

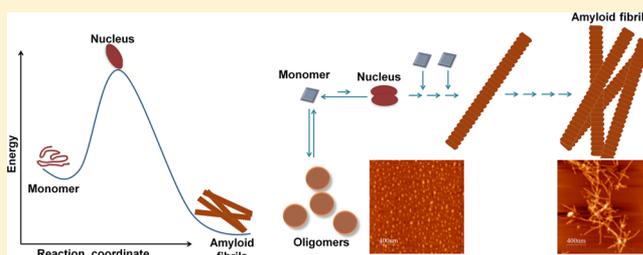
# Amyloid Fibril Formation by the Chain B Subunit of Monellin Occurs by a Nucleation-Dependent Polymerization Mechanism

A. T. Sabareesan and Jayant B. Udgaonkar\*

National Centre for Biological Sciences, Tata Institute of Fundamental Research, Bangalore 560065, India

## Supporting Information

**ABSTRACT:** Proteins possessing very different structures, or even no structure, form amyloid fibrils that are very similar in internal structure. This suggests that the mechanisms by which amyloid fibrils form might be very similar, irrespective of whether the fibrils are associated with disease or with normal cellular function, or even if they have no physiological importance. In this context, it is important to have a model protein system whose amyloid fibril formation is robust in its reproducibility, which can reveal the fundamentals of the amyloid fibril reaction that may be applicable to all proteins. In this study, the aggregation mechanism of amyloid fibril formation by chain B of the heterodimeric protein monellin has been elucidated in detail. It is shown that the aggregation reaction meets all the stringent kinetic criteria of a homogeneous nucleation-dependent polymerization mechanism, which is valid over a wide range of protein concentrations. Quantitative analyses of the kinetic data using one approach based on features of the entire kinetic curve, and another based on only the initial rate of aggregation, indicate that the thermodynamic nucleus is a dimer. Spherical oligomers are observed by atomic force microscopy to form transiently early during fibril formation but are off-pathway to the direct fibril formation pathway. It is shown that amyloid fibril formation can be prevented by the addition of chain A of monellin at early stages of chain B aggregation: the two free chains combine to form native monellin, which leads to the dissociation of early aggregates.



Amyloid fibrils are composed of protein molecules that have self-assembled into polymeric structures possessing a characteristic cross- $\beta$  spine, not into their functional native structures.<sup>1</sup> It appears that any protein is capable of forming amyloid fibrils under suitable conditions, which are usually destabilizing in the case of globular proteins and mildly stabilizing in the case of intrinsically disordered proteins.<sup>2</sup> The structures of amyloid fibrils appear to be strikingly similar, whether they are formed from purely  $\alpha$  proteins,<sup>3,4</sup> predominantly  $\beta$  proteins,<sup>5</sup>  $\alpha+\beta$  proteins,<sup>6,7</sup> or natively unstructured proteins.<sup>8</sup> In fact, the structural homogeneity between the fibrils of different proteins is sufficient in some cases to allow cross seeding of fibril formation, even though the original native structures are vastly different.<sup>9,10</sup> Recent X-ray crystallographic studies have, however, brought out the complexity and heterogeneity in amyloid fibril structures,<sup>11,12</sup> particularly at the level of the bilayer  $\beta$ -sheet motif, which extends to form the characteristic cross- $\beta$  spine.<sup>13–15</sup>

Originally, amyloid fibrils gained prominence as protein aggregates associated with certain human diseases, including Alzheimer's disease, associated with the aggregation of  $A\beta$ , and Parkinson's disease, linked to the aggregation of  $\alpha$ -synuclein.<sup>16,17</sup> More recently, it has been observed that many living systems use amyloid fibrils for various structural and regulatory functions.<sup>18,19</sup> In these cases, amyloid fibril formation would be a desirable process occurring in competition with protein folding. Amyloid fibrils have also evoked interest as nanobiomaterial with potential biotechnological application.<sup>20</sup> Not

surprisingly, the mechanism of formation of amyloid fibrils is the focus of much intense study.

Amyloid fibril formation reactions commonly display sigmoidal kinetics characteristic of nucleation-dependent polymerization (NDP),<sup>21–24</sup> but such kinetics may also be observed<sup>25</sup> for isodesmic (linear) polymerization.<sup>26,27</sup> While there are stringent criteria for distinguishing between isodesmic polymerization and NDP,<sup>27</sup> it is difficult in practice to distinguish between them for two reasons. First, aggregation reactions are inherently complex<sup>28–31</sup> and often characterized by the transient appearance of multiple oligomeric intermediates<sup>27</sup> that could be spherical, rodlike, or curvilinear.<sup>32</sup> Such oligomeric intermediates are often assumed to be on-pathway,<sup>33,34</sup> but some proteins have been shown to be off-pathway.<sup>35</sup> Second, fibril formation reactions are not easy to reproduce,<sup>36,37</sup> making it difficult to study the dependence of aggregation kinetics on protein concentration, which is so necessary for determining the mechanism. Not surprisingly, only a few amyloid fibril formation reactions<sup>38–40</sup> have been explicitly shown to meet all the defining criteria<sup>25,27</sup> of a NDP reaction. In this context, the utility of a model protein system, whose amyloid fibril formation reaction is robust in its

Received: October 29, 2013

Revised: January 27, 2014

Published: February 4, 2014

reproducibility, is that it can reveal principles of the amyloid fibril reaction that may be applicable to all proteins.

Monellin is an intensely sweet heterodimeric protein isolated from the berries of an African plant, *Dioscoreophyllum cuminsii*.<sup>41</sup> It is a dimeric protein made of two polypeptide chains, A and B. It is a member of the  $\beta$ -grasp family of proteins, several of whose members form amyloid fibrils that are associated with human disease.<sup>42,43</sup> Monellin is an attractive model system for the study of protein aggregation because much is known about its folding and unfolding pathways.<sup>44–49</sup> In particular, much is known about how it unfolds in multiple stages,<sup>45,47</sup> making it potentially valuable for identifying the partially unfolded monomeric intermediate from which aggregation commences. Aggregation studies so far have shown that monellin can form amyloid fibrils,<sup>7,50</sup> but only after the protein has dissociated into its constituent chains. It appears that it is chain B and not chain A that forms the amyloid fibrils, but there has so far been no quantitative study of the aggregation mechanism.

In this study, the kinetics of fibril formation by chain B derived from double-chain monellin (dcMN) has been studied extensively over a wide range of protein concentrations. The aggregation reaction has been shown to meet all the defining criteria of a NDP reaction: the existence of a lag phase dependent on protein concentration, the elimination of this lag phase upon addition of preformed fibrils (seed), and the existence of a critical concentration below which aggregation does not occur. Phenomenological scaling was used to show that the same mechanism describes fibril formation over the entire range of protein concentrations studied. The aggregation kinetics were analyzed quantitatively using two different approaches, one based on analysis of the entire kinetic curve and the other based on analysis of the commencement of the aggregation reaction. Both approaches show that the critical aggregation nucleus is a dimer. Although the transient formation of spherical oligomers is seen, it is shown to be off-pathway to the direct fibril formation reaction.

## MATERIALS AND METHODS

**Protein Expression, Purification, and Characterization.** dcMN was expressed and purified using a previously described method.<sup>51</sup> The individual subunits, chain A and chain B, were purified from dcMN using reverse phase chromatography, as described previously.<sup>51</sup> Before the start of each aggregation experiment, to remove any preformed aggregates, the chains were incubated in 6 M guanidine hydrochloride for 2 h, after which the solutions were desalted using a HiTrap desalting column (GE). Freshly processed chains, at a maximal stock concentration of 250  $\mu$ M, were used for all the experiments. The same protocol was followed to ensure consistency and reproducibility in all the aggregation experiments.

The purity of dcMN as well as of the individual chains was determined using electrospray ionization mass spectrometry and sodium dodecyl sulfate–polyacrylamide gel electrophoresis. Circular dichroism (CD) and fluorescence spectra of dcMN and of both chains were acquired to confirm the spectroscopic signatures of the proteins.

**Buffers, Solutions, and Experimental Conditions.** All reagents used were of the highest-purity grade from Sigma, unless otherwise specified. At the start of each experiment, proteins were in either 10 mM phosphate buffer (pH 7.0) (for dcMN) or pH-adjusted (pH 6.0) Milli-Q water (for chains).

The proteins were diluted into aggregation buffer. The final aggregation reaction took place in 10 mM glycine-HCl (pH 3.0), 200 mM NaCl, and 1 mM DTT (250  $\mu$ M for CD experiments) at 25 °C. The time of addition of protein to the aggregation buffer was considered as time zero of aggregation. The variability in the kinetics was reduced by using the same heating block and pipetting all aliquots thrice, thoroughly and precisely using the same 200  $\mu$ L pipet.

**Thioflavin T Fluorescence Assay.** Thioflavin T (ThT) fluorescence was measured at pH 8.0 in 10 mM Tris-HCl buffer. Final concentrations of 1  $\mu$ M protein and 10  $\mu$ M ThT were used. For every measurement, a calculated volume of protein, according to the concentration used for the experiment, was taken from the sample and added to the ThT assay solution. Fluorescence readings were taken within 20 s of the addition of protein to the ThT solution using a Fluoromax-3 spectrofluorimeter (Jobin Yvon) with the following parameters: excitation wavelength, 440 nm; emission wavelength, 482 nm; excitation bandwidth, 1 nm; emission bandwidth, 10 nm. The signal was averaged for 30 s with a response time of 2 s.

**Static Light Scattering Assay.** Static light scattering measurements were taken using the Fluoromax-3 spectrofluorimeter (Jobin Yvon); 100  $\mu$ L of the aggregating protein sample was removed at different time points, and the scattering intensity was measured using a 10 mm path length cuvette, with the following parameters: excitation and emission wavelengths, 800 nm; excitation bandwidth, 1 nm; emission bandwidth, 10 nm. Data were averaged for 120 s.

**Sedimentation Assay for Fibril Formation.** At different time points during aggregation, a 100  $\mu$ L aliquot was transferred into an Amicon Ultra 0.5 mL centricon filter unit with an Ultracel-3 membrane with a cutoff of 3 kDa (Millipore Inc.). The centricon filters were centrifuged at 20000g for 15 min at 25 °C. When free monomeric chain B was subjected to the same treatment, all of it was found to be present in the filtrate. The amount of protein present in the filtrate was measured using the tryptophan fluorescence at 357 nm, upon excitation at 295 nm. The amount of protein present in the aggregates was calculated by subtracting the filtrate protein concentration from the starting monomer concentration.

**Seeding Experiments.** For experiments with sonicated and unsonicated seeds, 5  $\mu$ M B chain aggregates [aggregates obtained at 3 $\tau$  (120 h) of the ThT-monitored kinetics] were used as the seed. For seeding experiments with unsonicated seed, the seed was used directly. The volume of seed suspension replaced an equal volume of the reaction mixture, so that the desired percentage (v/v) of seed concentration was obtained. For experiments with sonicated seed, the seed suspension was kept on ice and sonicated using a micro probe with the following parameters: amplitude, 25%; pulse, 5 s on and 4 s off; total time, 5 min. Across seeding experiments, there was a variation in the final ThT fluorescence value, because of inherent problems associated with sonication, but fractional change plots rectified this difference and were therefore used for comparison with unseeded reactions. The initial rate of polymerization was determined from the slope of a linear fit to the initial part (until 20% of aggregation finishes) of the aggregation curve monitored using ThT fluorescence.

**Circular Dichroism Measurements.** A Jasco J-815 CD spectrometer was used for the far-UV CD measurements, with a cuvette with a path length of 1 mm. The instrument settings were as follows: digital integration time, 1 s; bandwidth, 1 nm; wavelength, 222 or 216 nm; scan averaging time, 120 s. A

sample from the desired concentration of the aggregation reaction mixture was used directly for all the final amplitude measurements. In case of the spectral measurements, a protein concentration of 20 μM was used with a wavelength scan of 200–250 nm.

**Atomic Force Microscopy (AFM).** For the AFM studies, 100 μL of the sample was withdrawn at different time points from a 25 μM aggregation reaction mixture. The samples were applied on freshly cleaved mica and incubated for 5 min, and the mica surface was then rinsed twice with double-filtered Milli-Q water (pH adjusted to 3.0). Samples were dried under vacuum for 1 h, before being scanned. The AFM images were obtained using a PicoPlus AFM instrument (Molecular Imaging Inc.) in the noncontact mode. The height and length of the aggregates were determined using the profile option of WSXM.<sup>52</sup>

**Effect of Chain A on Chain B Aggregation Kinetics.** The effect of addition of chain A on the aggregation of chain B was studied using the 25 μM chain B aggregation reaction mixture. At defined time points during the aggregation of 25 μM chain B, a desired volume of aggregation suspension was removed and replaced with a stock solution of chain A, in the same aggregation buffer, to yield a final chain A concentration of 25 μM (chain B concentration will come down to 23 μM). The reaction mixture was mixed well using a 200 μL pipet before each measurement.

**Data Analysis and Curve Fitting.** The kinetic curves measured by monitoring either ThT fluorescence or light scattering were fit to the equation

$$S = S_0 + \frac{S_\infty - S_0}{1 + e^{-[(t-t_{50})/\tau]}} \quad (1)$$

where  $S_0$  is the spectroscopic signal at time zero,  $S_\infty$  is the final signal,  $t$  is the time,  $t_{50}$  is the time at which the change in signal is 50%, and  $\tau$  is a characteristic time constant. The lag time ( $t_{lag}$ ) was calculated with the equation  $t_{lag} = t_{50} - 2\tau$  as described previously.<sup>53</sup> A similar value for  $t_{lag}$  was obtained by determining the time taken to complete 10% of the reaction. The value of the signal at time zero and the final amplitude of the change in the signal determined from fitting the data to eq 1 were used to calculate the fractional change ( $f$ ) at each time point using the equation

$$f = \frac{S - S_0}{S_\infty - S_0} \quad (2)$$

The apparent elongation rate constant was determined by fitting the kinetic data points, excluding the initial 10% of points, to the single-exponential equation

$$S = S_0 + a[1 - e^{-(t/\tau_{el})}] \quad (3)$$

where  $a$  is the amplitude of the signal and  $\tau_{el}$  is the time constant of elongation. The values obtained for  $\tau$  and  $\tau_{el}$  do not differ by more than 10%.

To determine whether the kinetic curves of aggregation at different protein concentrations collapse upon phenomenological scaling as observed for tubulin polymerization,<sup>54</sup> the signal changes and times of a kinetic curve were normalized to  $S_\infty$  and  $t_{50}$ , respectively.

The kinetic data for the aggregation of chain B were analyzed using nucleation-dependent polymerization (NDP) models based on two different approaches: (a) analysis of the entire kinetic curve<sup>55</sup> and (b) linear perturbation analysis of only the

initial part of the kinetic curve.<sup>22,56</sup> Both analyses are based on the premise that an equilibrium (defined by equilibrium constant  $K_{n^*}$ ) is established between a high-energy oligomeric nucleus (containing  $n^*$  monomers) and the monomer.

The consequence of the first analysis is that  $t_{lag}$  (for a 10% change) and  $t_{50}$  are related to the total protein monomer concentration,  $[X]_{tot}$ , by the equation

$$\log t_{50} = \frac{-(n^* + 1)}{2} \log [X]_{tot}$$

or

$$\log t_{lag} = \frac{-(n^* + 1)}{2} \log [X]_{tot} \quad (4)$$

It should be noted that both equations are valid regardless of whether off-pathway aggregates are formed during the course of fibril formation.<sup>55</sup>

In the second analysis, based on examination of the early part of the kinetic curve of aggregation,<sup>22</sup> it is also assumed, as in the first analysis, that the concentration of prenuclear oligomers is much smaller than either the free monomer concentration  $c(t)$  or the concentration of polymerized monomers  $\Delta(t)$ . Therefore

$$\Delta(t) = c_0 - c(t) \quad (5)$$

where  $c_0$  is the total monomer concentration.  $\Delta(t)$  changes by monomer addition or loss at polymer ends, and this process is assumed to be length-independent for long polymers. Therefore

$$\frac{d\Delta}{dt} = (k_+c - k_-)c_p \quad (6)$$

where  $k_+$  and  $k_-$  are the polymerization and depolymerization rates, respectively, and  $c_p$  is the concentration of polymers.  $k_+$  and  $k_-$  are related by the critical concentration  $c_s$  as  $k_- = k_+c_s$ .

The rate of homogeneous polymer formation can be described as

$$\frac{dc_p}{dt} = k_+cc_{(n^*)} - k_-^*c_{(n^*+1)} \quad (7)$$

The polymer size is  $n^* + 1$  or larger;  $k_+$  and  $k_-^*$  are the rate constants for monomer addition and loss from the nucleus, respectively.  $k_+$  is the same for the addition of the monomer to both the nucleus and the polymer. Equation 7 then becomes

$$\frac{dc_p}{dt} = k_+cc_{n^*} = k_+K_{n^*}c^{n^*+1} \quad (8)$$

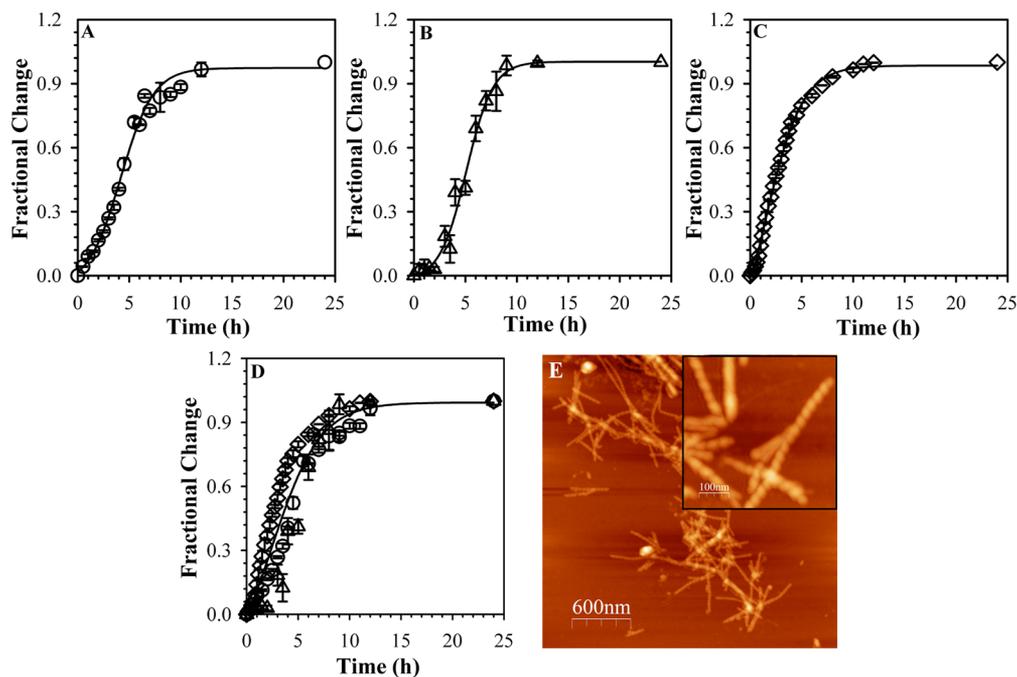
The initial (up to 5%) amplitude of the kinetic curve of aggregation was analyzed using the following equation for homogeneous nucleation:<sup>22,56</sup>

$$\Delta = A[1 - \cos(Bt)] \quad (9)$$

where  $A$  is a parameter that directs the apparent shape of the kinetic curve and  $B$  is an effective rate constant for nucleation. Only the initial 5% of the kinetic data was used because of the oscillatory behavior of the cos function.

A plot of  $\log B^2A/[c_0(c_0 - c_s)]$  versus  $\log c_0$  gives  $n^*$ :

$$\log \frac{B^2A}{c_0(c_0 - c_s)} = \log k_+^2K_{n^*} + n^* \log c_0 \quad (10)$$



**Figure 1.** Amyloid fibril formation by chain B at pH 3 and 25 °C. The aggregation of 25 μM chain B was monitored by the ThT fluorescence assay (A), by direct measurement of the fraction of the total protein in aggregates, as determined by a centrifugation assay (B), and by measurement of the intensity of light scattered at 800 nm (C). The fractional progress of the aggregation reaction is plotted in panel D, where the data from panels A–C have been normalized between values of 0 and 1 using eq 2: ThT fluorescence (○), fibril concentration (△), and light scattering intensity (◇). In panels A–D, the solid lines through the data are nonlinear least-squares fits to eq 1, and the error bars represent the standard deviations from three independent experiments using three different batches of protein preparations. Panel E shows an AFM image of the amyloid fibrils formed by 25 μM chain B at the 3τ (12 h) time point of the aggregation kinetics. The inset in panel E shows the fibrils at a higher magnification.

The value of  $c_s$  was determined experimentally to be 4 μM (see Results).

For experiments conducted in the presence of preformed fibrils (seed) at fixed concentration  $c_p$ , eqs 5 and 6 can be combined to give the first-order equation

$$\frac{d\Delta}{dt} = (k_+c_0 - k_-)c_p - k_+c_p\Delta \quad (11)$$

At time zero, the initial rate of fibril extension from the seeds is given by

$$\frac{d\Delta}{dt} = k_+c_0c_p \quad (12)$$

When  $c_0$  is constant, the initial rate of fibril extension is proportional to  $c_p$ , and when  $c_p$  is constant, the initial rate is proportional to  $c_0$ .

All fits were performed using the regression wizard option of SigmaPlot 12, which uses the Levenberg–Marquardt algorithm to find the best fit between the data and the equation.

## RESULTS

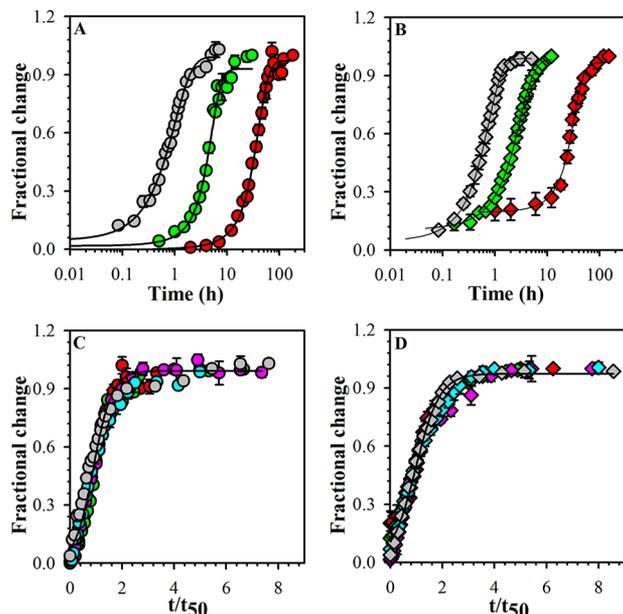
Previous studies of amyloid fibril formation of dcMN had indicated that aggregation proceeded only after dissociation of dcMN into its constituent chains, and only chain B was shown to be capable of forming amyloid fibrils.<sup>7,50</sup> In that study, the mechanism of aggregation was not established quantitatively. Here, the mechanism of amyloid fibril formation by chain B, at pH 3 and 25 °C in the absence of any agitation, has been characterized in detail in a quantitative manner.

**Amyloid Fibril Formation by Chain B Was Monitored Using Multiple Probes.** Figure 1A shows that the kinetics of amyloid fibril formation are sigmoidal in nature when

determined using the ThT assay, which is a standard assay for the β-sheet conformational change accompanying amyloid fibril formation. When the amount of insoluble material at any time during aggregation was assayed by a sedimentation assay (see Materials and Methods), similar sigmoidal kinetics were again observed (Figure 1B). Visible precipitation was observed only after the reaction is more than 50% complete. Amyloid fibril formation was also monitored by measurement of the intensity of scattering at 90° of 800 nm light. It should be noted that measurements of the scattering intensity, which is proportional to the number of particles and to the square of their mass, are biased toward the measurement of large particles. This probe also showed sigmoidal kinetics (Figure 1C), but two features of the signal change it reports distinguish it from the other two probes. First, the scattering intensity shows an initial increase immediately upon transfer of chain B to the aggregation buffer, and second, the kinetics appears to be slightly faster than that measured by the other two probes. The latter feature is more apparent in Figure 1D, where the kinetic curves measured by the three probes are compared. Very importantly, for quantitative analysis of the amyloid fibril formation reaction, the kinetic curves are highly reproducible for all probes used, as reflected in the small errors in measurement. Figure 1E shows a representative AFM image of the aggregate formed at a time corresponding to 3 times the time constant of the elongation phase of the sigmoidal kinetic curve, determined using eq 1. Only fibrils are seen, and no amorphous aggregates are seen in any AFM image. The fibrils are invariably seen to clump together. The fibrils at the borders of clumps have a height of  $7 \pm 2$  nm, as determined from Gaussian fits to the distributions of their measured heights. The height of each fibril was found to vary by  $\pm 0.6$  nm with a

periodicity of ~50 nm all along its length (not shown). This regular repetitive variation in height gives the fibrils the appearance of being composed of a pair of helical filaments tightly wound around each other and may be similar to that seen in the case of fibrils formed by tau.<sup>35,57</sup>

**The Kinetics of Aggregation Can Be Explained by a Single Mechanism at All Protein Concentrations.** The kinetics of amyloid fibril formation was studied over a range of protein concentrations, from 10 to 100  $\mu\text{M}$  (Figure 2). At all

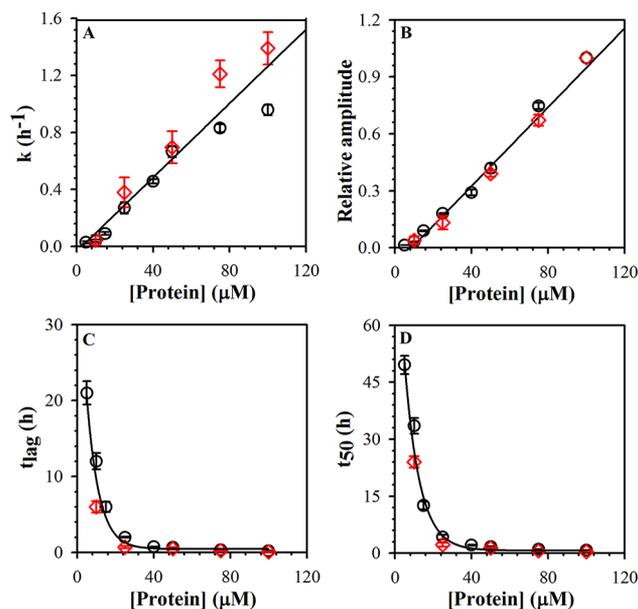


**Figure 2.** Phenomenological scaling of the kinetic curves of aggregation. Panels A and B show kinetic curves of aggregation obtained for three different concentrations of chain B, monitored by ThT fluorescence and scattering intensity at 800 nm, respectively. A logarithmic time axis is used to show widely different chain B concentrations. Panels C and D show the entire data set of kinetic curves obtained at different chain B concentrations, monitored by ThT fluorescence and light scattering intensity, respectively, in which the progress of each reaction is shown relative to a value of 1 at completion and the time scale for each reaction is normalized to the value of  $t_{50}$  for that reaction. The data are for the following protein concentrations: 10 (red), 25 (green), 50 (pink), 75 (cyan), and 100  $\mu\text{M}$  (gray). In each panel, the solid line through each kinetic curve is a nonlinear least-squares fit to eq 1. Error bars represent the standard deviations determined from three independent experiments using three different batches of chain B.

protein concentrations, the reproducibility of the aggregation kinetics was confirmed by using three different preparations of chain B. Panels A and B of Figure 2 show the aggregation kinetics at three different concentrations, i.e., 10, 25, and 100  $\mu\text{M}$ , as monitored using ThT fluorescence and light scattering at 800 nm, respectively. The aggregation kinetics remains sigmoidal at all protein concentrations, with the lag time decreasing and the elongation rate increasing with an increase in protein concentration. In aggregation studies, it is important to show that the mechanism of aggregation does not change over the range of protein concentrations used. All the kinetic curves obtained at chain B concentrations in the range of 10–100  $\mu\text{M}$ , with either of the two probes, were found to collapse into a single kinetic curve (Figure 2C,D), upon phenomenological scaling,<sup>54</sup> when the signal from the spectroscopic probe of each kinetic curve was normalized to the final signal and the

time scale for each curve was normalized to the time at which the signal had reached 50% of the final signal (see Materials and Methods). This result suggests that the same mechanism must describe aggregation over the entire range of protein concentrations studied.

**Amyloid Fibril Formation Does Not Occur below a Critical Concentration of ~4  $\mu\text{M}$ .** Panels A and B of Figure 3



**Figure 3.** Dependence of aggregation kinetics on chain B concentration. The kinetics of fibril formation was monitored by measurement of ThT fluorescence ( $\circ$ ) and light scattering at 800 nm ( $\diamond$ ). Panels A and B show the dependencies of the elongation rate and final amplitude of the reaction, respectively, on chain B concentration. Linear fits through the data extrapolate to intercept the  $x$ -axis at a critical concentration of  $4 \pm 0.5 \mu\text{M}$ . Panels C and D show the dependencies of the lag time ( $t_{\text{lag}}$ ) and  $t_{50}$  on chain B concentration, respectively. The solid lines through the data are least-squares fits to a single-exponential equation. In each panel, the error bars represent the standard deviations determined from three independent experiments using three different batches of protein.

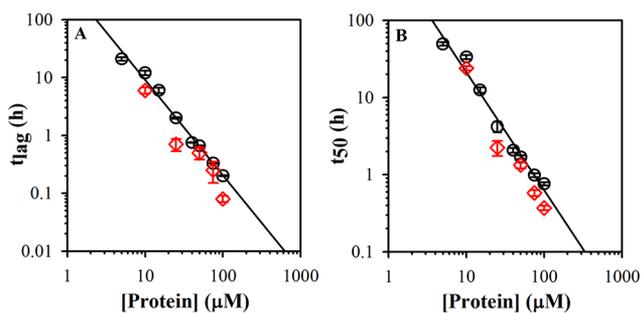
show how the apparent rate constant of elongation as well as the amplitude of the aggregation reaction, corresponding to the formation of mature fibrils, depends upon protein concentration. ThT fluorescence and scattering intensity measurements were used to monitor the reaction. The intercepts on the abscissa, of the straight line fits to both the ThT fluorescence and scattering data, suggest that a critical concentration of ~4  $\mu\text{M}$  exists for the formation of mature fibrils. The dependence of the apparent rate constant measured by ThT fluorescence on protein concentration becomes nonlinear at protein concentrations of >50  $\mu\text{M}$ , although the relative amplitude retains a linear dependence on protein concentration (Figure 3B).

If there indeed exists a critical concentration below which the protein does not aggregate, then it is expected that when the aggregation reaction has reached equilibrium, the monomer should still be present at a concentration corresponding to the critical concentration. The final monomer concentration present at equilibrium for aggregation reactions conducted at different protein concentrations was found to be  $5 \pm 1.0 \mu\text{M}$  (data not shown). It was also found that no amyloid fibril formation could be detected for a long time, when the monomer concentration was 2  $\mu\text{M}$  (data not shown). These

observations confirmed the existence of a critical concentration of  $\sim 4 \mu\text{M}$ .

**Establishment of a NDP Mechanism and Characterization of the Size of the Nucleus.** The observation of sigmoidal aggregation kinetics and a critical protein concentration only above which fibril formation occurs suggests that the fibrils form via a NDP mechanism. Figure 3C shows that the lag time,  $t_{\text{lag}}$ , decreases exponentially with an increase in protein concentration, which is another characteristic feature of a NDP mechanism.<sup>23</sup>

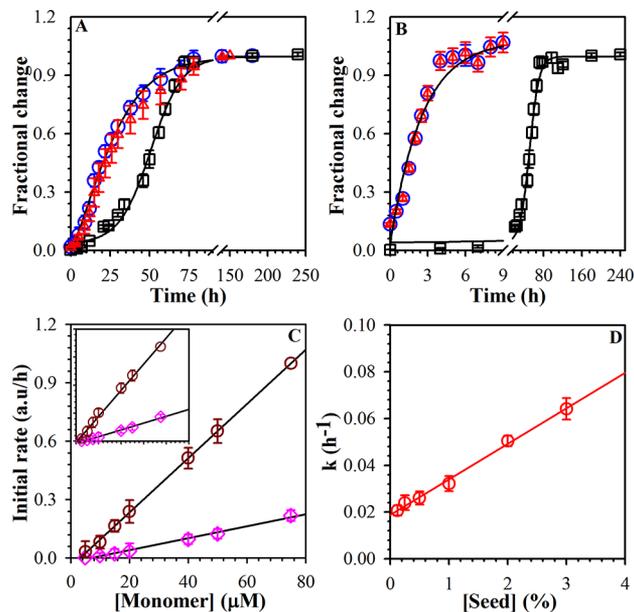
The strong dependence of the rate of fibril formation on protein concentration can be expressed in terms of  $t_{50}$ , the time at which fibril formation is 50% complete (Figure 3D). Figure 4A is a log–log plot of  $t_{\text{lag}}$  versus protein concentration. The



**Figure 4.** Estimation of nucleus size. The kinetics of fibril formation was monitored by measurement of ThT fluorescence (○) and light scattering at 800 nm (◇). Panels A and B are log–log plots of the dependencies of  $t_{\text{lag}}$  and  $t_{50}$ , respectively, on chain B concentration. The straight lines through the data in each panel are linear fits, with slopes of 1.6 (A) and 1.5 (B). In each panel, the error bars represent the standard deviations determined from three independent experiments using three different batches of protein.

slope of this plot was used to calculate the size of the nucleus,  $n^*$  (see eq 4), according to a simple NDP mechanism.<sup>24,55</sup>  $n^*$  was found to have a value of 2. A log–log plot of  $t_{50}$  versus protein concentration (Figure 4B) also yields a value of 2 for  $n^*$  (see eq 4). Hence, the critical nucleus for amyloid fibril formation is a dimer, which is always in equilibrium with the monomer. It should be noted that in the analyses used in this study<sup>22,55,56</sup> the nucleus has been considered to be the least stable oligomer on the aggregation pathway. In other analyses of homogeneous nucleation,<sup>27,58</sup> the nucleus was considered to be the first stable oligomer that is on-pathway to the fibril formation.<sup>23,59</sup> If the latter definition of the nucleus is used to analyze our data, its size is found to be one monomer unit more than when the former definition of the nucleus is used.

**Chain B Aggregation Follows a Monomer Addition Polymerization Mechanism.** A critical test of the NDP mechanism is that the lag phase can be abolished by the addition of preformed fibrils (seed). Panels A and B of Figure 5 show that the addition of 3 and 10% seed, respectively, either sonicated or unsonicated, completely abolishes the lag phase of aggregation of  $5 \mu\text{M}$  chain B. At either seed concentration, sonicated and unsonicated seeds had exactly the same effect (Figure 5A,B). At either seed concentration, for unsonicated and for sonicated seed, the initial rate of aggregation is found to be linearly dependent on the monomeric chain B concentration (Figure 5C). This result suggests that monomeric chain B directly adds to the seed during its growth into a fibril. Figure 5D shows that the elongation rate is directly proportional to

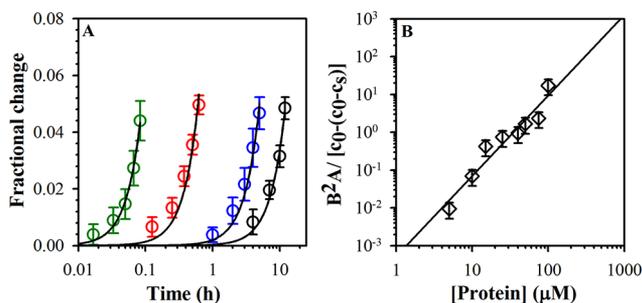


**Figure 5.** Effect of seeding on aggregation kinetics. Panels A and B show the ThT fluorescence-monitored kinetics of amyloid fibril formation by  $5 \mu\text{M}$  chain B in the presence of 3 and 10% unsonicated seeds (○) and sonicated seeds (△), respectively, and in the absence of seed (□). Panel C shows the dependence of the initial rate of the reaction on starting chain B concentration in the presence of 3% (◇) and 10% (○) unsonicated seeds. The inset shows the same plot with sonicated seeds. Linear fits through the data extrapolate to intersect the  $x$  axis at a critical concentration of  $2.5 \pm 0.5 \mu\text{M}$ . The slopes of the two lines are 0.0139 (10% seed) and 0.0037 (3% seed)  $\mu\text{M}^{-1} \text{h}^{-1}$ . In panels A and B, the solid lines through the data are least-squares fits to eq 1 (for unseeded reactions) or eq 3 (for seeded reactions). Panel D shows the dependence of the elongation rate of the reaction on the starting seed concentration. In panels C and D, the straight lines through the data in each panel are linear fits. For all the plots, the error bars represent the spread in data determined from two or more independent experiments using two different batches of protein.

seed concentration even at the highest seed concentrations used. It appears that if an oligomeric conformation and not the monomer were the form to add onto the fibril ends, then there would be the need for a rapid equilibrium between it and monomeric chain B. It should be noted that AFM images of both sonicated and unsonicated fibrils show the former to be shorter in length but similar in diameter (Figure S1 of the Supporting Information). Figure S2 of the Supporting Information compares the length distributions of unsonicated and sonicated fibrils. Because both unsonicated and sonicated fibrils tend to be found in clumps, separate length distributions were determined for fibrils inside and outside clumps. In both cases, the sonicated fibrils were found to be shorter in length. Importantly, the length distributions of the sonicated fibrils, inside as well as outside clumps, were found not to change over a 24 h incubation under aggregation conditions. Hence, it appears that the fragments of chain B fibrils formed upon sonication cannot rejoin like fragments of fibrils formed by other proteins can.<sup>60</sup>

**Quantitative Analysis of the Commencement of the Aggregation Reaction.** Linear perturbation analysis<sup>22,56</sup> of a homogeneous nucleation-dependent polymerization reaction suggests that the initial part of the kinetic curve of aggregation should be able to be described by a  $t^2$  function or a  $\cos t$  function. Figure 6A shows that the initial 5% of the aggregation

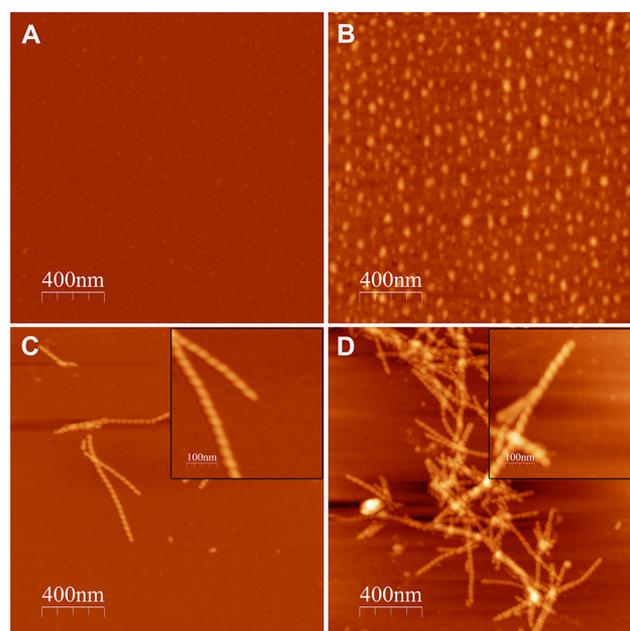
reaction at any protein concentration fits well to a  $\cos t$  function (eq 9). A  $t^2$  function works equally well (fits not shown).



**Figure 6.** Initial rates of amyloid fibril formation. Panel A shows the initial 5% of the kinetic curves of aggregation, monitored by ThT fluorescence, of 5 (black), 15 (blue), 50 (red), and 100 μM (green) chain B. The black solid lines are least-squares fits to eq 9. Panel B shows data plotted in accordance with eq 10. The straight line through the data points is a least-squares fit. The slope of line that represents the nucleus size,  $n^*$ , is  $2.13 \pm 0.15$ . Error bars represent the standard deviations from three independent experiments using three different batches of protein preparations.

Because the initial 5% of the kinetic curves of aggregation at all protein concentrations fit well to eq 9, the values of parameters A and B in eq 9 could be determined at each protein concentration. This in turn allowed the use of eq 10, obtained from perturbation analysis, to determine the size of the nucleus,  $n^*$  (Figure 6B). A value of 2 is obtained for  $n^*$ , in agreement with the value obtained from the dependence of  $t_{50}$  or  $t_{lag}$  on protein concentration (Figure 4). Hence, the dimer is the largest on-pathway oligomer that remains in equilibrium with the monomer.

**AFM Reveals the Presence of Off-Pathway Oligomers during the Course of Aggregation.** The observation that the kinetics of aggregation appeared to be marginally faster when measured by light scattering intensity than when measured by ThT fluorescence (Figure 1) suggested off-pathway oligomers might form transiently during the course of fibril formation. AFM was therefore used to follow the aggregation reaction. AFM images of aggregating protein were collected at  $t_5$ ,  $t_{20}$ ,  $t_{50}$ , and  $t_{100}$ , the times at which the aggregation reaction was 5, 20, 50, and 100% complete, respectively, as monitored by the ThT signal. At  $t_5$ , no oligomers were observed (Figure 7A); the mica surface coated with the 5% aggregation sample looked no different from the mica surface coated either with buffer or with a monomeric protein solution (not shown). Hence, no oligomers appear to form during the initial 5% of the reaction, for which the linear perturbation analysis was conducted (see above). At  $t_{20}$ , which is the end of the lag phase, only spherical oligomers are observed by AFM (Figure 7B). The diameter of the oligomers as determined from Gaussian fits to the distributions of their measured heights (not shown) is  $2.0 \pm 0.5$  nm. No fibrillar structures are seen at this time. At  $t_{30}$ , no oligomers were observed, and only very few fibrils were observed (data not shown). At  $t_{50}$ , fibrils could be observed, and again no spherical oligomers were seen (Figure 7C). Mostly, the fibrils were found to occur individually and not in clumps and had a height of  $5 \pm 2$  nm, as determined from Gaussian fits to the distributions of their measured heights (not shown), and lengths of 300–400

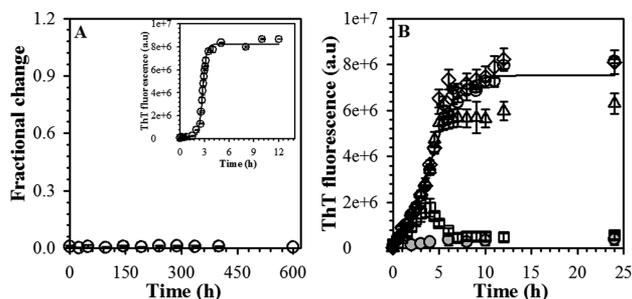


**Figure 7.** Structural characterization of the amyloid fibril formation reaction of 25 μM chain B. AFM images were obtained at 1 (A), 3 (B), 5 (C), and 24 h (D) of the aggregation reaction. The images are shown in the topography mode. In panels C and D, the insets show a magnified part of the main image.

nm. At  $t_{100}$ , fibrils were observed to be 1–2 μm in length and to have a height of  $7 \pm 2$  nm (Figure 7D).

Hence, the AFM images suggest that oligomers form during the initial phase of the aggregation reaction and disappear before the time the reaction is 30% complete. During the second half of the aggregation reaction, fibrils grow progressively in size. It appears that this growth must occur by monomer addition because no oligomers are seen during the growth phase. The observation that spherical oligomers disappear well before fibril growth occurs indicates that they are not on-pathway intermediates in fibril growth, because in that case the kinetics of their disappearance would have matched the kinetics of fibril growth. It appears, instead, that the spherical oligomers are off-pathway with respect to the pathway of fibril growth.

**Addition of Chain A to Aggregating Chain B Halts or Modulates Fibril Growth.** At the temperature (25 °C) at which chain B forms amyloid fibrils within 6 h, 25 μM dcMN does not form amyloid fibrils even over a period of 600 h (Figure 8A). dcMN does, however, form amyloid fibrils within 6 h at 60 °C (under otherwise identical aggregation conditions), which is the midpoint of the thermally induced unfolding–dissociation transition of the protein. Hence, it is likely that the amyloid fibrils formed by dcMN are actually fibrils formed by chain B after it dissociates from dcMN. To confirm this, fibrils were pelleted down by centrifugation at 14000 rpm for 45 min, and both the pellet and the supernatant were analyzed by mass spectrometry and tryptophan fluorescence (only chain B, not chain A, contains a tryptophan residue). It was found that the supernatant contained only chain A, while the pellet contained predominantly (>85%) chain B (data not shown). When 25 μM chain A was incubated in aggregation buffer, it was found to form amorphous insoluble aggregates, not ThT fluorescence-positive amyloid fibrils (data not shown). When, however, 25 μM chain A was added to



**Figure 8.** Effect of chain A on chain B aggregation. Panel A shows the ThT fluorescence-monitored aggregation of 25  $\mu\text{M}$  dcMN under the same aggregation conditions that were used to study the aggregation of chain B. No increase in ThT fluorescence is seen. The inset shows the ThT fluorescence-monitored kinetic curve of aggregation of 25  $\mu\text{M}$  dcMN at pH 3.0 and 60  $^{\circ}\text{C}$ . Panel B shows the ThT fluorescence-monitored kinetic trace of aggregation of 25  $\mu\text{M}$  chain B chain aggregation kinetics in the absence of any added chain A ( $\circ$ ), along with the kinetic traces obtained when chain A was added to a final concentration of 25  $\mu\text{M}$ , after aggregation for 0 min ( $\bullet$ ) and 3 ( $\square$ ), 5 ( $\triangle$ ), and 12 h ( $\diamond$ ). The solid line through the data is a least-squares fit to eq 1. Error bars represent standard deviations determined from three independent experiments using three different batches of protein preparations.

aggregation buffer containing 25  $\mu\text{M}$  chain B, the chains complemented fully to form dcMN,<sup>49</sup> as judged by far-UV CD spectroscopy (data not shown).

Hence, to determine the time at which the amyloid fibril reaction becomes irreversible, 25  $\mu\text{M}$  chain A was added to 25  $\mu\text{M}$  aggregating chain B at different times ( $t_5$ ,  $t_{20}$ ,  $t_{50}$ , and  $t_{100}$ ) during the fibril formation reaction, monitored by ThT fluorescence (Figure 8B). When chain A was added at  $t_5$ , no further increase in ThT fluorescence was observed. When chain A was added at  $t_{20}$ , the ThT fluorescence decreased within 8 h to the value observed at the start of the reaction. In both cases, the far-UV CD spectrum of the solution at 24 h indicated the presence of 23  $\mu\text{M}$  dcMN. When chain A was added at  $t_{50}$ , the ThT fluorescence increased to a value that is 20% lower than the value observed at  $t_{100}$  in the absence of chain A. When chain A was added at  $t_{100}$ , the ThT fluorescence was not found to change even after 48 h.

It appears therefore that at a stage ( $t_{20}$ ) when only the monomer and off-pathway oligomers of chain B are present, the addition of chain A leads to dissociation of any aggregate present via a thermodynamic coupling mechanism whereby chain A binds to chain B, thereby pulling the equilibrium between free chain B and its off-pathway aggregate in the direction of the former. In future experiments, it will be important to test such a thermodynamic coupling mechanism. At the other two time points,  $t_{50}$  and  $t_{100}$ , fibrils were present. Therefore, at these two time points, the addition of chain A does not decrease the ThT fluorescence; rather, it increases at  $t_{50}$ . The increase in ThT fluorescence at  $t_{50}$  can be due to either the conformational rearrangement in the fibrils or the association of fibrils.

## DISCUSSION

Previous studies had suggested that the amyloid fibrils formed by dcMN at elevated temperatures are not composed of both of its constituent chains A and B, but only of chain B released by the unfolding of this heterodimeric protein.<sup>50</sup> In this respect, monellin is similar to other multimeric proteins, including

transthyretin<sup>61</sup> and superoxide dismutase-1,<sup>62</sup> which are also known to form amyloid fibrils only after they are dissociated into their constituent subunits. For such multimeric proteins, including monellin, it becomes difficult to determine whether fibril formation commences from the folded, partially unfolded, or completely unfolded subunit. In the case of monellin, because its two constituent chains are different, it is possible to study the aggregation of each chain by itself in the absence of any competitive reaction leading to the formation of the native multimer. In this study, amyloid fibril formation by chain B has been studied at pH 3 and 25  $^{\circ}\text{C}$ . Under these conditions, the CD spectrum of chain B, before it starts aggregating, is identical to that at pH 7, where it has been shown not to be fully unfolded but to possess partial residual structure.<sup>48</sup> Hence, it appears that the aggregation of monellin commences from a partially folded conformation, just like it appears to do in the case of intrinsically disordered proteins,<sup>63</sup> including  $\alpha$ -synuclein.<sup>64</sup>

**Mechanism of Formation of Fibrils.** Fibril formation by chain B appears to occur with all three characteristic features of a NDP reaction.<sup>21,23,27</sup> An initial lag phase is seen during fibril formation, during which mature fibrils have yet to form (Figures 1 and 2). Fibril formation does not appear to occur below a critical concentration of  $\sim 4$   $\mu\text{M}$  (Figure 3). Seeding abolishes the lag phase (Figure 5). A NDP mechanism for fibril formation is supported by the observation that at the end of the fibril formation reaction, mature fibrils and monomers are the two predominant forms of the protein present and that the concentration of monomeric protein is approximately the same as the critical concentration for fibril formation. Moreover, the lag time decreases exponentially with an increase in protein concentration, and the initial rate of fibril formation can be described as a  $\cos t$  or  $t^2$  function (Figure 6) as expected for a NDP mechanism. It should be noted that although the NDP model has been applied to the study of the amyloid fibril reactions of many proteins,<sup>37,38,65–72</sup> very few aggregating systems meet all the criteria of homogeneous NDP.<sup>37</sup>

The observation that two different analyses based on the mechanism being described as homogeneous NDP indicate that the nucleus is a dimer (Figure 4) suggests that the NDP description of fibril formation by chain B is a robust one. One analysis is based on treatment of the entire kinetic curve of aggregation,<sup>55</sup> while the other is based on analysis of the initial rate of aggregation;<sup>56</sup> however, both analyses require the same assumptions. The assumption that the nucleus remains in equilibrium with the monomer seems reasonable given that the nucleus is a dimer. The assumption that the nucleus size remains a dimer even at high protein concentrations seems to be valid given that the kinetic curves at all protein concentrations collapse into a single curve upon phenomenological scaling (Figure 2), which suggests that the mechanism does not change over the range of protein concentrations studied. Previous studies using monellin and its subunit chain B had suggested that a prenucleus might be formed during the aggregation.<sup>50</sup> However, no prenucleus formation during chain B aggregation was observed under the conditions used in this study. The assumption that fibril growth occurs by monomer addition is reasonable because it is unlikely that multistranded fibrils such as those of chain B grow either by annealing of large aggregates because the translational and rotational diffusion of large aggregates would be very slow or by annealing of short fragments because the concentration of short polymers would be much lower than that of monomers.

**Fibril Growth Appears To Occur by Monomer Addition.** In both methods of analysis of the aggregation curves, utilized in this study, whether of the entire kinetic curve<sup>55</sup> or of only the initial part of the kinetic curve,<sup>22,56</sup> it was assumed that fibrils grow by addition of the monomer. This assumption is supported by the observation that the initial rate of fibril formation has a linear dependence on starting monomer concentration, when aggregation is conducted in the presence of a fixed concentration of seed, as well as by the observation of a linear dependence on seed concentration when aggregation is conducted in the presence of a fixed initial concentration of monomeric chain B.<sup>73</sup> Certainly, at the times at which fibril growth occurs, at times between  $t_{30}$  and  $t_{100}$  of fibril growth (Figure 7), no oligomers can be observed either by AFM imaging (Figure 7) or by sedimentation assays (data not shown). Nevertheless, it is difficult to rule out the possibility that monomeric chain B is in rapid equilibrium with a very low concentration of the aggregation-competent oligomer, and that it is this oligomer that adds to the growing fibril ends.

**Seeding Experiments Rule Out the Presence of Secondary Pathways for Fibril Formation.** When secondary pathways of fibril formation are operative because of the fragmentation of fibrils formed on the primary pathway, the addition of sonicated seeds leads to an acceleration of the polymerization rate greater than that caused by addition of unsonicated seed.<sup>22,38</sup> The observation in this study that sonicated and unsonicated seeds accelerate fibril formation to the same extent suggests therefore that secondary pathways are not operative for chain B aggregation. Such a conclusion is supported by the observation that the basic mechanism of fibril formation does not change over the entire range of protein concentration. The observation that sonicated seeds have the same effect as unsonicated seeds suggests that while sonication does lead to the creation of more fibril ends by causing existing fibrils to fragment, this fragmentation occurs in a manner that does not allow newly created fibril ends to act as fibril growth centers. It would appear that sonication leads to new fibril ends that are structurally distinct from the ends of the original fibrils.

**The Thermodynamic Nucleus Is a Dimer.** The estimation of nucleus size is always challenging and controversial in protein aggregation studies.<sup>74</sup> The early application of NDP models to the self-assembly of some proteins suggested that the thermodynamic nuclei are large oligomers because of the observation of a very large dependence of aggregation kinetics on protein concentration.<sup>21,75–77</sup> In contrast, amyloid fibril reactions appear to be characterized by small nuclei, from as small as a monomer<sup>67</sup> to dimers, trimers, and other very small oligomers.<sup>35,39,53,71,72,78–83</sup> It is possible that the nucleus size for amyloid fibril formation may be underestimated either because of the presence of prenucleic aggregates<sup>24</sup> or because of the common usage of continuous mechanical agitation to accelerate fibril growth,<sup>53,84</sup> but in our study, the aggregating protein solution was not continuously agitated.

It is important to note that, in this study, the theoretical treatment based on features of the entire kinetic curve yields the same equation (eq 4) relating nucleus size to  $t_{50}$ , regardless of whether off-pathway aggregates are present.<sup>55</sup> The observed aggregation kinetics possesses a characteristic feature of simple NDP reactions: an increasing protein concentration causes the reaction to reach completion faster (Figure 2). For NDP reactions accompanied by off-pathway aggregation, it is expected that the reaction will reach completion more slowly

when the protein concentration becomes very high.<sup>55</sup> Such behavior, which would manifest itself in  $t_{50}$  values first decreasing with an increasing protein concentration and then increasing at very high protein concentrations, is not observed in this study:  $t_{50}$  values are seen only to decrease with an increasing protein concentration (Figure 4), as expected for a simple NDP mechanism. It should be noted that although off-pathway aggregates are seen to form, they form only during the initial part of the aggregation curve, between  $t_5$  and  $t_{20}$  (see Results). In this time frame, the fibril formation reaction cannot be observed (see above): the first fibrils are seen only at  $t_{30}$ , a time at which off-pathway aggregates have disappeared completely. In this study, to confirm that the formation of off-pathway aggregates does not affect the determination of the nucleus size, the data were also analyzed in a completely different manner. In the perturbation analysis method,<sup>22</sup> only the initial 5% of the aggregation curve was analyzed until  $t_5$  when off-pathway oligomers cannot be observed (see Results). Again a dimeric size for the nucleus was obtained. Hence, although in other cases the presence of off-pathway aggregates could affect the determination of the size of the critical nucleus, this is not the case in our study.

It should also be noted that both methods used in this study for the analysis of the fibril formation reaction according to a NDP mechanism<sup>22,55</sup> consider the fibril formation reaction to be irreversible. In this study, the reversibility of aggregation kinetics was checked by adding chain A at different times during the aggregation process. It was found that when chain A was added before fibril formation can be detected (before  $t_{20}$ ), the aggregation reaction leading to the formation of off-pathway aggregates was completely reversible (Figure 8), but once fibril formation commences to a detectable extent (at approximately  $t_{30}$ ), it appears to be irreversible. Thus, when chain A was added at  $t_{50}$  when fibril formation was well underway, no disaggregation of fibrils occurred: the ThT fluorescence signal did not decrease. When chain A was added at  $t_{100}$ , when the fibril formation reaction was complete, the ThT fluorescence signal did not decrease even over 24 h (Figure 8). Hence, the fibril formation reaction itself appears to be completely irreversible, and only the off-pathway aggregates that form before the fibril formation reaction commences can revert back to the monomer.

**Oligomer Formation and Amyloid Fibril Formation.** AFM imaging clearly shows that spherical oligomers,  $\sim 2$  nm in diameter, are populated transiently very early during the fibril formation reaction (Figure 7). It is likely that it is the formation of these oligomers that results in the change in light scattering intensity preceding that in ThT fluorescence (Figure 1). Spherical oligomers and curvilinear protofibrils have been observed to form early during amyloid fibril formation by other proteins, too,<sup>34,85,86</sup> and in some cases, it has been suggested that mature fibrils are assembled from such soluble aggregates.<sup>33,34,87</sup> In the study presented here, the observation that the oligomers disappear before significant fibril growth has occurred (Figure 7) indicates, however, that these spherical oligomers are not on-pathway but are instead off the direct pathway to fibril formation. Such off-pathway aggregates have also been observed during the aggregation of tau protein.<sup>35</sup> Understanding the transformation of prefibrillar oligomers and protofibrils into mature fibrils<sup>30</sup> has become important because of the realization that oligomers and/or protofibrils, and not mature fibrils, may be the toxic species in amyloid-linked diseases,<sup>33,88–90</sup> and hence, the transformation of spherical

oligomers and protofibrils into insoluble fibrils may be a mechanism employed to sequester the toxic material. In the case of monellin, it appears that the off-pathway oligomers dissociate to free chain B, which then becomes available for fibril formation. At present, the fraction of the protein that forms the off-pathway oligomers is not known. If this fraction is significant, then the oligomeric pool of protein may act as a buffering system to maintain the concentration of monomers at a constant level up to the start of the elongation phase. The ease of dissociation of the off-pathway oligomers is indicated by the observation that complete dissociation of all oligomers present to free chain B is effected upon the addition of chain A at  $t_{20}$ : free chains A and B then bind to each other and refold to form monellin (Figure 8).

In conclusion, amyloid fibril formation by the chain B of monellin has been shown to meet all the criteria of a NDP mechanism. The nucleus has been shown to be a dimer by two different analyses, one of the entire aggregation curve and the other of only the initial part of the aggregation curve. Although spherical oligomers are also observed during the course of the aggregation reaction, they are shown to disappear well before fibril formation occurs to any appreciable extent; hence, their formation appears to be off-pathway to the main pathway for fibril formation.

## ■ ASSOCIATED CONTENT

### ● Supporting Information

Effect of sonication on chain B amyloid fibrils (Figures S1 and S2). This material is available free of charge via the Internet at <http://pubs.acs.org>.

## ■ AUTHOR INFORMATION

### Corresponding Author

\*National Centre for Biological Sciences, Tata Institute of Fundamental Research, Bangalore 560065, India. E-mail: jayant@ncbs.res.in. Phone: + 91-80-6717-6150.

### Funding

This work is funded by the Tata Institute of Fundamental Research and the Department of Biotechnology, Government of India. J.B.U. is a recipient of a J. C. Bose National Fellowship from the Government of India.

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

We thank Gayathri Ramachandran and Jogender Singh for their comments on the manuscript.

## ■ ABBREVIATIONS

NDP, nucleation-dependent polymerization; dcMN, double-chain monellin; ThT, thioflavin T; CD, circular dichroism; AFM, atomic force microscopy.

## ■ REFERENCES

- (1) Chiti, F., and Dobson, C. M. (2006) Protein Misfolding, Functional Amyloid, and Human Disease. *Annu. Rev. Biochem.* 75, 333–366.
- (2) Rochet, J. C., and Lansbury, P. T., Jr. (2000) Amyloid fibrillogenesis: Themes and variations. *Curr. Opin. Struct. Biol.* 10, 60–68.
- (3) Pertinhez, T. A., Bouchard, M., Tomlinson, E. J., Wain, R., Ferguson, S. J., Dobson, C. M., and Smith, L. J. (2001) Amyloid fibril formation by a helical cytochrome. *FEBS Lett.* 495, 184–186.

- (4) Fandrich, M., Fletcher, M. A., and Dobson, C. M. (2001) Amyloid fibrils from muscle myoglobin. *Nature* 410, 165–166.
- (5) Guijarro, J. I., Sunde, M., Jones, J. A., Campbell, I. D., and Dobson, C. M. (1998) Amyloid fibril formation by an SH3 domain. *Proc. Natl. Acad. Sci. U.S.A.* 95, 4224–4228.
- (6) Staniforth, R. A., Giannini, S., Higgins, L. D., Conroy, M. J., Hounslow, A. M., Jerala, R., Craven, C. J., and Waltho, J. P. (2001) Three-dimensional domain swapping in the folded and molten-globule states of cystatins, an amyloid-forming structural superfamily. *EMBO J.* 20, 4774–4781.
- (7) Konno, T., Murata, K., and Nagayama, K. (1999) Amyloid-like aggregates of a plant protein: A case of a sweet-tasting protein, monellin. *FEBS Lett.* 454, 122–126.
- (8) Irvine, G. B., El-Agnaf, O. M., Shankar, G. M., and Walsh, D. M. (2008) Protein aggregation in the brain: The molecular basis for Alzheimer's and Parkinson's diseases. *Mol. Med.* 14, 451–464.
- (9) Petkova, A. T., Leapman, R. D., Guo, Z., Yau, W. M., Mattson, M. P., and Tycko, R. (2005) Self-propagating, molecular-level polymorphism in Alzheimer's  $\beta$ -amyloid fibrils. *Science* 307, 262–265.
- (10) Jones, E. M., and Surewicz, W. K. (2005) Fibril conformation as the basis of species- and strain-dependent seeding specificity of mammalian prion amyloids. *Cell* 121, 63–72.
- (11) Fandrich, M., Meinhardt, J., and Grigorieff, N. (2009) Structural polymorphism of Alzheimer  $A\beta$  and other amyloid fibrils. *Prion* 3, 89–93.
- (12) Jahn, T. R., and Radford, S. E. (2008) Folding versus aggregation: Polypeptide conformations on competing pathways. *Arch. Biochem. Biophys.* 469, 100–117.
- (13) Nelson, R., Sawaya, M. R., Balbirnie, M., Madsen, A. O., Riekel, C., Grothe, R., and Eisenberg, D. (2005) Structure of the cross- $\beta$  spine of amyloid-like fibrils. *Nature* 435, 773–778.
- (14) Wiltzius, J. J., Landau, M., Nelson, R., Sawaya, M. R., Apostol, M. I., Goldschmidt, L., Soriaga, A. B., Cascio, D., Rajashankar, K., and Eisenberg, D. (2009) Molecular mechanisms for protein-encoded inheritance. *Nat. Struct. Mol. Biol.* 16, 973–978.
- (15) Sawaya, M. R., Sambashivan, S., Nelson, R., Ivanova, M. I., Sievers, S. A., Apostol, M. I., Thompson, M. J., Balbirnie, M., Wiltzius, J. J., McFarlane, H. T., Madsen, A. O., Riekel, C., and Eisenberg, D. (2007) Atomic structures of amyloid cross- $\beta$  spines reveal varied steric zippers. *Nature* 447, 453–457.
- (16) Narhi, L., Wood, S. J., Steavenson, S., Jiang, Y., Wu, G. M., Anafi, D., Kaufman, S. A., Martin, F., Sitney, K., Denis, P., Louis, J. C., Wypych, J., Biere, A. L., and Citron, M. (1999) Both familial Parkinson's disease mutations accelerate  $\alpha$ -synuclein aggregation. *J. Biol. Chem.* 274, 9843–9846.
- (17) Selkoe, D. J. (2003) Folding proteins in fatal ways. *Nature* 426, 900–904.
- (18) Chien, P., Weissman, J. S., and DePace, A. H. (2004) Emerging principles of conformation-based prion inheritance. *Annu. Rev. Biochem.* 73, 617–656.
- (19) Chapman, M. R., Robinson, L. S., Pinkner, J. S., Roth, R., Heuser, J., Hammar, M., Normark, S., and Hultgren, S. J. (2002) Role of *Escherichia coli* curli operons in directing amyloid fiber formation. *Science* 295, 851–855.
- (20) Hamada, D., Yanagihara, I., and Tsumoto, K. (2004) Engineering amyloidogenicity towards the development of nanofibrillar materials. *Trends Biotechnol.* 22, 93–97.
- (21) Oosawa, F., and Kasai, M. (1962) A theory of linear and helical aggregations of macromolecules. *J. Mol. Biol.* 4, 10–21.
- (22) Bishop, M. F., and Ferrone, F. A. (1984) Kinetics of nucleation-controlled polymerization. A perturbation treatment for use with a secondary pathway. *Biophys. J.* 46, 631–644.
- (23) Goldstein, R. F., and Stryer, L. (1986) Cooperative polymerization reactions. Analytical approximations, numerical examples, and experimental strategy. *Biophys. J.* 50, 583–599.
- (24) Powers, E. T., and Powers, D. L. (2006) The Kinetics of Nucleated Polymerizations at High Concentrations: Amyloid Fibril Formation Near and Above the "Supercritical Concentration". *Biophys. J.* 91, 122–132.

- (25) Kumar, S., and Udgaonkar, J. B. (2010) Mechanisms of amyloid fibril formation by proteins. *Curr. Sci.* 98, 639–656.
- (26) Asakura, S., and Oosawa, F. (1958) Interaction between particles suspended in solutions of macromolecules. *J. Polym. Sci.* 33, 183–192.
- (27) Frieden, C. (2007) Protein aggregation processes: In search of the mechanism. *Protein Sci.* 16, 2334–2344.
- (28) Kumar, S., Mohanty, S. K., and Udgaonkar, J. B. (2007) Mechanism of formation of amyloid protofibrils of barstar from soluble oligomers: Evidence for multiple steps and lateral association coupled to conformational conversion. *J. Mol. Biol.* 367, 1186–1204.
- (29) Jain, S., and Udgaonkar, J. B. (2008) Evidence for stepwise formation of amyloid fibrils by the mouse prion protein. *J. Mol. Biol.* 382, 1228–1241.
- (30) Singh, J., Sabareesan, A. T., Mathew, M. K., and Udgaonkar, J. B. (2012) Development of the Structural Core and of Conformational Heterogeneity during the Conversion of Oligomers of the Mouse Prion Protein to Worm-like Amyloid Fibrils. *J. Mol. Biol.* 423, 217–231.
- (31) Zhai, J., Lee, T.-H., Small, D. H., and Aguilar, M.-I. (2012) Characterization of Early Stage Intermediates in the Nucleation Phase of A $\beta$  Aggregation. *Biochemistry* 51, 1070–1078.
- (32) Jain, S., and Udgaonkar, J. B. (2011) Defining the pathway of worm-like amyloid fibril formation by the mouse prion protein by delineation of the productive and unproductive oligomerization reactions. *Biochemistry* 50, 1153–1161.
- (33) Harper, J. D., Lieber, C. M., and Lansbury, P. T., Jr. (1997) Atomic force microscopic imaging of seeded fibril formation and fibril branching by the Alzheimer's disease amyloid- $\beta$  protein. *Chem. Biol.* 4, 951–959.
- (34) Poirier, M. A., Li, H., Macosko, J., Cai, S., Amzel, M., and Ross, C. A. (2002) Huntingtin spheroids and protofibrils as precursors in polyglutamine fibrillogenesis. *J. Biol. Chem.* 277, 41032–41037.
- (35) Ramachandran, G., and Udgaonkar, J. B. (2011) Understanding the kinetic roles of the inducer heparin and of rod-like protofibrils during amyloid fibril formation by tau protein. *J. Biol. Chem.* 286, 38948–38959.
- (36) Giehm, L., and Otzen, D. E. (2011) Strategies to increase the reproducibility of protein fibrillization in plate reader assays. *Anal. Biochem.* 400, 270–281.
- (37) Hellstrand, E., Boland, B., Walsh, D. M., and Linse, S. (2010) Amyloid  $\beta$ -Protein Aggregation Produces Highly Reproducible Kinetic Data and Occurs by a Two-Phase Process. *ACS Chem. Neurosci.* 1, 13–18.
- (38) Ramachandran, G., and Udgaonkar, J. B. (2012) Evidence for the Existence of a Secondary Pathway for Fibril Growth during the Aggregation of Tau. *J. Mol. Biol.* 421, 296–314.
- (39) Cellmer, T., Douma, R., Huebner, A., Prausnitz, J., and Blanch, H. (2007) Kinetic studies of protein L aggregation and disaggregation. *Biophys. Chem.* 125, 350–359.
- (40) Congdon, E. E., Kim, S., Bonchak, J., Songrug, T., Matzavinos, A., and Kuret, J. (2008) Nucleation-dependent Tau Filament Formation: The importance of dimerization and an estimation of elementary rate constants. *J. Biol. Chem.* 283, 13806–13816.
- (41) Morris, J. A., and Cagan, R. H. (1972) Purification of monellin, the sweet principle of *Dioscoreophyllum cumminsii*. *Biochim. Biophys. Acta* 261, 114–122.
- (42) Turk, V., Stoka, V., and Turk, D. (2008) Cystatins: Biochemical and structural properties, and medical relevance. *Front. Biosci.* 13, 5406–5420.
- (43) Esposito, V., Guglielmi, F., Martin, S. R., Pauwels, K., Pastore, A., Piccoli, R., and Temussi, P. A. (2010) Aggregation Mechanisms of Cystatins: A Comparative Study of Monellin and Oryzacystatin. *Biochemistry* 49, 2805–2810.
- (44) Kimura, T., Uzawa, T., Ishimori, K., Morishima, I., Takahashi, S., Konno, T., Akiyama, S., and Fujisawa, T. (2005) Specific collapse followed by slow hydrogen-bond formation of  $\beta$ -sheet in the folding of single-chain monellin. *Proc. Natl. Acad. Sci. U.S.A.* 102, 2748–2753.
- (45) Patra, A. K., and Udgaonkar, J. B. (2007) Characterization of the folding and unfolding reactions of single-chain monellin: Evidence for multiple intermediates and competing pathways. *Biochemistry* 46, 11727–11743.
- (46) Patra, A. K., and Udgaonkar, J. B. (2009) GroEL can unfold late intermediates populated on the folding pathways of monellin. *J. Mol. Biol.* 389, 759–775.
- (47) Jha, S. K., Dhar, D., Krishnamoorthy, G., and Udgaonkar, J. B. (2009) Continuous dissolution of structure during the unfolding of a small protein. *Proc. Natl. Acad. Sci. U.S.A.* 106, 11113–11118.
- (48) Aghera, N., Earanna, N., and Udgaonkar, J. B. (2011) Equilibrium Unfolding Studies of Monellin: The Double-Chain Variant Appears To Be More Stable Than the Single-Chain Variant. *Biochemistry* 50, 2434–2444.
- (49) Aghera, N., and Udgaonkar, J. B. (2012) Kinetic studies of the folding of heterodimeric monellin: Evidence for switching between alternative parallel pathways. *J. Mol. Biol.* 420, 235–250.
- (50) Konno, T. (2001) Multistep nucleus formation and a separate subunit contribution of the amyloidogenesis of heat-denatured monellin. *Protein Sci.* 10, 2093–2101.
- (51) Aghera, N., and Udgaonkar, J. B. (2011) Heterologous expression, purification and characterization of heterodimeric monellin. *Protein Expression Purif.* 76, 248–253.
- (52) Horcas, I., Fernandez, R., Gomez-Rodriguez, J. M., Colchero, J., Gomez-Herrero, J., and Baro, A. M. (2007) WSXM: A software for scanning probe microscopy and a tool for nanotechnology. *Rev. Sci. Instrum.* 78, 013705–013708.
- (53) Nielsen, L., Khurana, R., Coats, A., Frokjaer, S., Brange, J., Vyas, S., Uversky, V. N., and Fink, A. L. (2001) Effect of environmental factors on the kinetics of insulin fibril formation: Elucidation of the molecular mechanism. *Biochemistry* 40, 6036–6046.
- (54) Flyvbjerg, H., Jobs, E., and Leibler, S. (1996) Kinetics of self-assembling microtubules: An “inverse problem” in biochemistry. *Proc. Natl. Acad. Sci. U.S.A.* 93, 5975–5979.
- (55) Powers, E. T., and Powers, D. L. (2008) Mechanisms of Protein Fibril Formation: Nucleated Polymerization with Competing Off-Pathway Aggregation. *Biophys. J.* 94, 379–391.
- (56) Ferrone, F. (1999) Analysis of protein aggregation kinetics. In *Methods in Enzymology*, pp 256–274, Academic Press, San Diego.
- (57) Pollan, M. S., Markiewicz, P., Bergeron, C., and Goh, M. C. (1994) Twisted ribbon structure of paired helical filaments revealed by atomic force microscopy. *Am. J. Pathol.* 144, 869–873.
- (58) Wegner, A., and Engel, J. (1975) Kinetics of the cooperative association of actin to actin filaments. *Biophys. Chem.* 3, 215–225.
- (59) Erickson, H. P., and Pantaloni, D. (1981) The role of subunit entropy in cooperative assembly. Nucleation of microtubules and other two-dimensional polymers. *Biophys. J.* 34, 293–309.
- (60) Binger, K. J., Pham, C. L., Wilson, L. M., Bailey, M. F., Lawrence, L. J., Schuck, P., and Howlett, G. J. (2008) Apolipoprotein C-II amyloid fibrils assemble via a reversible pathway that includes fibril breaking and rejoining. *J. Mol. Biol.* 376, 1116–1129.
- (61) Colon, W., and Kelly, J. W. (1992) Partial denaturation of transthyretin is sufficient for amyloid fibril formation in vitro. *Biochemistry* 31, 8654–8660.
- (62) Ray, S. S., Nowak, R. J., Brown, R. H., Jr., and Lansbury, P. T., Jr. (2005) Small-molecule-mediated stabilization of familial amyotrophic lateral sclerosis-linked superoxide dismutase mutants against unfolding and aggregation. *Proc. Natl. Acad. Sci. U.S.A.* 102, 3639–3644.
- (63) Speed, M. A., Wang, D. I., and King, J. (1996) Specific aggregation of partially folded polypeptide chains: The molecular basis of inclusion body composition. *Nat. Biotechnol.* 14, 1283–1287.
- (64) Uversky, V. N., Li, J., and Fink, A. L. (2001) Evidence for a partially folded intermediate in  $\alpha$ -synuclein fibril formation. *J. Biol. Chem.* 276, 10737–10744.
- (65) Bhak, G., Choe, Y. J., and Paik, S. R. (2009) Mechanism of amyloidogenesis: Nucleation-dependent fibrillation versus double-concerted fibrillation. *BMB Rep.* 42, 541–551.
- (66) Bhattacharyya, A. M., Thakur, A. K., and Wetzel, R. (2005) Polyglutamine aggregation nucleation: Thermodynamics of a highly

unfavorable protein folding reaction. *Proc. Natl. Acad. Sci. U.S.A.* 102, 15400–15405.

(67) Chen, S., Ferrone, F. A., and Wetzel, R. (2002) Huntington's disease age-of-onset linked to polyglutamine aggregation nucleation. *Proc. Natl. Acad. Sci. U.S.A.* 99, 11884–11889.

(68) Fandrich, M. (2007) Absolute Correlation between Lag Time and Growth Rate in the Spontaneous Formation of Several Amyloid-like Aggregates and Fibrils. *J. Mol. Biol.* 365, 1266–1270.

(69) Fay, N., Inoue, Y., Bousset, L., Taguchi, H., and Melki, R. (2003) Assembly of the Yeast Prion Ure2p into Protein Fibrils. *J. Biol. Chem.* 278, 30199–30205.

(70) Hills, R. D., Jr., and Brooks, C. L., III (2007) Hydrophobic Cooperativity as a Mechanism for Amyloid Nucleation. *J. Mol. Biol.* 368, 894–901.

(71) Ignatova, Z., and Gierasch, L. M. (2005) Aggregation of a Slow-Folding Mutant of a  $\beta$ -Clam Protein Proceeds through a Monomeric Nucleus. *Biochemistry* 44, 7266–7274.

(72) Xue, W.-F., Homans, S. W., and Radford, S. E. (2008) Systematic analysis of nucleation-dependent polymerization reveals new insights into the mechanism of amyloid self-assembly. *Proc. Natl. Acad. Sci. U.S.A.* 105, 8926–8931.

(73) Naiki, H., and Gejyo, F. (1999) Kinetic analysis of amyloid fibril formation. *Methods Enzymol.* 309, 305–318.

(74) Vitalis, A., and Pappu, R. V. (2011) Assessing the contribution of heterogeneous distributions of oligomers to aggregation mechanisms of polyglutamine peptides. *Biophys. Chem.* 159, 14–23.

(75) Ferrone, F. A., Hofrichter, J., Sunshine, H. R., and Eaton, W. A. (1980) Kinetic studies on photolysis-induced gelation of sickle cell hemoglobin suggest a new mechanism. *Biophys. J.* 32, 361–380.

(76) Goddette, D. W., and Frieden, C. (1986) Actin polymerization. The mechanism of action of cytochalasin D. *J. Biol. Chem.* 261, 15974–15980.

(77) Tobacman, L. S., and Korn, E. D. (1983) The kinetics of actin nucleation and polymerization. *J. Biol. Chem.* 258, 3207–3214.

(78) Lomakin, A., Chung, D. S., Benedek, G. B., Kirschner, D. A., and Teplow, D. B. (1996) On the nucleation and growth of amyloid  $\beta$ -protein fibrils: Detection of nuclei and quantitation of rate constants. *Proc. Natl. Acad. Sci. U.S.A.* 93, 1125–1129.

(79) Friedhoff, P., von Bergen, M., Mandelkow, E. M., Davies, P., and Mandelkow, E. (1998) A nucleated assembly mechanism of Alzheimer paired helical filaments. *Proc. Natl. Acad. Sci. U.S.A.* 95, 15712–15717.

(80) Wood, S. J., Wypych, J., Steavenson, S., Louis, J.-C., Citron, M., and Biere, A. L. (1999)  $\alpha$ -Synuclein Fibrillogenesis Is Nucleation-dependent. *J. Biol. Chem.* 274, 19509–19512.

(81) Frankenfield, K. N., Powers, E. T., and Kelly, J. W. (2005) Influence of the N-terminal domain on the aggregation properties of the prion protein. *Protein Sci.* 14, 2154–2166.

(82) Ferrone, F. A., Kheterpal, I., and Ronald, W. (2006) Nucleation: The Connections Between Equilibrium and Kinetic Behavior. In *Methods in Enzymology*, pp 285–299, Academic Press, San Diego.

(83) Cohen, S. I., Vendruscolo, M., Welland, M. E., Dobson, C. M., Terentjev, E. M., and Knowles, T. P. (2011) Nucleated polymerization with secondary pathways. I. Time evolution of the principal moments. *J. Chem. Phys.* 135, 065105.

(84) Collins, S. R., Douglash, A., Vale, R. D., and Weissman, J. S. (2004) Mechanism of prion propagation: Amyloid growth occurs by monomer addition. *PLoS Biol.* 2, e321.

(85) Walsh, D. M., Hartley, D. M., Kusumoto, Y., Fezoui, Y., Condron, M. M., Lomakin, A., Benedek, G. B., Selkoe, D. J., and Teplow, D. B. (1999) Amyloid  $\beta$ -protein fibrillogenesis. Structure and biological activity of protofibrillar intermediates. *J. Biol. Chem.* 274, 25945–25952.

(86) Hoyer, W., Antony, T., Cherny, D., Heim, G., Jovin, T. M., and Subramaniam, V. (2002) Dependence of  $\alpha$ -synuclein aggregate morphology on solution conditions. *J. Mol. Biol.* 322, 383–393.

(87) Goldsburly, C., Frey, P., Olivieri, V., Aebi, U., and Muller, S. A. (2005) Multiple assembly pathways underlie amyloid- $\beta$  fibril polymorphisms. *J. Mol. Biol.* 352, 282–298.

(88) Conway, K. A., Lee, S. J., Rochet, J. C., Ding, T. T., Harper, J. D., Williamson, R. E., and Lansbury, P. T., Jr. (2000) Accelerated oligomerization by Parkinson's disease linked  $\alpha$ -synuclein mutants. *Ann. N.Y. Acad. Sci.* 920, 42–45.

(89) Lashuel, H. A., Hartley, D., Petre, B. M., Walz, T., and Lansbury, P. T., Jr. (2002) Neurodegenerative disease: Amyloid pores from pathogenic mutations. *Nature* 418, 291.

(90) Volles, M. J., and Lansbury, P. T., Jr. (2002) Vesicle permeabilization by protofibrillar  $\alpha$ -synuclein is sensitive to Parkinson's disease-linked mutations and occurs by a pore-like mechanism. *Biochemistry* 41, 4595–4602.