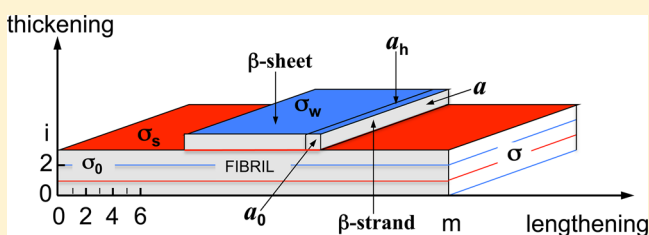


## Amyloid Fibril Nucleation: Effect of Amino Acid Hydrophobicity

Stefan Auer\*

School of Chemistry, University of Leeds, Leeds LS2 9JT, United Kingdom

**ABSTRACT:** We consider the nucleation of amyloid fibrils when the process occurs by direct polymerization of fully extended peptides (i.e.,  $\beta$ -strands) into fibrils composed of successively layered  $\beta$ -sheets with alternating weak and strong hydrophobic surfaces. We extend our recently developed nucleation model (Kashchiev, D.; Cabriolu, R.; Auer, S. *J. Am. Chem. Soc.* **2013**, 135, 1531–1539) to derive general expressions for the work to form such fibrils, the fibril solubility, the nucleation work, the equilibrium concentration of nuclei, and the fibril nucleation rate as explicit functions of the supersaturation of the protein solution. Analysis of these expressions illustrates the effect of increased asymmetry between the weak and strong hydrophobic  $\beta$ -sheet surfaces on the thermodynamics and kinetics of the polymerization process. In particular, the application of our theoretical framework to a simple model peptide system shows that lowering the hydrophobicity of one  $\beta$ -sheet surface can hamper protein fibrillation because the threshold concentration below which the fibril nucleation is practically arrested, and above which the process occurs vigorously—because then each monomer in the solution acts as a fibril nucleus—is shifted to higher concentrations. This effect is entirely due to the effect of asymmetry of the two hydrophobic  $\beta$ -sheet surfaces on the fibril solubility. In addition, with increasing asymmetry, the nucleation rate of one fibril polymorph becomes increasingly dominant, illustrating that there is a morphological selection between the two possible polymorphs.



## ■ INTRODUCTION

Proteins seem to have an intrinsic tendency to assemble into highly ordered fibrillar aggregates, including amyloid fibrils, which are associated with various neurodegenerative disorders such as Alzheimer's and Parkinson's diseases (e.g., refs 1, 2). Structural studies (e.g., refs 3–5) have shown that amyloid fibrils share a common cross- $\beta$  structure formed by intertwined layers of  $\beta$ -sheets extending in a direction parallel to the fibril axis. The formation of such amyloid fibrils requires strong molecular interactions between the proteins with a preference in direction of the fibril axis.<sup>6</sup> This might be attributed to the strongly directional protein backbone hydrogen bonding that enables  $\beta$ -strands to assemble into  $\beta$ -sheets arranged parallel to the fibril axis and to the weaker side-chain/side-chain interactions that cause the fibril to thicken and to form protofilaments.<sup>7</sup> The strength of such bonds depends (among other factors) on the side-chain hydrophobicity of the amino acids in the  $\beta$ -strand. As the orientation of side-chains within a  $\beta$ -strand alternates (they successively orient along the positive and negative direction of the fibril thickening axis), the hydrophobicity of the two surfaces of a  $\beta$ -sheet is generally different, except for homopolypeptides when all amino acids within a  $\beta$ -strand are composed of the same type of amino acid. In this article, we address the question about how does an asymmetry in the hydrophobicity of the two  $\beta$ -sheet surfaces affect the thermodynamics and the kinetics of amyloid fibril nucleation?

The nucleation of amyloid fibrils refers to the process of random generation of such nanofibrils that have the ability to irreversibly grow. Unless the nanofibril size exceeds the size  $n^*$

of the fibril nucleus, the nanofibril is more likely to dissolve than to grow (here  $n^*$  is the number of peptide monomers in the fibril nucleus). In classical nucleation theory (CNT) (e.g., refs 8, 9), at a given supersaturation, the nucleus size  $n^*$  has a unique value that can be obtained from thermodynamic considerations of the energy gained when  $n^*$  monomers in a supersaturated solution assemble into the nucleus and the work needed to create the nucleus surface. Our recent simulations revealed,<sup>10</sup> however, that the strong interaction anisotropy that characterizes amyloid fibrils leads to a peculiar kind of nucleation not entirely complying with the paradigm of standard nucleation theory. The simulations showed that amyloid fibril nuclei are actually aggregates with ambiguous dimensionality, and the work of creating the nucleus/solution interface does not scale in a definite way with the nucleus size (in CNT, this work scales as  $n^{*2/3}$  for 3D nuclei and as  $n^{*1/2}$  for 2D nuclei).<sup>9</sup> As a consequence, at a given supersaturation, the size  $n^*$  does not have a unique value, and amyloid nanofibrils of different size and shape can act as fibril nuclei. This called for a reconsideration of the theoretical description of amyloid fibril nucleation presented elsewhere,<sup>11–13</sup> and a new theoretical approach<sup>14</sup> provided remarkably good agreement with the simulation data<sup>10</sup> for the nucleation of nanosized amyloid fibril composed of successively layered  $\beta$ -sheets with surfaces symmetric (i.e., equivalent) with respect to their hydrophobicity. The objective of this article is to extend our new

Received: November 19, 2013

Revised: April 22, 2014

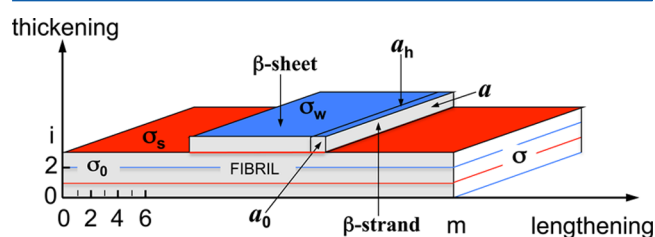
Published: May 2, 2014

nucleation model<sup>14</sup> to address the question about the effect that the asymmetry in the hydrophobicity of the two  $\beta$ -sheet surfaces can have on the thermodynamics and the kinetics of amyloid fibril nucleation.

The usefulness of the derived general expressions for the fibril nucleation rate,  $J$ , as an explicit function of the supersaturation,  $s$ , is 2-fold. First, they are a first step toward first-principle predictions of fibril nucleation rates based on the amino acid sequence of the proteins. Although existing models based on protein physiochemical properties are able to predict changes in aggregation propensities, they cannot provide information about the fibril nucleation barrier and rate (e.g., refs 15–18). Second, understanding the  $J(s)$  dependence of such fibrils is an indispensable ingredient in the set of rate equations describing the overall kinetics of unseeded amyloid fibrillation (see, e.g., refs 19–21).

## THEORETICAL METHODS

**Model.** For describing the arrangement of virtually fully extended  $\beta$ -strands in a nanosized amyloid fibril (protofibril) of successively layered  $\beta$ -sheets with alternating weakly hydrophobic (WH) and strongly hydrophobic (SH) surfaces (Figure 1), we extend our recently developed lattice model.<sup>12</sup>



**Figure 1.** Schematic of nanosized amyloid fibril (protofibril) with thickness of  $i$   $\beta$ -sheets and length of  $m$   $\beta$ -strands. The  $\sigma$ 's are the specific surface energies of the fibril surfaces, and the  $a$ 's are the areas of the  $\beta$ -strand faces. The red and blue areas indicate the  $\beta$ -sheet SH and WH surfaces, respectively. In our model, we assume that for a 1 $\beta$ -sheet, the SH surface is always on top (indicated by the red line) whereas the WH surface is at the bottom (indicated by the blue line). In addition, a  $\beta$ -strand can bind itself to a WH  $\beta$ -sheet surface only with its WH side (blue binds to blue) and to an SH  $\beta$ -sheet surface only with its SH side (red binds to red). Thus, the hydrophobicity of the surface of a nanofibril alternates with an increasing number of  $\beta$ -sheets (red, blue, red, blue, etc.).

Each  $\beta$ -strand (a segment of a protein chain composed of several amino acids) is represented by a right rectangular prism. Because of their strong hydrogen bonds, the  $\beta$ -strands can arrange themselves laterally into  $\beta$ -sheets. The sheets consist of different numbers  $m$  of  $\beta$ -strands ( $m = 1, 2, 3, \dots$ ) and are parallel to the fibril lengthening axis (the  $m$  axis in Figure 1). Along its thickening axis (the  $i$  axis in Figure 1), the fibril is built up of  $i$   $\beta$ -sheets ( $i = 1, 2, 3, \dots$ ) that are held together by, for example, relatively weak hydrophobicity-mediated bonds between the  $\beta$ -strands.<sup>7</sup> The strength of such bonds depends on the side-chain hydrophobicity of the amino acids in the  $\beta$ -strand. Because the orientation of side-chains within a  $\beta$ -strand alternates (they successively orient along the positive and negative direction of the  $m$  axis), the hydrophobicity of the two surfaces of a  $\beta$ -sheet is generally different, except for the case when all amino acids within a  $\beta$ -strand are identical. Because the fibril width is fixed and equal to the  $\beta$ -strand length, the fibril can be considered as a 2D aggregate in the  $m, i$  plane, with

building blocks (the  $\beta$ -strands) arranged in a 2D lattice with simple rectangular symmetry (Figure 1). The areas occupied by a  $\beta$ -strand at the fibril faces perpendicular to the  $m$  and  $i$  axes are  $a$  ( $\text{m}^2$ ) and  $a_i$  ( $\text{m}^2$ ), respectively, and the area occupied by the  $\beta$ -strand in the  $m, i$  plane is  $a_0$  ( $\text{m}^2$ ) (Figure 1). These areas are given by  $a_0 = d_h d$ ,  $a = d_0 d$  and  $a_i = d_0 d_h$ , where  $d_h$  (m) is the distance between the  $\beta$ -strands in a  $\beta$ -sheet;  $d$  (m) is the intersheet distance in the fibril; and  $d_0$  (m) is the  $\beta$ -strand length, that is, the  $\beta$ -sheet width.

The ontogenesis of the smallest (nanosized) amyloid fibrils in the  $m, i$  plane can be described by 2D crystal nucleation and growth theories.<sup>8,9</sup> Essential parameters in these theories are the specific surface energies of the different crystal faces. In our model for amyloid fibril nucleation, three of the four fibril specific surface energies are of immediate importance: the first one is the specific surface energy  $\sigma$  ( $\text{J}/\text{m}^2$ ) of the fibril face perpendicular to the  $m$  axis, and the second and third ones are the strong and weak specific surface energy  $\sigma_s$  ( $\text{J}/\text{m}^2$ ) and  $\sigma_w$  ( $\text{J}/\text{m}^2$ ) of the fibril faces perpendicular to the  $i$  axis, respectively. The fourth fibril-specific surface energy,  $\sigma_0$  ( $\text{J}/\text{m}^2$ ), characterizes the fibril face parallel to the  $m, i$  plane and enters only implicitly, via the supersaturation, the description of the fibril energetics.<sup>11</sup> To a first approximation, the surface energy is proportional to the energy of the broken bonds at the respective surface and, for that reason,  $\sigma$ ,  $\sigma_s$ , and  $\sigma_w$  are largely determined by the strength of the bonds between the neighboring  $\beta$ -strands in the fibril. Because fibril elongation is primarily driven by the formation of strong hydrogen bonds between the  $\beta$ -strands along the  $m$  axis and fibril thickening along the  $i$  axis is due to the much weaker bonds between the  $\beta$ -strands arising, for example, from the hydrophobic effect, for the amyloid fibrils, we have the important inequality  $\sigma \gg \sigma_s, \sigma_w$ .

For the analysis to follow, we introduce the dimensionless specific surface energies  $\psi$  and  $\psi_w$ ,  $\psi_s$  of the fibril faces perpendicular to the  $m$  axis and the  $i$  axis, respectively, which are given by

$$\psi = a\sigma/kT = E/2kT \quad (1)$$

$$\psi_s = a_i\sigma_s/kT = E_s/2kT \quad (2)$$

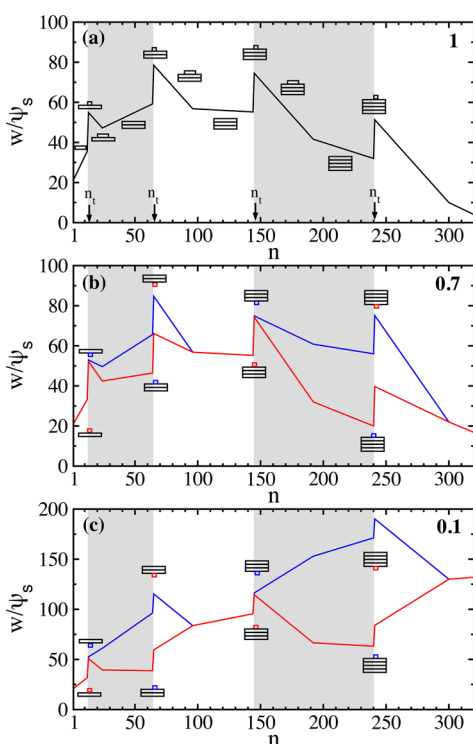
$$\psi_w = a_i\sigma_w/kT = E_w/2kT \quad (3)$$

where  $k$  is the Boltzmann constant and  $T$  (K) is the absolute temperature. The second equality in eqs 1–3 results from using the approximate relation<sup>12</sup>  $\sigma = E/2kT$  between  $\sigma$  and the binding energy,  $E$ , of the bond between two nearest-neighbor  $\beta$ -strands in a  $\beta$ -sheet as well as the approximate relation  $\sigma_s = E_s/2kT$  (or  $\sigma_w = E_w/2kT$ ) between  $\sigma_s$  (or  $\sigma_w$ ) and the strong (or weak) hydrophobicity-mediated binding energy  $E_s$  (or  $E_w$ ) of two nearest-neighbor  $\beta$ -strands in adjacent  $\beta$ -sheets.

Importantly, the dimensionless specific surface energy,  $\psi$ , perpendicular to the  $m$  axis can contain contributions from both nearest-neighbor hydrogen bonding and hydrophobicity-mediated bonds and is given by  $\psi = \psi_{\text{hb}} + c_s\psi_s + c_w\psi_w$ . Here,  $\psi_{\text{hb}}$  is the dimensionless specific surface energy due to hydrogen bonding, and  $c_s$  and  $c_w$  are parameters determining the contributions of hydrophobicity-mediated bonds to  $\psi$ . For most of our illustrations, we either set  $c_s = c_w = 0$ , which means that there is no hydrophobic contribution to  $\psi$ , or we set  $c_s = c_w = 0.5$ , as in our previous work.<sup>22</sup>

## RESULTS AND DISCUSSION

**Nucleation Work.** According to atomistic nucleation theory,<sup>12</sup> an  $n$ -sized nanofibril composed of  $i$   $\beta$ -sheets forms by a layer-by-layer nucleation mechanism. First, a  $1\beta$ -sheet forms, which at a transition size  $n_t$  changes its shape to that of a  $2\beta$ -sheet with a partially or fully built-up second  $\beta$ -sheet on one of its sides. The so-formed  $2\beta$ -sheet lengthens further by attachment of  $\beta$ -strands to the fibril ends until a second transition size  $n_t$  is reached, at which the  $2\beta$ -sheet changes its shape to that of a  $3\beta$ -sheet with again a partially or fully built-up third  $\beta$ -sheet on one of its two sides. This process continues until the  $n$ -sized nanofibril is composed of  $i$   $\beta$ -sheets (see the schematic illustration of nanofibrils in Figure 2a). In our model



**Figure 2.** Work,  $w$ , for nanofibril formation as a function of the nanofibril size  $n$  with  $c_s = c_w = 0$  at  $\psi = 10$  and  $\psi_s = 1$ . In panel a,  $\psi_w/\psi_s = 1$  (symmetric hydrophobicity) at  $s/\psi_s = 0.7$ ; in panel b,  $\psi_w/\psi_s = 0.7$  (weakly asymmetric hydrophobicity) at  $s/\psi_s = 0.6$ ; and in panel c,  $\psi_w/\psi_s = 0.1$  (strongly asymmetric hydrophobicity) at  $s/\psi_s = 0.12$ . The transition sizes  $n_t = 12, 64, 144, 240$  are the same for all three  $\psi_w/\psi_s$  ratios (as indicated). Successive nanofibril shapes are schematically shown to visualize the nanofibril transitions from  $1\beta$ - to  $2\beta$ -sheet, from  $2\beta$ - to  $3\beta$ -sheet, and so on. The smallest rectangle represents a single  $\beta$ -strand, and it is shown in red when it binds itself to an SH  $\beta$ -sheet surface and in blue when it binds itself to a WH one. The red and blue lines correspond to the formation work,  $w$ , of nanofibrils that form along a low- and high-energy pathway, respectively.

(Figure 1), we assume that for a  $1\beta$ -sheet, the SH surface is always on top (indicated by the red line), whereas the WH surface is at the bottom (indicated by the blue line). In addition, a  $\beta$ -strand can bind itself to a WH  $\beta$ -sheet surface only with its WH side, and to an SH  $\beta$ -sheet surface only with its SH side. Consequently, the hydrophobicity of the surface of a nanofibril alternates with increasing number of layers.

To write down a formula for the dimensionless work,  $w$ , for nanofibril formation, we need to consider two different cases. The first is the formation of a fibril with an odd number  $i$  of  $\beta$ -

sheets which, because of the assumed growth mechanism described above, always has one WH and one SH surface. Then  $w$  for sizes  $n = 1, 2, \dots, n_t$  can be written as ( $i = 1, 3, 5, \dots$ )

$$w(i, n) = -[s - (\psi_w + \psi_s)/i]n + 2\psi i \quad (4)$$

where  $s$  is the dimensionless supersaturation. The second case is the formation of a fibril with an even number  $i$  of  $\beta$ -sheets that can have either two WH or two SH surfaces. Then  $w$  for  $n = 1, 2, \dots, n_t$  can be written as ( $i = 2, 4, 6, \dots$ )

$$w(i, n) = -(s - 2\psi_w/i)n + 2\psi i \quad (5)$$

$$w(i, n) = -(s - 2\psi_s/i)n + 2\psi i \quad (6)$$

for a fibril with two WH or two SH surfaces, respectively. In eqs 4, 5, and 6, the term  $-sn$  is the energy gained by assembling  $n$  monomers from the solution into an  $n$ -size nanofibril, and the  $\psi_w$ ,  $\psi_s$ , and  $\psi$  terms represent the work done on creating the total surface area of a nanofibril with the shape of an  $i\beta$ -sheet without a  $\beta$ -strand on one of its two hydrophobic sides (in this case, the nanofibril length  $m = n/i$  is the same for all  $i\beta$ -sheets).

To write the formula for  $w$  for  $n = n_t + 1, n_t + 2, \dots, n_t + m_t$  (where  $m_t = n_t/i$  is the number of  $\beta$ -strands in any of the  $i$  equally long  $\beta$ -sheets constituting the  $i\beta$ -sheet-shaped nanofibril when the first  $\beta$ -strand of the subsequent  $\beta$ -sheet is attached to the nanofibril), we first consider the case of an  $i\beta$ -sheet with an odd number of  $\beta$ -sheets (with one WH surface and one SH surface) that transforms into an  $(i + 1)\beta$ -sheet where the new  $\beta$ -sheet forms by  $\beta$ -strand binding to the  $i\beta$ -sheet's either WH or SH side. Then  $w$  is given by ( $i = 1, 3, 5, \dots$ )

$$w(i, n) = -sn + \psi_s m_t + \psi_w [m_t - (n - im_t)] + \psi_s (n - im_t) + 2\psi(i + 1) \quad (7)$$

$$w(i, n) = -sn + \psi_w m_t + \psi_s [m_t - (n - im_t)] + \psi_w (n - im_t) + 2\psi(i + 1) \quad (8)$$

for the binding to the WH or SH side, respectively. In eqs 7 and 8, the third and fourth terms on the right side describe the change in the surface energy when the  $\beta$ -strands of the subsequent  $\beta$ -sheet attach to the WH or SH side of the nanofibril, respectively (the term  $(n - im_t)$  is the number of  $\beta$ -strands in the subsequent layer). When a new layer forms, the hydrophobicity of the surface of a nanofibril changes from weak to strong (or strong to weak; see Figure 1). In the second case, an  $i\beta$ -sheet composed of an even number of layers (with two either WH or SH surfaces) transforms into an  $(i + 1)\beta$ -sheet in which the new  $\beta$ -sheet forms by  $\beta$ -strand binding to one of the two  $i\beta$ -sheet sides, either WH or SH. Then  $w$  is given by ( $i = 2, 4, 6, \dots$ )

$$w(i, n) = -sn + \psi_w m_t + \psi_w [m_t - (n - im_t)] + \psi_s (n - im_t) + 2\psi(i + 1) \quad (9)$$

$$w(i, n) = -sn + \psi_s m_t + \psi_s [m_t - (n - im_t)] + \psi_w (n - im_t) + 2\psi(i + 1) \quad (10)$$

for binding to a WH or SH side, respectively.

Figure 2 displays the  $n$  dependence of  $w$  from eqs 4–10 for nanofibril formation with exemplifying transition sizes  $n_t = 12, 64, 144, 240$ . The lines are drawn with  $c_s = c_w = 0$  at  $\psi = 10$  and  $\psi_s = 1$  for  $\psi_w = 1, 0.7$ , and  $0.1$ ; that is, at three different ratios,



$\psi_w/\psi_s = 1, 0.7$ , and  $0.1$ , of the nanofibril hydrophobicity asymmetry. As seen in Figure 2a for the symmetric case ( $\psi_w/\psi_s = 1$ ), at  $s = 0.7\psi_s$ , according to eq 4,  $w$  for a single  $\beta$ -sheet ( $i = 1$ ) increases linearly until  $n$  reaches the exemplifying transition size  $n_t = 12$  corresponding to transition length  $m_t = 12$ . The transformation from a  $1\beta$ -sheet to a  $1\beta$ -sheet plus one  $\beta$ -strand on top of it corresponds to a sharp increase of  $w$  between  $n = 12$  and  $13$ . After this shape transformation, according to eq 7,  $w$  diminishes linearly with  $n$  up to  $n = n_t + m_t = 24$ . At this size, the nanofibril is composed of two equally long  $\beta$ -sheets and, according to eq 5,  $w$  for further elongation increases linearly again until the next exemplifying transition size,  $n_t = 64$ . The transformation from a  $2\beta$ -sheet to a  $2\beta$ -sheet plus one  $\beta$ -strand on top of it corresponds to a sharp increase of  $w$  between  $n = 64$  and  $65$ . After this shape transformation, according to eq 9,  $w$  diminishes linearly with  $n$  up to  $n = n_t + m_t = 96$ . After this size, the  $3\beta$ -sheet can elongate unlimitedly, because at  $s = 0.7\psi_s$ , according to eq 4, the work  $w(3,n)$  also decreases with  $n$ . In standard nucleation theory, the nanofibril of size  $n^* = n_t + 1 = 65$ , which requires the maximum formation work, is the fibril nucleus.

The case of weak hydrophobicity asymmetry is illustrated in Figure 2b at  $\psi_w/\psi_s = 0.7$  and  $s = 0.6\psi_s$ . As before, according to eq 4,  $w$  for a single  $\beta$ -sheet ( $i = 1$ ) increases linearly until  $n$  reaches the exemplifying transition size  $n_t = 12$ , corresponding to  $m_t = 12$ . The sharp increase in  $w$  between  $n = 12$  and  $13$  now depends on whether the  $1\beta$ -sheet transforms into a  $1\beta$ -sheet plus one  $\beta$ -strand on top or on bottom of the sheet. Binding of a  $\beta$ -strand with its SH side on top requires less work than binding of a  $\beta$ -strand with its WH side on bottom. Although the difference in  $w$  is small at  $n = n_t + 1 = 13$ , according to eqs 7 and 8, it does increase as  $w$  diminishes linearly with  $n$  (with different slopes) up to  $n = n_t + m_t = 24$ , until the second  $\beta$ -sheet is fully formed. At this size, the nanofibril is composed of two equally long  $\beta$ -sheets with two either SH or WH surfaces. According to eqs 5 and 6, during further nanofibril elongation,  $w$  increases linearly with different slopes until the next exemplifying transition size,  $n_t = 64$ , because it requires more work to form a fibril with two SH surfaces than a fibril with two WH surfaces. There is a sharp increase in  $w$  between  $n = 64$  and  $65$ , when the  $2\beta$ -sheet (with two either WH or SH sides) transforms into a  $2\beta$ -sheet plus one  $\beta$ -strand attached either on top or on bottom, and according to eqs 9 and 10,  $w$  diminishes linearly (with different slopes) with  $n$  up to  $n = n_t + m_t = 96$ . The  $3\beta$ -sheets formed along these two different pathways are now energetically identical because they both have the same fibril structure, with one WH and one SH surface. The  $3\beta$ -sheet can elongate unlimitedly, because according to eq 4, at  $s = 0.6\psi_s$ , the work  $w(3,n)$  also decreases with  $n$ . In this case, the nanofibril of size  $n^* = n_t + 1 = 65$ , which requires the maximum formation work, is also the fibril nucleus.

The  $w(n)$  dependence for the strongly asymmetric case, with  $\psi_w/\psi_s = 0.1$  at  $s = 0.12\psi_s$ , is illustrated in Figure 2c. The differences in  $w$  when fibrils with two SH or two WH surfaces form is now more pronounced. As can be seen in the figure, at the chosen  $s$  value, according to eq 5,  $w(2,n)$  decreases with  $n$  for a fully formed  $2\beta$ -sheet with two WH surfaces, and it can elongate unlimitedly. In contrast, according to eq 6,  $w(2,n)$  increases with  $n$  for a fully formed  $2\beta$ -sheet with two SH surfaces. Furthermore, according to eq 4,  $w(3,n)$  for a fully formed  $3\beta$ -sheet also increases with  $n$ . Similarly, a fully formed  $4\beta$ -sheet with two WH surfaces can elongate unlimitedly, whereas  $w(4,n)$  for a  $4\beta$ -sheet with two SH surfaces and  $w(5,n)$

for a fully formed  $5\beta$ -sheet increase with  $n$ . This implies that in a protein solution under such conditions, preferably  $2\beta$ -sheets with two WH surfaces nucleate and grow because  $2\beta$ -sheets with two SH surfaces and  $3\beta$ -sheets dissolve. The growing  $2\beta$ -sheets can transform into a  $4\beta$ -sheet with, again, two WH surfaces, but for this to happen, a potentially high nucleation barrier needs to be overcome (see Figure 2c).

The transformation of fibrils from  $2\beta$ -sheets to  $4\beta$ -sheets, and  $4\beta$ -sheets to  $6\beta$ -sheets was reported by Irback et al.<sup>23</sup> and observed in Monte Carlo simulations of essentially the same model peptide system. The underlying physical reason for such a behavior is that increasing the hydrophobicity asymmetry of  $\beta$ -sheet surfaces can make  $i\beta$ -sheets less soluble than  $(i + 1)\beta$ -sheets, that is, thinner nanofibrils are less soluble than thicker ones, which cannot be the case for the symmetric case (see, e.g., the simulations by Zhang et al.<sup>24</sup> on the nucleation and elongation of amyloid fibrils). This will be discussed in the following section.

**Fibril Solubility.** The coexistence supersaturations  $s_{i\beta}$  at which fibrils of fixed thickness of  $i\beta$ -sheets coexist with the protein solution are readily obtained by setting equal to zero the bracketed factor in eqs 4 to 6 because this factor is the driving force for fibril growth or dissolution. This is so, because at these  $s_{i\beta}$  values  $w(i,n)$  at fixed  $i$  do not change with  $n$ . Doing that leads to the formulas ( $i = 1, 3, 5, \dots$ )

$$s_{i\beta} = (\psi_w + \psi_s)/i \quad (11)$$

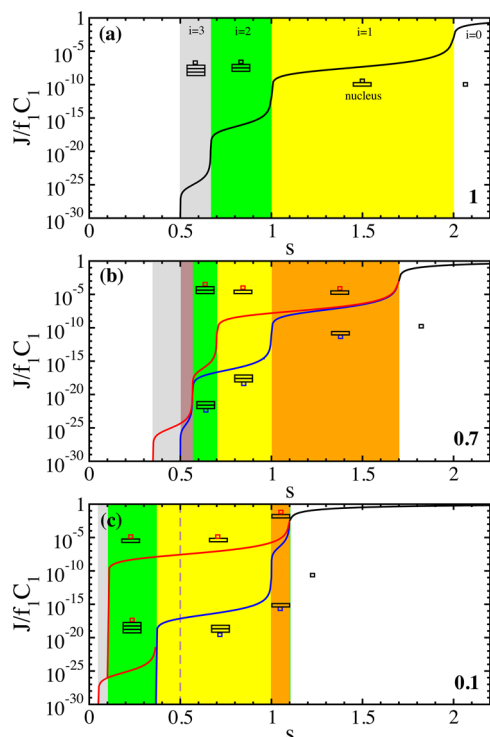
for  $i\beta$ -sheets with an odd number of layers and ( $i = 2, 4, 6, \dots$ )

$$s_{i\beta,w} = 2\psi_w/i \quad (12)$$

$$s_{i\beta,s} = 2\psi_s/i \quad (13)$$

for  $i\beta$ -sheets with an even number of  $\beta$ -sheets with two either WH or SH surfaces, respectively. For the symmetric case ( $\psi_w/\psi_s = 1$ ), the above equations yield  $s_{i\beta} = s_{i\beta,w} = s_{i\beta,s} = 2\psi_s/i$ ; that is, then the coexistence supersaturations for the fibrils of thickness  $i = 1, 2, 3$ , and  $4$  are given by  $s_{i\beta} = 2, 1, 2/3$ , and  $1/2$ , respectively. As indicated in Figure 3a, the  $s > s_{1\beta} = 2$  range (range  $i = 0$  in the figure) corresponds to the metanucleation regime in which each  $\beta$ -strand in the solution acts as a fibril nucleus. Similarly, when  $s > s_{2\beta} = 1$ , the  $2\beta$ -sheets can lengthen irreversibly. Thus, in the  $s_{2\beta} < s < s_{1\beta}$  range (range  $i = 1$  in the figure),  $1\beta$ -sheets tend to dissolve, and the fibril nucleus is a  $1\beta$ -sheet plus one  $\beta$ -strand attached to one of its two sides. The situation is analogous in the  $s_{3\beta} < s < s_{2\beta}$  range (range  $i = 2$ ) and the  $s_{4\beta} < s < s_{3\beta}$  range (range  $i = 3$ ) in which the fibril nucleus is a  $2\beta$ -sheet plus one  $\beta$ -strand attached sideways and a  $3\beta$ -sheet plus one  $\beta$ -strand attached sideways, respectively.

For the weakly asymmetric case (Figure 3b),  $\psi_w/\psi_s = 0.7$ , the coexistence supersaturations for  $i\beta$ -sheets of thickness  $i = 1$  and  $3$  are obtained from eq 11 and are  $s_{i\beta} = 1.7$  and  $0.57$ , respectively. The coexistence supersaturations for the fibrils of thickness  $i = 2$  and  $4$  with two WH surfaces are obtained from eq 12 and are  $s_{i\beta,w} = 0.7$  and  $0.35$ , respectively. Similarly, for fibrils with two SH surfaces, they are given by eq 13 and are  $s_{i\beta,s} = 1$  and  $0.5$ , respectively. At this weak asymmetry, there exists a hierarchy of the coexistence supersaturations,  $s_{4\beta,w} < s_{4\beta,s} < s_{3\beta} < s_{2\beta,w} < s_{2\beta,s} < s_{1\beta}$ , such that  $i\beta$ -sheets are more soluble than  $(i + 1)\beta$ -sheets. As indicated in Figure 3b, the  $s > s_{1\beta} = 1.7$  range (white area) corresponds to the metanucleation regime in which each  $\beta$ -strand in the solution acts as a fibril nucleus. When  $s > s_{2\beta,s} = 1$ , the  $2\beta$ -sheets with two either WH or two SH surfaces can lengthen irreversibly, and in the  $s_{2\beta,s} < s < s_{1\beta}$  range



**Figure 3.** Rate,  $J$ , of fibril nucleation as a function of the supersaturation,  $s$ , with  $c_s = c_w = 0$  at  $\psi = 10$  and  $\psi_s = 1$  for  $\psi_w/\psi_s = 1$  (panel a),  $\psi_w/\psi_s = 0.7$  (panel b), and  $\psi_w/\psi_s = 0.1$  (panel c). Shapes of the fibril nuclei are schematically shown in the different supersaturation ranges indicated by the white, orange, yellow, green, brown, and gray areas. The red and blue lines correspond to the nucleation rates obtained for fibril nuclei with a  $\beta$ -strand attached on the nucleus SH (red monomer) or WH (blue monomer) sides, respectively.

(orange area), the  $1\beta$ -sheet plus one  $\beta$ -strand attached to either side is a fibril nucleus. The situation changes when  $s > s_{2\beta w} = 0.7$ , because then only the  $2\beta$ -sheets with two WH surfaces can lengthen irreversibly, whereas the one with two SH surfaces dissolves. Thus, in the  $s_{2\beta w} < s < s_{2\beta s}$  range (yellow area) the  $1\beta$ -sheet plus one  $\beta$ -strand attached to the  $\beta$ -sheet SH side is the fibril nucleus. In addition, we need to consider that in this supersaturation range  $3\beta$ -sheets can also nucleate (although with a low rate) where the fibril nucleus is a  $2\beta$ -sheet with two SH sides plus one  $\beta$ -strand attached sideways is a fibril nucleus. When  $s > s_{3\beta} = 0.57$ , the  $3\beta$ -sheet can lengthen irreversibly, and in the  $s_{3\beta} < s < s_{2\beta w}$  range (green area), the  $2\beta$ -sheet with two either WH or SH surfaces plus one  $\beta$ -strand attached sideways is the fibril nucleus. The situation when  $s > s_{4\beta s} = 0.5$  is analogous to the one when  $s > s_{2\beta s}$ , and the situation when  $s > s_{4\beta w} = 0.35$  is analogous to the one when  $s > s_{2\beta w}$ .

For the strongly asymmetric case (Figure 3c),  $\psi_w/\psi_s = 0.1$ , the coexistence supersaturations for the fibrils of thickness  $i = 1$  and  $3$  are also given by eq 11 and are  $s_{i\beta} = 1.1$  and  $0.37$ , respectively. The coexistence supersaturations for the fibrils of thickness  $i = 2$  and  $4$  with two WH surfaces are given by eq 12 and are  $s_{i\beta w} = 0.1$  and  $0.05$ , respectively, and with two SH surfaces, they are given by eq 13 and are  $s_{i\beta s} = 1$  and  $0.5$ , respectively. At this strong asymmetry, the hierarchy of the coexistence supersaturations,  $s_{4\beta w} < s_{2\beta w} < s_{3\beta} < s_{4\beta s} < s_{2\beta s} < s_{1\beta}$ , is such that  $i\beta$ -sheets can be less soluble than  $(i + 1)\beta$ -sheets. In particular, a  $2\beta$ -sheet with two WH surfaces is less soluble than a  $3\beta$ -sheet and a  $4\beta$ -sheet with two SH surfaces. As can be seen

in Figure 3c, although the threshold supersaturations have been shifted to lower values, for  $s > s_{3\beta} = 0.37$ , the situation is analogous to that of the weakly asymmetric case in the sense that the fibril nuclei are the same in the white, orange, and yellow  $s$  ranges. When  $s > s_{2\beta w} = 0.1$ , however, a  $2\beta$ -sheet and a  $4\beta$ -sheet, both with two WH surfaces, can lengthen irreversibly. Thus, in the  $s_{2\beta w} < s < s_{3\beta}$  range (green area), the  $1\beta$ -sheet plus one  $\beta$ -strand attached to the sheet SH side and the  $3\beta$ -sheet plus one  $\beta$ -strand attached to the sheet SH side are the fibril nuclei. In this asymmetry ratio,  $s_{4\beta s} = 0.5$  (indicated by the dashed line) does not play a role. When  $s > s_{4\beta w} = 0.05$ , the  $4\beta$ -sheets with two WH surfaces can lengthen irreversibly, and in the  $s_{4\beta w} < s < s_{2\beta w}$  range, (gray area) the  $3\beta$ -sheet plus one  $\beta$ -strand attached to the SH side is the fibril nucleus.

The threshold value of the asymmetry ratio above which there exists a hierarchy of the coexistence supersaturations, that is,  $s_{4\beta w} < s_{4\beta s} < s_{3\beta} < s_{2\beta w} < s_{2\beta s} < s_{1\beta}$ , where a  $(i + 1)\beta$ -sheet is less soluble than a  $i\beta$ -sheet, is of practical relevance because it allows the sheet to preferably nucleate and grow fibrils with a particular morphology, that is,  $2\beta$ -sheets with two weak hydrophobic surfaces (see Nucleation Work). This threshold value can readily be obtained by setting equal the coexistence supersaturations  $s_{2\beta w}$  and  $s_{3\beta}$  from eqs 11 and 12, which gives  $\psi_w/\psi_s = 0.5$ .

**Length Distribution of Nuclei.** According to CNT in its self-consistent formulation, the equilibrium concentration  $C_m^*$  of nuclei is given by the Boltzmann-type formula  $C_m^* = C_1 e^{w(1,1) - w^*}$ , where  $w(1,1)$  is the dimensionless work to form the smallest nanofibril of one  $\beta$ -strand, and  $w^* \equiv w(i, n^*)$  is the dimensionless nucleation work. From eq 4, we have  $w(1,1) = -s + \psi_s + \psi_w + 2\psi$ . Setting  $n = n^*$  and using  $n^* = n_t + 1 = im_t + 1$  in eqs 7 and 8, we obtain the following general formulas for the nucleation work of fibril nuclei composed of an odd number of  $i\beta$ -sheets plus one  $\beta$ -strand attached to their SH ( $i = 1, 3, 5, \dots$ ;  $m_t = 1, 2, 3, \dots$ ;  $2\psi_w/(i + 1) < s < (\psi_s + \psi_w)/i$ ),

$$w^* = (\psi_w + \psi_s - is)m_t - \psi_s + \psi_w - s + 2\psi(i + 1) \quad (14)$$

or WH ( $i = 1, 3, 5, \dots$ ;  $m_t = 1, 2, 3, \dots$ ;  $2\psi_s/(i + 1) < s < (\psi_s + \psi_w)/i$ ) side.

$$w^* = (\psi_w + \psi_s - is)m_t - \psi_w + \psi_s - s + 2\psi(i + 1) \quad (15)$$

Similarly, from eqs 9 and 10, we obtain the nucleation work for fibril nuclei with an even number of  $i\beta$ -sheets plus one  $\beta$ -strand attached to either one of the  $\beta$ -sheet two WH ( $i = 2, 4, 6, \dots$ ;  $m_t = 1, 2, 3, \dots$ ;  $(\psi_s + \psi_w)/(i + 1) < s < 2\psi_w/i$ ),

$$w^* = (2\psi_w - is)m_t - \psi_w + \psi_s - s + 2\psi(i + 1) \quad (16)$$

or two SH ( $i = 2, 4, 6, \dots$ ;  $m_t = 1, 2, 3, \dots$ ;  $(\psi_s + \psi_w)/(i + 1) < s < 2\psi_s/i$ ) sides.

$$w^* = (2\psi_s - is)m_t - \psi_s + \psi_w - s + 2\psi(i + 1) \quad (17)$$

Using  $w(1,1)$  from above, and eqs 14–17, we now obtain the corresponding formulas for the equilibrium concentration of fibril nuclei composed of an odd number of  $i\beta$ -sheets plus one  $\beta$ -strand attached to the  $\beta$ -sheet WH ( $i = 1, 3, 5, \dots$ ;  $m_t = 1, 2, 3, \dots$ ;  $2\psi_s/(i + 1) < s < (\psi_s + \psi_w)/i$ ),

$$C_{m_t}^* = m_t C_1 e^{[is - (\psi_w + \psi_s)]m_t - 2\psi i + 2\psi_w} \quad (18)$$

or SH ( $i = 1, 3, 5, \dots; m_t = 1, 2, 3, \dots; 2\psi_w/(i+1) < s < (\psi_s + \psi_w)/i$ ) side,

$$C_{m_t}^* = m_t C_1 e^{[is - (\psi_w + \psi_s)]m_t - 2\psi_i + 2\psi_s} \quad (19)$$

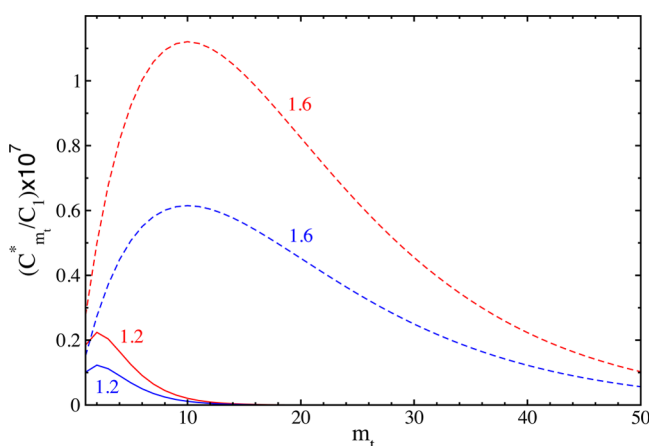
and ( $i = 2, 4, 6, \dots; m_t = 1, 2, 3, \dots; (\psi_s + \psi_w)/(i+1) < s < 2\psi_s/i$ ),

$$C_{m_t}^* = 2m_t C_1 e^{[is - 2\psi_w]m_t - 2\psi_i + 2\psi_w} \quad (20)$$

for fibril nuclei with an even number of  $i\beta$ -sheets plus one  $\beta$ -strand attached to either one of the  $\beta$ -sheet two WH sides or ( $i = 2, 4, 6, \dots; m_t = 1, 2, 3, \dots; (\psi_s + \psi_w)/(i+1) < s < 2\psi_s/i$ ),

$$C_{m_t}^* = 2m_t C_1 e^{[is - 2\psi_s]m_t - 2\psi_i + 2\psi_s} \quad (21)$$

when the  $\beta$ -strand attached to either one or the other of the  $\beta$ -sheet two SH sides. The factors  $m_t$  and  $2m_t$  in these expressions take into account the fact that the  $\beta$ -strand can attach itself to  $m_t$  or  $2m_t$  equivalent sites. Figure 4 displays  $C_{m_t}^*$  from eqs 18 and



**Figure 4.** Length distribution of fibril nuclei with thickness of one  $\beta$ -sheet: lines 1.2 and 1.6 are at  $s/\psi_s = 1.2$  and 1.6, respectively;  $i = 1$ ;  $\psi = 10$ ;  $\psi_s = 1$ ; and  $\psi_w = 0.7$ . The red and blue lines correspond to fibril nuclei with a  $\beta$ -strand on the SH (red) or WH (blue) side of the  $\beta$ -sheet, respectively.

19 at  $\psi = 10$ ,  $\psi_s = 1$ ,  $\psi_w = 0.7$ , and  $s/\psi_s = 1.2$  and 1.6 (as indicated). As can be seen from the figure, there exists a broad distribution of transition lengths,  $m_t$ , at which a  $1\beta$ -sheet transforms into a  $2\beta$ -sheet. In addition,  $C_{m_t}^*$  passes through a maximum at  $m_t = m_t^*$ . We note that the length  $m_t = m_t^*$  of the most numerous nuclei and their concentration  $C_{m_t^*}^*$  can easily be calculated from the condition for maximum of  $C_{m_t^*}^*$ , but we do not show the formulas here. As already mentioned earlier, such a distribution is in contrast to CNT, which assumes that the nucleus has a uniquely defined size. The concentration of fibril nuclei with thickness of one  $\beta$ -sheet plus one  $\beta$ -strand attached to the  $\beta$ -sheet SH side is higher than that when the  $\beta$ -strand is attached to the  $\beta$ -sheet WH side, whereas  $m_t^*$  is similar for both cases. Increasing  $s$  shifts  $m_t^*$  to larger sizes and broadens the distribution.

To calculate the fibril nucleation rate, we need to know the total concentration  $C^* = \sum_{m_t=1}^{\infty} C_{m_t}^*$  of nuclei of all lengths. From eqs 18–21, with the aid of the exact formula  $\sum_{p=1}^{\infty} px^p = x/(1-x)^2$ , it follows that  $C^*$  is given by ( $i = 1, 3, 5, \dots; m_t = 1, 2, 3, \dots; 2\psi_w/(i+1) < s < (\psi_s + \psi_w)/i$ ),

$$C^* = C_1 e^{i(s-2\psi)-\psi_w+\psi_s}/(1 - e^{is-(\psi_w+\psi_s)})^2 \quad (22)$$

for fibril nuclei composed of an odd number of  $\beta$ -sheets plus one  $\beta$ -strand attached to the  $\beta$ -sheet SH or ( $i = 1, 3, 5, \dots; m_t = 1, 2, 3, \dots; 2\psi_s/(i+1) < s < (\psi_s + \psi_w)/i$ )

$$C^* = C_1 e^{i(s-2\psi)-\psi_s+\psi_w}/(1 - e^{is-(\psi_w+\psi_s)})^2 \quad (23)$$

two WH side, respectively, and by ( $i = 2, 4, 6, \dots; m_t = 1, 2, 3, \dots; (\psi_s + \psi_w)/(i+1) < s < 2\psi_s/i$ ,

$$C^* = 2C_1 e^{i(s-2\psi)}/(1 - e^{is-2\psi_w})^2 \quad (24)$$

for fibril nuclei with an even number of  $\beta$ -sheets plus one  $\beta$ -strand attached to either one of the  $\beta$ -sheet two WH or ( $i = 2, 4, 6, \dots; m_t = 1, 2, 3, \dots; (\psi_s + \psi_w)/(i+1) < s < 2\psi_w/i$ ):

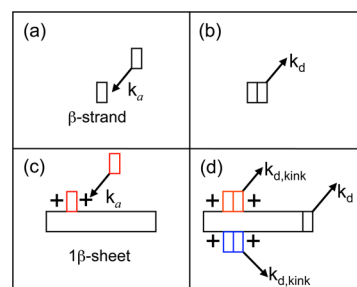
$$C^* = 2C_1 e^{i(s-2\psi)}/(1 - e^{is-2\psi_s})^2 \quad (25)$$

two SH sides.

**Nucleation Rate.** The fibril nucleation rate,  $J$  ( $\text{m}^{-3} \text{s}^{-1}$ ), is the frequency of appearance of supernucleus fibrils per unit solution volume. We first determine  $J$  in the metanucleation range  $s > \psi_w + \psi_s$ , in which there is no nucleation barrier for fibril formation and a single  $\beta$ -strand in the solution is the fibril nucleus. Then  $J$  is the net number  $f_1 - g_2$  of monomers that attach per unit time to one of the two strong bonding sides of a monomer ( $s > \psi_w + \psi_s$ ) (ref 14).

$$J = (f_1 - g_2)C_1 \quad (26)$$

Here, the monomer-to-monomer attachment frequency  $f_1 = 2k_a$  is twice the frequency  $k_a$  ( $\text{s}^{-1}$ ) of monomer attachment to one of the two strong-bond sides of a monomer (Figure 5a), and



**Figure 5.** Schematic illustration of the monomer attachment and detachment frequencies  $k_a$ ,  $k_d$  and  $k_{d,kink}$ . The red and blue monomers indicate the attachment to the fibril SH and WH side, respectively. The pluses indicate the nanofibril kink sites.

the detachment frequency,  $g_2 = 2k_d$ , is twice the frequency  $k_d$  ( $\text{s}^{-1}$ ) of monomer detachment from one of the two strong-bond sides of a monomer (Figure 5b). Here,  $k_a$  is expected to increase exponentially with  $s$  according to (ref 10)

$$k_a(s) = k_e e^s \quad (27)$$

where  $k_e$  is the value of  $k_a$  at equilibrium (then  $s = 0$ ). To obtain an expression for the frequency of monomer detachment from one of the two strong-bond sides of a monomer, we derive a general expression for the detachment frequency  $k_d(b, b_w, b_s)$  of a monomer from the fibril periphery. As known from Boltzmann statistics, the probability for transition from an initial to a final state decreases exponentially with the work spent on the transition. The detachment of a peripheral monomer requires the work  $2b\psi + 2b_w\psi_w + 2b_s\psi_s$ , if the detaching  $\beta$ -strand is bound to neighboring  $\beta$ -strands with  $b$ ,



$b_w$  and  $b_s$  bonds of energies  $2\psi$ ,  $2\psi_w$ , and  $2\psi_s$  per bond, respectively. Therefore, the frequency of monomer detachment  $k_d(b, b_w, b_s)$  from the fibril periphery can be written as  $k_d(b, b_w, b_s) = k_{\text{ref}} \exp(-2b\psi - 2b_w\psi_w - 2b_s\psi_s)$ , where  $k_{\text{ref}}$  is a reference frequency. Elimination of  $k_{\text{ref}}$  can be achieved with the aid of the relation  $k_{d,\text{kink}} = k_a(0)$  for the equality between the frequency for monomer detachment from a kink site and the frequency of monomer attachment at equilibrium. Fibrils with one SH and one WH surface, however, have two such kink sites (Figure 5d), with detachment frequencies  $k_d(1,1,0)$  and  $k_d(1,0,1)$ , and the work required to remove a monomer from them is different. Thus, for the frequency for monomer detachment from a kink site, we use the geometric average  $k_{d,\text{kink}} = (k_d(1,1,0) k_d(1,0,1))^{1/2} = k_{\text{ref}} \exp(-2\psi - \psi_s - \psi_w)$ . Using this equation and eq 27 in  $k_{d,\text{kink}} = k_a(0)$  leads to  $k_{\text{ref}} = k_e \exp(2\psi + \psi_w + \psi_s)$  so that the general equation for the detachment frequency becomes

$$k_d(b, b_w, b_s) = k_e e^{2\psi(1-b) + \psi_w(1-2b_w) + \psi_s(1-2b_s)} \quad (28)$$

The expression for monomer detachment from one of the two strong-bond sides of a monomer can then easily be obtained from eq 28 and is

$$k_d(1, 0, 0) = k_e e^{\psi_w + \psi_s} \quad (29)$$

Using the above expressions for  $f_1$ ,  $g_2$ ,  $k_a$ , and  $k_d$  in eq 26 yields the general formula for the fibril metanucleation rate ( $s > \psi_w + \psi_s$ )

$$J = f_1 C_1 (1 - e^{\psi_s + \psi_w - s}) \quad (30)$$

Similarly, we can determine  $J$  in the nucleation regime using

$$J = (f^* - g^*) C^* \quad (31)$$

where  $f^*$  and  $g^*$  are the frequencies of monomer attachment and detachment to and from the nucleus two kink sites (Figure 5c,d), respectively, and  $C^*$  is the total concentration of nuclei. As before, we set  $f^* = f_1 = 2k_a$ ,  $k_a$  being given by eq 27. Concerning  $g^* = 2k_{d,\text{eff}}$  we need to take into account that this frequency depends on the nucleus thickness, that is, the number  $i \geq 1$  of  $\beta$ -sheets constituting the nucleus. If the nucleus is composed of a  $1\beta$ -sheet plus one  $\beta$ -strand on its SH side,  $k_{d,\text{eff}} = (k_d(1,0,0) k_d(1,0,1))^{1/2} = k_e e^{\psi_w}$  is the geometric average of the detachment frequencies  $k_d(1,0,0)$  from one of the two strong-bond sides of a fibril and the detachment frequency  $k_d(1,0,1)$  of a monomer from the fibril top kink site. Similarly, if the nucleus is composed of a  $3\beta$ -sheet plus one  $\beta$ -strand on its SH surface,  $k_{d,\text{eff}} = (k_d(1,0,0) k_d(1,0,1) k_d(1,1,0) k_d(1,0,1))^{1/4} = k_e e^{\psi_w/2}$  is the geometric average of the detachment frequencies  $k_d(1,0,0)$  from one of the two strong-bond sides of a fibril,  $k_d(1,0,1)$  from the fibril top kink site,  $k_d(1,1,0)$  from the fibril bottom kink site, and  $k_d(1,0,1)$  from the fibril top kink site. The general formula of the monomer detachment frequency from a nucleus composed of an odd number  $i = 1, 3, \dots$  of  $\beta$ -sheets plus one  $\beta$ -strand on its top SH side is  $k_{d,\text{eff}} = k_e e^{2\psi_w/(i+1)}$ . Using the above expressions for  $f^*$ ,  $g^*$ ,  $k_a$ , and  $k_{d,\text{eff}}$  in eq 31 yields the corresponding general formula for the fibril nucleation rate ( $i = 1, 3, 5, \dots$ ;  $2\psi_w/(i+1) < s < (\psi_s + \psi_w)/i$ ).

$$J = C_1 f_1 e^{i(s-2\psi) - \psi_w + \psi_s} \frac{1 - e^{[2\psi_w/(i+1)] - s}}{(1 - e^{i s - (\psi_w + \psi_s)})^2} \quad (32)$$

Analogously, it can be shown that when the nucleus is composed of an odd number ( $i = 1, 3, 5, \dots$ ) of  $\beta$ -sheets plus one  $\beta$ -strand bound to the  $\beta$ -sheet WH side,  $J$  is given by ( $i = 1, 3, 5, \dots$ ;  $2\psi_s/(i+1) < s < (\psi_s + \psi_w)/i$ )

$$J = C_1 f_1 e^{i(s-2\psi) - \psi_s + \psi_w} \frac{1 - e^{[2\psi_s/(i+1)] - s}}{(1 - e^{i s - (\psi_w + \psi_s)})^2} \quad (33)$$

Next, we consider the case when the nucleus is composed of a  $2\beta$ -sheet plus one  $\beta$ -strand on either of its two SH or WH sides. In both cases,  $k_{d,\text{eff}} = (k_d(1,0,0) k_d(1,1,0) k_d(1,0,1))^{1/3} = k_e e^{(\psi_w + \psi_s)/2}$ , and the general formula for the effective monomer detachment frequency from nuclei is  $k_{d,\text{eff}} = k_e e^{(\psi_w + \psi_s)/(i+1)}$ . Using the above expressions for  $f^*$ ,  $g^*$ ,  $k_a$ ,  $k_{d,\text{eff}}$  in eq 31 yields the general formula for the fibril nucleation rate ( $i = 2, 4, 6, \dots$ ;  $(\psi_s + \psi_w)/(i+1) < s < 2\psi_w/i$ )

$$J = 2C_1 f_1 e^{i(s-2\psi)} \frac{1 - e^{[(\psi_s + \psi_w)/(i+1)] - s}}{(1 - e^{i s - 2\psi_w})^2} \quad (34)$$

when the fibril nucleus is composed of an even number  $i$  of  $\beta$ -sheets plus one  $\beta$ -strand on either of its two WH sides, or by ( $i = 2, 4, 6, \dots$ ;  $(\psi_s + \psi_w)/(i+1) < s < 2\psi_s/i$ )

$$J = 2C_1 f_1 e^{i(s-2\psi)} \frac{1 - e^{[(\psi_s + \psi_w)/(i+1)] - s}}{(1 - e^{i s - 2\psi_s})^2} \quad (35)$$

when the  $\beta$ -strand is on either of its two SH sides. Equations 30 and 32–35 are the central result of the present study.

Figure 3 shows the dependence of the fibril nucleation rate,  $J$ , on supersaturation,  $s$ , obtained from eqs 30 and 32–35 with  $c_s = c_w = 0$  at  $\psi = 10$ ,  $\psi_s = 1$  for asymmetry ratios  $\psi_w/\psi_s = 1, 0.7, 0.1$  (as indicated). As shown in Fibril Solubility, for the symmetric case,  $\psi_w/\psi_s = 1$ , the coexistence supersaturations for the fibrils of thickness  $i = 1, 2, 3$ , and  $4$  are given by  $s_{i\beta} = 2, 1, 2/3$ , and  $1/2$ , respectively. As indicated in Figure 3a, the rate  $J$  (eq 30) in the  $s > s_{1\beta}$  range (range  $i = 0$  in the figure) corresponds to the metanucleation regime in which each  $\beta$ -strand in the solution acts as a fibril nucleus. In the nucleation regime,  $J$  from eqs 32–35 are identical and exhibit jumps at the transition supersaturations  $s_{i\beta} = 2, 1, 2/3$ , and  $1/2$ . This is so because at these supersaturations, the nucleus changes from, for example, a monomer to a  $1\beta$ -sheet plus one  $\beta$ -strand at the nucleation/metanucleation border  $s_{1\beta} = 2$  or in general from an  $i\beta$ -sheet plus one  $\beta$ -strand to an  $(i+1)\beta$ -sheet plus one  $\beta$ -strand at  $s_{i\beta}$ .

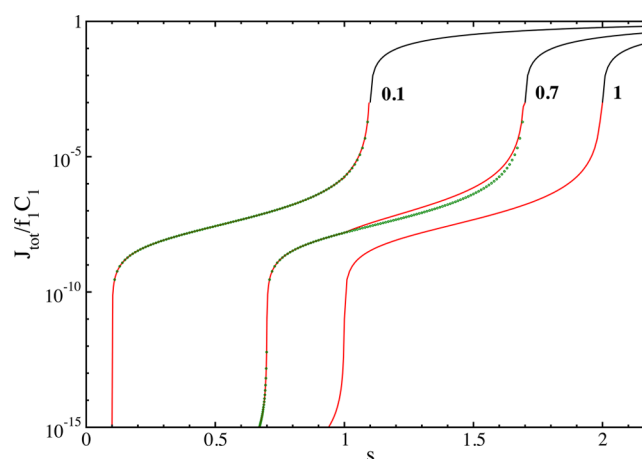
For the weakly asymmetric case,  $\psi_w/\psi_s = 0.7$ , as shown in Fibril Solubility, there exists a hierarchy of the coexistence supersaturations,  $s_{4\beta w} < s_{4\beta s} < s_{3\beta} < s_{2\beta w} < s_{2\beta s} < s_{1\beta}$ , such that  $i\beta$ -sheets are more soluble than  $(i+1)\beta$ -sheets. In the metanucleation range,  $s > s_{1\beta} = 1.7$ , the nucleation rate  $J$  is again described by eq 30. In the  $s_{2\beta s} < s < s_{1\beta}$  range (orange area in Figure 3b), when the fibril nucleus is either a  $1\beta$ -sheet plus one  $\beta$ -strand attached to the  $\beta$ -sheet either WH or SH side,  $J$  is described by eqs 32 and 33, respectively. In this supersaturation range, the difference between the two rates is small, mainly because the work to form the corresponding fibril nuclei is similar. The situation changes in the  $s_{2\beta w} < s < s_{2\beta s}$  range (yellow area in Figure 3b) because in this range, the fibril nuclei are  $1\beta$ -sheets plus one  $\beta$ -strand attached to their SH side and  $2\beta$ -sheets with two WH surfaces plus one  $\beta$ -strand attached to either surface. For the former,  $J$  is described by eq 32 and is many orders of magnitude higher than  $J$  for the latter (described by eq 34). In the  $s_{3\beta} < s < s_{2\beta w}$  range (green area in Figure 3b), the fibril nuclei are  $2\beta$ -sheets plus one  $\beta$ -strand

attached to either one of the two WH or SH sides, and  $J$  is described by eqs 34 and 35, respectively. In this supersaturation range, the difference between the two rates is small again because the work to form the two different types of nuclei is similar. In the  $s_{4\beta s} < s < s_{3\beta}$  range (brown area in Figure 3b),  $J$  is again described by eqs 32 and 33, as in the  $s_{2\beta s} < s < s_{1\beta}$  range, and the difference between  $J$  for fibrils with two either WH or SH sides becomes large again.

The strongly asymmetric case,  $\psi_w/\psi_s = 0.1$ , is analogous to the weakly asymmetric case for supersaturations  $s > s_{3\beta}$  (the yellow, orange, and white areas in Figure 3c). The main effect of increasing the hydrophobicity asymmetry is that the nucleation rate resulting from a fibril nucleus composed of a  $1\beta$ -sheet plus one  $\beta$ -strand attached to the SH side is orders of magnitude higher than that of fibrils in which the fibril nucleus is a  $1\beta$ -sheet plus one  $\beta$ -strand attached to the WH side or a  $2\beta$ -sheet with two WH surfaces plus one  $\beta$ -strand attached to either surface. In the  $s_{2\beta w} < s < s_{3\beta}$  range (green area in Figure 3c), as discussed in Fibril Solubility, the fibril nuclei are composed of a  $1\beta$ -sheet plus one  $\beta$ -strand attached to the SH side and a  $3\beta$ -sheet plus one  $\beta$ -strand attached to the SH side. In this  $s$  range, the nucleation rate,  $J$ , resulting from a fibril nucleus composed of a  $1\beta$ -sheet plus one  $\beta$ -strand attached to the SH side is orders of magnitude higher than the one due to nuclei with the shape of a  $3\beta$ -sheet plus one  $\beta$ -strand attached to the SH side. The prediction of  $J$  quantifies our previous suggestion in Nucleation Work (Figure 2c) that in a protein solution with strongly asymmetric hydrophobicity of  $\beta$ -strands, over a wide supersaturation range (the green and yellow areas in Figure 3c), preferably  $2\beta$ -sheets with WH surfaces nucleate and grow.

Finally, we note that experimentally, it is not possible to directly measure nucleation rates for fibrils with different structures but, rather, the overall nucleation rate. For example, in the case illustrated in Figure 3b, in the orange range, it is not possible to directly measure the nucleation rate resulting from  $2\beta$ -sheets with two WH surfaces and to distinguish it from that due to  $2\beta$ -sheets with two SH ones. Instead, the overall nucleation rate  $J_{\text{tot}}$  is measured, which is the sum of the two rates. To illustrate the main effect of increasing the hydrophobicity asymmetry, we compare in Figure 6 the  $s$  dependence of  $J_{\text{tot}}$  obtained in each supersaturation regime with  $c_s = c_w = 0$  at  $\psi = 10$ ,  $\psi_s = 1$  and  $\psi_w/\psi_s = 1, 0.7, 0.1$  (as indicated). The main effect is that lowering  $\psi_w$  can increase the overall nucleation rate by many orders of magnitude at a given  $s$  value. As can be seen in the Figure, in particular in the low supersaturation range,  $0.1 < s < 0.7$ , whereas  $J_{\text{tot}}$  for the symmetric and weakly asymmetric case is practically zero, for the strongly asymmetric case  $J_{\text{tot}}$  is high enough such that it can be detected experimentally. As we have found in a previous work,<sup>14</sup> the main reason for this is the dependence of the threshold supersaturation  $s_{1\beta} = e^{\psi_w + \psi_s}$  at the nucleation/metanucleation border on both  $\psi_s$  and  $\psi_w$ . Thus, at a given  $\psi_s$  value, lowering  $\psi_w$  shifts  $s_{1\beta}$  to lower values, thereby narrowing the nucleation range and broadening the metanucleation one. Furthermore, in this figure, we also show the predictions of the dominating  $J$  in each supersaturation range (red lines in Figures 3b and 3c), showing that they are almost identical to  $J_{\text{tot}}$ .

To illustrate and understand the practical implications of the effect of asymmetric hydrophobicity of the  $\beta$ -sheet surfaces, in the next section, we apply the developed formulas to a model



**Figure 6.** Overall nucleation rate,  $J_{\text{tot}}$ , for fibril formation as a function of the supersaturation,  $s$ , with  $c_s = c_w = 0$  at  $\psi = 10$ ,  $\psi_s = 1$  for  $\psi_w = 1, 0.7, 0.1$  corresponding to asymmetry ratios  $\psi_w/\psi_s = 1, 0.7, 0.1$  (as indicated). The black and red lines indicate the rate in the metanucleation and nucleation regimes, respectively. The dotted green line indicates the highest rate from eq 38 in the corresponding  $s$  range (red lines from Figure 3b and 3c).

peptide system and predict absolute values for the fibril nucleation rate.

**Application to a Model Peptide System.** Experiments on protein aggregation are often performed at fixed temperature,  $T$ . Then the supersaturation can be controlled by the monomer protein concentration,  $C_1$ , because<sup>11</sup>

$$s = \ln(C_1/C_e) \quad (36)$$

where  $C_e$  is the fibril solubility. This enables us to derive formulas for the  $C_1$  dependence of the nucleation rate,  $J$ . Using eq 36 in eq 27 and combining it with the relation  $f_1 = 2k_a$  yields the monomer attachment frequency,  $f_1 = 2k_e C_1/C_e$ , as a function of  $C_1$ . Then the  $C_1$  dependence of the metanucleation rate obtained from eq 30 is ( $C_1 > C_e e^{\psi_w + \psi_s}$ )

$$J = A_1 C_1^2 (1 - A_2 C_1^{-1}) \quad (37)$$

where  $A_1 = 2k_e/C_e$  and  $A_2 = C_e e^{\psi_w + \psi_s}$ .

Similarly, we obtain the formula for  $J$  in the nucleation regime when the fibril nuclei are composed of an odd number  $i$  of  $\beta$ -sheets plus one  $\beta$ -strand ( $i = 1, 3, 5, \dots$ ;  $C_e e^{2\psi_w/(i+1)} < C_1 < C_e e^{(\psi_w + \psi_s)/i}$  or  $C_e e^{2\psi_w/(i+1)} < C_1 < C_e e^{(\psi_w + \psi_s)/i}$ ).

$$J = A_1 C_1^{i+2} \frac{1 - A_2 C_1^{-1}}{(1 - A_3 C_1)^2} \quad (38)$$

Here,  $A_1 = (2k_e/C_e^{i+1})e^{-2\psi_i - \psi_w + \psi_s}$  or  $A_1 = (2k_e/C_e^{i+1})e^{-2\psi_i - \psi_s + \psi_w}$  when the  $\beta$ -strand is on the nucleus SH or WH side, respectively, and  $A_2$  and  $A_3$  are given by  $A_2 = C_e e^{2\psi_w/(i+1)}$  and  $A_3 = C_e e^{-(\psi_w + \psi_s)}$  in both cases.

Furthermore, when the fibril nuclei are composed of an even number  $i$  of  $\beta$ -sheets plus one  $\beta$ -strand, the fibril nucleation rate is given again by eq 38, but with  $i = 2, 4, 6, \dots$ ;  $C_e e^{(\psi_w + \psi_s)/(i+1)} < C_1 < C_e e^{2\psi_w/i}$  or  $C_e e^{(\psi_w + \psi_s)/(i+1)} < C_1 < C_e e^{2\psi_w/i}$ ,  $A_1 = (4k_e/C_e^{i+1})e^{-2\psi_i}$  and  $A_2 = C_e e^{(\psi_w + \psi_s)/(i+1)}$ . As to the constant  $A_3$ , it is different when the  $\beta$ -strand is on one of the nucleus two WH or SH sides:  $A_3 = C_e e^{-2\psi_w}$  or  $A_3 = C_e e^{-2\psi_s}$  in the former or the latter case, respectively.

Equations 37 and 38 are also a central result in the present study, and in the following, we apply them to a model peptide



system. The peptides are composed of 10 amino acids that can arrange in their fully extended conformation as  $\beta$ -strands in a nanosized amyloid fibril (protofilament) of successively layered  $\beta$ -sheets with alternating WH and SH surfaces (Figure 1). The  $\beta$ -strand length is approximately  $d_0 = 0.4$  nm (e.g., ref 7), the known value for the distance between the  $\beta$ -strands in a  $\beta$ -sheet is  $d_h = 0.5$  nm (e.g., ref 7), and the intersheet distance in the fibril is  $d = 1$  nm. Taking these values together with  $\sigma = 20$  mJ/m<sup>2</sup> (estimated value, ref 25) and  $\sigma_s = 4$  mJ/m<sup>2</sup> (e.g., ref 26), it follows from eqs 1 and 2 that  $\psi = 19$ ,  $\psi_s = 1.9$ , and  $\psi/\psi_s = 10$  for this model peptide at  $T = 300$  K. Importantly, because the binding energy of  $\beta$ -strands within the fibrils is different for each  $\psi_w = 1.9$ , 1.33, and 0.19 corresponding to asymmetry ratios  $\psi_w/\psi_s = 1$ , 0.7, and 0.1, respectively, we expect that the fibril solubility  $C_e$  is different. Because the effect of asymmetry of the two hydrophobic  $\beta$ -sheet surfaces on  $C_e$  has not been determined experimentally, we estimate it theoretically by making use of the van't Hoff equation and the Haas–Drenth lattice model for protein crystals.<sup>27</sup> The integrated van't Hoff equation is given by  $C_e = C_r e^{-\lambda}$ , where  $C_r$  is a practically temperature-independent reference concentration and  $\lambda = L/kT$  is the dimensionless latent heat of peptide aggregation into  $\beta$ -sheets. Here,  $L$  is the latent heat of peptide aggregation into such aggregates. In the Haas–Drenth lattice model for protein crystals,  $\lambda$  is half the binding energy of peptides in the aggregates, which is equivalent to the broken bond energy  $\lambda = 2\psi + \psi_s + \psi_w$  at the periphery of a fibril in the  $m, i$  plane. The fibril solubility is then given by

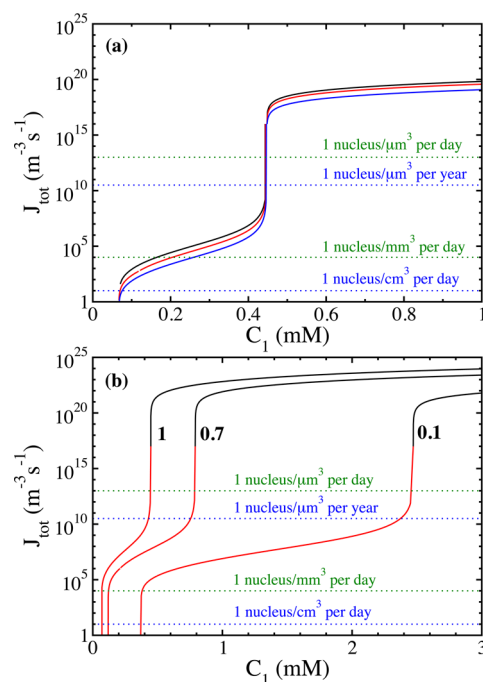
$$C_e = C_r e^{-2\psi - \psi_s - \psi_w} \quad (39)$$

where  $\psi = \psi_{hb} + c_s \psi_s + c_w \psi_w$ . Assuming that  $C_e = 6.0 \times 10^{21}$  m<sup>-3</sup> ( $= 10$   $\mu$ M) (e.g., ref 28) for fibrils composed of short peptides and that the coefficients are  $c_s = c_w = 0$ , we can calculate  $C_r$  to be  $8.6 \times 10^{39}$  m<sup>-3</sup>. Using this  $C_r$  value and  $\psi = 19$ ,  $\psi_s = 1.9$  in eq 39, the solubilities for fibrils with  $\psi_w = 1.33$  and 0.19 (corresponding to asymmetry ratio 0.7 and 0.1) are  $C_e = 1.1 \times 10^{22}$  m<sup>-3</sup> ( $= 18$   $\mu$ M) and  $C_e = 3.3 \times 10^{22}$  m<sup>-3</sup> ( $= 55$   $\mu$ M), respectively. Thus, lowering  $\psi_w$  at constant  $\psi_s$  shifts  $C_e$  to higher concentrations. Assuming that  $k_e = 10^{-4}$  s<sup>-1</sup> (e.g., ref 29) enables us to calculate the  $J(C_1)$  dependence from eq 38 and, thus,  $J_{\text{tot}}(C_1)$  with  $c_s = c_w = 0$  at  $\psi = 19$  and  $\psi_s = 1.9$  for  $\psi_w = 1.9$ , 1.33, 0.19 (Figure 7a). The characteristic feature of the  $J_{\text{tot}}(C_1)$  dependence is the sharp rise (10 orders of magnitude) at the nucleation/metanucleation border over a very narrow concentration range. Below this threshold concentration, fibrils can be nucleated within a day only in volumes of 1 mm<sup>3</sup> or larger, used in in vitro experiments. In volumes of  $\sim 1$   $\mu$ m<sup>3</sup> or smaller, comparable to that of biological cells, the fibril nucleation rate is practically negligible (less than 1 nucleus per year). This illustrates the important role of the threshold concentration  $C_{1\beta}$  in amyloid fibril nucleation, which by combining eq 11 with eq 36 is given by

$$C_{1\beta} = C_e e^{\psi_s + \psi_w} \quad (40)$$

Combining eqs 39 and 40 shows that the threshold concentration  $C_{1\beta}$  does not change when lowering  $\psi_w$  at constant  $\psi_s$  because the shift of  $C_e$  to higher concentrations compensates the shift of  $C_{1\beta}$  lower ones.

The important role of the fibril solubility in amyloid fibril nucleation can be further illustrated when we consider that the binding energy in the direction of the fibril axis is a combination of hydrogen bonding energy and the hydrophobic



**Figure 7.** (a) Concentration dependence of the overall nucleation rate  $J_{\text{tot}}$  for the model peptide system with  $c_s = c_w = 0$  at  $\psi = 19$ ,  $\psi_s = 1.9$  for  $\psi_w = 1.9$  (black line), 1.33 (red line), and 0.19 (blue line) corresponding to asymmetry ratios  $\psi_w/\psi_s = 1$ , 0.7, 0.1. (b) Concentration dependence of the overall nucleation rate  $J_{\text{tot}}$  for the model peptide system with  $c_s = c_w = 0.5$  at  $\psi = 19$ ,  $\psi_s = 1.9$  for  $\psi_w = 1.9$  (asymmetry ratio  $\psi_w/\psi_s = 1$ ), at  $\psi = 18.72$ ,  $\psi_s = 1.9$  for  $\psi_w = 1.33$  (asymmetry ratio  $\psi_w/\psi_s = 0.7$ ), and at  $\psi = 18.15$ ,  $\psi_s = 1.9$  for  $\psi_w = 0.19$  (asymmetry ratio  $\psi_w/\psi_s = 0.1$ ), as indicated. The black and red lines indicate the rate in the metanucleation and nucleation regimes, respectively.

energies, that is,  $\psi = \psi_{hb} + (\psi_s + \psi_w)/2$ , where we set  $c_s = c_w = 0.5$ . Using again that  $\psi = 19$ ,  $\psi_s = 1.9$  for this model peptide at  $T = 300$  K and considering that for the symmetric case  $\psi_w = 1.9$ ,  $\psi_w = 1.9$ , we calculate that  $\psi_{hb} = 17.1$ . Thus, the  $\psi$  values with  $c_s = c_w = 0$  at  $\psi_{hb} = 17.1$ ,  $\psi_w = 19$  for  $\psi_w = 1.33$  and 0.19 (corresponding to asymmetry ratio 0.7 and 0.1) are  $\psi = 18.7$  and 18.2, respectively. Assuming  $m^{-3} (= 10$   $\mu$ M) (e.g., ref 28) for fibrils composed of short peptides, we calculate from eq 39 that  $C_r = 8.6 \times 10^{39}$  m<sup>-3</sup>, as above. Substituting this  $C_r$  value into eq 39, the solubilities for such fibrils with  $\psi_w = 1.33$  and 0.19 (corresponding to asymmetry ratio 0.7 and 0.1) are  $C_e = 1.9 \times 10^{22}$  m<sup>-3</sup> ( $= 31$   $\mu$ M) and  $C_e = 1.8 \times 10^{23}$  m<sup>-3</sup> ( $= 306$   $\mu$ M), respectively. Thus, considering the contribution of hydrophobicity to  $\psi$  when lowering  $\psi_w$  shifts  $C_e$  to substantially higher concentrations, as when they are neglected. With  $k_e = 10^{-4}$  s<sup>-1</sup> (e.g., ref 29), we again calculate the  $J(C_1)$  dependence from eq 38 and, thus,  $J_{\text{tot}}(C_1)$  with  $c_s = c_w = 0.5$  for asymmetry ratios  $\psi_w/\psi_s = 1$ , 0.7, 0.1 (Figure 7b). As can be seen from this figure, the main effect of increasing the asymmetry (decreasing  $\psi_w$  at constant  $\psi_s$ ) is to shift  $C_{1\beta}$  to a higher concentration and to hamper protein fibrillation because metanucleation commences at higher  $C_1$  values. Using the  $C_e$  values calculated in eq 40, the threshold concentrations for fibrils with asymmetry ratios  $\psi_w/\psi_s = 1$ , 0.7, and 0.1 are  $C_{1\beta} = 2.7 \times 10^{23}$  m<sup>-3</sup> ( $= 45$   $\mu$ M),  $C_{1\beta} = 4.8 \times 10^{23}$  m<sup>-3</sup> ( $= 0.79$  mM), and  $C_{1\beta} = 1.5 \times 10^{24}$  m<sup>-3</sup> ( $= 2.47$  mM), respectively.

The central result of this discussion is that increasing the hydrophobicity asymmetry (decreasing  $\psi_w$  at constant  $\psi_s$ ) can

shift  $C_{1\beta}$  to higher concentrations, thereby hampering protein fibrillation, because metanucleation commences at higher  $C_1$  values. This effect is entirely due to the effect of asymmetry of the two hydrophobic  $\beta$ -sheet surfaces on the fibril solubility, because it shifts  $C_e$  to higher concentrations. This effect becomes more pronounced if we consider that the binding energy in the direction of the fibril axis is a combination of hydrogen bonding energy and the hydrophobic energies.

## CONCLUSIONS

In this study, we provide insight into the thermodynamics and kinetics of the direct polymerization of peptides into nanofibrils composed of successively layered  $\beta$ -sheets with alternating WH and SH surfaces.

Analysis of the work,  $w$ , for fibril formation reveals that an  $n$ -sized nanofibril can form by two polymerization pathways. Along the first pathway, a  $1\beta$ -sheet forms first, which then changes its shape to that of a  $2\beta$ -sheet with two SH surfaces, which then changes its shape to a  $3\beta$ -sheet with one WH and one SH surface, and then transforms into a  $4\beta$ -sheet with two SH surfaces. This process continues in the same way during the further fibril thickening. Along the second pathway, the  $1\beta$ -sheet changes its shape to that of a  $2\beta$ -sheet with two WH surfaces, which changes its shape to a  $3\beta$ -sheet with one WH and one SH surface, and then transforms into a  $4\beta$ -sheet with two WH surfaces. Because the surface tension of fibrils with two WH surfaces is lower than that of fibrils with two SH surfaces, the corresponding nucleation work,  $w^*$ , is lower, and the fibrils preferably nucleate along the second pathway. The larger the asymmetry of the surface energies of the two hydrophobic surfaces (i.e., the smaller the ratio  $\psi_w/\psi_s$ ), the less likely it is that fibrils form by the first pathway.

Regarding the thermodynamics, analysis of the work  $w$  for fibril formation provides expressions for the threshold supersaturations  $s_{i\beta}$  above which  $i\beta$ -sheets can elongate irreversibly. This analysis shows that there exists a threshold hydrophobicity asymmetry ratio  $\psi_w/\psi_s = 0.5$  above which the hierarchy of the coexistence supersaturations is such that thicker  $i\beta$ -sheets are less soluble than thinner ones (i.e.,  $s_{3\beta} < s_{2\beta w} < s_{2\beta s} < s_{1\beta}$ ). In contrast, below this threshold ratio, this hierarchy becomes distorted such that thicker  $i\beta$ -sheets can be more soluble than thinner ones. In particular, we show that  $s_{2\beta w} < s_{3\beta} < s_{4\beta s} < s_{2\beta s} < s_{1\beta}$  at  $\psi_w/\psi_s = 0.1$ , implying that a  $2\beta$ -sheet with two WH surfaces is less soluble than a  $3\beta$ -sheet with one WH and one SH surface and a  $4\beta$ -sheet with two SH surfaces. Importantly, the order of the threshold supersaturations  $s_{i\beta}$  determines the type of fibril nuclei in a given  $s$  range. For example, in the range  $s_{2\beta s} < s < s_{1\beta}$ , the  $1\beta$ -sheet tends to dissolve, whereas a  $2\beta$ -sheet with two WH surfaces can irreversibly elongate, and we found that the  $1\beta$ -sheet plus one  $\beta$ -strand attached to the sheet either WH or SH surface is the fibril nucleus.

In each of the supersaturation ranges, there exists an equilibrium concentration,  $C_{m_i}^*$ , of nuclei with a broad distribution of transition lengths,  $m_i$ , at which an  $i\beta$ -sheet transforms into an  $(i + 1)\beta$ -sheet. In the example above, the  $1\beta$ -sheet can transform into a  $2\beta$ -sheet at any transition length,  $m_i$ . Such a distribution is in contrast to classical nucleation theory, which assumes that the nucleus has a uniquely defined size. Because the asymmetry of the surface energies of the two hydrophobic surfaces, there exist two such distribution functions in a given  $s$  range. Because the nucleation work,  $w^*$ , of fibrils with two WH surfaces is less than that of fibrils

with two SH surfaces, the concentration of the first kind of fibril nuclei in the protein solution is much higher than that of the second kind of nuclei.

Regarding the kinetics of fibril formation, we obtain general formulas for the  $J(s)$  dependence given by eqs 30 and 32–35. In each  $s$  range, there exist two nucleation rates that are due to the fact that fibrils can nucleate by the two pathways described above. A comparison between the  $J(s)$  dependences of fibrils that nucleate along the first and the second pathways, and the overall nucleation rate  $J_{\text{tot}}$  (the sum of the two rates) shows that  $J_{\text{tot}}$  is almost identical to  $J$  for the second pathway over the whole supersaturation range. So kinetically, there is a segregation (or morphological selection) between the two types of fibrils, and only one type is relevant. The characteristic feature of the  $J_{\text{tot}}(s)$  dependence is a sharp drop at the nucleation/metanucleation border. We show that the main effect of increasing the hydrophobicity asymmetry, that is, of lowering  $\psi_w$  at a given  $\psi_s$  value, is a shift of  $s_{1\beta}$  to lower  $s$  values and, thereby, a stimulation of protein fibrillation. However, application of our considerations to a model peptide system (and to derivation of a general formula for the  $J(C_1)$  dependence, eq 38) shows that increasing the hydrophobicity asymmetry (decreasing  $\psi_w$  at constant  $\psi_s$ ) can shift  $C_{1\beta}$  to higher concentrations and hamper protein fibrillation because metanucleation commences at higher  $C_1$  values. This effect is entirely due to the effect of asymmetry of the two hydrophobic  $\beta$ -sheet surfaces on the fibril solubility because it shifts  $C_e$  to higher concentrations. This effect becomes more pronounced if we consider that the binding energy in direction of the fibril axis is a combination of hydrogen bonding energy and the hydrophobic energies.

Finally, we emphasize that the results obtained above apply to one-step fibril nucleation, that is, when the monomeric proteins polymerize directly into fibrils, and that the analysis treats homogeneous nucleation of amyloid fibrils occurring when nucleation-active foreign particles or substrates are absent from the solution. The derived general expressions for the fibril nucleation rate  $J$  as an explicit function of the supersaturation are a step toward first-principle predictions of fibril nucleation rates based on the amino acid sequence of the proteins. The application to specific protein systems, however, requires knowledge of the fibrils' structure, the fibril solubility, and the broken bond energies of amino acids (or alternatively the surface tensions of the fibril lengthening and thickening axis). It should be mentioned that entropic effects such as the loss of entropy when a  $\beta$ -strand is attached to the fibril or the entropy due to vibrations of the  $\beta$ -strand within fibrils are not explicitly considered. However, entropic effects on the fibril solubility and entropic contributions to the broken bond energies are automatically accounted for when experimental data for  $C_e$  and  $\psi$ ,  $\psi_h$  are used.

The results obtained in this study highlight the importance of the fibril solubility in amyloid fibril nucleation. Currently, the focus of experimental studies on amyloids is mostly on their structure, assembly mechanism, and their interactions with the biological environment. Not so many experiments focus on determining the fibril solubility and how it changes with the fibril structure and amino acid sequence. Similarly, computational and theoretical studies of amyloids focus less on determining the fibril solubility, and a better understanding is urgently needed. Nevertheless, the analysis made offers answers to the fundamental question of the effect of asymmetry between the weak and strong hydrophobic  $\beta$ -sheet surfaces on

the thermodynamics and kinetics of the polymerization process, which could be a helpful guide in studying the intriguing phenomenon of amyloid fibril nucleation.

## AUTHOR INFORMATION

### Corresponding Author

\*E-mail: s.auer@leeds.ac.uk.

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

The author thanks Dr. Richard Bingham and Professor Dimo Kashchiev for stimulating discussions during the course of this study and for their comments on the manuscript.

## REFERENCES

- (1) Chiti, F.; Dobson, C. M. Protein Misfolding, Functional Amyloid, and Human Disease. *Annu. Rev. Biochem.* **2006**, *75*, 333–366.
- (2) Karran, E.; Mercken, M.; De Strooper, B. The Amyloid Cascade Hypothesis for Alzheimer's Disease: An Appraisal for the Development of Therapeutics. *Nat. Rev. Drug Discovery* **2011**, *10*, 698–712.
- (3) Makin, O. S.; Atkins, E.; Sikorski, P.; Johansson, J.; Serpell, L. C. Molecular Basis for Amyloid Fibril Formation and Stability. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 315–320.
- (4) Sawaya, M. R.; Sambashivan, S.; Nelson, R.; Ivanova, M. I.; Sievers, S. A.; Apostol, M. I.; Thompson, M. J.; Balbirnie, M.; Wiltzius, J. J. W.; McFarlane, H. T.; Madsen, A. O.; Riekel, C.; Eisenberg, D. Atomic Structures of Amyloid Cross- $\beta$  Spines Reveal Varied Steric Zippers. *Nature* **2007**, *447*, 453–457.
- (5) Tycko, R. Molecular Structure of Amyloid Fibrils: Insights from Solid-State NMR. *Q. Rev. Biophys.* **2006**, *39*, 1–55.
- (6) Dobson, C. M. Protein Misfolding, Evolution and Disease. *Trends Biochem. Sci.* **1999**, *24*, 329–332.
- (7) Cheng, P. N.; Pham, J. D.; Nowick, J. S. The Supramolecular Chemistry of  $\beta$ -Sheets. *J. Am. Chem. Soc.* **2013**, *135*, 5477–5492.
- (8) Abraham, F. F. *Homogeneous Nucleation Theory*; Academic: New York, 1974.
- (9) Kashchiev, D. *Nucleation: Basic Theory with Applications*; Butterworth-Heinemann: Oxford, 2000.
- (10) Cabriolu, R.; Kashchiev, D.; Auer, S. Breakdown of Nucleation Theory for Crystals with Strongly Anisotropic Interactions Between Molecules. *J. Chem. Phys.* **2012**, *137*, 204903.
- (11) Kashchiev, D.; Auer, S. Nucleation of Amyloid Fibrils. *J. Chem. Phys.* **2010**, *132*, 215101.
- (12) Cabriolu, R.; Kashchiev, D.; Auer, S. Atomistic Theory of Amyloid Fibril Nucleation. *J. Chem. Phys.* **2010**, *133*, 225101.
- (13) Cabriolu, R.; Kashchiev, D.; Auer, S. Size Distribution of Amyloid Nanofibrils. *Biophys. J.* **2011**, *101*, 2232–2241.
- (14) Kashchiev, D.; Cabriolu, R.; Auer, S. Confounding the Paradigm: Peculiarities of Amyloid Fibril Nucleation. *J. Am. Chem. Soc.* **2013**, *135*, 1531–1539.
- (15) Chiti, F.; Stefani, M.; Taddei, N.; Ramponi, G.; Dobson, C. M. Rationalization of the Effects of Mutations on Peptide and Protein Aggregation Rates. *Nature* **2003**, *424*, 805–808.
- (16) Trovato, A.; Chiti, F.; Maritan, A.; Seno, F. Insight into the Structure of Amyloid Fibrils from the Analysis of Globular Proteins. *PLOS Comput. Biol.* **2006**, *2*, 1608–1618.
- (17) Fernandez-Escamilla, A. M.; Rousseau, F.; Schymkowitz, J.; Serrano, L. Prediction of Sequence-Dependent and Mutational Effects on the Aggregation of Peptides and Proteins. *Nat. Biotechnol.* **2004**, *22*, 1302–1306.
- (18) Tartaglia, G. G.; Pawar, A. P.; Campioni, S.; Dobson, C. M.; Chiti, F.; Vendruscolo, M. Prediction of Aggregation-Prone Regions in Structured Proteins. *J. Mol. Biol.* **2008**, *380*, 425–436.
- (19) Gillam, J. E.; Macphree, C. E. Modelling Amyloid Fibril Formation Kinetics: Mechanisms of Nucleation and Growth. *J. Phys.: Condens. Matter* **2013**, *25*, 373101.
- (20) Cohen, S. I. A.; Vendruscolo, M.; Dobson, C. M.; Knowles, T. P. J. From Macroscopic Measurements to Microscopic Mechanisms of Protein Aggregation. *J. Mol. Biol.* **2012**, *421*, 160–171.
- (21) Morris, A. M.; Watzky, M. A.; Finke, R. G. Protein Aggregation Kinetics, Mechanism, And Curve-Fitting: A Review of the Literature. *Biochim. Biophys. Acta, Proteins Proteomics* **2009**, *1794*, 375–397.
- (22) Cabriolu, R.; Auer, S. Amyloid Fibrillation Kinetics: Insight from Atomistic Nucleation Theory. *J. Mol. Biol.* **2011**, *411*, 275–285.
- (23) Irback, A.; Jonsson, S. A. E.; Linnemann, N.; Linse, B.; Wallin, S. Aggregate Geometry in Amyloid Fibril Nucleation. *Phys. Rev. Lett.* **2013**, *110*, 058101.
- (24) Zhang, J. N.; Muthukumar, M. Simulations of Nucleation and Elongation of Amyloid Fibrils. *J. Chem. Phys.* **2009**, *130*, 035102.
- (25) Auer, S.; Ricchiuto, P.; Kashchiev, D. Two-Step Nucleation of Amyloid Fibrils: Omnipresent or Not? *J. Mol. Biol.* **2012**, *422*, 723–730.
- (26) Auer, S. Phase diagram of polypeptide chains. *J. Chem. Phys.* **2011**, *135*, 175103.
- (27) Haas, C.; Drenth, J. The Interaction Energy between 2 Protein Molecules Related to Physical-Properties of Their Solution and Their Crystals and Implications for Crystal-Growth. *J. Cryst. Growth* **1995**, *154*, 126135.
- (28) Aggeli, A.; Nyrkova, I. A.; Bell, M.; Harding, R.; Carrick, L.; McLeish, T. C. B.; Semenov, A. N.; Boden, N. Hierarchical Self-Assembly of Chiral Rod-Like Molecules as a Model for Peptide  $\beta$ -Sheet Tapes, Ribbons, Fibrils, And Fibers. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 11857–11862.
- (29) Knowles, T. P. J.; Shu, W. M.; Devlin, G. L.; Meehan, S.; Auer, S.; Dobson, C. M.; Welland, M. E. Kinetics and Thermodynamics of Amyloid Formation from Direct Measurements of Fluctuations in Fibril Mass. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 10016–10021.