chen, S., and R. Wetzel. 2001. Solubilization and disaggregation of polyglutamine peptides. *Protein Sci.* 10:887–891.
14. Chen, S., F. A. Ferrone, and R. Wetzel. 2002. Huntington's disease age-of-onset linked to polyglutamine aggregation nucleation. *Proc. Natl. Acad. Sci. USA.* 99:11884–11889.
15. Bhattacharyya, A. M., A. K. Thakur, and R. Wetzel. 2005. polyglutamine aggregation nucleation: thermodynamics of a highly unfavorable protein folding reaction. *Proc. Natl. Acad. Sci. USA.* 102:15400–15405.
16. Bhattacharyya, A., A. K. Thakur, …, R. Wetzel. 2006. Oligoproline effects on polyglutamine conformation and aggregation. *J. Mol. Biol.* 355:524–535.
17. Wetzel, R. 2006. Kinetics and thermodynamics of amyloid fibril assembly. *Acc. Chem. Res.* 39:671–679.
18. Thakur, A. K., M. Jayaraman, …, R. Wetzel. 2009. Polyglutamine disruption of the huntingtin exon 1 N terminus triggers a complex aggregation mechanism. *Nat. Struct. Mol. Biol.* 16:380–389.
19. Jayaraman, M., A. K. Thakur, …, R. Wetzel. 2011. Assays for studying nucleated aggregation of polyglutamine proteins. *Methods.* 53:246–254.
20. Kar, K., M. Jayaraman, …, R. Wetzel. 2011. Critical nucleus size for disease-related polyglutamine aggregation is repeat-length dependent. *Nat. Struct. Mol. Biol.* 18:328–336.
21. Jayaraman, M., R. Kodali, …, R. Wetzel. 2012. Slow amyloid nucleation via α-helix-rich oligomeric intermediates in short polyglutamine-containing huntingtin fragments. *J. Mol. Biol.* 415:881–899.
22. Jayaraman, M., R. Mishra, …, R. Wetzel. 2012. Kinetically competing huntingtin aggregation pathways control amyloid polymorphism and properties. *Biochemistry.* 51:2706–2716.
23. Wetzel, R. 2012. Physical chemistry of polyglutamine: intriguing tales of a monotonous sequence. *J. Mol. Biol.* 421:466–490.
24. Walters, R. H., K. H. Jacobson, …, R. M. Murphy. 2012. Elongation kinetics of polyglutamine peptide fibrils: a quartz crystal microbalance with dissipation study. *J. Mol. Biol.* 421:329–347.
25. Tam, S., C. Spiess, …, J. Frydman. 2009. The chaperonin TRiC blocks a huntingtin sequence element that promotes the conformational switch to aggregation. *Nat. Struct. Mol. Biol.* 16:1279–1285.
26. Ignatova, Z., A. K. Thakur, …, L. M. Gierasch. 2007. In-cell aggregation of a polyglutamine-containing chimera is a multistep process initiated by the flanking sequence. *J. Biol. Chem.* 282:36736–36743.
27. Caron, N. S., C. L. Hung, …, R. Truant. 2014. Live cell imaging and biophotonic methods reveal two types of mutant huntingtin inclusions. *Hum. Mol. Genet.* 23:2324–2338.
28. Garai, K., and C. Frieden. 2013. Quantitative analysis of the time course of Aβ oligomerization and subsequent growth steps using tetramethylrhodamine-labeled Aβ. *Proc. Natl. Acad. Sci. USA.* 110:3321–3326.
29. Williamson, T. E., A. Vitalis, …, R. V. Pappu. 2010. Modulation of polyglutamine conformations and dimer formation by the N-terminus of huntingtin. *J. Mol. Biol.* 396:1295–1309.
30. Vitalis, A., and R. V. Pappu. 2009. ABSINTH: a new continuum solvation model for simulations of polypeptides in aqueous solutions. *J. Comput. Chem.* 30:673–699.
31. Radhakrishnan, A., A. Vitalis, …, R. V. Pappu. 2012. Improved atomistic Monte Carlo simulations demonstrate that poly-L-proline adopts heterogeneous ensembles of conformations of semi-rigid segments interrupted by kinks. *J. Phys. Chem. B.* 116:6862–6871.
32. Vitalis, A., and R. V. Pappu. 2009. Methods for Monte Carlo simulations of biomacromolecules. *Annu. Rep. Comput. Chem.* 5:49–76.
33. Mao, A. H., S. L. Crick, …, R. V. Pappu. 2010. Net charge per residue modulates conformational ensembles of intrinsically disordered proteins. *Proc. Natl. Acad. Sci. USA.* 107:8183–8188.

34. Vitalis, A., and A. Caflisch. 2010. Micelle-like architecture of the monomer ensemble of Alzheimer's amyloid-β peptide in aqueous solution and its implications for Aβ aggregation. *J. Mol. Biol.* 403:148–165.

35. Wuttke, R., H. Hofmann, …, B. Schuler. 2014. Temperature-dependent solvation modulates the dimensions of disordered proteins. *Proc. Natl. Acad. Sci. USA.* 111:5213–5218.

36. Das, R. K., S. L. Crick, and R. V. Pappu. 2012. N-terminal segments modulate the α-helical propensities of the intrinsically disordered basic regions of bZIP proteins. *J. Mol. Biol.* 416:287–299.

37. Das, R. K., and R. V. Pappu. 2013. Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. *Proc. Natl. Acad. Sci. USA.* 110:13392–13397.

38. Lyle, N., R. K. Das, and R. V. Pappu. 2013. A quantitative measure for protein conformational heterogeneity. *J. Chem. Phys.* 139:121901–121912.

39. Mittal, A., N. Lyle, …, R. V. Pappu. 2014. Hamiltonian Switch Metropolis Monte Carlo simulations for improved conformational sampling of intrinsically disordered regions tethered to ordered domains of proteins. *J. Chem. Theory Comput.* http://dx.doi.org/10.1021/ct5002297.

40. Meng, W., N. Lyle, …, R. V. Pappu. 2013. Experiments and simulations show how long-range contacts can form in expanded unfolded proteins with negligible secondary structure. *Proc. Natl. Acad. Sci. USA.* 110:2123–2128.

41. Halfmann, R., S. Alberti, …, S. Lindquist. 2011. Opposing effects of glutamine and asparagine govern prion formation by intrinsically disordered proteins. *Mol. Cell.* 43:72–84.

42. England, J. L., and D. Kaganovich. 2011. Polyglutamine shows a urea-like affinity for unfolded cytosolic protein. *FEBS Lett.* 585:381–384.

43. Sacanna, S., D. J. Pine, and G. R. Yi. 2013. Engineering shape: the novel geometries of colloidal self-assembly. *Soft Matter.* 9:8096–8106.

44. Sciortino, F., and E. Zaccarelli. 2011. Reversible gels of patchy particles. *Curr. Opin. Solid State Mater. Sci.* 15:246–253.

45. Semenov, A. N. 1990. Dynamic correlation-function of polymer density-fluctuations in concentrated solutions. *Physica A.* 166:263–287.

46. Raos, G., and G. Allegra. 1996. A cluster of chains can be smaller than a single chain: new interpretation of kinetics of collapse experiments. *Macromolecules.* 29:8565–8567.

47. Raos, G., and G. Allegra. 1997. Macromolecular clusters in poor-solvent polymer solutions. *J. Chem. Phys.* 107:6479–6490.

48. Pappu, R. V., X. Wang, …, S. L. Crick. 2008. A polymer physics perspective on driving forces and mechanisms for protein aggregation. *Arch. Biochem. Biophys.* 469:132–141.

49. Wang, Y., and G. A. Voth. 2010. Molecular dynamics simulations of polyglutamine aggregation using solvent-free multiscale coarse-grained models. *J. Phys. Chem. B.* 114:8735–8743.

50. Marchut, A. J., and C. K. Hall. 2006. Side-chain interactions determine amyloid formation by model polyglutamine peptides in molecular dynamics simulations. *Biophys. J.* 90:4574–4584.

51. Marchut, A. J., and C. K. Hall. 2007. Effects of chain length on the aggregation of model polyglutamine peptides: molecular dynamics simulations. *Proteins.* 66:96–109.

52. Dama, J. F., A. V. Sinitskiy, …, G. A. Voth. 2013. The theory of ultra-coarse-graining. 1. General principles. *J. Chem. Theory Comput.* 9:2466–2480.

53. Witten, T. A., and L. M. Sander. 1981. Diffusion-limited aggregation, a kinetic critical phenomenon. *Phys. Rev. Lett.* 47:1400–1403.

54. Meakin, P., and T. A. Witten. 1983. Growing interface in diffusion-limited aggregation. *Phys. Rev. A.* 28:2985–2989.

55. Witten, T. A., and L. M. Sander. 1983. Diffusion-limited aggregation. *Phys. Rev. B.* 27:5686–5697.

56. Sander, L. M. 2000. Diffusion-limited aggregation: a kinetic critical phenomenon? *Contemp. Phys.* 41:203–218.

57. Lapidus, L. J. 2013. Understanding protein aggregation from the view of monomer dynamics. *Mol. Biosyst.* 9:29–35.

58. Meakin, P. 1984. Diffusion-limited aggregation in 3 dimensions: results from a new cluster cluster aggregation model. *J. Colloid Interface Sci.* 102:491–504.

59. Meakin, P., I. Majid, …, H. E. Stanley. 1984. Topological properties of diffusion limited aggregation and cluster cluster aggregation. *J. Phys. A.* 17:L975–L981.

60. Meakin, P. 1985. Structure of the active zone in diffusion-limited aggregation, cluster-cluster aggregation, and the screened-growth model. *Phys. Rev. A.* 32:453–459.

61. Meakin, P. 1987. Computer simulations of diffusion-limited aggregation processes. *Faraday Discuss. Chem. Soc.* 83:113–124.

62. Honda, K., H. Toyoki, and M. Matsushita. 1986. A theory of fractal dimensionality for generalized diffusion-limited aggregation. *J. Phys. Soc. Jpn.* 55:707–710.

63. Puertas, A. M., A. Fernández-Barbero, and F. J. de Las Nieves. 2005. Structure factor scaling in colloidal charge heteroaggregation. *Eur. Phys. J. E Soft Matter.* 18:335–341.

64. Sandkühler, P., M. Lattuada, …, M. Morbidelli. 2005. Further insights into the universality of colloidal aggregation. *Adv. Colloid Interface Sci.* 113:65–83.

65. Porod, G. 1949. Theory of the diffuse small angle x-ray scattering in colloidal systems. *Z. Naturforsch. A.* 4:401–414.

66. Lin, M. Y., H. M. Lindsay, …, P. Meakin. 1989. Universality in colloid aggregation. *Nature.* 339:360–362.

67. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: visual molecular dynamics. *J. Mol. Graph.* 14:33–38, 27–28.

68. Lee, C. C., R. H. Walters, and R. M. Murphy. 2007. Reconsidering the mechanism of polyglutamine peptide aggregation. *Biochemistry.* 46:12810–12820.

69. Smith, M. I., J. S. Sharp, and C. J. Roberts. 2008. Insulin fibril nucleation: the role of prefibrillar aggregates. *Biophys. J.* 95:3400–3406.

70. Weiss, 4th, W. F., T. K. Hodgdon, …, C. J. Roberts. 2007. Nonnative protein polymers: structure, morphology, and relation to nucleation and growth. *Biophys. J.* 93:4392–4403.

71. Beam, M., M. C. Silva, and R. I. Morimoto. 2012. Dynamic imaging by fluorescence correlation spectroscopy identifies diverse populations of polyglutamine oligomers formed in vivo. *J. Biol. Chem.* 287:26136–26145.

72. Arrasate, M., and S. Finkbeiner. 2012. Protein aggregates in Huntington's disease. *Exp. Neurol.* 238:1–11.

A Coarse-Grained Model for Polyglutamine Aggregation Modulated by Amphipathic Flanking Sequences

Kiersten M. Ruff,[1,2] Siddique J. Khan,[1] and Rohit V. Pappu[1,*]

[1]Department of Biomedical Engineering and Center for Biological Systems Engineering, and [2]Division of Biology and Biomedical Sciences, Computational and Systems Biology Program, Washington University in St. Louis, St. Louis, Missouri

SUPPORTING MATERIAL

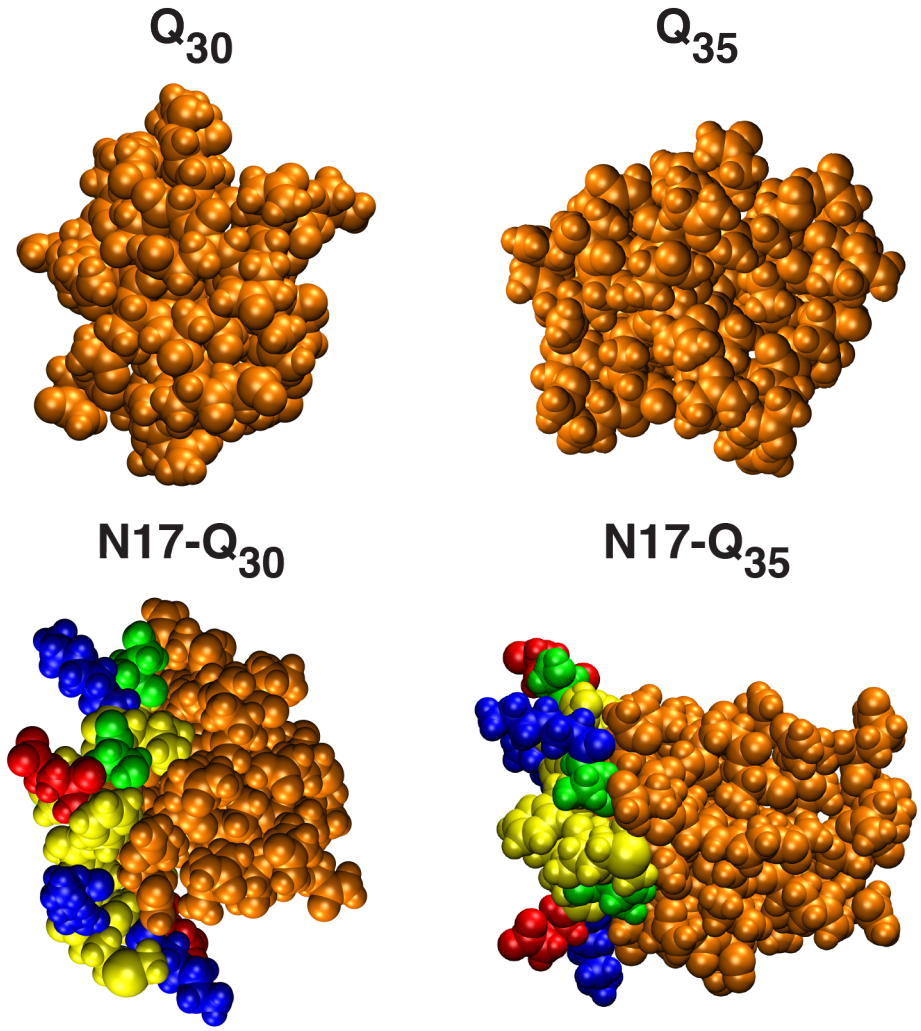

Fig. S1: Representative snapshots for Q_{30} and Q_{35} (top row) and N17-Q_{30}, N17-Q_{35} (bottom row) drawn from the simulated conformational ensembles of Williamson et al. (1). The color coding in the space filling models is as follows: orange spheres are for glutamine residues, yellow for hydrophobic, green for polar, blue for positively charged residues, and red for negatively charged residues. The drawings were made using the Visual Molecular Dynamics package (2).

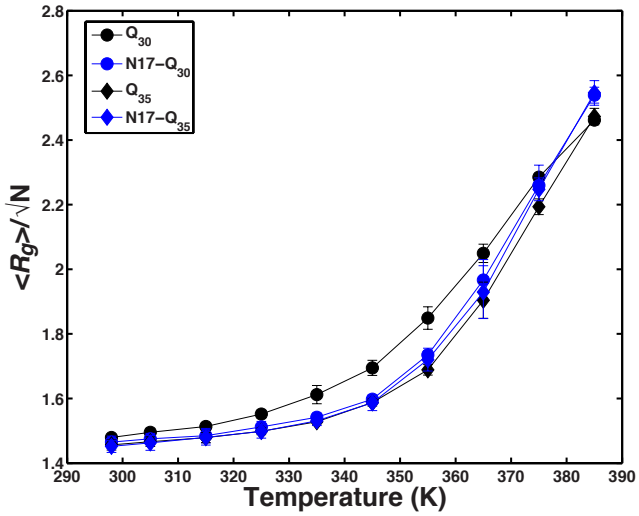

Fig. S2: Plot of the temperature dependence of the normalized average radii of gyration for Q_n and N17-Q_n (n=30, 35). In the temperature range of 298 K – 315 K, the normalized R_g values are similar for all four peptides implying that the densities of the corresponding globules are, on average, similar to each other.

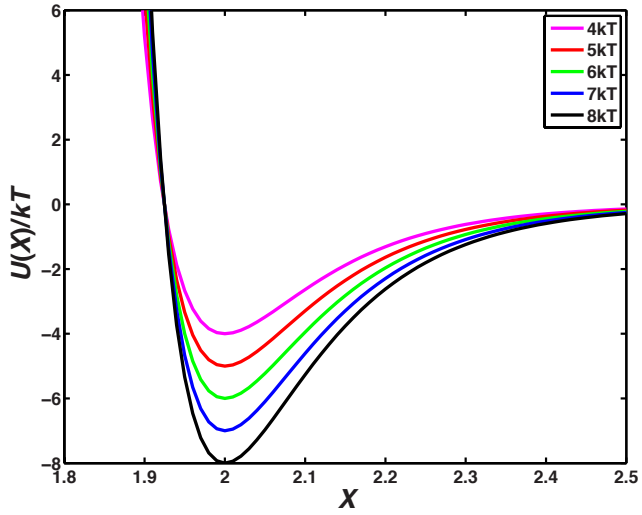

Fig. S3: Schematic of the calibration potentials. The figure plots the single well potentials based on Eq. (2) of the main text for different well depths. These were used in simulations of spherical particles to calibrate the overall energy scale that yields reliable aggregation and the steady state coexistence of small spherical aggregates with large, aspherical / linear aggregates.

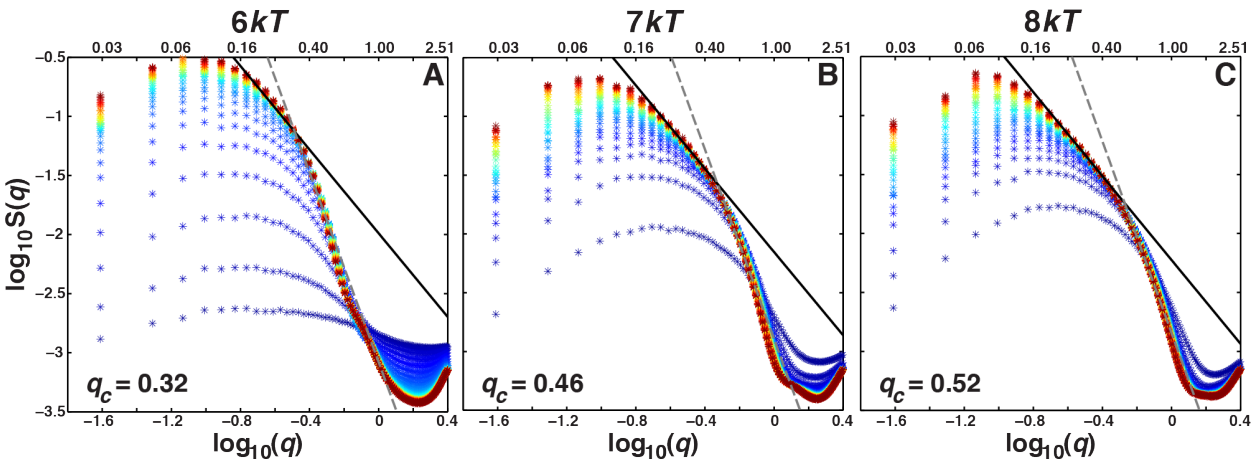

Fig. S4: Temporal evolution of the structure factors for the single well potentials shown in Fig. S3. These plots are shown for the three well depths that produce reliable small- and large-scale aggregation. As the well depth increases the crossover length scale q_c at steady state shifts to higher values. The values of q_c change minimally for well depths that are larger than $8kT$ and therefore, we choose an energy scale of $8kT$ in our simulations because this represents a sufficient level of supersaturation for a volume fraction of $f_v = 0.005$.

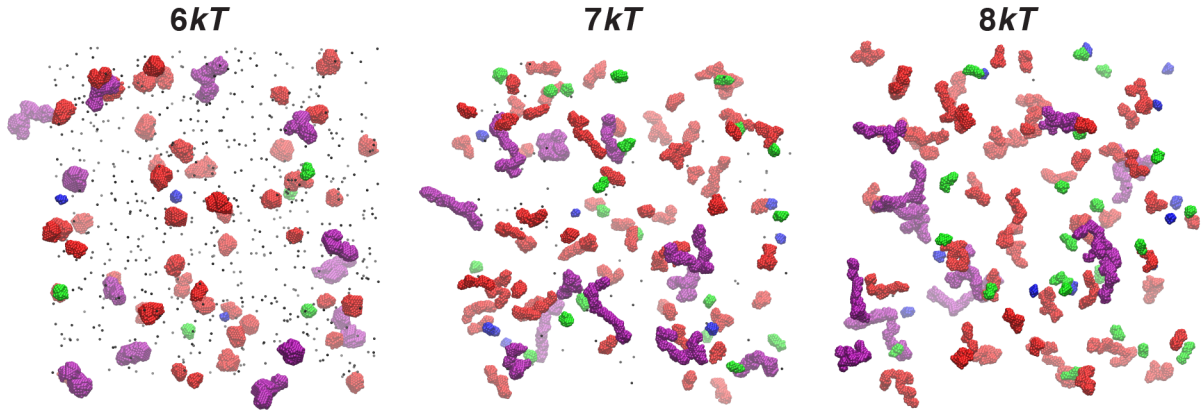

Fig. S5: Snapshots from the steady states for systems interacting via single well potentials with three different overall energy scales viz., $6kT$, $7kT$, and $8kT$, respectively. As the well depth increases, the steady state aggregate sizes become larger and show greater consistency with fibrillar morphologies. Molecules in the snapshots are colored according to the size of the corresponding aggregates *i.e.*, the number of molecules n within the aggregate to which they belong. The scheme is as follows: Black $\equiv n \leq 10$; Blue $\equiv 10 < n \leq 50$; Green $\equiv 50 < n \leq 100$; Red $\equiv 100 < n \leq 400$; Purple $\equiv n > 400$. The snapshots were generated using the Visual Molecular Dynamics package (2).

Section S1: Identifying the linear regimes in log-log plots of steady state structure factors
We identified the high-q and low-q regimes from log-log plots of the steady state structure factors with slopes of ca. –4 and ca. –1.8, respectively. These regions were determined using

sliding windows, each of length ten q-units. Additionally a restriction was added so as to not calculate the slope over very high q regions corresponding to length scales of only a few particles. Specifically, the minimum of the $S(q)$ versus q log-log plot was determined and a cutoff was defined such that the scanning windows ranged from the lowest q value (q=0.0245) to the q value that corresponded to minimum shifted by the cutoff (q_{cutoff}). Then the slopes were calculated over each of the sliding windows and the windows that corresponded to slopes with the minimum difference from –4 and –1.8 were selected for the high-q and low-q regimes, respectively. This procedure yields robust estimates for the slopes and regimes corresponding to the two slopes. It also enables reliable identification of the crossover length scales as evidenced by the low errors in our estimates of q_c.

Table 1: Linear regime details. The table shows the q intervals within which we identify the high- and low-q linear regimes. The values in the parentheses in the fourth and fifth columns denote the slopes found in each regime.

	q_{min}	q_{cutoff}	high-q regime	low-q regime
DPE-1	1.82	1.57	$1.18 \le q \le 1.40$ (–3.73)	$0.22 \le q \le 0.44$ (–1.81)
DPE-2	2.21	1.96	$1.08 \le q \le 1.30$ (–4.00)	$0.44 \le q \le 0.66$ (–1.82)
DNB	1.64	1.40	$0.96 \le q \le 1.18$ (–3.99)	$0.20 \le q \le 0.42$ (–1.82)
DWB	1.74	1.50	$0.98 \le q \le 1.20$ (–3.51)	$0.17 \le q \le 0.39$ (–1.82)

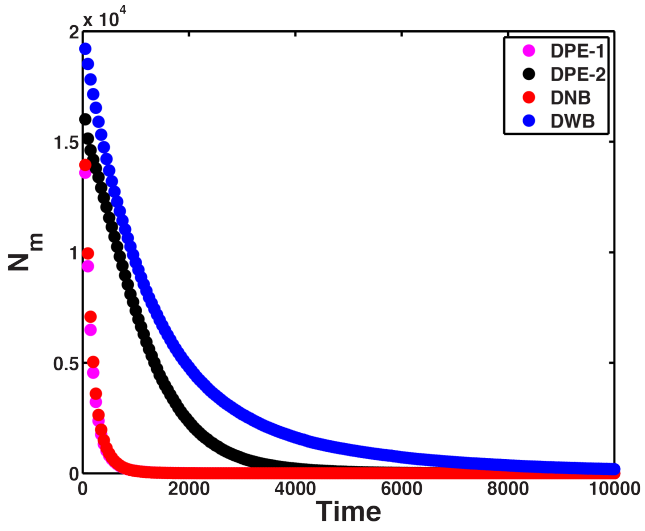

Fig. S6: Rates of loss of monomers into aggregates. These kinetic profiles are shown for the four models, DPE-1, DPE-2, DNB, and DWB. It is worth reiterating that loss of monomers into aggregates is a necessary albeit insufficient condition for achieving loss of fluorescence in the TMR assay. The additional requirement is for the formation of TMR dimers, which requires a high local concentration of TMR molecules and this is achieved in molecular aggregates with aggregation number that is generally larger than at least a 10-mer (3). Accordingly, the rates of monomer loss are less informative of the kinetics of aggregation of smaller species when compared to the development of structure in the high-q regime. Nevertheless, we see that the rates of monomer loss are faster for DPE-2 vis-à-vis DWB.

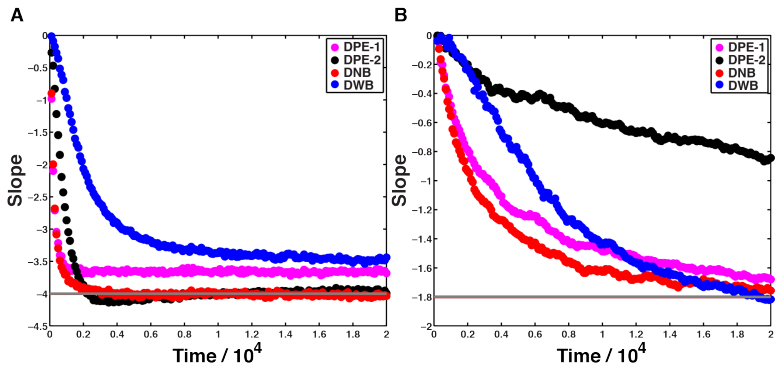

Fig. S7: Equivalent plot to Fig. 6 in the main text. This plot shows comparative assessments of the rate of change of slopes within the following high-q (left) and low-q regimes: $1.08 \leq q \leq 1.30$ and $0.17 \leq q \leq 0.39$, respectively. The former corresponds to the steady state high-q regime for the DPE-2 system and the latter corresponds to the steady state low-q regime for the DWB system.

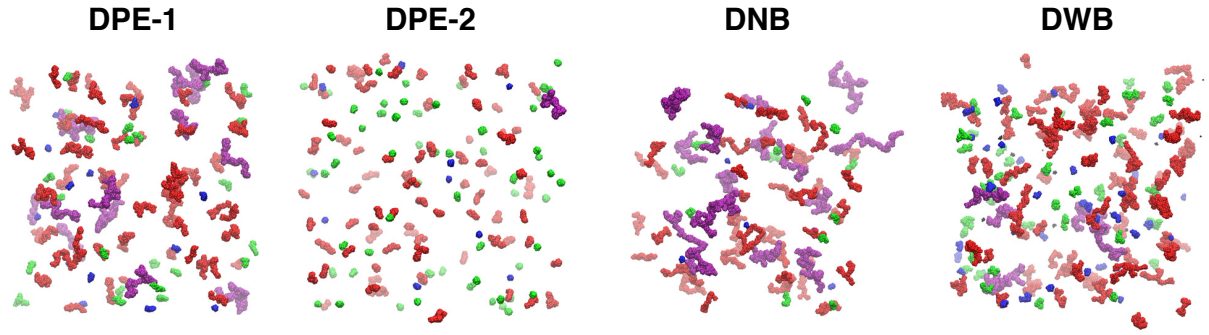

Fig. S8: Representative snapshots of the steady state aggregates that result from the simulations based on DPE-1, DPE-2, DNB, and DWB, respectively. Molecules in the snapshots are colored according to the size of the corresponding aggregates *i.e.*, the number of molecules n within the aggregate to which they belong. The scheme is as follows: Black $\equiv n \leq 10$; Blue $\equiv 10 < n \leq 50$; Green $\equiv 50 < n \leq 100$; Red $\equiv 100 < n \leq 400$; Purple $\equiv n > 400$. The snapshots were generated using the Visual Molecular Dynamics package (2).

References

1. Williamson, T. E., A. Vitalis, S. L. Crick, and R. V. Pappu. 2010. Modulation of Polyglutamine Conformations and Dimer Formation by the N-Terminus of Huntingtin. J. Mol. Biol. 396:1295-1309.
2. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: visual molecular dynamics. J. Mol. Graphics Modell. 14:33-38.
3. Crick, S. L., K. M. Ruff, K. Garai, C. Frieden, and R. V. Pappu. 2013. Unmasking the roles of N- and C-terminal flanking sequences from exon 1 of huntingtin as modulators of polyglutamine aggregation. Proc. Natl. Acad. Sci. USA 110:20075-20080.