Published in final edited form as:
Biophys Chem. 2011 November ; 159(1): 14–23. doi:10.1016/j.bpc.2011.04.006.

Assessing the contribution of heterogeneous distributions of oligomers to aggregation mechanisms for polyglutamine peptides

Andreas Vitalis[1,3] and **Rohit V. Pappu**[1,2,*]

[1]Department of Biomedical Engineering, Washington University in St. Louis, One Brookings Drive, St. Louis, MO 63130

[2]Hope Center for Neurological Disorders, Washington University in St. Louis, One Brookings Drive, St. Louis, MO 63130

Abstract

Polyglutamine aggregation is associated with neurodegeneration in nine different disorders. The effects of polyglutamine length and peptide concentration on the kinetics of aggregation were previously analyzed using a homogeneous nucleation model that assumes the presence of a single bottleneck along the free energy profile $G(n)$, where n denotes the number of polyglutamine molecules. The observation of stable, soluble oligomers as intermediates along aggregation pathways is refractory to the assumptions of homogeneous nucleation. Furthermore, the analysis of *in vitro* kinetic data using a specific variant of homogeneous nucleation leads to confounding observations such as fractional and / or negative values for estimates of the critical nucleus size. Here, we show that the homogeneous nucleation model is inherently robust and is unlikely to yield fractional values if the underlying process is strictly homogeneous with a free energy profile $G(n)$ that displays a sharp maximum at $n=n^*$, where n^* corresponds to the critical nucleus. Conversely, a model that includes oligomers of different size and different potentials for supporting turnover into fibrils yields estimates of fractional and / or negative nucleus sizes when the kinetic data are analyzed using the assumption of a homogeneous process. This model provides a route to reconcile independent observations of heterogeneous distributions of oligomers and other non-fibrillar aggregates with results obtained from analysis of aggregation kinetics using the assumption of a homogeneous nucleation model. In the new model, the mechanisms of fibril assembly are governed by the relative stabilities of two types of oligomers *viz.*, fibril-competent and fibril-incompetent oligomers, the size of the smallest fibril competent oligomer, and rates for conformational conversion within different oligomers.

1. Introduction

Nine different neurodegenerative disorders including Huntington's disease [1] are associated with polyglutamine expansions [2]. In these diseases, the age-of-onset and severity at onset are inversely correlated with polyglutamine expansion lengths [3]. Expanded polyglutamine tracts have all the hallmarks of being intrinsically disordered regions [4-9] although this

*Corresponding author. Tel: +1 314-935-7958, pappu@wustl.edu.
[3]*Current address*: Department of Biochemistry, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland

does not translate into their classification as canonical random coils [10]. Polyglutamine regions destabilize their host protein thereby increasing its susceptibility to proteolysis [11-21]. Products of proteolysis are rich in unprocessed polyglutamine regions [22, 23], which are prone to aggregation [6, 24-30] as well as recruitment and sequestration [6, 31, 32] of polyglutamine-rich regions from other proteins, specifically transcription factors [33-46]. Together, aggregation as well as recruitment and sequestration form a gain-of-function / loss-of-function route to toxicity. In the gain-of-function paradigm, soluble, oligomeric intermediates are currently implicated, albeit in unresolved ways, with consequences for trafficking [47], interactions with membranes [48, 49], and protein quality control [50-57]. The loss-of-function paradigm implicates unwarranted interactions with polyglutamine expansions leading to the loss of essential proteins. The proteins involved in polyglutamine expansion disorders are different from each other in terms of their sequences and functions. Their only shared attribute is the association with disease on account of mutant polyglutamine expansions. As a result, much effort has been focused on understanding the conformational properties and aggregation mechanisms of polyglutamine-rich sequences.

Constructs for biophysical investigation

Studies based on homopolymeric or quasi-homopolymeric polyglutamine constructs are directly relevant to the situation *in vivo* because mechanisms formulated to explain the behavior of synthetic peptides serve as references for interpreting the effects of naturally occurring flanking sequences and heterotypic interactions in the cellular milieu. Reports from molecular simulations [8, 9, 58-63] have focused on analyzing the conformational properties and oligomerization [64, 65] of homopolymeric polyglutamine. In biophysical studies, the poor solubility of homopolymeric polyglutamine [25] necessitates the use of flanking charged residues including one [10] or two [6, 26-28, 66-70] pairs of flanking lysines or pairs of oppositely charged flanking residues [71]. The implicit assumption is that these charges do not alter the intrinsic conformational preferences and intermolecular associations. Recent experimental investigations [69] and computational studies [72] call this assumption into question. At this juncture, the effects of naturally occurring flanking sequences remain unresolved because of conflicting interpretations of *in vitro* data [67, 73] and these interpretations being called into question by *in silico* results [72].

The need for a thermodynamic framework for aggregation

The goal of understanding the “gatekeeping” [74, 75] or other modulating effects of flanking sequences on polyglutamine aggregation is reminiscent of thermodynamic linkage models [76] for analyzing the effects of ligand on the self-assembly of macromolecules into large aggregates. One can develop a linkage analysis for the effect of ligands in *cis*, namely flanking sequences on the polyglutamine-length dependent driving forces for aggregation [77]. This is likely to be feasible by drawing from the rich literature of the physics of phase transitions – a topic that has been reviewed recently [78]. Specifically, phase equilibria of polymer solutions modeled as quasi-binary mixtures provide thermodynamic constraints on aggregation mechanisms. At this juncture, systematic thermodynamic characterization of driving forces for polyglutamine aggregation remains unavailable. Instead, Scherzinger et al. [29, 30] used *in vitro* aggregation experiments to show that the overall rate of aggregation in polyglutamine-rich systems increases with polyglutamine length and peptide concentration. One of the necessary (but insufficient) hallmarks of a nucleation-dependent mechanism is the presence of a lag phase that can be eliminated by adding pre-formed aggregates to the reaction mixture [79, 80]. Scherzinger et al. showed that lag times could be reduced / eliminated with the addition of pre-formed aggregates. Subsequently, Wetzel and coworkers established the necessary set of protocols to disaggregate synthetic polyglutamine peptides

of the form K_2-Q_N-K_2 [6, 81, 82]. A key step was ensuring the absence of seeds such that the starting peptide concentration c_t can be reliably confined to the monomer pool.

Homogeneous nucleation

Chen et al. [83] adapted the homogeneous nucleation model of Ferrone [79] to analyze their data for the time course of aggregation of K_2-Q_N-K_2 peptides. In this model, the Gibbs free energy of a species with n molecules can be written as follows: $G(n+1) - G(n) > 0$ for $n < n^*$ whereas, $G(n+1) - G(n) < 0$ for $n > n^*$. Here, n^* is the size of the critical nucleus that forms a single, well-defined bottleneck that kinetically partitions the reaction mixture between a pool that is dominantly monomeric (because the monomer is the thermodynamically favored state for pre-nuclear species) and large aggregates that for polyglutamine are assumed to be fibrillar. By assuming that pre-nuclear species are in rapid pre-equilibrium with each other and postulating that the rates of elongation of nuclei and fibrils are equivalent, Chen et al. analyzed the time courses for aggregation of K_2-Q_N-K_2 peptides. Of particular relevance is the expression for $\Delta(t)$, which measures the concentration of monomeric material that is incorporated into the growing fibril. It can be shown that $\Delta(t) \sim t^2$ for the early stages of homogeneous nucleation. If one assumes that the only relevant species are monomers and large aggregates, then the time course for monomer loss can be used as a proxy for $\Delta(t)$ and analyzed at different overall peptide concentrations to extract an estimate for the critical nucleus size n^*. For brevity, we shall use the acronym ETM (for Early Time course Model) when using the variant of homogeneous nucleation that is based on the t^2 dependence of $\Delta(t)$.

In studies of K_2-Q_N-K_2 peptides, Chen et al. [83] used ETM and obtained values of 0.98, 0.68, and 0.59, as estimates of n^* for N=28, 36, and 47, respectively. They rounded these estimates up to the nearest integer to conclude that monomers form the critical nucleus for polyglutamine aggregation. Folding of individual molecules is rate-limiting because it is presumed to be thermodynamically unfavorable by ca. 12-15 kcal/mol, and becomes less unfavorable as polyglutamine length increases [66]. Parameters from the model were used to predict the age-of-onset of Huntington's disease by connecting the accumulation of large fibrillar species to the onset of disease.

Issues with homogeneous nucleation

Bernacki and Murphy [84] have shown that the t^2 dependence for the early time course of $\Delta(t)$ is not restricted to homogeneous nucleation. They also raised concerns about certifying a specific mechanism by following only one side of the reaction *viz.*, monomer loss kinetics because this imposes strong assumptions on the model that is inferred from the analysis. The nature of the proposed rate-limiting unimolecular conformational transition remains speculative, although mutagenesis studies [28] and inferences drawn from analysis of morphologies of high molecular weight aggregates [27] were extrapolated to suggest that the transition involves a conversion to β-sheets. Within smaller species, specifically for monomers, a heterogeneous ensemble of collapsed structures is preferred for homopolymeric polyglutamine irrespective of chain length. Computational studies of homopolymeric constructs show that the preference for heterogeneity at the monomer level leads to significant entropic bottlenecks for forming β-sheet conformations [64]. It was shown that disorder prevails even when the entropic penalties for sampling the requisite dihedral angles have been pre-paid [65]. Conversely, ordered β-sheets emerge as one of the possible outcomes of intermolecular interfaces within aggregates of growing size [65]. This idea is consistent with results [83] demonstrating that the rate of elongation of K_2-Q_N-K_2 peptides was considerably slower than diffusion suggesting that the ordered template that forms upon nucleation needs to be deformed to promote elongation by a so-called “dock-and-lock” mechanism, which requires the mixing / entanglement of chain degrees-of-

freedom at intermolecular interfaces. Several studies report the presence of stable oligomers [57, 85-92] that form early and are either on-pathway to the fibril or act as off-pathway sinks. Additionally, recent attention in the aggregation literature has focused on the effects of fragmentation of larger species [93] and the positive cooperativity between fragmentation and assembly because it increases the concentration of growing ends [94-96].

The presence of heterogeneous distributions of species of different sizes and morphologies can confound the analysis of aggregation kinetics that relies on the strict assumptions of homogeneous nucleation. Of particular interest are two sets of observations that result from the application of ETM to processes that are presumed to follow homogeneous nucleation. These are the observations of fractional values for n^*, which was reproduced in recent analysis of kinetic data for shorter K_2Q_N-K_2 constructs [97]. Secondly, when Thakur et al. [67] applied the ETM analysis to monomer loss kinetics for synthetic peptide mimics of the exon 1 encoded region of huntingtin, they obtained negative values as estimates for n^*. Here, we ask if fractional and / or negative values for n^* point to inherent non-robustness of the ETM analysis or if it might point to contributions from underlying heterogeneities such as oligomers of different sizes and potential for growing into fibrils.

In the analysis that follows, we first introduce the details of the analysis based on ETM. The effects of specific approximations in ETM are quantified by comparing it to the classical homogeneous nucleation model developed by Oosawa and Kasai [98] or OKM that makes fewer approximations. We re-derive the t^2 dependence for $\Delta(t)$ as a limiting form of OKM. We then fix the nucleus size and assess the contributions due to experimental errors on the robustness of the procedure used to estimate n^* from real data. In particular, we show that if the underlying process adheres to homogeneous nucleation with an integer-sized nucleus, then fractional values for n^* are inconsistent with the inherent robustness we demonstrate for the ETM analysis. Following the introduction of a schematic that postulates a mechanism for fibril formation through rapid accumulation of stable oligomers, we proceed to show how the homogeneous nucleation model can be generalized to investigate the effects of heterogeneities. When such a model is analyzed using ETM, we show that it yields estimates for n^* that can be fractional and also negative under certain circumstances. Overall, the analysis presented here might serve two purposes: First, it shows how the ETM analysis itself can serve as a useful diagnostic for testing the assumption of homogeneous nucleation. Secondly, our analysis provides a route for incorporating the effects of heterogeneous distributions of oligomers.

2. The homogeneous nucleation model

To facilitate our analysis, we first describe the main features of homogeneous nucleation. We consider a generic, step-wise polymerization reaction for species A:

$$\begin{aligned} A+A &\underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} A_2;\ \ K_2=\frac{k_2}{k_{-2}}=\frac{[A_2]}{[A]^2} \\ &\vdots \\ A_i+A &\underset{k_{-(i+1)}}{\overset{k_{i+1}}{\rightleftharpoons}} A_{i+1};\ \ K_{i+1}=\frac{k_{i+1}}{k_{-(i+1)}}=\frac{[A_{i+1}]}{[A_i][A]} \end{aligned} \tag{1}$$

In Equation (1), the $K's$ are equilibrium constants, the $k's$ are rate constants, and square brackets denote concentrations of different species or more precisely, their activities. Only the initial (dimerization) step and the generalization for later steps are shown. In the Ferrone approach, the accumulation of species up to the critical nucleus of size n^* are assumed to be in a rapid pre-equilibrium because formation of the nucleus is rate limiting. Accordingly,

$$n^*A \underset{k_{-n*}}{\overset{k_{n*}}{\rightleftharpoons}} A_{n*};\ \ K_{n*}=\frac{k_{n*}}{k_{-n*}}=\frac{[A_{n*}]}{[A]^{n*}} \tag{2}$$

Equation (2) identifies one of the important parameters in the homogeneous nucleation model, *viz.*, K_{n*}, the equilibrium constant that describes the thermodynamic stability of the critical nucleus. The net rate of formation of growing polymer ends is defined using:

$$\frac{dc_p}{dt}=k_+^*c_*[A]-k_-^*c_*=K_{n*}\cdot[A]^{n*}\cdot(k_+^*[A]-k_-^*) \tag{3}$$

In Equation (3), c_* denotes the concentration of nuclei and c_p the concentration of growing ends. The pre-equilibrium in Equation (2) was used to obtain an expression in terms of [*A*]. If we assume rate constants that are independent of aggregate size for species larger than the nucleus, then the rate of incorporation of monomer into the growing polymer is governed by:

$$\frac{d(c_t-[A])}{dt}=\frac{d\Delta}{dt}=c_p\cdot(k_+[A]-k_-) \tag{4}$$

Here, c_t is the total monomer concentration at $t = 0$. During the early stages of the polymerization reaction, one can assume that c_t and [*A*] are similar. Additionally, because post-nucleation processes favor downhill polymerization, the rates for monomers dissociating from either a polymer or from a species larger than *n** are assumed to be equivalent and negligibly small. These assumptions are summarized in Equation (5).

$$k_-^*=k_-\approx 0\,;\,k_+^*=k_+\,;\,[A]\approx c_t=\text{const.} \tag{5}$$

We can combine Equations (3) and (4) to obtain:

$$\frac{d^2\Delta}{dt^2}=\frac{d}{dt}(c_pk_+c_t)=k_+c_t\frac{dc_p}{dt}=K_{n*}k_+^2c_t^{n*+2} \tag{6}$$

Equation (6) can be integrated to estimate the rate of incorporation of monomers into the growing polymer during the initial stages of a polymerization reaction in which a well-defined, homogeneous nucleation event is followed by irreversible, kinetically uniform, downhill addition of monomers. This yields:

$$\Delta(t)=\frac{1}{2}K_{n*}k_+^2c_t^{n*+2}t^2 \tag{7}$$

and

$$\frac{d\ln\left(\frac{d^2\Delta}{dt^2}\right)}{d\ln c_t}=(n^*+2)\cdot K_{n*}k_+^2c_t^{n*+1}\cdot\frac{c_t}{K_{n*}k_+^2c_t^{n*+2}}=n^*+2 \quad (8)$$

For a given value of c_t, Equation (7) suggests the existence of a time interval where $\Delta(t)$ increases linearly with t^2. Regression analysis in this region yields an estimate for the slope $s(c_t)=\frac{K_{n*}k_+^2c_t^{n*+2}}{2}$ [83]. Performing this regression analysis using measurements for the early time course to quantify $\Delta(t)$ for different values of c_t allows the determination of the nucleus size, n^* by estimating the slope of the plot of $\ln[s(c_t)]$ against $\ln(c_t)$, which should be n^*+2. Using mass conservation and the assumption of a homogeneous process, *i.e.*, the absence of intermediates such as oligomers, Chen et al. [83] proposed that the rate of change of $\Delta(t)$ can be inferred directly from the rate of loss of monomers.

2.1 General expression for Δ(*t*) given the assumption of homogeneous nucleation

The expressions developed thus far, particularly Equations (6) – (8) are limited to the analysis of early time courses for the rate at which monomers are incorporated into growing polymers. Oosawa and Kasai [98] showed that for $k_+[A] \gg k_-$ *i.e.*, when nucleus formation and elongation are irreversible and the initial concentration of nuclei is zero, the rate at which monomers are incorporated into polymers can be written in terms of k_{n*}^{eff}, the effective rate of formation of nuclei as:

$$\begin{gathered}\ln\left(\frac{1+x}{1-x}\right)=t\sqrt{2n^*k_+k_{n*}^{\text{eff}}c_t^{n*}}\\ x^2=1-\left(\frac{[A]}{c_t}\right)^{n*}=1-\left(1-\frac{\Delta}{c_t}\right)^{n*}\approx\frac{n^*\Delta}{c_t}\text{ for small values of }\left(\frac{\Delta}{c_t}\right)\end{gathered} \quad (9)$$

For $x^2 < 1$, which corresponds to the early time limit, one can expand the logarithmic term in Equation (9) as a power series in x. Truncation of this series at the first-order term and subsequent algebraic manipulation yields an expression for $\Delta(t)$ shown below:

$$\Delta(t)=\frac{1}{2}k_{n*}^{\text{eff}}k_+c_t^{n*+1}t^2 \quad (10)$$

Comparing the expressions in Equations (7) and (10), we find that the two models can be matched if we set $k_{n*}^{\text{eff}}=K_{n*}c_tk_+$, which is consistent with the definition of k_{n*}^{eff}. The limiting form of the Oosawa-Kasai result yields an expression that shows the t^2 dependence for $\Delta(t)$ during the early stages of the reaction.

2.2 Comparative analysis of the two models for homogeneous nucleation

Equation (9) allows us to assess the validity of the approximations made in arriving at Equation (7) and the conclusions drawn from analysis of measurements that leads to estimates for n^* from plots of $\ln[s(c_t)]$ against $\ln(c_t)$. Using Equation (10) and the definition of k_{n*}^{eff}, we generated artificial data over a full time course (defined as the time taken to reach equilibrium) and this allows us to test the robustness of the analysis outlined in Equation (7). It is worth emphasizing that the Oosawa-Kasai model is also built on the assumption of homogeneous nucleation. The motivation for our comparative analysis is as follows: If the estimates for n^* are overly sensitive to the approximations that lead up to Equations (7) and

(8), then experimental errors may be large and a re-analysis using a more complete model of aggregation is in order. Conversely, if we find that the estimate of n^* is fairly robust to errors and approximations, then we can make inferences regarding the validity of the assumption of a homogeneous nucleation process. In all applications of the homogeneous nucleation model to study aggregation of polyglutamine containing systems, the estimates for n^* end up being fractional values and they also show considerable variability with polyglutamine length and flanking sequences. At issue is whether such behavior can be better explained by the presence of a heterogeneous distribution of oligomers with different intrinsic abilities for turning over into fibrils. Contributions from such species are ignored in the assumption of a homogeneous process.

We generated data for the time course of $\Delta(t)$ using parameters reported in the literature [83]. Specifically, we set $k_+ = 10^4\ M^{-1}s^{-1}$, $K^{n^*} = 2.7 \times 10^{-9}$, and n^*=1. Panel A of Figure 1 shows plots of $\Delta(t)$ versus t for values of c_t ranging from 10 μM – 100 μM. As expected, the expression in Equation (7) applies only over a limited time scale. However, panel B of Figure 1 shows that the approximation for $\Delta(t)$ is reasonable during the early stages of the reaction. Given this agreement, we tested the quality of the proposed linear relationship between $\Delta(t)$ and t^2 by comparing the simulated data from the Oosawa-Kasai model that plots $\Delta(t)$ against t^2 to linear fits to these data (Figure 2). The quality of the agreement between the simulated data and linear fits suggests that the analysis proposed in Equation (8), *i.e.*, calculating the slope of the double logarithmic plot of $\ln[s(c_t)]$ against $\ln(c_t)$ should provide a reliable estimate of the nucleus size. That this is the case is shown in Figure 3, which shows that the ETM reproduces the value of (n^*+2)= 3, which was used to generate the simulation data for $\Delta(t)$ at different c_t values.

We further assessed the robustness of the analysis by quantifying the effects of noise on the estimate of n^*. Of particular relevance is the delineation of the time interval that corresponds to the initial stages of the reaction, a time interval we designate as t_c. For different values of c_t we varied t_c from 10% to 60% of the total time course for $\Delta(t)$ and found that the analysis of the double logarithmic plot yields a normal distribution of values for the estimate of n^* +2, with a mean value of 3.0 and standard deviation of 0.06. Additionally, if arbitrariness in defining the initial phase is removed, *i.e.*, a similar criterion is applied for choosing the initial phase for all values of c_t, then the slope of the double logarithmic plot is exactly 3.0. These two results suggest that for a truly homogeneous process, errors involved in choosing a value for t_c lead to negligible variations in the estimate for n^*+2.

2.3 Implications of comparative analysis for polyglutamine aggregation mechanisms

The preceding discussion highlights the robustness of the approach shown in Equation (8) for a strictly homogeneous process with a well-defined nucleus size. Our assessment of the robustness of the ETM analysis focused on quantifying the effects of noise / errors in defining the early stages of the reaction process. Chen et al. [83] applied the methods described above to analyze kinetic data for the rate of loss of soluble material. Measurements were performed at different values of c_t for peptides of the form K_2-Q_N-K_2 with N=28, 36, and 47, respectively. For these three constructs, their analysis of the double logarithmic plots yielded slopes of 2.98, 2.68, and 2.59, from which they estimated a value of n^*=1 for all constructs as the nearest integer value to the fractional values they obtained. As noted above, for a strictly homogeneous process the estimates for n^* should be robust and reasonably insensitive to errors in defining the interval t_c for the extent of reaction used to extract $\ln[s(c_t)]$.

The fractional values for n^* are at odds with the expected robustness of the analysis for estimating n^*, providing aggregation follows homogeneous nucleation. There are other errors that might arise from difficulties in accurately measuring the fraction of unaggregated

material that remains in the supernatant. Such errors are unlikely to be systematic during the early stages of the reaction and can be overcome through multiple independent measurements. An additional source of error might pertain to the number of c_t values for which the time courses of $\Delta(t)$ are measured. Chen et al. used four different concentrations that span a reasonable range of concentrations.

Bernacki and Murphy [84] demonstrated that kinetic data using monomer loss as the only experimental readout for aggregation may be fit with a variety of different mechanistic models including the homogeneous nucleation. Distinguishing between models was shown to be difficult in the absence of more detailed data, in particular the complimentary readout of fibril numbers and sizes. Similarly, Morris et al. [99] demonstrated that a mechanism termed “the Finke-Watzky model of nucleation followed by autocatalytic surface growth” fits several independent sets of protein aggregation data reasonably well with only two free parameters. Neither approach accounted for the effects of heterogeneities in the (effective) nucleation and / or the elongation steps. Heterogeneous distributions of oligomers can act as defects for homogeneous nucleation specifically if different oligomers vary in their intrinsic potential for turning over into fibrils. In the following sections we will review the evidence from experimental work and atomistic simulations that jointly suggest a role for heterogeneous distributions of oligomers in polyglutamine aggregation.

3. A scheme for incorporating heterogeneous distributions of oligomers into polyglutamine aggregation mechanisms

Figure 4 summarizes an alternative “two-stage” mechanism that integrates findings from atomistic simulations and biophysical experiments [65]. In this proposal, disordered globules associate reversibly to form a distribution of molten oligomers. If $p(n)$ denotes the distribution of oligomer sizes, then the peak and width of this distribution will in all likelihood vary with polyglutamine length and the nature of the sequences that flank the polyglutamine expansion on the N- and C-termini. Inter-peptide interfaces are thermodynamically favored to peptide-solvent interfaces because chains within oligomers are solvated by other chains. Since intra- and intermolecular interactions are thermodynamically equivalent, individual chains within clusters can expand / contract freely and this facilitates the requisite conformational rearrangements needed to converge upon energetically favorable conformations with high β-content in large aggregates. As a result, oligomers of the same size can be distinguished by the conformational preferences of individual chains within the oligomer leading to additional heterogeneities in the reaction mixture.

3.1 Development of a model that explicitly accounts for a heterogeneous distribution of oligomers

We consider a model in which monomers in solution rapidly associate to give rise to a distribution of oligomers. This model is motivated by the schematic shown in Figure 4. For simplicity, we consider oligomer formation through monomer addition alone. At equilibrium:

$$\begin{aligned} A+A &\rightleftharpoons A_2 \; ; \; K_2 = \frac{[A_2]}{[A]^2} \\ A_i+A &\rightleftharpoons A_{i+1} \; ; \; K_{i+1} = \frac{[A_{i+1}]}{[A_i][A]} \end{aligned} \tag{11}$$

In Equation (11), A is the aggregating species, square brackets denote concentrations (more precisely activities), and the K_i are equilibrium constants that quantify the stability of

species i. We can generate equations for the populations of individual species, f_i, by observing conservation of mass:

$$f_i = \frac{[A_i]}{c_t} = \frac{[A]}{c_t} \cdot \prod_{j=2}^{i} K_j [A]$$
$$c_t = \sum_{i=1}^{\infty} i[A_i] \tag{12}$$

Here, c_t is the total concentration of aggregating material and Equation (12) provides an implicit relationship between the concentration of free monomer and the total monomer concentration. We shall consider the case where the infinite sum in Equation (2) is approximated by a finite sum; this is reasonable since we focus on soluble oligomers. We define a maximum oligomer size $i_{\max}$ by setting K_i to be zero for $i > i_{\max}$. Equation (12) then becomes:

$$f_i = \begin{cases} \frac{[A]}{c_t} \cdot \prod_{j=2}^{i} K_j [A] & \text{for } i < i_{\max} \\ 0.0 & \text{for } i \geq i_{\max} \end{cases}$$
$$c_t = \sum_{i=1}^{i_{\max}} i[A_i] \tag{13}$$

Next, we assume a slow and irreversible conversion of large oligomers A_i with $i \geq i_{\min}$ to fibrillar or more accurately to protofibrillar species:

$$\begin{aligned} A_{i_{\min}} &\xrightarrow{k_{i_{\min}}} F_{i_{\min}} \\ A_{i_{\min}+1} &\xrightarrow{k_{i_{\min}+1}} F_{i_{\min}+1} \\ &\vdots \\ A_{i_{\max}} &\xrightarrow{k_{i_{\max}}} F_{i_{\max}} \end{aligned} \tag{14}$$

In Equation (14), F_i denotes a fibrillar species composed of i monomers. We propose that fibrillar species form through a unimolecular process, which requires an internal re-arrangement as sketched in Figure 4. Due to the slowness of the process we can assume that the oligomer distribution re-equilibrates rapidly (pre-equilibrium assumption). We write a rate equation for fibril formation nucleated by unimolecular conformational conversion within any of the oligomers of size $i \geq i_{\min}$ as:

$$\frac{d[F_i]}{dt} = k_i [A_i] = k_i [A] \cdot \prod_{j=2}^{i} K_j [A]$$
$$\frac{dc_p}{dt} = \sum_{i=i_{\min}}^{i_{\max}} k_i [A_i] = \sum_{i=i_{\min}}^{i_{\max}} k_i [A] \cdot \prod_{j=2}^{i} K_j [A] \tag{15}$$

In Equation (15), c_p is the total concentration of growing ends (independent of the size of the particular fibril nucleus) and each k_i denotes the rate of nucleation within an oligomer of size i. Consequently, we can treat elongation via monomer addition as:

$$F_i + A \underset{k_-}{\overset{k_+}{\rightleftharpoons}} F_{i+1}$$
$$\frac{dc_f}{dt} = k_+ c_p[A] - k_- c_p$$
$$\frac{d^2 c_f}{dt^2} = k_+ c_p \frac{d[A]}{dt} + (k_+[A] - k_-) \cdot \sum_{i=i_{\min}}^{i_{\max}} k_i[A] \cdot \prod_{j=2}^{i} K_j[A] \quad (16)$$

Here, k_+ is the elongation rate constant for monomer addition, k_- is the rate constant for monomer loss from a growing fibril, and c_f is the total concentration of monomers incorporated into fibrils.

If we focus on the initial rate of aggregation and treat fibril elongation as being irreversible and homogeneous we may assume, at least for certain values of K_i, that the free monomer concentration [A] is static and determined by its pre-equilibrium value in the absence of any fibrillar aggregates. This concentration a_0 is obtained implicitly via Equation (13). Then, we can integrate Equation (16) similar to Equation (7):

$$c_f(t) = \frac{1}{2} k_+ t^2 a_0 \cdot \sum_{i=i_{\min}}^{i_{\max}} k_i a_0 \cdot \prod_{j=2}^{i} K_j a_0 \quad (17)$$

Through the generalizations introduced leading up to Equation (16) we are in a position to assess the role of oligomers and rate limiting steps within oligomers on the concentration dependence of the initial rate of aggregation. In the proposed model $i_{\max}$ is the size of the largest soluble oligomer and $i_{\min}$ is the size of the smallest oligomer that can support conversion into a fibril / protofibrils.

From Equation (17) we see that the slope of a double logarithmic plot still convolves elongation and nucleation processes. Secondly, we recognize that the effective dependence on total concentration is not guaranteed to yield a well-defined integer as the slope of a double logarithmic plot since it depends on the relationship between a_0 and c_t, which is given by Equation (13). This relationship convolves contributions from the significant population of off-pathway, *i.e.*, oligomers of size $i < i_{\min}$ that are incapable of converting to fibrils and the heterogeneity of the nucleation mechanism itself, which determines the rate of conversion of different oligomers to fibrils. The slope of a double logarithmic plot convolves elongation and nucleation processes, as is the case with the original analysis. Additionally, the slope is now directly proportional to the free monomer concentration and to the fraction of fibril competent species. Because the latter quantity depends on a_o and includes terms that range from 1 to $a_o^{i_{\min}}$, we can expect slopes *i.e.*, n^*+2 values ranging from unity (yielding negative values for n^*) to $i_{\min}+1$. Importantly, Equation (17) recovers the ETM expression *i.e.*, Equation (7) if both $i_{\min}$ and $i_{\max}$ are equal to two. The only difference is that the pre-equilibrium constant K_{n^*} becomes a pre-equilibrium dimerization constant K_2 and that the bimolecular nucleus elongation rate constant k^*_+ becomes a unimolecular nucleation rate constant k_2.

3.2 Analysis of the effects of heterogeneous oligomer distributions using the ETM approach

We demonstrate the influence of oligomers that have different potentials for turning over to fibrils by quantifying their effects on estimates of apparent nucleus sizes, n^*. Simulated data were generated using values for k_i that were sampled from a normal distribution, the parameters of which were held constant throughout (mean and standard deviation are

$10^{-2}s^{-1}$; values are adjusted to zero if drawn as negative). Additionally, we set $i_{max} = 20$ in all of our simulations. Figure 5 shows histograms of apparent nucleus size estimates for a minimum fibril-forming oligomer of size $i_{min} = 10$ and for different values of the mean of the distribution of K_i. The variance for the underlying distribution was fixed at 10^5M^{-1}. Every histogram represents 10^3 simulated datasets. Within each dataset, 450 free monomer concentrations up to the high μM-range were used to obtain species distributions and total concentrations via Equation (13). Linear regression of the logarithm of the constant terms in Equation (17) plotted against $\ln(c_t)$ yielded estimates for n^*+2 according to Equation (8). To mimic experimental conditions, only values for c_t of 1-500 μM were considered for the linear regression. Figure 6 shows a plot that is analogous to the plot in Figure 5. Here, i_{min} is varied and the mean of the K_i is fixed to $4\times10^5M^{-1}$.

Figures 5 and 6 show that applying the methods of Chen et al. to processes with built-in heterogeneities such as the presence of oligomers of varying sizes, stabilities, and potential for transforming to fibrils will yield estimates for the nucleus size that depend strongly on the parameters that give rise to heterogeneities. If most of the aggregating material exists in fibril-incompetent forms (small K_i and / or large i_{min}), the apparent nucleus size is large but smaller than i_{min} and fractional values for n^* are readily observed. If, however, large fractions of the aggregating material are present in soluble oligomers that are large enough to promote fibril formation (large K_i and / or small i_{min}), then the smallest apparent nucleus size is close to zero (corresponding to a slope of 2.0) and can be less than zero. This finding is also consistent with the autocatalytic surface growth mechanism proposed by Morris et al., whose analysis of the initial rate dependence on total concentration would similarly yield a slope of 2.0 [99].

4. Conclusions

Our investigations have focused on understanding the implications of heterogeneous distributions of oligomers for aggregation kinetics that are analyzed using the ETM variant of homogeneous nucleation. Accounting for the presence of oligomers and multiple routes to fibrillar / protofibrillar forms provides a plausible explanation for fractional and / or negative values for n^*. By comparing the approximate solution used by Chen et al. for $\Delta(t)$ to the full solution derived from the classical model of Oosawa and Kasai we were able to delineate the time interval $0 < t \leq t_c$ over which the t^2 dependence of $\Delta(t)$ is reasonable. We then investigated the effects of uncertainty in defining t_c*i.e.*, the extent of uncertainty in delineating the reaction for which the approximation of $[A] \approx c_t$ is valid. This analysis demonstrated that the estimate of n^* is rather robust to the introduction of uncertainties in defining t_c. Consequently, for a truly homogeneous nucleation process, estimates of fractional values for n^* might provide a reasonable basis to question the assumption of a homogeneous process [100].

It is worth emphasizing that the assumption of heterogeneous distributions of oligomers is not the only way to obtain fractional values of n^* using the ETM analysis. In fact, for homogeneous nucleation the specific shape of the free energy profile $G(n)$ can lead to fractional values of n^*. This is because the concentration of super nuclei, *i.e.*, species of size $n > n^*$ is proportional to $\exp[-z^2(n - n^*)^2]$ [100] where z is the so-called Zeldovich factor that quantifies the curvature of the free energy profile $G(n)$ at $n = n^*$. The smaller the value of z, the broader the free energy maximum, and the greater the likelihood that the nucleus size will be a fractional value around n^* because clusters with aggregate numbers in the vicinity of n^* have similar free energy values. Conversely, for higher values of z, the free energy profile becomes sharper and the nucleus size will converge upon n^* with high precision. It is also worth emphasizing that focusing on the curvature of the free energy profile cannot yield a negative estimate for the nucleus size. Consequently, the formalism

presented in this work is different from this “blended homogeneous nucleation model” because we sought to develop a model that rationalizes the observations of oligomers in experiments [68, 69, 90].

We generalized the homogeneous process to be a two-stage process in conformity with the scheme introduced in Figure 4. In this scheme, a distribution of oligomers can form in a thermodynamically downhill fashion. Fibril formation requires a unimolecular conversion of individual chains within one or more of oligomers into a conformation that is competent for templated growth. Oligomers can be classified as being fibril competent or incompetent based on their ability to support the requisite unimolecular conversion of individual chains into conformations that can support templated growth. Incorporation of heterogeneities into the model followed by application of the ETM analysis to the resultant kinetics shows how the methodology leads to an estimate of fractional as well as negative values for n^*. The relative stabilities of different oligomeric species (fibril competent / incompetent species) and the size of the smallest fibril competent oligomer (i_{min}) contribute to determining the estimated values for n^*. The model outlined in this work provides a quantitative framework for the proposal of “nucleated conformational conversion within oligomers” put forth by Serio et al. [101].

Acknowledgments

This work was supported by grant 5RO1NS056114 from the National Institutes of Health. We appreciate the opportunity to present this work as part of the special issue celebrating the 25th anniversary of the Gibbs Conference on Biothermodynamics. As regular participants at this conference, we have benefited immensely from stimulating discussions with colleagues at these meetings. The standards set by pioneers of this meeting continue to inspire and motivate us to apply rigorous statistical thermodynamics concepts to relevant biological problems. We are also grateful to the anonymous reviewer who made us think about alternative models that result in fractional estimates for n^*.

References

1. MacDonald ME, Ambrose CM, Duyao MP, Myers RH, Lin C, Srinidhi L, Barnes G, Taylor SA, James M, Groot N, MacFarlane H, Jenkins B, Anderson MA, Wexler NS, Gusella JF, Bates GP, Baxendale S, Hummerich H, Kirby S. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. Cell. 1993; 72:971–983. [PubMed: 8458085]
2. Cummings CJ, Zoghbi HY. Fourteen and counting: unraveling trinucleotide repeat diseases. Hum Mol Genet. 2000; 9:909–916. [PubMed: 10767314]
3. Walker FO. Huntington's disease. The Lancet. 2007; 369:218–228.
4. Altschuler EL, Hud NV, Mazrimas JA, Rupp B. Random coil conformation for extended polyglutamine stretches in aqueous soluble monomeric peptides. J Pept Res. 1997; 50:73–75. [PubMed: 9273890]
5. Altschuler EL, Hud NV, Mazrimas JA, Rupp B. Structure of polyglutamine. FEBS Lett. 2000; 472:166–167. [PubMed: 10781826]
6. Chen S, Berthelier V, Yang W, Wetzel R. Polyglutamine aggregation behavior in vitro supports a recruitment mechanism of cytotoxicity. J Mol Biol. 2001; 311:173–182. [PubMed: 11469866]
7. Masino L, Kelly G, Leonard K, Trottier Y, Pastore A. Solution structure of polyglutamine tracts in GST-polyglutamine fusion proteins. FEBS Lett. 2002; 513:267–272. [PubMed: 11904162]
8. Wang XL, Vitalis A, Wyczalkowski MA, Pappu RV. Characterizing the conformational ensemble of monomeric polyglutamine. Proteins: Struct Funct Bioinform. 2006; 63:297–311.
9. Vitalis A, Wang X, Pappu RV. Quantitative characterization of intrinsic disorder in polyglutamine: insights from analysis based on polymer theories. Biophys J. 2007; 93:1923–1937. [PubMed: 17526581]

10. Crick SL, Jayaraman M, Frieden C, Wetzel R, Pappu RV. Fluorescence correlation spectroscopy shows that monomeric polyglutamine molecules form collapsed structures in aqueous solutions. Proc Natl Acad Sci USA. 2006; 103:16764–16769. [PubMed: 17075061]
11. Bader R, Seeliger MA, Kelly SE, Ilag LL, Meersman F, Limones A, Luisi BF, Dobson CM, Itzhaki LS. Folding and fibril formation of the cell cycle protein Cks1. J Biol Chem. 2006; 281:18816–18824. [PubMed: 16675442]
12. Berke SJS, Schmied FAF, Brunt ER, Ellerby LM, Paulson HL. Caspase-mediated proteolysis of the polyglutamine disease protein ataxin-3. J Neurochem. 2004; 89:908–918. [PubMed: 15140190]
13. Masino L, Musi V, Menon RP, Fusi P, Kelly G, Frenkiel TA, Trottier Y, Pastore A. Domain architecture of the polyglutamine protein ataxin-3: a globular domain followed by a flexible tail. FEBS Lett. 2003; 549:21–25. [PubMed: 12914917]
14. Perez MK, Paulson HL, Pittman RN. Ataxin-3 with an altered conformation that exposes the polyglutamine domain is associated with the nuclear matrix. Hum Mol Genet. 1999; 8:2377–2385. [PubMed: 10556285]
15. Ratovitski T, Gucek M, Jiang H, Chighladze E, Waldron E, D'Ambola J, Hou Z, Liang Y, Poirer MA, Hirschhorn RR, Graham R, Hayden MR, Cole RN, Ross CA. Mutant Huntingtin N-terminal fragments of specific size mediate aggregation and toxicity in neuronal cells. J Biol Chem. 2009; 284:10855–10867. [PubMed: 19204007]
16. Ratovitski T, Nakamura M, D'Ambola J, Chighladze E, Liang Y, Wang W, Graham R, Hayden MR, Borchelt DR, Hirschhorn RR, Ross CA. N-terminal proteolysis of full-length mutant huntingtin in an inducible PC12 cell model of Huntington's disease. Cell Cycle. 2007; 6:2970–2981. [PubMed: 18156806]
17. Sun B, Fan W, Balciunas A, Cooper JK, Bitan G, Steavenson S, Denis PE, Young Y, Adler B, Daugherty L, Manoukian R, Elliott G, Shen WY, Talvenheimo J, Teplow DB, Haniu M, Haldankar R, Wypych J, Ross CA, Citron M, Richards WG. Polyglutamine repeat length-dependent proteolysis of huntingtin. Neurobiol Dis. 2002; 11:111–122. [PubMed: 12460551]
18. Walsh R, Storey E, Stefani D, Kelly L, Turnbull V. The roles of proteolysis and nuclear localisation in the toxicity of the polyglutamine diseases. A review. Neurotox Res. 2005; 7:43–57. [PubMed: 15639797]
19. Ellerby LM, Andrusiak RL, Wellington CL, Hackam AS, Propp SS, Wood JD, Sharp AH, Margolis RL, Ross CA, Salvesen GS, Hayden MR, Bredesen DE. Cleavage of atrophin-1 at caspase site aspartic acid 109 modulates cytotoxicity. J Biol Chem. 1999; 274:8730–8736. [PubMed: 10085113]
20. Haacke A, Broadley SA, Boteva R, Tzvetkov N, Hartl FU, Breuer P. Proteolytic cleavage of polyglutamine-expanded ataxin-3 is critical for aggregation and sequestration of non-expanded ataxin-3. Hum Mol Genet. 2006; 15:555–568. [PubMed: 16407371]
21. Haacke A, Hartl FU, Breuer P. Calpain inhibition is sufficient to suppress aggregation of polyglutamine-expanded ataxin-3. J Biol Chem. 2007; 282:18851–18856. [PubMed: 17488727]
22. Holmberg CI, Staniszewski KE, Mensah KN, Matouschek A, Morimoto RI. Inefficient degradation of truncated polyglutamine proteins by the proteasome. EMBO J. 2004; 23:4307–4318. [PubMed: 15470501]
23. Venkatraman P, Wetzel R, Tanaka M, Nukina N, Goldberg AL. Eukaryotic proteasomes cannot digest polyglutamine sequences and release them during degradation of polyglutamine-containing proteins. Mol Cell. 2004; 14:95–104. [PubMed: 15068806]
24. Bennett MJ, Huey-Tubman KE, Herr AB, West AP, Ross SA, Bjorkman PJ. A linear lattice model for polyglutamine in CAG-expansion diseases. Proc Natl Acad Sci USA. 2002; 99:11634–11639. [PubMed: 12193654]
25. Krull LH, Wall JS. Synthetic Polypeptides Containing Side-Chain Amide Groups. Water-Soluble Polymers. Biochemistry (Mosc). 1966; 5:1521–1527.
26. Chen S, Ferrone FA, Wetzel R. Huntington's disease age-of-onset linked to polyglutamine aggregation nucleation. Proc Natl Acad Sci U S A. 2002; 99:11884–11889. [PubMed: 12186976]
27. Chen SM, Berthelier V, Hamilton JB, O'Nuallain B, Wetzel R. Amyloid-like features of polyglutamine aggregates and their assembly kinetics. Biochemistry (Mosc). 2002; 41:7391–7399.

28. Thakur AK, Wetzel R. Mutational analysis of the structural organization of polyglutamine aggregates. Proc Natl Acad Sci USA. 2002; 99:17014–17019. [PubMed: 12444250]
29. Scherzinger E, Lurz R, Turmaine M, Mangiarini L, Hollenbach B, Hasenbank R, Bates GP, Davies SW, Lehrach H, Wanker EE. Huntingtin-encoded polyglutamine expansions form amyloid-like protein aggregates in vitro and in vivo. Cell. 1997; 90:549–558. [PubMed: 9267034]
30. Scherzinger E, Sittler A, Schweiger K, Heiser V, Lurz R, Hasenbank R, Bates GP, Lehrach H, Wanker EE. Self-assembly of polyglutamine-containing huntingtin fragments into amyloid-like fibrils: Implications for Huntington's disease pathology. Proc Natl Acad Sci USA. 1999; 96:4604–4609. [PubMed: 10200309]
31. Holbert S, Denghien I, Kiechle T, Rosenblatt A, Wellington C, Hayden MR, Margolis RL, Ross CA, Dausset J, Ferrante RJ, Neri C. The Gln-Ala repeat transcriptional activator CA150 interacts with huntingtin: Neuropathologic and genetic evidence for a role in Huntington's disease pathogenesis. Proc Natl Acad Sci USA. 2001; 98:1811–1816. [PubMed: 11172033]
32. Schiffer NW, Ceraline J, Hartl FU, Broadley SA. N-terminal polyglutamine-containing fragments inhibit androgen receptor transactivation function. Biol Chem. 2008; 389:1455–1466. [PubMed: 18844449]
33. Luthi-Carter R, Strand AD, Hanson SA, Kooperberg C, Schilling G, La Spada AR, Merry DE, Young AB, Ross CA, Borchelt DR, Olson JM. Polyglutamine and transcription: gene expression changes shared by DRPLA and Huntington's disease mouse models reveal context-independent effects. Hum Mol Genet. 2002; 11:1927–1937. [PubMed: 12165555]
34. Van Roon-Mom WMC, Reid SJ, Faull RLM, Snell RG. TATA-Binding protein in neurodegenerative disease. Neuroscience. 2005; 133:863–872. [PubMed: 15916858]
35. McCampbell A, Taylor JP, Taye AA, Robitschek J, Li M, Walcott J, Merry D, Chai YH, Paulson H, Sobue G, Fischbeck KH. CREB-binding protein sequestration by expanded polyglutamine. Hum Mol Genet. 2000; 9:2197–2202. [PubMed: 10958659]
36. Hughes RE, Lo RS, Davis C, Strand AD, Neal CL, Olson JM, Fields S. Altered transcription in yeast expressing expanded polyglutamine. Proc Natl Acad Sci USA. 2001; 98:13201–13206. [PubMed: 11687606]
37. McCampbell A, Taye AA, Whitty L, Penney E, Steffan JS, Fischbeck KH. Histone deacetylase inhibitors reduce polyglutamine toxicity. Proc Natl Acad Sci USA. 2001; 98:15179–15184. [PubMed: 11742087]
38. Nucifora FC, Sasaki M, Peters MF, Huang H, Cooper JK, Yamada M, Takahashi H, Tsuji S, Troncoso J, Dawson VL, Dawson TM, Ross CA. Interference by Huntingtin and atrophin-1 with CBP-mediated transcription leading to cellular toxicity. Science. 2001; 291:2423–2428. [PubMed: 11264541]
39. Steffan JS, Bodai L, Pallos J, Poelman M, McCampbell A, Apostol BL, Kazantsev A, Schmidt E, Zhu YZ, Greenwald M, Kurokawa R, Housman DE, Jackson GR, Marsh JL, Thompson LM. Histone deacetylase inhibitors arrest polyglutamine-dependent neurodegeneration in Drosophila. Nature. 2001; 413:739–743. [PubMed: 11607033]
40. Li FS, Macfarlan T, Pittman RN, Chakravarti D. Ataxin-3 is a histone-binding protein with two independent transcriptional corepressor activities. J Biol Chem. 2002; 277:45004–45012. [PubMed: 12297501]
41. Mantamadiotis T, Lemberger T, Bleckmann SC, Kern H, Kretz O, Villalba AM, Tronche F, Kellendonk C, Gau D, Kapfhammer J, Otto C, Schmid W, Schutz G. Disruption of CREB function in brain leads to neurodegeneration. Nature Genet. 2002; 31:47–54. [PubMed: 11967539]
42. Taylor JP, Taye AA, Campbell C, Kazemi-Esfarjani P, Fischbeck KH, Min KT. Aberrant histone acetylation, altered transcription, and retinal degeneration in a Drosophila model of polyglutamine disease are rescued by CREB-binding protein. Genes Dev. 2003; 17:1463–1468. [PubMed: 12815067]
43. Kalkhoven E. CBP and p300: HATs for different occasions. Biochem Pharmacol. 2004; 68:1145–1155. [PubMed: 15313412]
44. Rouaux C, Loeffler JP, Boutillier AL. Targeting CREB-binding protein (CBP) loss of function as a therapeutic strategy in neurological disorders. Biochem Pharmacol. 2004; 68:1157–1164. [PubMed: 15313413]

45. Jiang HB, Poirier MA, Liang YD, Pei Z, Weiskittel CE, Smith WW, DeFranco DB, Ross CA. Depletion of CBP is directly caused by mutant huntingtin. Neurobiol Dis. 2006; 23:543–551. [PubMed: 16766198]
46. Schaffar G, Breuer P, Boteva R, Behrends C, Tzvetkov N, Strippel N, Sakahira H, Siegers K, Hayer-Hartl M, Hartl FU. Cellular toxicity of polyglutamine expansion proteins: Mechanism of transcription factor deactivation. Mol Cell. 2004; 15:95–105. [PubMed: 15225551]
47. Orr AL, Li S, Wang CE, Li H, Wang J, Rong J, Xu X, Mastroberardino PG, Greenamyre JT, Li XJ. N-terminal mutant huntingtin associates with mitochondria and impairs mitochondrial trafficking. J Neurosci. 2008; 28:2783–2792. [PubMed: 18337408]
48. Taylor JP, Tanaka F, Robitschek J, Sandoval CM, Taye A, Markovic-Plese S, Fischbeck KH. Aggresomes protect cells by enhancing the degradation of toxic polyglutamine-containing protein. Hum Mol Genet. 2003; 12:749–757. [PubMed: 12651870]
49. Rockabrand E, Slepko N, Pantalone A, Nukala VN, Kazantsev A, Marsh JL, Sullivan PG, Steffan JS, Sensi SL, Thompson LM. The first 17 amino acids of Huntingtin modulate its sub-cellular localization, aggregation and effects on calcium homeostasis. Hum Mol Genet. 2007; 16:61–77. [PubMed: 17135277]
50. Morley JF, Brignull HR, Weyers JJ, Morimoto RI. The threshold for polyglutamine-expansion protein aggregation and cellular toxicity is dynamic and influenced by aging in Caenorhabditis elegans. Proc Natl Acad Sci USA. 2002; 99:10417–10422. [PubMed: 12122205]
51. Berke SJS, Paulson HL. Protein aggregation and the ubiquitin proteasome pathway: gaining the UPPer hand on neurodegeneration. Curr Opin Genet Dev. 2003; 13:253–261. [PubMed: 12787787]
52. Morimoto RI. Proteotoxic stress and inducible chaperone networks in neurodegenerative disease and aging. Genes Dev. 2008; 22:1427–1438. [PubMed: 18519635]
53. Williams AJ, Knutson TM, Gould VFC, Paulson HL. In vivo suppression of polyglutamine neurotoxicity by C-terminus of Hsp70-interacting protein (CHIP) supports an aggregation model of pathogenesis. Neurobiol Dis. 2009; 33:342–353. [PubMed: 19084066]
54. Behrends C, Langer CA, Boteva R, Bottcher UM, Stemp MJ, Schaffar G, Rao BV, Giese A, Kretzschmar H, Siegers K, Hartl FU. Chaperonin TRiC promotes the assembly of polyQ expansion proteins into nontoxic oligomers. Mol Cell. 2006; 23:887–897. [PubMed: 16973440]
55. Broadley SA, Hartl FU. The role of molecular chaperones in human misfolding diseases. FEBS Lett. 2009; 583:2647–2653. [PubMed: 19393652]
56. Wacker JL, Huang SY, Steele AD, Aron R, Lotz GP, Nguyen Q, Giorgini F, Roberson ED, Lindquist S, Masliah E, Muchowski PJ. Loss of Hsp70 Exacerbates Pathogenesis But Not Levels of Fibrillar Aggregates in a Mouse Model of Huntington's Disease. J Neurosci. 2009; 29:9104–9114. [PubMed: 19605647]
57. Wacker JL, Zareie MH, Fong H, Sarikaya M, Muchowski PJ. Hsp70 and Hsp40 attenuate formation of spherical and annular polyglutamine oligomers by partitioning monomer. Nature Struct Mol Biol. 2004; 11:1215–1222. [PubMed: 15543156]
58. Chopra M, Reddy AS, Abbott NL, De Pablo JJ. Folding of polyglutamine chains. J Chem Phys. 2008; 129
59. Khare SD, Ding F, Gwanmesia KN, Dokholyan NV. Molecular origin of polyglutamine aggregation in neurodegenerative diseases. PLoS Comp Biol. 2005; 1:230–235.
60. Marchut AJ, Hall CK. Side-chain interactions determine amyloid formation by model polyglutamine peptides in molecular dynamics simulations. Biophys J. 2006; 90:4574–4584. [PubMed: 16565057]
61. Marchut AJ, Hall CK. Effects of chain length on the aggregation of model polyglutamine peptides: Molecular dynamics simulations. Proteins: Struct Funct Bioinform. 2007; 66:96–109.
62. Zanuy D, Gunasekaran K, Lesk AM, Nussinov R. Computational study of the fibril organization of polyglutamine repeats reveals a common motif identified in beta-helices. J Mol Biol. 2006; 358:330–345. [PubMed: 16503338]
63. Esposito L, Paladino A, Pedone C, Vitagliano L. Insights into structure, stability, and toxicity of monomeric and aggregated polyglutamine models from molecular dynamics simulations. Biophys J. 2008; 94:4031–4040. [PubMed: 18234827]

64. Vitalis A, Wang X, Pappu RV. Atomistic simulations of the effects of polyglutamine chain length and solvent quality on conformational equilibria and spontaneous homodimerization. J Mol Biol. 2008; 384:279–297. [PubMed: 18824003]
65. Vitalis A, Lyle N, Pappu RV. Thermodynamics of beta-Sheet Formation in Polyglutamine. Biophys J. 2009; 97:303–311. [PubMed: 19580768]
66. Bhattacharyya AM, Thakur AK, Wetzel R. Polyglutamine aggregation nucleation: Thermodynamics of a highly unfavorable protein folding reaction. Proc Natl Acad Sci USA. 2005; 102:15400–15405. [PubMed: 16230628]
67. Thakur AK, Jayaraman M, Mishra R, Thakur M, Chellgren VM, Byeon IJL, Anjum DH, Kodali R, Creamer TP, Conway JF, Gronenborn AM, Wetzel R. Polyglutamine disruption of the huntingtin exon 1 N terminus triggers a complex aggregation mechanism. Nature Struct Mol Biol. 2009; 16:380–389. [PubMed: 19270701]
68. Lee CC, Walters RH, Murphy RM. Reconsidering the mechanism of polyglutamine peptide aggregation. Biochemistry (Mosc). 2007; 46:12810–12820.
69. Walters RH, Murphy RM. Examining Polyglutamine Peptide Length: A Connection between Collapsed Conformations and Increased Aggregation. J Mol Biol. 2009; 393:978–992. [PubMed: 19699209]
70. Singh VR, Lapidus LJ. The intrinsic stiffness of polyglutamine peptides. J Phys Chem B. 2008; 112:13172–13176. [PubMed: 18817433]
71. Perutz MF. Glutamine repeats and neurodegenerative diseases: molecular aspects. Trends Biochem Sci. 1999; 24:58–63. [PubMed: 10098399]
72. Williamson TE, Vitalis A, Crick SL, Pappu RV. Modulation of polyglutamine conformations and dimer formation by the N-terminus of Huntingtin. J Mol Biol. 2010; 396:1295–1309. [PubMed: 20026071]
73. Tam S, Spiess C, Auyeung W, Joachimiak L, Chen B, Poirier MA, Frydman J. The chaperonin TRiC blocks a huntingtin sequence element that promotes the conformational switch to aggregation. Nature Struct Mol Biol. 2009; 16:1279–1286. [PubMed: 19915590]
74. Duennwald ML, Jagadish S, Muchowski PJ, Lindquist S. Flanking sequences profoundly alter polyglutamine toxicity in yeast. Proc Natl Acad Sci USA. 2006; 103:11045–11050. [PubMed: 16832050]
75. Rousseau F, Schymkowitz J, Serrano L. Protein aggregation and amyloidosis: confusion of the kinds? Current Opinion In Structural Biology. 2006; 16:118–126. [PubMed: 16434184]
76. Wyman, J.; Gill, SJ. Binding and Linkage: Functional chemistry of biological macromolecules. First. University Science Books; Mill Valley, CA: 1990.
77. Darnell GD, Derryberry J, Kurutz JW, Meredith SC. Mechanism of Cis-Inhibition of PolyQ Fibrillation by PolyP: PPII Oligomers and the Hydrophobic Effect. Biophys J. 2009; 97:2295–2305. [PubMed: 19843462]
78. Crick, SL.; Pappu, RV. Thermodynamic and kinetic models for the aggregation of intrinsically disordered proteins. In: Schweitzer-Stenner, R., editor. Peptide Folding, Misfolding, and Unfolding. John Wiley & Sons; Hoboken, NJ: 2010.
79. Ferrone F. Analysis of protein aggregation kinetics. Methods in Enzymology. 1999; 309:256–274. [PubMed: 10507029]
80. Powers ET, Powers DL. The kinetics of nucleated polymerizations at high concentrations: Amyloid fibril formation near and above the “supercritical concentration”. Biophys J. 2006; 91:122–132. [PubMed: 16603497]
81. Wetzel R. Kinetics and thermodynamics of amyloid fibril assembly. Acc Chem Res. 2006; 39:671–679. [PubMed: 16981684]
82. Wetzel R. Nucleation of huntingtin aggregation in cells. Nat Chem Biol. 2006; 2:297–298. [PubMed: 16710335]
83. Chen SM, Ferrone FA, Wetzel R. Huntington's disease age-of-onset linked to polyglutamine aggregation nucleation. Proc Natl Acad Sci USA. 2002; 99:11884–11889. [PubMed: 12186976]
84. Bernacki JP, Murphy RM. Model Discrimination and Mechanistic Interpretation of Kinetic Data in Protein Aggregation Studies. Biophys J. 2009; 96:2871–2887. [PubMed: 19348769]

85. Nagai Y, Fujikake N, Ohno K, Higashiyama H, Popiel HA, Rahadian J, Yamaguchi M, Strittmatter WJ, Burke JR, Toda T. Prevention of polyglutamine oligomerization and neurodegeneration by the peptide inhibitor QBP1 in Drosophila. Hum Mol Genet. 2003; 12:1253–1259. [PubMed: 12761040]

86. Sanchez I, Mahlke C, Yuan JY. Pivotal role of oligomerization in expanded polyglutamine neurodegenerative disorders. Nature. 2003; 421:373–379. [PubMed: 12540902]

87. Ossato G, Digman MA, Aiken C, Lukacsovich T, Marsh JL, Gratton E. A Two-Step Path to Inclusion Formation of Huntingtin Peptides Revealed by Number and Brightness Analysis. Biophys J. 2010; 98:3078–3085. [PubMed: 20550921]

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

89. Tada M, Kerppola TK, Decker SJ, Todi SV, Scaglione MK, Costa MDP, Paulson H. Detection of Polyglutamine Protein Oligomers in Living Cells Using Protein-Fragment Complementation Assays. Neurology. 2010; 74:A83–A83.

90. Legleiter J, Mitchell E, Lotz GP, Sapp E, Ng C, DiFiglia M, Thompson LM, Muchowski PJ. Mutant Huntingtin Fragments Form Oligomers in a Polyglutamine Length-dependent Manner in Vitro and in Vivo. J Biol Chem. 2010; 285:14777–14790. [PubMed: 20220138]

91. Suopanki J, Gotz C, Lutsch G, Schiller J, Harjes P, Herrmann A, Wanker EE. Interaction of huntingtin fragments with brain membranes - clues to early dysfunction in Huntington's disease. J Neurochem. 2006; 96:870–884. [PubMed: 16405500]

92. Poirier MA, Li HL, Macosko J, Cai SW, Amzel M, Ross CA. Huntingtin spheroids and protofibrils as precursors in polyglutamine fibrilization. J Biol Chem. 2002; 277:41032–41037. [PubMed: 12171927]

93. Knowles TPJ, Waudby CA, Devlin GL, Cohen SIA, Aguzzi A, Vendruscolo M, Terentjev EM, Welland ME, Dobson CM. An Analytical Solution to the Kinetics of Breakable Filament Assembly. Science. 2009; 326:1533–1537. [PubMed: 20007899]

94. Xue WF, Homans SW, Radford SE. Systematic analysis of nucleation-dependent polymerization reveals new insights into the mechanism of amyloid self-assembly. Proc Natl Acad Sci USA. 2008; 105:8926–8931. [PubMed: 18579777]

95. Andrews JM, Roberts CJ. A Lumry-Eyring nucleated polymerization model of protein aggregation kinetics: 1. Aggregation with pre-equilibrated unfolding. J Phys Chem B. 2007; 111:7897–7913. [PubMed: 17571872]

96. Li Y, Roberts CJ. Lumry-Eyring Nucleated-Polymerization Model of Protein Aggregation Kinetics. 2. Competing Growth via Condensation and Chain Polymerization. J Phys Chem B. 2009; 113:7020–7032. [PubMed: 19368365]

97. Kar K, Jayaraman M, Sahoo B, Kodali R, Wetzel R. Critical nucleus size for disease-related polyglutamine aggregation is repeat-length dependent. Nature Struct Mol Biol. 2011; 18:10. [PubMed: 21666678]

98. Oosawa F, Kasai M. A theory of linear and helical aggregations of macrmolecules. J Mol Biol. 1962; 4:10–21. [PubMed: 14482095]

99. Morris AM, Watzky MA, Agar JN, Finke RG. Fitting neurological protein aggregation kinetic data via a 2-step, minimal/“Ockham's razor” model: The Finke-Watzky mechanism of nucleation followed by autocatalytic surface growth. Biochemistry (Mosc). 2008; 47:2413.

100. Kaschiev, D. Nucleation: Basic theory with applications. Butterworth-Heinemann; Oxford, UK: 2000.

101. Serio TR, Cashikar AG, Kowal AS, Sawicki GJ, Moslehi JJ, Serpell L, Arnsdorf MF, Lindquist SL. Nucleated Conformational Conversion and the Replication of Conformational Information by a Prion Determinant. Science. 2000; 289:1317–1321. [PubMed: 10958771]

102. Lomakin A, Asherie N, Benedek GB. Liquid-solid transition in nuclei of protein crystals. Proc Natl Acad Sci USA. 2003; 100:10254–10257. [PubMed: 12925745]

103. Bitan G, Kirkitadze MD, Lomakin A, Vollers SS, Benedek GB, Teplow DB. Amyloid beta-protein (Abeta) assembly: Abeta40 and Abeta42 oligomerize through distinct pathways. Proc Natl Acad Sci USA. 2003; 100:330–335. [PubMed: 12506200]

104. Krishnan R, Lindquist SL. Structural insights into a yeast prion illuminate nucleation and strain diversity. Nature. 2005; 435:765–772. [PubMed: 15944694]

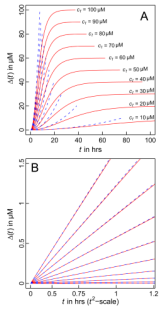

Figure 1. Simulated time course for aggregation

Panel A shows the entire time course of aggregation according to Equation (7) (blue dashed curves) and Equation (9) (solid red curves). Panel B shows a magnification focusing on the early time course and demonstrates the validity of the t^2 dependence for $\Delta(t)$ during the early stages. In the interest of clarity, $\Delta(t)$ is plotted against t^2 in panel B and in both panels, each pair of curves (red and blue) corresponds to a specific value for c_t and these values are shown in panel A.

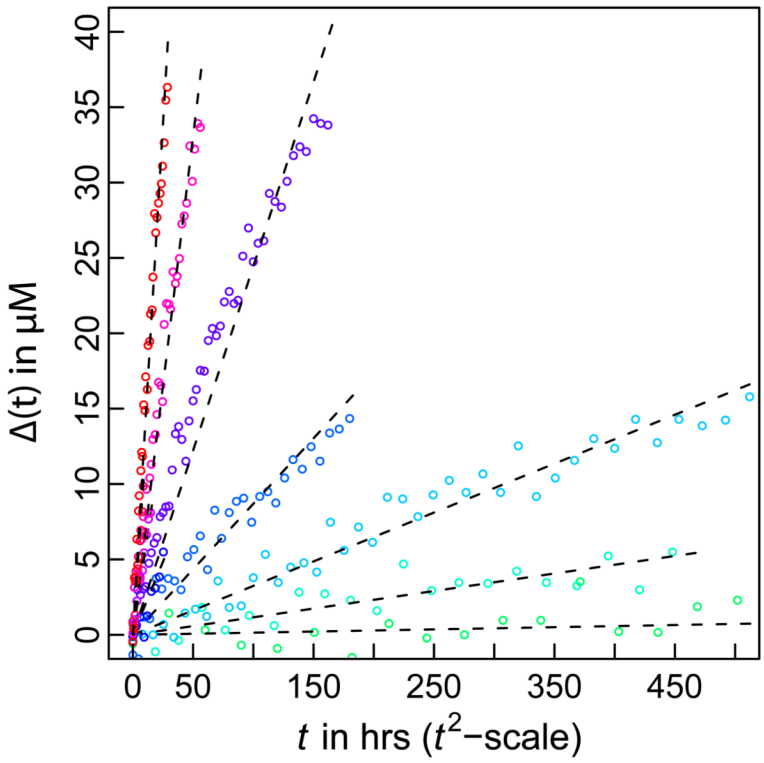

Figure 2. Assessment of the linear relationship between Δ(*t*) and t^2

The red circles denote solutions to the Oosawa-Kasai model summarized in Equation (9) and the black dashed lines denote the results of linear regression, which show reasonable agreement with the full solution.

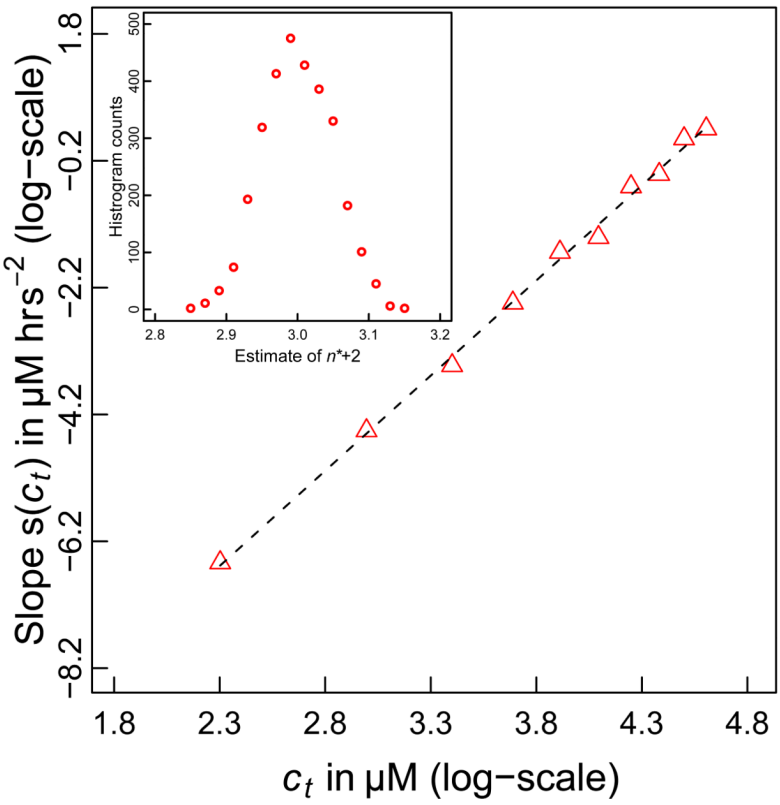

Figure 3. Estimating *n using the double logarithmic plot**

This figure plots ln[s(c_t)] against ln(c_t). The slope of the resultant straight line equals n*+2, from which we can estimate n*. For the current plot, we obtain a slope of 2.971, which is concordant with the value of n*=1 that was used to generate the data in Figures 1-2. The inset shows a histogram of estimates for n*+2 based on 3,000 independent trials and analysis such as shown in Figures 1 and 2. This distribution has a mean of 3.01 and a standard deviation of 0.048.

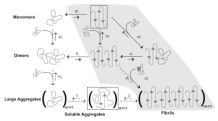

Figure 4. Schematic of possible aggregation pathways for polyglutamine *in vitro*

Here, n denotes the number of polyglutamine molecules within a disordered aggregate and n_F denotes the number of polyglutamine molecules within an ordered amyloid fibril. The ordered amyloid fibril rich in β-sheets is shown in the bottom right corner of the schematic. The gray shaded region encompasses steps (a), (e), (c), and (d) and depicts the homogeneous nucleation proposal of Chen et al. [83]. Studies of the thermodynamics of step (a), indicate that the formation of ordered conformations is thermodynamically unfavorable. Monomeric polyglutamine prefers disordered, collapsed conformations, left of step (a). Step (b) pertains to the thermodynamics of interactions between chains that adopt ordered conformations. However, the likelihood that chains will sample the associations shown in step (b) is very small because this is tied to the equilibria in step (a), which requires the population of the conformations with high β-content. Similarly, step (f) shows that disordered dimers are thermodynamically favored to dimers with high β-contents in individual chains. This is the result of linkage to step (a) as discussed above. The aggregates achieved in step (h) are likely to be large (in terms of n) and exhibit spherical, "liquid-like" [102-104] organization of polyglutamine chains around each other. Step (i) depicts a slow conformational conversion of individual / small numbers of chains to β-sheets. This slow step is likely to lead to the creation of an ordered template for fibril formation via monomer or oligomer addition and elongation to yield the ordered amyloid fibril.

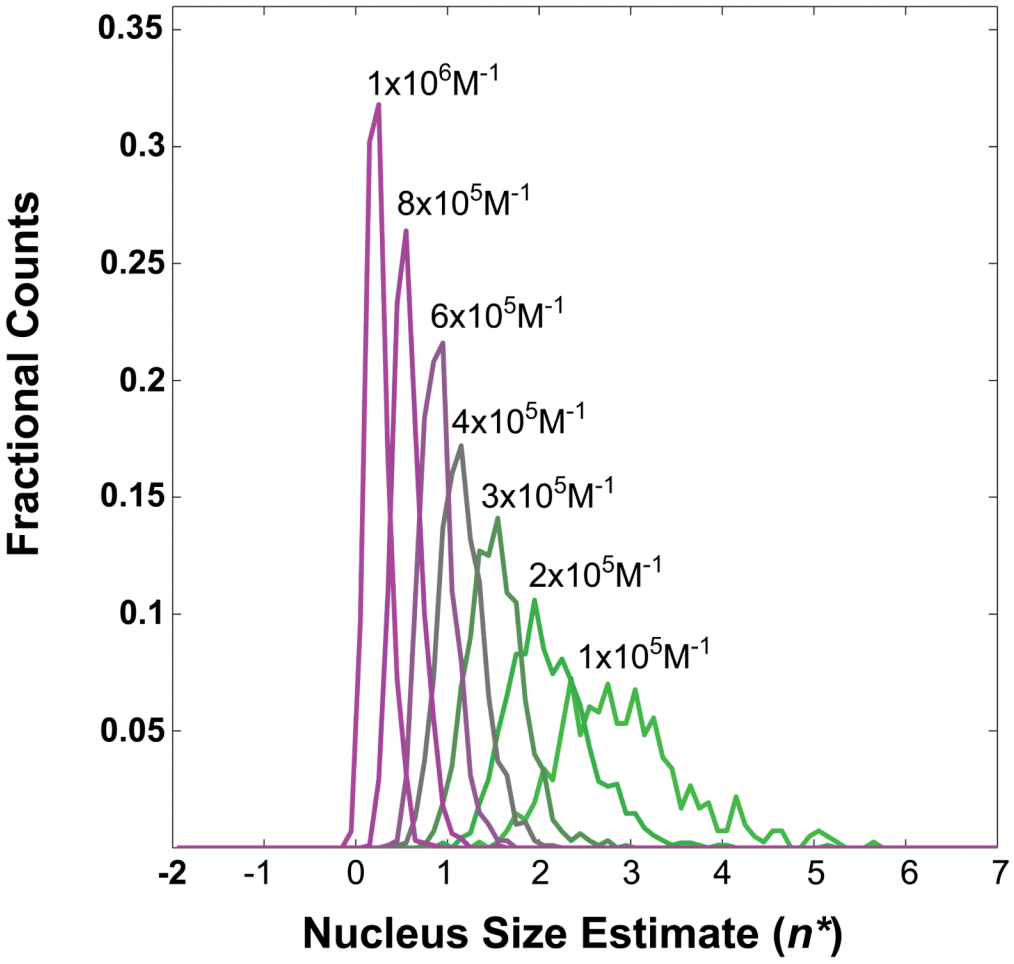

Figure 5. Distribution of apparent nucleus sizes obtained from analysis of simulated data for a heterogeneously nucleated process – effect of oligomer stabilities

While the model includes the presence of heterogeneous distributions of oligomers, the data were analyzed using the homogeneous nucleation model. The plot shows histograms of apparent nucleus sizes for a minimum fibril-forming oligomer with 10 monomers, *i.e.*, i_{min}=10. Each histogram corresponds to different values for the mean of the distribution of K_i and each curve is labeled with the appropriate value used.

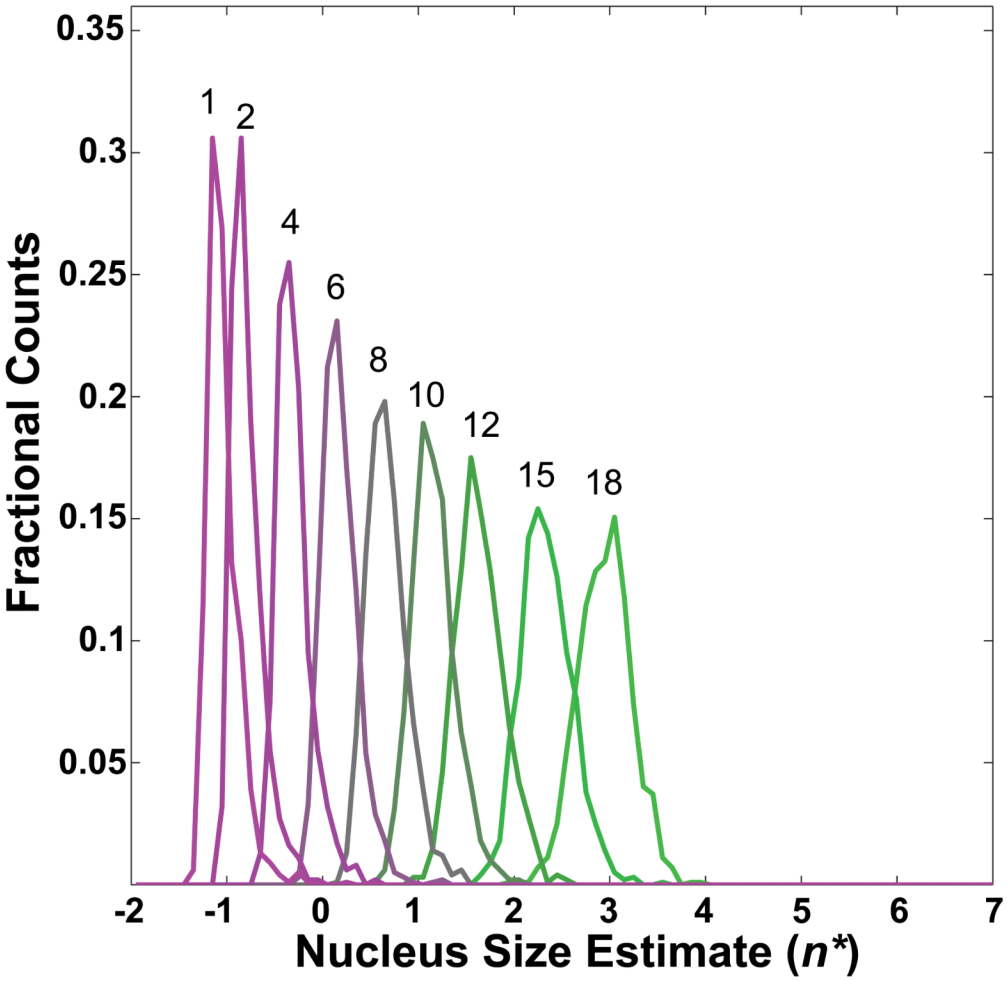

Figure 6. Distribution of apparent nucleus sizes obtained from analysis of simulated data for a heterogeneously nucleated process – effect of varying i_{min}

In contrast to Figure 5, here we fix all $K_i = 4\times10^5 M^{-1}$ and obtain the distribution of n^* values for different values of i_{min}. Each curve is labeled with the value of i_{min} used in the calculations. As i_{min} decreases, so does the estimated value for n^*.