JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

pubs.acs.org/JACS

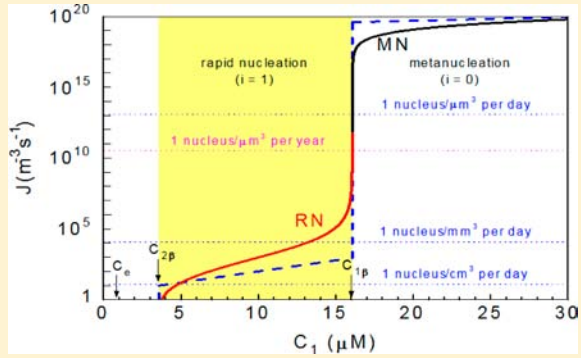
Confounding the Paradigm: Peculiarities of Amyloid Fibril Nucleation

Dimo Kashchiev,[†] Raffaela Cabriolu,[‡] and Stefan Auer*[,‡]

[†]Institute of Physical Chemistry, Bulgarian Academy of Sciences, ul. Acad. G. Bonchev 11, Sofia 1113, Bulgaria

[‡]Centre for Molecular Nanoscience, University of Leeds, Leeds LS2 9JT, United Kingdom

ABSTRACT: Fibrils of amyloid proteins are currently of great interest because of their involvement in various amyloid-related diseases and nanotechnological products. In a recent kinetic Monte Carlo simulation study (Cabriolu, R.; Kashchiev, D.; Auer, S. *J. Chem. Phys.* **2012**, *137*, 204903), we found that our simulation data for the rate of amyloid fibril nucleation occurring by direct polymerization of monomeric protein could not be described adequately by nucleation theory. It turned out that the process occurred in a peculiar way, thus confounding the nucleation paradigm and demanding a new theoretical treatment. In the present study, we reconsider the theoretical approach to nucleation of amyloid fibrils and derive new expressions for the stationary rate of the process. As these expressions provide a remarkably good description of the simulation data, by using them we propose a theoretical dependence of the amyloid-β_{40} fibril nucleation rate on the concentration of monomeric protein in the solution. This dependence reveals the existence of a threshold concentration below which the fibril nucleation in small enough solution volumes is practically arrested, and above which the process occurs vigorously, because then each monomeric protein in the solution acts as fibril nucleus. The presented expressions for the threshold concentration and for the dependence of the fibril nucleation rate on the concentration of monomeric protein can be a valuable guide in designing new therapeutic and/or technological strategies for prevention or stimulation of amyloid fibril formation.



INTRODUCTION

Amyloid fibrils are involved in dozens of amyloid-related diseases such as Alzheimer's, Parkinson's, type II diabetes, and cataract.[1] The nucleation of such fibrils refers to the process of random generation of those nanoscale fibrils (or protofilaments) that have the ability of irreversible growth. Unless the nanofibril size exceeds the size n^* of the fibril nucleus, the nanofibril is more likely to dissolve than to grow (n^* is the number of protein monomers in the fibril nucleus, known also as critical nucleus, which is that particular nanofibril that requires maximum work for its formation). As only a fibril supernucleus, i.e., a nanofibril bigger than the fibril nucleus, can grow irreversibly into a macroscale amyloid fibril, the fibril nucleation rate J is defined as the number of fibril supernuclei that form per unit time and per unit volume of the protein solution. Finding J is a central problem in nucleation of amyloid fibrils, because J has a strong impact on many important quantities, such as the fibril size distribution, the number of fibrils formed, and the lag time of the fibrillation process. Also, J is an indispensable ingredient in the set of rate equations describing the overall kinetics of unseeded amyloid fibrillation (e.g., refs 2−11).

In classical nucleation theory (CNT) (e.g., refs 12, 13), at a given supersaturation the nucleus size n^* has a unique value obtainable from thermodynamic considerations about the work gained from assembling n^* monomers of the supersaturated parent phase into the nucleus and the work spent on creating the nucleus surface, periphery, or ends for three-dimensional (3D), two-dimensional (2D), or one-dimensional (1D) nuclei, respectively. The latter work scales as $n^{*2/3}$ for 3D nuclei, as $n^{*1/2}$ for 2D nuclei, and is n^*-independent for 1D nuclei.[13] Due to this independence, regardless of the supersaturation, $n^* = 1$ for the appearance of 1D aggregates, the energy barrier to the process is nil, and, for that reason, 1D nucleation does not exist (ref 13, p 42). Hence, nucleation cannot be a process involved in the formation of amyloid fibrils if they are strictly 1D aggregates. Contrary to that, however, numerous experiments provide evidence for nucleation-mediated formation of the amyloid fibrils. This evidence can be accommodated within CNT when the amyloid fibril nucleus is treated as a 2D aggregate, because then $n^* \geq 1$ and there exists an energy barrier to nucleation. Yet, considering the amyloid fibril nucleus as a 2D aggregate is not entirely justified, because the amyloid fibrils are more prone to grow in one dimension (i.e., to elongate) than in two dimensions (i.e., to simultaneously elongate and thicken). Thus, as the amyloid fibril nuclei are actually aggregates with ambiguous dimensionality, the work done on creating the nucleus/solution interface does not scale in a definite way with the nucleus size. As a consequence, at a given supersaturation the size n^* of the fibril nucleus does not have a unique value and amyloid nanofibrils of different size and shape can act as fibril nuclei. This peculiarity of the amyloid fibril nucleation makes it impossible for the process to be fully

Received: November 15, 2012
Published: January 10, 2013

   dx.doi.org/10.1021/ja311228d | *J. Am. Chem. Soc.* 2013, 135, 1531−1539

described by CNT or related nucleation theories which we shall collectively call the standard nucleation theory.

At a fixed temperature, the nucleation rate J can be varied by changing the concentration C_1 of monomeric protein in the solution. Building on concepts from CNT,[12,13] various expressions for the $J(C_1)$ dependence have been proposed and used in the literature (e.g., refs 2−11 and 14−16), but they can be unitedly represented by the formula $J = aC_1^b$. While the factor a depends on several quantities among which the frequency of attachment of monomeric protein to the nanofibrils, the power b (a positive number) is only related to the nucleus size n^*. For example, $b = n^* - 1$ (ref 14) or $b = n^*$ (refs 7−9) or $b = n^* + 1$ (refs 2−6, 10, 15, 16). The different theories consider the factor a as C_1-dependent or independent. Similarly, the power b is treated as changing or fixed when C_1 is varied. What is important for the present study, however, is that according to all theories ln J is a concavely or linearly increasing function of ln C_1. The latter is readily seen from the above $J(C_1)$ formula at C_1-independent a and b.

It seems that, hitherto, there has been no quantitative verification of theoretical $J(C_1)$ dependences with the help of directly obtained experimental $J(C_1)$ data for amyloid fibrils. A major reason for that are the considerable experimental difficulties in the direct determination of J as a function of C_1. Information about the change of J with C_1 is usually obtained only indirectly from the effect of C_1 on the initial course of the fibrillation progress curve or on the lag time of the overall process of fibrillation. In such a situation, computer simulations prove very useful in gaining insight into various aspects of the nucleation of amyloid fibrils. Recently,[17] by means of kinetic Monte Carlo simulations, we have directly obtained $J(s)$ data for nucleation of crystals with strongly anisotropic interactions between molecules (s is the dimensionless supersaturation of the nucleating system). These data in fact represent the $J(s)$ dependence for nucleation of amyloid fibrils with solubility C_e when the process occurs by direct polymerization of practically fully extended protein segments, i.e., β-strands, because then s is given by $s = \ln(C_1/C_e)$ and because the fibrils can be modeled as crystal-like aggregates of β-strands with strong hydrogen bonds along the fibril elongation axis and weak hydrophobicity-mediated bonds perpendicular to this axis.[15] Surprisingly, our simulations[17] revealed that at the strong interaction anisotropy that characterizes amyloid fibrils the dependence of ln J on s (i.e., on C_1) is steplike, with sharp jumps at certain s values. Such a change of J with s (or C_1) is paradoxical from the viewpoint of CNT[12,13] and impossible to describe by the $J(C_1)$ formula given above. This behavior of J is thus a clear sign that the nucleation of amyloid fibrils is a peculiar kind of nucleation, not entirely complying with the paradigm of standard nucleation theory.

Evidently, our finding that ln J is a steplike function of s in nucleation of amyloid fibrils calls for a reconsideration of the theoretical description of amyloid fibril nucleation presented elsewhere.[15,16] This reconsideration and the derivation of a $J(s)$ formula describing our simulation $J(s)$ data are the objective of the present study. The results obtained are applicable to homogeneous nucleation which occurs when the protein solution is free of foreign nucleation-active particles or is sufficiently strongly supersaturated.

RESULTS

Concentration Ranges. Both theoretical considerations[15] and a computer-simulated peptide solubility diagram[18,19] reveal that for the irreversible elongation of differently thick amyloid fibrils thermodynamics requires different ranges of the concentration C_1 of monomeric β-strands (peptides or protein segments) in the solution. Figure 1 illustrates schematically

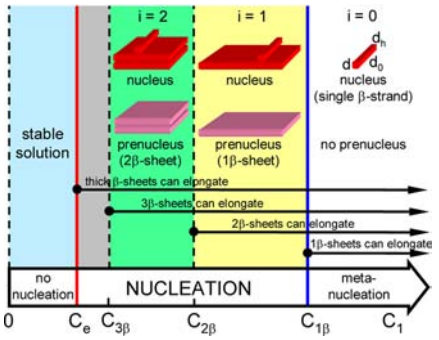


Figure 1. Protein concentration ranges in which the protein solution is in stable ($0 \le C_1 \le C_e$) or metastable ($C_1 > C_e$) thermodynamic equilibrium at a given temperature. In the metanucleation range ($i = 0$) the nucleus is a single β-strand. In the first ($i = 1$) and the second ($i = 2$) nucleation ranges the nucleus is a 1β- or 2β-sheet, respectively, with one β-strand attached sidewise.

these ranges at a fixed absolute temperature T at which the β-strands are in practically fully extended conformation. These ranges are limited by the equilibrium concentration (or solubility) C_e of the bulk fibrillar phase and the increasingly higher equilibrium concentrations (or solubilities) $C_{1\beta}$, $C_{2\beta}$, $C_{3\beta}$, etc., of the fibrils constituted of one, two, three, etc., equally long β-sheets, respectively (these fibrils are the so-called $i\beta$-sheets[15,18]). The solubilities are merely the C_1 values at which the respective $i\beta$-sheets neither lengthen nor dissolve. The $i\beta$-sheet solubility $C_{i\beta}$ is related to C_e by the expression[15] ($i = 1, 2, 3, \ldots$)

$$C_{i\beta} = C_e e^{2d_0 d_h \sigma_h / ikT} \quad (1)$$

where d_0 is the length of a β-strand in the β-sheet, i.e., the β-sheet width (Figure 1), d_h is the distance occupied by a β-strand along a β-sheet (Figure 1), σ_h (J/m^2) is the specific surface energy of the β-sheet sides, i is the number of β-sheets in the $i\beta$-sheet, and k is the Boltzmann constant. For example, for the amyloid-β_{40} (Aβ_{40}) protein $d_0 \approx 5$ nm (ref 20), $d_h \approx 0.5$ nm (ref 20), and $\sigma_h \approx 2.6$ J/m^2 (estimated value[21]).

As indicated in Figure 1, the $C_1 > C_{1\beta}$ range (range $i = 0$ in the figure) corresponds to metanucleation,[15] a process of fibril formation without energy barrier, because then each protein monomer (i.e., single β-strand) in the solution acts as fibril nucleus: attachment of another monomer to it allows irreversible elongation of the so-formed 1β-sheet with length of two β-strands. Similarly, when $C_1 > C_{2\beta}$, the 2β-sheets can lengthen irreversibly. Importantly, however, in the $C_{2\beta} < C_1 < C_{1\beta}$ range (range $i = 1$ in Figure 1), the 1β-sheets tend to dissolve and their appearance is due to fluctuations. In this range the fibril nucleus is a 1β-sheet plus one β-strand attached to one of the two 1β-sheet sides so that a fibril prenucleus is any of the randomly formed, different length 1β-sheets in the solution (see Figure 1). The situation is analogous with the 3β-sheets when $C_1 > C_{3\beta}$, because then these sheets can elongate

dx.doi.org/10.1021/ja311228d | J. Am. Chem. Soc. 2013, 135, 1531−1539

irreversibly. However, in the $C_{3\beta} < C_1 < C_{2\beta}$ range (range $i = 2$ in Figure 1), the 2β-sheets are subject to dissolution and can only appear by fluctuations. Hence, in this range the fibril nucleus is a 2β-sheet with one β-strand attached sidewise, and the respective prenucleus is a 2β-sheet of any length (Figure 1).

The general rule is, therefore, that in the ith supersaturation range, defined by ($i = 0, 1, 2, ...$)

$$2/(i+1) < s/\psi_h < 2/i \tag{2}$$

all different length $i\beta$-sheets are fibril prenuclei, and these sheets plus one β-strand attached to one of their two sides are the nuclei of the $(i+1)\beta$-sheet-thick fibrils that can lengthen irreversibly. With $i = 0$, inequality 2 is the condition for metanucleation: then prenuclei (0β-sheets) do not exist, and the single β-strands are the fibril nuclei. Thus, $i = 0$ and $i \geq 1$ correspond to fibril metanucleation and nucleation, respectively. Inequality 2 is obtained by using eq 1 to replace $C_{i\beta}$ by C_e, and in it the dimensionless supersaturation s and the dimensionless specific surface energy ψ_h of the β-sheet sides are given by[15,16]

$$s = \ln(C_1/C_e) \tag{3}$$

$$\psi_h = d_0 d_h \sigma_h / kT = E_h / 2kT \tag{4}$$

The second equality in eq 4 results from using the approximate relation[16] $\sigma_h = E_h/2d_0d_h$ between σ_h and the weak (hydrophobicity-mediated) binding energy E_h of two nearest-neighbor β-strands positioned in adjacent β-sheets.

Nucleation Work. For the analysis to follow, it is necessary to know the work $W(i, n)$ to form an n-sized nanofibril which at a given transition size n_t changes its shape from that of $i\beta$-sheet to that of $i\beta$-sheet with a partially or fully built-up subsequent β-sheet on one of its two sides. Let the number n of β-strands constituting the nanofibril be in the range $1 \leq n \leq n_t + m_t$, where $m_t = n_t/i$ is the nanofibril transition length (m_t is the number of β-strands in any of the i equally long β-sheets constituting the $i\beta$-sheet-shaped nanofibril when the first β-strand of the subsequent β-sheet is attached to the nanofibril). Then the dimensionless work w for nanofibril formation can be written down as ($i = 1, 2, 3, ...$)

$$w(i, n) = -(s - 2\psi_h/i)n + 2\psi i \tag{5}$$

for $n = 1, 2, ..., n_t$ (ref 15) and as

$$w(i, n) = -sn + 2\psi_h m_t + 2\psi(i+1) \tag{6}$$

for $n = n_t + 1, n_t + 2, ..., n_t + m_t$. Here, $w(i,n) \equiv W(i,n)/kT$ and the dimensionless specific surface energy ψ of the $i\beta$-sheet ends is given by[15,16]

$$\psi = d_0 d\sigma/kT = E/2kT \tag{7}$$

In this expression d is the thickness of a single β-sheet (Figure 1), σ (J/m^2) is the dimensional specific surface energy of the $i\beta$-sheet ends, and E is the energy of the strong (primarily hydrogen) bond between two nearest-neighbor β-strands in a β-sheet (the second equality in eq 7 follows from using the approximate relation[16] $\sigma = E/2d_0d$ between σ and E). The amyloid fibrils are characterized by the inequality $E \gg E_h$ and, hence, $\sigma \gg \sigma_h$ and $\psi \gg \psi_h$, because the d and d_h values are not too different. For example, for the Aβ_{40} protein $d \approx 1$ nm (ref 20) and $\sigma \approx 18$ J/m^2 (estimated value[21]), so that with these values and the d_0, d_h, and σ_h values given above, it follows from eqs 4 and 7 that $\psi = 21$, $\psi_h = 1.5$, and $\psi/\psi_h = 14$ for this protein at $T = 310$ K.

In eqs 5 and 6, the term $-sn$ is the energy gained by assembling n monomers from the solution into an n-sized nanofibril, and the ψ_h and ψ terms represent the work done on creating the total surface area of a nanofibril with the shape of $i\beta$-sheet without (eq 5) or with (eq 6) a subsequent β-sheet on one of the $i\beta$-sheet sides. The solid line 12 in Figure 2 displays

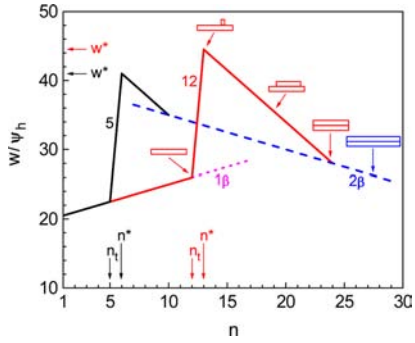


Figure 2. Work w for fibril formation as a function of the nanofibril size n at $\psi = 10\psi_h$ and $s = 1.5\psi_h$: lines 5 and 12, eqs 5 and 6 for nanofibrils with transition size $n_t = 5$ and 12, respectively; line 1β, eq 5 at $i = 1$; line 2β, eq 5 at $i = 2$. In addition to n_t, the nucleus size n^* and the nucleation work w^* are indicated, and successive nanofibril shapes are schematically shown to visualize the nanofibril transition from 1β- to 2β-sheet (the smallest rectangle represents a single β-strand).

the n dependence of w from eqs 5 and 6 for a single β-sheet ($i = 1$) which reaches exemplifying transition size $n_t = 12$ and length $m_t = 12$ and polymerizes $m_t = 12$ more monomers as a second β-sheet on one of its sides. The line is drawn with $\psi = 10\psi_h$ and $s = 1.5\psi_h$. The s value is chosen in the middle of the supersaturation range defined by inequality 2 with $i = 1$ (range $i = 1$ in Figure 1) where the 1β-sheets are fibril prenuclei. As seen in Figure 2, w increases linearly with the 1β-sheet size or, equivalently, length until the sheet becomes $n_t = m_t = 12$ monomers long. Importantly, as visualized by the dotted line in Figure 2, w would keep increasing in the same way if the 1β-sheet would continue lengthening (cf. eq 5). This means that at the chosen $s = 1.5\psi_h$ and, more generally, in the whole supersaturation range $\psi_h < s < 2\psi_h$ (range $i = 1$ in Figure 1), none of the differently long 1β-sheets is thermodynamically competent to lengthen unlimitedly. Any of the 1β-sheets can do that only after randomly attaching sidewise one monomer and thus transforming itself into a $(1\beta + 1)$-shaped nanofibril, i.e., a 1β-sheet plus one β-strand on it. This event corresponds to a sharp increase of w, manifested by the jump in the solid line 12 in Figure 2 between $n = 12$ and 13. With this costly shape-transformation work done, thermodynamics allows the reshaped nanofibril to grow irreversibly, because w then diminishes linearly with n up to $n = n_t + m_t = 24$ (cf. eq 6). At this size, the nanofibril is already a 2β-sheet and can elongate unlimitedly, because at $s = 1.5\psi_h$ the work $w(2, n)$ for 2β-sheet formation also decreases with n (see the dashed line in Figure 2, which is drawn according to eq 5 with $i = 2$).

Figure 2 reveals that, in general, the $(i\beta + 1)$-shaped nanofibril of size $n^* = n_t + 1$ requires maximum work for its formation, which is why, as in standard nucleation theory,[12,13] this nanofibril is the nucleus of the fibrillar phase appearing in the protein solution. Unlike in this theory, however, the fibril nucleus does not have a uniquely specified size n^* diminishing

dx.doi.org/10.1021/ja311228d | J. Am. Chem. Soc. 2013, 135, 1531−1539

continuously[15] or stepwise[16] with increasing supersaturation. Instead, any $(i\beta + 1)$-shaped, $(n_t + 1)$-sized nanofibril can be a nucleus, merely because any $i\beta$-sheet is able to attach a monomer to its side, acting thereby as a prenucleus of arbitrary length m_t and size $n_t = im_t$ if s is in the ith supersaturation range determined by inequality 2. The possibility for different fibril prenucleus and nucleus sizes is illustrated in Figure 2 in which the solid line 5 exhibits w from eqs 5 and 6 at the above s, ψ, and i values, but at transition size $n_t = 5$. Then the prenucleus is 1β-sheet of length $m_t = n_t = 5$ and the $(1\beta + 1)$-shaped, 6-sized nanofibril is the nucleus because of its energetically most costly formation. As seen in Figure 2, the dimensionless nucleation work $w^* \equiv w(i, n^*)$ is smaller for the nucleus of size $n^* = n_t + 1 = 6$ than for the nucleus of size $n^* = n_t + 1 = 13$. This means that thermodynamics favors the formation of shorter and, thereby, smaller fibril nuclei. It thus acts against the appearance of bigger nuclei, a process due to the fact that the respective prenuclei are longer and offer more monomer attachment sites on their sides. As a result, as shown below, at some intermediate transition length the fibril nuclei in the solution prove to be most numerous.

Setting $n = n^*$ and using $n^* = n_t + 1 = im_t + 1$ in eq 6, we readily arrive at the following general formula for w^* in the ith s range in which all different length $(i\beta + 1)$-shaped nanofibrils are fibril nuclei $(i = 1, 2, 3, ...; m_t = 1, 2, 3, ...; 2/(i + 1) < s/\psi_h < 2/i)$:

$$w^* = (2\psi_h - is)m_t - s + 2(i + 1)\psi \tag{8}$$

As seen, w^* diminishes linearly with increasing s and, due to the above relation between n^* and m_t, it obeys the nucleation theorem of standard nucleation theory in the form[13,22,23] $dw^*/ds = -n^*$.

Concentration and Length Distribution of Nuclei. Using eq 8, we can determine the equilibrium concentration $C^*_{m_t}$ of nuclei in the ith s range in which they are m_t-long $i\beta$-sheets with one monomer attached sidewise. According to CNT in its self-consistent formulation,[13] $C^*_{m_t}$ and w^* are related by the Boltzmann-type formula $C^*_{m_t} = C_1 e^{w(1,1)-w^*}$, where $w(1,1)$ is the dimensionless work for monomer formation (the monomer is formally considered as the smallest nanofibril of one β-strand). As this formula is of general statistical character and as from eq 5 we have $w(1, 1) = -s + 2\psi_h + 2\psi$, upon employing w^* from eq 8, we obtain $(i = 1, 2, 3, ...; m_t = 1, 2, 3, ...; 2/(i + 1) < s/\psi_h < 2/i)$

$$C^*_{m_t} = 2m_t C_1 e^{(is-2\psi_h)m_t - 2\psi i + 2\psi_h} \tag{9}$$

The factor $2m_t$ in this expression is introduced to take into account that a nucleus stemming from an m_t-long prenucleus appears in $2m_t$ energetically equivalent configurations corresponding to the $2m_t$ monomer attachment sites on the two sides of the prenucleus.

Equation 9 represents the length distribution of nuclei in the ith s range: it gives the concentration of nuclei with length of m_t β-strands (all these nuclei have the same number i of β-sheets). While the pre-exponential factor in this equation increases with increasing m_t, the exponential factor diminishes. As a result, $C^*_{m_t}$ passes through a maximum at $m_t = m^*_{max}$. This is seen in Figure 3 which displays $C^*_{m_t}$ from eq 9 at $\psi_h = 1$, $\psi = 10$, and $s = 1.5$ or 1.7 (as indicated). The length m^*_{max} of the most numerous

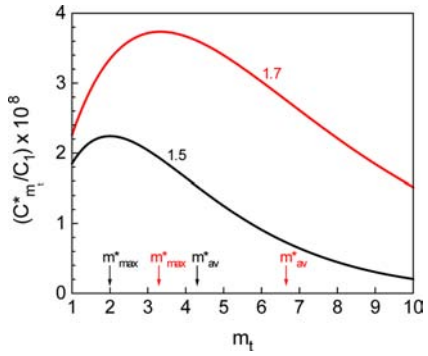


Figure 3. Length distribution of fibril nuclei with thickness of one β-sheet: lines 1.5 and 1.7, eq 9 at $s = 1.5$ and 1.7, respectively, $i = 1$, $\psi_h = 1$, and $\psi = 10$. The arrows indicate the corresponding values of m^*_{max} and m^*_{av} from eqs 10 and 14.

nuclei, their size $n^*_{max} = im^*_{max} + 1$, and their concentration $C^*_{max} \equiv C^*_{m^*_{max}}$ are given by $(i = 1, 2, 3, ...; 2/(i + 1) < s/\psi_h < 2/i)$

$$m^*_{max} = 1/(2\psi_h - is) \tag{10}$$

$$n^*_{max} = [i/(2\psi_h - is)] + 1 \tag{11}$$

$$C^*_{max} = [2C_1/(2\psi_h - is)]e^{-2(\psi i - \psi_h)-1} \tag{12}$$

Equations 10 and 12 follow from eq 9 and the condition for maximum of $C^*_{m_t}$, namely $dC^*_{m_t}/dm_t = 0$ at $m_t = m^*_{max}$. Exemplifying again the Aβ_{40} protein, with $\psi_h = 1.5$ and $i = 1$, from eq 10 we find that when the supersaturation is in the range in which the fibril nuclei are 1β-sheet-thick, the most numerous of them are with length $m^*_{max} = 0.7$, 3.3, and 6.7 β-strands at $s = \psi_h = 1.5$, $s = 1.8\psi_h = 2.7$, and $s = 1.9\psi_h = 2.85$, respectively. This result means that in the s range next to the metanucleation one, the most numerous Aβ_{40} nuclei are of size $n^*_{max} = m^*_{max} + 1 = 2$ to 8, except for s values quite close to that of the metanucleation border $(s = 2\psi_h = 3)$ when these nuclei are already above a dozen or more β-strands in length and, hence, size.

From eq 9, with the aid of the exact formula $\sum_{p=1}^{\infty} px^p = x/(1 - x)^2$ related to the geometric series, it follows that in the ith s range the total concentration $C^* = \sum_{m_t=1}^{\infty} C^*_{m_t}$ of nuclei of all lengths is given by $(i = 1, 2, 3, ...; 2/(i + 1) < s/\psi_h < 2/i)$

$$C^* = 2C_1 e^{i(s-2\psi)}/(1 - e^{is-2\psi_h})^2 \tag{13}$$

This important relation shows that, given the nucleus (and prenucleus) thickness i, just like m^*_{max}, n^*_{max}, and C^*_{max} from eqs 10, 11, and 12, C^* increases strongly with s and diverges at the upper limit $2\psi_h/i$ of the ith s range specified by inequality 2. This unphysical divergence is a mathematical expression of the fact that beyond this limit, rather than being in equilibrium with the solution, the $i\beta$-sheets are thermodynamically allowed to lengthen irreversibly (see Figure 1). We note as well that, owing to the factor $e^{-2i\psi}$ in eq 13, the nuclei in the ith s range are many orders of magnitude more numerous than those in the $(i + 1)$th s range. As will be seen below, this translates into a stark quantitative difference between the fibril nucleation rates in the different s ranges. Of all s ranges in which nucleation occurs (for them $i \geq 1$), the nucleation rate is highest when $i = 1$, which is why the s range with this i value will be called the range of rapid nucleation.

dx.doi.org/10.1021/ja311228d | J. Am. Chem. Soc. 2013, 135, 1531−1539

Equation 9 makes it possible to determine the average length m^*_{av} and size $n^*_{av} = im^*_{av} + 1$ of the $(i\beta + 1)$-shaped nuclei in the ith s range. As the average length is defined by $m^*_{av}C^* = \sum_{m_t=1}^{\infty} m_t C^*_{m_t}$, using eqs 9and 13 and the exact formula $\sum_{p=1}^{\infty} p^2 x^p = x(1 + x)/(1 - x)^3$ also related to the geometric series, we obtain $(i = 1, 2, 3, ...; 2/(i + 1) < s/\psi_h < 2/i)$

$$m^*_{av} = (1 + e^{is-2\psi_h})/(1 - e^{is-2\psi_h}) \quad (14)$$

$$n^*_{av} = [i(1 + e^{is-2\psi_h})/(1 - e^{is-2\psi_h})] + 1 \quad (15)$$

These expressions say that, similar to m^*_{max} and n^*_{max}, the average nucleus length m^*_{av} and size n^*_{av} increase with s and diverge at the upper limit $2\psi_h/i$ of the ith s range specified by inequality 2. The values of m^*_{av} and n^*_{av} are about twice those of m^*_{max} and n^*_{max}. Indeed, with $i = 1$ and the ψ_h value of 1.5 exemplifying the Aβ_{40} protein, from eq 14 we find that, in the range of rapid nucleation in which the fibril nuclei are 1β-sheet-thick, they are with average length $m^*_{av} = 1.6$, 6.7, and 13.4 β-strands at $s = \psi_h = 1.5$, $s = 1.8\psi_h = 2.7$, and $s = 1.9\psi_h = 2.85$, respectively. Thus, in the rapid-nucleation range the average size of the $(1\beta + 1)$-shaped Aβ_{40} nuclei is $n^*_{av} = m^*_{av} + 1 = 3$ to 14, except for s values close enough to that of the metanucleation border $(s = 2\psi_h = 3)$ when, on average, these nuclei are already more than a score of β-strands in length and, hence, size. Worth nothing also is that the average length and size of the fibril nuclei as well as the length and size of the most numerous of them are independent of ψ; i.e., they are invariants with respect to the energy of the strong (primarily hydrogen) bonds between nearest-neighbor β-strands in a β-sheet. As seen from eqs 10, 11, 14, and 15, at given i and s, the values of all these quantities depend only on ψ_h, i.e., they are dictated solely by the energy of the weak (hydrophobicity-mediated) interstrand bonds that hold the β-sheets together in a fibril.

Nucleation Rate. In line with standard nucleation theory,[12,13] the nucleation rate J (m^{-3} s^{-1}) is the frequency of appearance of supernucleus fibrils per unit solution volume. The peculiarity that the fibril nuclei have different number i of β-sheets in different s ranges (Figure 1) makes it difficult to obtain a unified formula for the $J(s)$ dependence at any $s > 0$. Finding this dependence for a given s range specified by inequality 2 is, however, quite straightforward, and one way of doing that is described below.

Let us first determine J in the metanucleation range $s > 2\psi_h$ (range $i = 0$ in Figure 1) in which there is no energy barrier to nucleation, because every monomer, i.e., single β-strand, in the solution plays the role of fibril nucleus. Then J, which is more correct to call the metanucleation rate, is merely the product of the monomer concentration C_1 and the net number $f_1 - g_2$ of monomers that a given monomer, acting as nucleus, attaches per unit time to its two strong-bond sides, i.e., $(C_1 > C_e e^{2\psi_h})$

$$J = (f_1 - g_2)C_1 \quad (16)$$

Here the monomer attachment frequency $f_1 = 2k_a$ is twice the frequency k_a (s^{-1}) of monomer attachment to one of the two strong-bond sides of a given monomer (Figure 4a). Similarly, the monomer detachment frequency $g_2 = 2k_d$ is twice the frequency k_d (s^{-1}) of monomer detachment from one of the two ends of a dimer (Figure 4b). While k_a is expected to increase linearly with C according to[15,17]

$$k_a = k_e C_1/C_e = k_e e^s \quad (17)$$

k_d is C_1-independent and can be represented as[17]

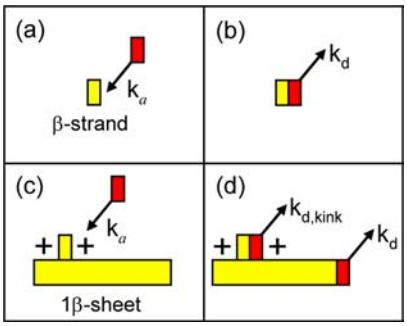


Figure 4. Schematic illustration of the monomer attachment and detachment frequencies k_a, k_d, and $k_{d,kink}$. The pluses indicate the nanofibril kink sites.

$$k_d = k_{d,kink} e^{2d_0 d_h \sigma_h/kT} = k_e e^{2\psi_h} \quad (18)$$

In eq 17, k_e is the value of k_a at equilibrium, i.e., at $C_1 = C_e$ or $s = 0$, and in eq 18, $k_{d,kink}$ is the detachment frequency of a monomer from a kink on the fibril surface (Figure 4d). Importantly, $k_{d,kink} = k_e$, because the fibril/solution equilibrium is maintained via monomer attachment and detachment to and from kinks. As to the exponential factor in eq 18, it takes into account that monomer detachment from a dimer (Figure 4b) is energetically less costly and, hence, more frequent than that from a kink (Figure 4d), because while in the former case only the area $2d_0 d$ is created and the work $2d_0 d\sigma$ is done, in the latter case the greater work $2d_0 d\sigma + 2d_0 d_h \sigma_h$ is done on creating the greater area $2d_0 d + 2d_0 d_h$.

Using now the above expressions for f_1, g_2, k_a, and k_d in eq 16, we readily obtain the following general formula for the time-independent (or stationary) fibril metanucleation rate $(s > 2\psi_h)$:

$$J = f_1 C_1 (1 - e^{2\psi_h - s}) \quad (19)$$

In experiments on protein aggregation at fixed temperature T, according to eq 3, the supersaturation s can be controlled by means of the monomer concentration C_1. Then $f_1 = 2k_e C_1/C_e$ and the metanucleation rate from eq 19 is the following function of C_1 $(C_1 > C_e e^{2\psi_h})$

$$J(C_1) = A_1 C_1^2 (1 - A_2 C_1^{-1}) \quad (20)$$

the factors A_1 and A_2 being given by $A_1 = 2k_e/C_e$ and $A_2 = C_e e^{2\psi_h}$.

In the other experimentally important case of supersaturation controlled by the temperature T at fixed C_1, s is of the form[15]

$$s = L(T_e - T)/kT_e T \quad (21)$$

where L (in Joules) is the latent heat or enthalpy (per β-strand) of fibril formation, $T_e - T > 0$ is the undercooling, and T_e is the absolute equilibrium temperature, i.e., the T value at which the bulk fibrillar phase and the solution with monomer concentration C_1 are in coexistence. Also, f_1 can be expressed as[15] $f_1 = f_0 e^{-E_a/kT}$, where the virtually T-independent frequency factor f_0 (s^{-1}) is proportional to C_1, and E_a is the activation energy for the monomer-to-monomer attachment schematized in Figure 4a. Then, combining eqs 4, 19, and 21, we find that the T-dependence of the stationary metanucleation rate is of the form $(T < T_e - 2d_0 d_h \sigma_h T_e/L)$

$$J(T) = A_1 e^{-E_a/kT}(1 - A_2 e^{-E_0/kT}) \quad (22)$$

dx.doi.org/10.1021/ja311228d | J. Am. Chem. Soc. 2013, 135, 1531−1539

where the energy E_0 and the factors A_1 and A_2 are given by $E_0 = L-2d_0d_h\sigma_h$, $A_1 = f_0C_1$, and $A_2 = e^{L/kT_e}$. Importantly, the above $J(T)$ dependence is in force provided that in the entire T range studied the β-strands remain in the same extended conformation, because then all four parameters A_1, A_2, E_a, and E_0 in eq 22 can be treated as practically T-independent.

We now turn to the determination of the fibril nucleation rate J at a given $i \geq 1$ when all $i\beta$-sheets are prenuclei, all $(i\beta + 1)$-shaped nanofibrils are nuclei, and the supersaturation is in the ith range specified by inequality 2 (cf. Figure 1). Then, in analogy with the case of metanucleation, J is the product of the total concentration C^* of nuclei and the net number $f^* - g^*$ of monomers attached per unit time to the nucleus two kink sites (indicated by the pluses in Figure 4c), i.e.

$$J = (f^* - g^*)C^* \tag{23}$$

Here f^* is the frequency of monomer attachment to the nucleus two kink sites (Figure 4c), and g^* is the frequency of monomer detachment from the so-formed dimer on the nucleus side (Figure 4d). As in CNT, we can consider the attachment frequency f^* as independent of the kind of the site to which a monomer is attached. Within this approximation, we can therefore set $f^* = f_1 = 2k_a$, k_a being given by eq 17. Concerning the detachment frequency g^*, however, following CNT, it is necessary to take into account that this frequency depends on the nucleus thickness, i.e., on the number i of β-sheets constituting the nucleus. From Figure 4d, we see that while the detachment frequency from the dimer on the 1β-sheet side is $k_{d,kink}$, the detachment frequency from the end of the 1β-sheet is k_d. Hence, as an effective detachment frequency $k_{d,eff}$ we can use the geometric average of $k_{d,kink}$ and k_d, i.e., $k_{d,eff} = (k_dk_{d,kink})^{1/2}$. Generalizing this result for nuclei with thickness of any number $i \geq 1$ of β-sheets, we obtain $k_{d,eff} = (k_dk_{d,kink}^i)^{1/(i+1)}$ so that employing $k_d = k_e e^{2\psi_h}$ and $k_{d,kink} = k_e$ yields $k_{d,eff} = k_e e^{2\psi_h/(i+1)}$.

Taking now into account that $g^* = 2k_{d,eff}$ because of the nucleus two ends, with the help of eqs 13, 17, 23 and the above expressions for $k_{d,eff}$ and f^*, we arrive at the following general formula for the stationary fibril nucleation rate in the ith s range in which the nuclei are built-up of $i = 1, 2, 3, \ldots$ β-sheets $(2/(i + 1) < s/\psi_h < 2/i)$:

$$J = 2f_1C_1e^{i(s-2\psi)}\frac{1 - e^{[2\psi_h/(i+1)]-s}}{(1 - e^{is-2\psi_h})^2} \tag{24}$$

As already noted, experimentally, it is of interest to have the dependence of J on C_1 at fixed T and the dependence of J on T at fixed C_1. In the former case, due to eq 3, eq 17 and the relation $f_1 = 2k_a$, eq 24 leads to $(i = 1, 2, 3, \ldots; C_e e^{2\psi_h/(i+1)} < C_1 < C_e e^{2\psi_h/i})$

$$J(C_1) = A_1C_1^{i+2}\frac{1 - A_2C_1^{-1}}{(1 - A_3C_1^i)^2} \tag{25}$$

where $A_1 = (4k_e/C_e^{i+1})e^{-2i\psi}$, $A_2 = C_e e^{2\psi_h/(i+1)}$, and $A_3 = C_e^{-i}e^{-2\psi_h}$. In the latter case, as then $f_1 = f_0e^{-E_a/kT}$, from eqs 4, 7, 21, and 24 it follows that $(i = 1, 2, 3, \ldots; T_e - 2d_0d_h\sigma_hT_e/iL < T < T_e - 2d_0d_h\sigma_hT_e/(i + 1)L)$

$$J(T) = A_1e^{-E_1/kT}\frac{1 - A_2e^{-E_2/kT}}{(1 - A_3e^{E_3/kT})^2} \tag{26}$$

Here the three pre-exponential factors and the three energies are given by $A_1 = 2f_0C_1e^{-iL/kT_e}$, $A_2 = e^{L/kT_e}$, $A_3 = e^{-iL/kT_e}$, $E_1 = E_a - i(L - 2d_0d\sigma)$, $E_2 = L - 2d_0d_h\sigma_h/(i + 1)$, and $E_3 = iL - 2d_0d_h\sigma_h$. Similar to eq 22, eq 26 is applicable when the β-strands are in the same extended conformation in the whole temperature range investigated. This is so, because then all six parameters in eq 26 have practically T-independent values.

Equations 19–21 and 24–26 are central results in the present study. Inspection of eqs 24–26 reveals that at $i = 0$ they describe rather accurately the metanucleation rate from eq 19–21, respectively (the metanucleation rate is overestimated only by a factor of about 2 if $\psi_h > 1$, a condition expected to be satisfied by amyloid proteins). This is important because of making it possible to employ eq 25 for analysis of experimental $J(C_1)$ data even without knowing the i value of the concentration range investigated. Upon fitting the data by the $J(C_1)$ dependence from eq 25 at fixed $i = 0$, 1, 2, etc., from the best of all fits the value of i could be obtained and, thereby, the thickness of the fibril nuclei (and prenuclei) determined. The expectation is that in most cases either $i = 0$ (metanucleation) or $i = 1$ (rapid nucleation) would provide the best fit, because when $i \geq 2$, J would be too low for the nucleation process to be detectable.

DISCUSSION

Comparison with Computer Experiment. In a recent paper,[17] we have reported kinetic Monte Carlo simulation data for the s dependence of the J/f_1C_1 ratio. The simulation code is described there, and here we only note that in the simulations we determined the probability P_2 of dimer (1β-sheet of two β-strands) growing into a macroscale fibril. We thus directly obtained J in units of f_1C_1, since J and P_2 are related by the exact formula[24] $J = f_1C_1P_2$. The so-obtained data for J are ideal for a most stringent verification of eqs 19 and 24, because in the computer experiment the supersaturation s and the energies ψ and ψ_h are precisely known and these equations contain no free parameters.

The symbols in Figure 5 exhibit the simulation data[17] which were obtained at fixed $\psi_h = 1$ and at $\psi = 5$, 8, 10, and 14 (as indicated), the ψ values being great enough to be relevant for amyloid fibrils. The lines in Figure 5 display the J/f_1C_1 ratio from eqs 19 and 24 at $i = 1$ or 2 and the above ψ and ψ_h values. As seen in the figure, the theoretical and the simulation nucleation rates are in a remarkably good agreement, especially in the metanucleation range ($i = 0$, $s > 2$) and in the rapid-nucleation range (then $i = 1$, $1 < s < 2$, and both the prenuclei and the nuclei are 1β-sheet-thick).

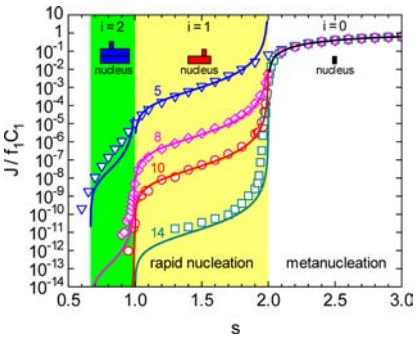


Figure 5. Dependence of the fibril nucleation rate J on the supersaturation s at $\psi_h = 1$ and $\psi = 5$, 8, 10, and 14 (as indicated): symbols, kinetic Monte Carlo simulation data;[17] line in area $i = 0$, eq 19; lines in areas $i = 1$ and $i = 2$, eq 24 at $i = 1$ and 2, respectively.

dx.doi.org/10.1021/ja311228d | J. Am. Chem. Soc. 2013, 135, 1531–1539

The conformity between theory and computer experiment at $\psi = 10$ is illustrated again in Figure 6 which, however, visualizes

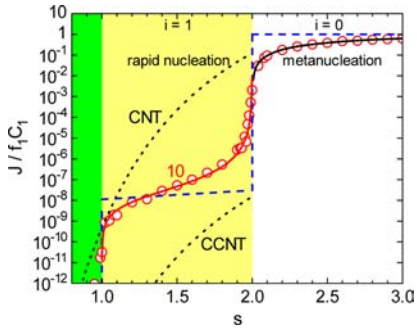


Figure 6. Dependence of the fibril nucleation rate J on the supersaturation s at $\psi_h = 1$ and $\psi = 10$: symbols, kinetic Monte Carlo simulation data;[17] solid line in area $i = 0$, eq 19; solid line in area $i = 1$, eq 24 at $i = 1$; lines CNT and CCNT, eqs 27 and 28 of ref 15; dashed line, eq 27 at $i = 0$ and 1.

the strong discrepancy between the $J(s)$ dependence from the simulation and from both CNT and the corrected CNT (CCNT) (eqs 27 and 28 in ref 15), a discrepancy which persists for the other ψ values used in the simulation.[17] As CCNT and the atomistic nucleation theory (ANT) are in good agreement (see Figure 5 in ref 16), this discrepancy implies that ANT cannot describe the simulation $J(s)$ data either. Neither can this be done by other theories proposing the formula $J = aC_1^b$, because as already pointed out, this formula predicts a concave or linear increase of ln J with ln C_1 and, hence, s.

Figure 6 shows also that CNT and CCNT only provide, respectively, upper and lower limits of the simulation $J(s)$ dependence. While CNT overestimates J because of disregarding the big work needed for sidewise attachment of a monomer to an $i\beta$-sheet, CCNT accounts for this work, but underestimates J by considering the transition length m_t as fixed and thereby ignoring the fact that not only the $i\beta$-sheets of this fixed length, but also the $i\beta$-sheets of all other lengths act as prenuclei. Interestingly, the simple approximate formula ($i = 0$, 1, 2, ...; $2/(i + 1) < s/\psi_h < 2/i$)

$$J = (i + 1)f_1 C_1 e^{i(s-2\psi)} \tag{27}$$

describes qualitatively the steplike course of the simulation $J(s)$ dependences in Figure 5, and even provides a rough estimate of the magnitude of the jumps in J as well as of the J values at supersaturations far enough from the borders of the successive s ranges. This is seen in Figure 6 where the dashed line illustrates eq 27 at $\psi = 10$ and $i = 0$ or 1. This equation follows from eq 24 upon neglecting the terms $1 - e^{[2\psi_h/(i+1)]-s}$ and $(1 - e^{is-2\psi_h})^2$ which account, respectively, for the vanishing net number of monomers attaching to the nucleus at s approaching the left end $2\psi_h/(i + 1)$ of the ith s range and for the sharp rise of the concentration of nuclei at s tending to the right end $2\psi_h/i$ of this range. As to the factor $i + 1$ in eq 27, it is introduced to take into account that the $(i\beta + 1)$-shaped nucleus has i more attachment sites at one of its ends than the single β-strand (cf. panels a and c of Figure 4).

Recalling that $f_1 = 2k_eC_1/C_e$ when C_1 is changed at fixed T and that $f_1 = f_0 e^{-E_a/kT}$ when T is changed at fixed C_1, upon using eqs 3, 21, and 27 we find that the dependence of J on C_1 at fixed T and on T at fixed C_1 is approximately given by ($i = 0$, 1, 2, ...; $C_e e^{2\psi_h/(i+1)} < C_1 < C_e e^{2\psi_h/i}$)

$$J = aC_1^b \tag{28}$$

in the former case and by ($i = 0$, 1, 2,...; $T_e - 2d_0d_h\sigma_hT_e/iL < T < T_e - 2d_0d_h\sigma_hT_e/(i + 1)L$)

$$J = Ae^{-E_1/kT} \tag{29}$$

in the latter case. The factors a and A, the power b, and the activation energy E_1 in these equations are given by $a = 2(i + 1)k_e e^{-2i\psi}/C_e^{i+1}$, $b = i + 2$, $A = (i + 1)f_0C_1e^{-iL/kT_e}$ and $E_1 = E_a - i(L - 2d_0d\sigma)$. Importantly, eq 28 reveals the physical meaning of the parameters a and b in the formula $J = aC_1^b$ already discussed. The equation says that this formula is applicable only approximately, besides with different a and b values in successive C_1 ranges. It shows as well that rather than determined by the nucleus size n^* (which is impossible merely because the n^* value is not a unique), the power b is controlled solely by the number i of β-sheets constituting the nucleus, i.e., by the nucleus thickness. As amyloid fibrils are most likely to form mainly in metanucleation ($i = 0$) and rapid-nucleation ($i = 1$) regimes, based on eq 28, we can expect the most often encountered values of the power b to be 2 and 3 or, perhaps, about twice greater because of the approximate character of eq 28 (cf. the solid and dashed lines in Figure 6). This expectation is supported by b values inferred from experimental data for the kinetics of amyloid fibrillation (e.g., Table 1 in ref 25).

Application to Aβ_{40} Fibril Nucleation. The successful description of the simulation $J(s)$ data by eqs 19 and 24 is gratifying and makes it meaningful to employ eqs 20 and 25 for predicting the $J(C_1)$ dependence in stationary homogeneous nucleation of wild-type Aβ_{40} protein at $T = 310$ K, a dependence not yet available from experiments. Using $k_e = 10^{-4}$ s^{-1} (ref 15), $C_e = 4.8 \times 10^{20}$ m^{-3} (=0.8 μM) (ref 26), and the already mentioned $\psi = 21$ and $\psi_h = 1.5$, from the formulas for the constants A_1 and A_2 in eq 20 we obtain $A_1 = 4.167 \times 10^{-25}$ m^3 s^{-1} and $A_2 = 9.641 \times 10^{21}$ m^{-3}. With the same parameter values, at $i = 1$, for the constants A_1, A_2, and A_3 in eq 25 we find $A_1 = 9.982 \times 10^{-64}$ m^6 s^{-1}, $A_2 = 2.151 \times 10^{21}$ m^{-3}, and $A_3 = 1.037 \times 10^{-22}$ m^3.

Lines MN and RN in Figure 7 display the metanucleation and rapid-nucleation ($i = 1$) rates obtained from eqs 20 and 25, respectively, with the above two sets of constants. We observe that in a rather narrow concentration interval from about 10 to 20 μM J increases nearly 16 orders of magnitude. The

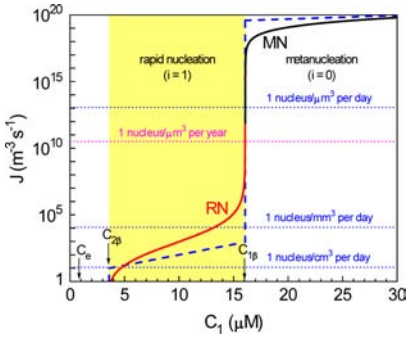


Figure 7. Theoretically predicted dependence of the Aβ_{40} fibril nucleation rate on the protein concentration at $T = 310$ K: solid line MN, eq 20 at $i = 0$; solid line RN, eq 25 at $i = 1$; dashed line, eq 28 at $i = 0$ and 1; dotted lines, selected values of the nucleation rate. The threshold concentration $C_{1\beta}$ is the border between fibril rapid nucleation and metanucleation.

dx.doi.org/10.1021/ja311228d | J. Am. Chem. Soc. 2013, 135, 1531−1539

spectacularly sharp rise of J occurs at the nucleation/metanucleation border

$$C_{1\beta} = C_e e^{2\psi_h} = C_e e^{2d_0 d_h \sigma_h / kT} \quad (30)$$

of 16 μM. Thus, $C_{1\beta}$ from this equation (which follows from eq 1 at $i = 1$) appears as a threshold concentration, below which the Aβ_{40} fibrils can be nucleated within a day solely in volumes of 1 mm^3 or larger, used for *in vitro* protein fibrillation. In volumes of about 1 μm^3 or smaller, as those of biological cells, the fibril nucleation is arrested for days and even years when $C_1 < C_{1\beta}$ (see Figure 7), and the Aβ_{40} fibrils can only come into being by metanucleation at $C_1 > C_{1\beta}$ when the fibril nuclei are single β-strands. We note as well that the Aβ_{40} nucleation rate in the rapid-nucleation range ($i = 1$) is more than 10 orders of magnitude higher than that in the next nucleation range ($i = 2$), which is why the latter is not displayed in Figure 7. In general, therefore, amyloid fibril nucleation at $i = 2$, 3, etc. (i.e., with 2β-sheets, 3β-sheets, etc., as prenuclei), can hardly be of significance when $\psi > 10$. As to the approximate eq 28, in Figure 7 it is illustrated by the dashed line drawn with $a = 4.167 \times 10^{-25}$ m^3 s^{-1} and $b = 2$ when $i = 0$ (metanucleation) and with $a = 9.982 \times 10^{-64}$ m^6 s^{-1} and $b = 3$ when $i = 1$ (rapid nucleation), values calculated from the formulas for these two parameters with the values of k_e, C_e, and ψ given above. We see that eq 28 can be used for approximating simply, but not very accurately, the $J(C_1)$ dependence and for estimating, albeit roughly, the magnitude of the sharp jump of J at the metanucleation/rapid-nucleation border.

The dramatic difference in the values of the rapid-nucleation and metanucleation rates suggests that the threshold concentration $C_{1\beta}$ is a crucial parameter in controlling the formation of amyloid fibrils both *in vivo* and *in vitro*. If the protein solution is to contain fibrils, $C_{1\beta}$ should be as small as possible. Vice versa, for the solution to be practically fibril-free, $C_{1\beta}$ has to be as large as possible. Equation 30 shows that for therapeutic or technological purposes $C_{1\beta}$ can be tuned by changing the protein solubility C_e, the β-strand length d_0 (i.e., the β-sheet width), the distance d_h occupied by a β-strand along a β-sheet, and/or the hydrophobicity-mediated specific surface energy σ_h of the β-sheet sides. For example, σ_h can be diminished by the presence in the solution of substances adsorbing on the β-sheet sides. Then $C_{1\beta}$ decreases exponentially with σ_h and the protein fibrillation is stimulated, because metanucleation commences at smaller C_1 values. As to C_e, it can increase, e.g., owing to protein point-mutations.[27] Then $C_{1\beta}$ increases proportionally to C_e and fibrillation is hampered, because metanucleation is shifted to higher protein concentrations. This result is consistent with experiments on protein fibrillation kinetics,[28] which reveal that the lag time before detectable fibrillation is longer for two Aβ_{40} mutants than for the wild-type Aβ_{40}, and that one Aβ_{40} mutant is not detectable at all. By affecting the C_e and σ_h values and, thereby, the threshold concentration $C_{1\beta}$, the solution pH can also be a significant factor in the fibril nucleation process.

Final Remarks. We emphasize that all results obtained above apply only when the fibril nucleation occurs in one step, i.e., when the monomeric protein polymerizes directly into fibrils. They are not applicable to the case of two-step fibril nucleation in which the monomers first aggregate into nonfibrillar oligomers that then convert into fibrils.[29] The operativeness of these two nucleation mechanisms depends on both thermodynamics and kinetics,[21] and experimentally, it may not be easy to distinguish them.

Another important point to make is about the analysis of experimental $J(s)$ data for one-step fibril nucleation at fixed T. In standard nucleation theory,[12,13] J is of the form $J = Ae^{-w^*}$ and as for condensed phases the factor A increases approximately proportionally to e^s, due to the nucleation theorem, the relation $n^* \approx d(\ln J)/ds - 1$ holds.[13,22,23] This means that for such phases the slope of an experimental ln J versus s dependence gives directly the nucleus size n^*. Amyloid fibril nucleation, however, is a very peculiar, nonstandard kind of nucleation in the sense that, as already noted, the fibril nucleus does not have a unique length and, thereby, size. For that reason, no information about the nucleus size in one-step amyloid fibril nucleation can be obtained from the slope of an experimental ln J versus s dependence. Mathematically, this is obvious from eqs 8 and 24 which show that the $J(s)$ dependence for amyloid fibrils is not of the standard form $J = Ae^{-w^*}$.

Finally, we note again that our analysis treats only the homogeneous nucleation of amyloid fibrils, which can occur when the protein solution is sufficiently highly supersaturated or when the solution container has walls whose surfaces are nucleation-inactive or when nucleation-active foreign nano- and/or microparticles are absent from the solution. When such particles are present in the solution, however, as convincingly shown in recent studies on human β_2-microglobulin fibrillation,[30,31] heterogeneous nucleation on the particle surfaces can take place. Hence, as preparing sufficiently pure solutions and solution containers with nucleation-inactive surfaces is rather difficult, heterogeneous nucleation is very likely to occur in practice. Nonetheless, a clear understanding and description of the homogeneous amyloid fibril nucleation remains an important problem *per se* for at least two reasons: (i) albeit with a considerable effort, the conditions (sufficiently high purity or supersaturation of the solution) can be realized for the occurrence of this process in experiments, and (ii) the knowledge about this process is of great value, because it provides a solid basis for theoretical treatment of the heterogeneous nucleation of amyloid fibrils.

CONCLUSION

Summing up, we can conclude that, despite the peculiarity of the one-step homogeneous nucleation of amyloid fibrils, the stationary rate J of the process depends in a rather simple way on the supersaturation s. The general $J(s)$ dependence in successive s ranges is given by eqs 19 and 24 according to which of practical importance are solely the rates in the metanucleation and the rapid-nucleation ranges (then the nucleus is, respectively, any single β-strand or any single β-sheet with one β-strand attached sidewise). The metanucleation rate is many orders of magnitude higher than that in the rapid-nucleation range. The crossover between these two rates is at the threshold monomer concentration $C_{1\beta}$ which, as seen from eq 30, can be manipulated by changing the protein solubility C_e, the β-strand geometrical parameters d_0 and d_h, and/or the specific surface energy σ_h of the β-sheet sides. This finding might prove valuable in designing new therapeutic and/or technological strategies for stymying or stimulating amyloid fibril formation.

AUTHOR INFORMATION

Corresponding Author
S.Auer@leeds.ac.uk

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the Engineering and Physical Sciences Research Council Grant EP/G026165/1.

■ REFERENCES

(1) Chiti, F.; Dobson, C. M. *Annu. Rev. Biochem.* **2006**, *75*, 333.
(2) Ferrone, F. A.; Hofrichter, J.; Sunshine, H. R.; Eaton, W. A. *Biophys. J.* **1980**, *32*, 361.
(3) Bishop, M. F.; Ferrone, F. A. *Biophys. J.* **1984**, *46*, 631.
(4) Ferrone, F. A.; Hofrichter, J.; Eaton, W. A. *J. Mol. Biol.* **1985**, *183*, 611.
(5) Ferrone, F. *Methods Enzymol.* **1999**, *309*, 256.
(6) Ferrone, F. A. *Methods Enzymol.* **2006**, *412*, 285.
(7) Andrews, J. M.; Roberts, C. J. *J. Phys. Chem. B* **2007**, *111*, 7897.
(8) Knowles, T. P. J.; Waudby, C. A.; Devlin, G. L.; Cohen, S. I. A.; Aguzzi, A.; Vendruscolo, M.; Terentjev, E. M.; Welland, M. E.; Dobson, C. M. *Science* **2009**, *326*, 1533.
(9) Cohen, S. I. A.; Vendruscolo, M.; Welland, M. E.; Dobson, C. M.; Terentjev, E. M.; Knowles, T. P. J. *J. Chem. Phys.* **2011**, *135*, 065105.
(10) Vitalis, A.; Pappu, R. V. *Biophys. Chem.* **2011**, *159*, 14.
(11) Crespo, R.; Rocha, F. A.; Damas, A. M.; Martins, P. M. *J. Biol. Chem.* **2012**, *287*, 30585.
(12) Abraham, F. F. *Homogeneous Nucleation Theory*; Academic: New York, 1974.
(13) Kashchiev, D. *Nucleation: Basic Theory with Applications*; Butterworth-Heinemann: Oxford, 2000.
(14) Hofrichter, J.; Ross, P. D.; Eaton, W. A. *Proc. Natl. Acad. Sci. U.S.A.* **1974**, *71*, 4864.
(15) Kashchiev, D.; Auer, S. *J. Chem. Phys.* **2010**, *132*, 215101.
(16) Cabriolu, R.; Kashchiev, D.; Auer, S. *J. Chem. Phys.* **2010**, *133*, 225101.
(17) Cabriolu, R.; Kashchiev, D.; Auer, S. *J. Chem. Phys.* **2012**, *137*, 204903.
(18) Auer, S.; Kashchiev, D. *Phys. Rev. Lett.* **2010**, *104*, 168105.
(19) Auer, S. *J. Chem. Phys.* **2011**, *135*, 175103.
(20) Nguyen, H. D.; Hall, C. K. *J. Am. Chem. Soc.* **2006**, *128*, 1890.
(21) Auer, S.; Ricchiuto, P.; Kashchiev, D. *J. Mol. Biol.* **2012**, *422*, 723.
(22) Kashchiev, D. *J. Chem. Phys.* **1982**, *76*, 5098.
(23) Kashchiev, D. *J. Chem. Phys.* **2006**, *125*, 014502.
(24) ter Horst, J. H.; Kashchiev, D. *J. Chem. Phys.* **2005**, *123*, 114507.
(25) Kar, K.; Jayaraman, M.; Sahoo, B.; Kodali, R.; Wetzel, R. *Nat. Struct. Mol. Biol.* **2011**, *18*, 328.
(26) O'Nuallain, B.; Shivaprasad, S.; Kheterpal, I.; Wetzel, R. *Biochemistry* **2005**, *44*, 12709.
(27) Cabriolu, R.; Auer, S. *J. Mol. Biol.* **2011**, *411*, 275.
(28) Christopeit, T.; Hortschansky, P.; Schroeckh, V.; Gührs, K.; Zandomeneghi, G.; Fändrich, M. *Protein Sci.* **2005**, *14*, 2125.
(29) Lee, J.; Culyba, E. K.; Powers, E. T.; Kelly, J. W. *Nat. Chem. Biol.* **2011**, *7*, 602.
(30) Linse, S.; Cabaleiro-Lago, C.; Xue, W.-F.; Lynch, I.; Lindman, S.; Thulin, E.; Radford, S. E.; Dawson, K. A. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 8691.
(31) Cabaleiro-Lago, C.; Quinlan-Pluck, F.; Lynch, I.; Lindman, S.; Minogue, A. M.; Thulin, E.; Walsh, D. M.; Dawson, K. A.; Linse, S. *J. Am. Chem. Soc.* **2008**, *130*, 15437.