

Kinetic theory of amyloid fibril templating

Jeremy D. Schmit

Citation: *J. Chem. Phys.* **138**, 185102 (2013); doi: 10.1063/1.4803658

View online: <http://dx.doi.org/10.1063/1.4803658>

View Table of Contents: <http://jcp.aip.org/resource/1/JCPSA6/v138/i18>

Published by the [AIP Publishing LLC](#).

Additional information on *J. Chem. Phys.*

Journal Homepage: <http://jcp.aip.org/>

Journal Information: http://jcp.aip.org/about/about_the_journal

Top downloads: http://jcp.aip.org/features/most_downloaded

Information for Authors: <http://jcp.aip.org/authors>

ADVERTISEMENT



 **RUN YOUR GPU
CODE 2X FASTER.
TRY A TESLA K20 GPU
ACCELERATOR TODAY.
FREE.**

Kinetic theory of amyloid fibril templating

Jeremy D. Schmit^{a)}

Department of Physics, Kansas State University, Manhattan, Kansas 66506, USA

(Received 18 February 2013; accepted 18 April 2013; published online 14 May 2013)

The growth of amyloid fibrils requires a disordered or partially unfolded protein to bind to the fibril and adapt the same conformation and alignment established by the fibril template. Since the H-bonds stabilizing the fibril are interchangeable, it is inevitable that H-bonds form between incorrect pairs of amino acids which are either incorporated into the fibril as defects or must be broken before the correct alignment can be found. This process is modeled by mapping the formation and breakage of H-bonds to a one-dimensional random walk. The resulting microscopic model of fibril growth is governed by two timescales: the diffusion time of the monomeric proteins, and the time required for incorrectly bound proteins to unbind from the fibril. The theory predicts that the Arrhenius behavior observed in experiments is due to off-pathway states rather than an on-pathway transition state. The predicted growth rates are in qualitative agreement with experiments on insulin fibril growth rates as a function of protein concentration, denaturant concentration, and temperature. These results suggest a templating mechanism where steric clashes due to a single mis-aligned molecule prevent the binding of additional molecules. © 2013 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution 3.0 Unported License. [<http://dx.doi.org/10.1063/1.4803658>]

I. INTRODUCTION

Polypeptide chains are known to aggregate into filamentous structures called amyloids that have been the subject of intense research due to their potential as novel materials,¹ their detrimental effects on the stability of peptide pharmaceuticals,² and their role in diseases such as Alzheimer's and diabetes.³ In particular, the kinetics of fibril formation have been extensively studied^{4–9} to understand the stochasticity of disease onset and the lifetimes of metastable states that are thought to play a role in disease progression.¹⁰ These kinetics are difficult to study because the timescales for disease progression, typically years to decades, are prohibitive for direct biophysical studies under physiological conditions. Because of this, *in vitro* protocols utilize elevated concentrations and/or destabilizing conditions to accelerate the aggregation process. To extract the physiological relevance of these experiments it will be necessary to extrapolate these results toward *in vivo* conditions using theoretical models of the aggregation process. At present, the theory of aggregation is not sufficient for this purpose. Part of the problem is that the timescales of even *in vitro* aggregation exceed what is accessible by molecular simulation. Some of this timescale gap can be closed using coarse-grained models. Such models have revealed a subtle competition between H-bonds and hydrophobic interactions as well as a variety of potential assembly pathways.^{11–14} Yet, it can be difficult to generalize these results to the large phase space of concentrations and solution conditions that is explored in experiments. In this paper I present a microscopic model of the assembly process that explains the origin of the long aggregation timescales and shows

how temperature, denaturants, and protein size affect growth rates.

While fibrils can grow either by the attachment of single proteins to the fibril ends, or by the coalescence of small oligomers,^{15–18} this paper addresses only the former process. There are two reasons for this simplification. First, the low protein concentrations found *in vivo*^{19,20} will greatly favor monomeric pathways relative to the high concentration systems studied *in vitro*.²¹ Second, the reduced complexity of monomeric growth makes the first-order process an obvious starting point.

Nuclear magnetic resonance (NMR) and X-ray structures of amyloids show highly ordered structures, usually with the polypeptides forming parallel, in-register β -sheets that span the length of the fibril.^{22–24} This suggests a subtle competition between the specificity of the side chains and the periodicity of the backbone H-bonds. This paper posits that the nearly crystalline state is achieved by an exhaustive search of the potential alignments. The resulting theory has two timescales governing fibril growth; the diffusion time for an unbound polypeptide, and the time required for incorrectly aligned polypeptides to unbind from the fibril. However, the growth rate is much slower than either of these due to the low probability of a correctly aligned binding event.

II. MODEL

A. Fibril order is determined by a competition between sidechain specificity and backbone periodicity

Consider a polypeptide consisting of L amino acids that is part of an assembled amyloid fibril. If this molecule adopts the optimal alignment and conformation it will have a free

^{a)}schmit@phys.ksu.edu



energy $\epsilon_0 N_\beta$ reflecting the fact that N_β peptide groups are buried within the cross- β core while $L - N_\beta$ amino acids remain solvent exposed and are insignificantly perturbed from the soluble state. Here ϵ_0 is the total free energy change associated with adding a peptide group to the cross- β core, which includes the formation of the backbone H-bonds, the loss of conformational entropy, and sidechain packing interactions. Due to the translational symmetry of the peptide backbone, this molecule can also satisfy backbone H-bonding requirements by shifting the entire molecule by one peptide unit in either direction or by switching from a parallel to anti-parallel β -sheet. While these additional alignments are fully consistent with the β -sheet structure of the fibril, presumably they are higher in free energy due to sub-optimal sidechain packing interactions. These interactions increase the free energy per peptide group to ϵ_1 . Within this simple model, the partition function describing the various alignments for a single molecule within the fibril is

$$Q = e^{-N_\beta \epsilon_0 / RT} + (2(L - N_\beta) - 1)e^{-N_\beta \epsilon_1 / RT} + 4 \frac{e^{-N_\beta \epsilon_1 / RT} - e^{-\epsilon_1 / RT}}{e^{-\epsilon_1 / RT} - 1}. \quad (1)$$

Here the first term represents the optimal alignment, the second term represents the remaining ways of forming parallel and anti-parallel β -sheets that satisfy all N_β backbone H-bonds, and the final term represents alignments where some H-bond groups in the neighboring molecules in the fibril go unsatisfied.

If the system is allowed to reach equilibrium, the fraction of molecules in the optimal alignment is given by $e^{-N_\beta \epsilon_0 / RT} / Q$. In the presently available structures of amyloid fibrils, this number is very close to unity, meaning that

$$e^{-N_\beta \epsilon_0 / RT} \gg \left[(2(L - N_\beta) - 1)e^{-N_\beta \epsilon_1 / RT} + 4 \frac{e^{-N_\beta \epsilon_1 / RT} - e^{-\epsilon_1 / RT}}{e^{-\epsilon_1 / RT} - 1} \right]. \quad (2)$$

This places a lower bound on the free energy difference $\epsilon_1 - \epsilon_0$ required to achieve specificity. However, this lower bound would only be sufficient to produce ordered fibrils in the limit where the system is given an infinite amount of time to equilibrate.²⁵ On shorter timescales, soluble molecules will sample each of the alignments described in Eq. (1) and ordered assembly will only be attained if the “correct” alignment can be discriminated from the large ensemble of sub-optimal alignments on the timescale of fibril assembly.

There are two fundamental timescales governing the process by which incoming molecules attempt to attach to a growing fibril. The first of these is t_{diff} which is the time required for a soluble molecule to diffuse close enough to the fibril end to make a H-bond. A rough estimate of t_{diff} can be obtained from the reaction rate for particles approaching an absorbing sphere, $4\pi acD_p$, where a is the radius of the absorbing surface, c is the concentration far from the surface, and D_p is the diffusion constant of the particles. To describe the growth rates of fibrils bound to a surface^{26,27} this must be

reduced by half to account for the 2π solid angle restriction

$$t_{\text{diff}} = (2\pi acD_p)^{-1}, \quad (3)$$

where $a \simeq 2$ nm is the protofilament radius.^{28,29}

The second timescale is $t_{\text{bond}}(\epsilon)$, which is the average lifetime of a molecule attached to the fibril end before thermal fluctuations sever the bonds allowing it to diffuse away. This lifetime is a function of ϵ , the strength of the bonds holding the molecule onto the fibril end. A requirement for successful templating is that $t_{\text{bond}}(\epsilon_0) > t_{\text{bond}}(\epsilon_1)$. Section II B is devoted to calculating the residence time t_{bond} to determine how it depends on the binding affinity.

B. The residence time can be described as a random walk

Upon making contact, the polypeptide will form H-bonds between its backbone and the exposed peptide groups on the fibril end (Fig. 1). If, at a given time, the polypeptide makes x H-bonds with the fibril ($x = 2$ in the top panels of Fig. 1) the system can evolve in one of two ways; the polypeptide can form another H-bond with the fibril (Fig. 1, bottom panels) leading to $x + 1$ H-bonds, or the polypeptide can break a H-bond leading to $x - 1$ bonds. Thus, the variable x performs a random walk bounded between the extreme values of $x = 0$ (an unbound polypeptide) and $x = N_\beta$.³⁰

I define k_+ as the rate at which a polypeptide that is partially bound to the fibril end forms an additional H-bond (the random walk moves to the right) and k_- as the rate of H-bond breakage (the random walk moves to the left). These rates can be related by the detailed balance condition $p_x k_+ = p_{x+1} k_-$, where p_x is the probability that the random walk is at x in a hypothetical system that is allowed to reach equilibrium. These probabilities are given by Boltzmann weights, meaning the ratio of the rate constants can be expressed in terms of the free energy change upon the formation of the H-bond

$$k_+ / k_- = e^{-\epsilon / RT}, \quad (4)$$

where $\epsilon = \epsilon_0$ or ϵ_1 .

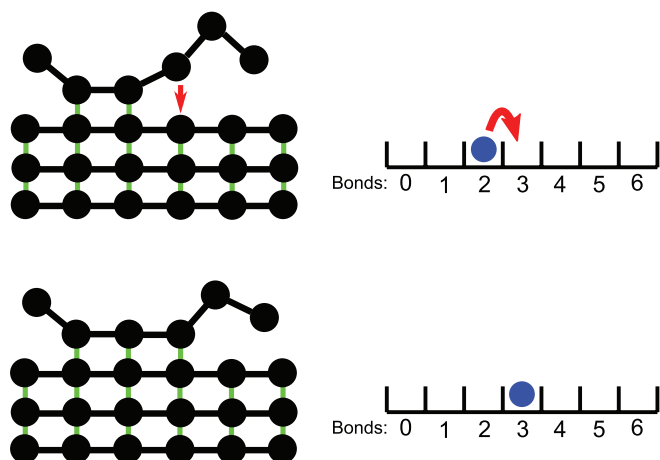


FIG. 1. A polypeptide can make or break H-bonds with the template established by the fibril end (left). The dynamics of this process can be described as a 1D random walk, where the position of the walker (right) describes the number of H-bonds. The interaction energy from the H-bonds imposes a drift velocity that tends to push the walker to the right.

The random walk in x is described by a diffusion constant $D = (k_+ + k_-)/2$ and a drift velocity $v = k_+ - k_-$. The statistics of this random walk are computed in the Appendix. Briefly, t_{bond} is equivalent to the mean first passage time for a particle starting at $x = 1$ to reach an absorbing boundary at $x = 0$ with a reflecting boundary at $x = N_\beta$. This time is³¹

$$t_{\text{bond}} = -\frac{1}{v} + \frac{D}{v^2} e^{vN_\beta/D} (1 - e^{-v/D}). \quad (5)$$

It will also be useful to calculate the time t_{attach} required for a newly bound molecule to form the full complement of N_β H-bonds with the fibril end, and the time t_{rel} which is the average lifetime of a bound molecule that fails to form the full complement of H-bonds. $t_{\text{attach}}(t_{\text{rel}})$ corresponds to the mean first passage time of a particle starting at $x = 1$ that reaches an absorbing boundary at $x = N_\beta$ ($x = 0$) without contacting the opposite boundary. These times are³¹

$$t_{\text{attach}} = \frac{N_\beta e^{-vN_\beta/D} + 1}{v} \frac{1 - e^{-v/D}}{1 - e^{-vN_\beta/D}} + \frac{1 - e^{-v/D}}{v}, \quad (6)$$

$$t_{\text{rel}} = \frac{1 - e^{-v/D} + e^{-vN_\beta/D} - 2P_+ N_\beta e^{-vN_\beta/D}}{v}, \quad (7)$$

where the probability that $x = N_\beta$ is reached first is³¹

$$P_+ = \frac{e^{-v/D} - 1}{e^{-vN_\beta/D} - 1}. \quad (8)$$

It is instructive to look at these times for typical amyloid parameters $N_\beta \simeq 25$, $\epsilon \simeq 1$ kJ/mol,³² and $k_+ \simeq 1$ ns.³³ The attachment and release times are both very fast $t_{\text{attach}} \simeq 50$ ns, $t_{\text{rel}} \simeq 2$ ns, indicating that the formation of H-bonds is a downhill process and that if a molecule fails to form a full set of N_β H-bonds it is because it detached before it had an opportunity to form more than one or two bonds. On the other hand, t_{bond} is much slower, occurring on the millisecond timescale, due to the low probability of simultaneously breaking all N_β H-bonds.

C. Ordered fibril growth requires incoming molecules to sample a large number of potential alignments

I assume that the first H-bond between a soluble molecule and fibril end is equally likely to involve any of the N_β peptide groups on the fibril end and the L amino acids on the soluble molecule. Therefore, there are LN_β permutations for the first bond. Of these, only the N_β permutations where the incoming amino acid bonds in-register with the matching residue on the fibril end can result in ordered growth. Ordered growth also requires that the polypeptide adopt the same orientation as the rest of the fibril. Without sacrificing generality, I assume the orientation to be parallel, rather than anti-parallel, β -sheets. The probability that the incoming polypeptide binds in the parallel, in-register alignment is $(2L)^{-1}$. Therefore, the large majority of the time the polypeptide will bind incorrectly to the fibril.³⁰ When this happens, there is a waiting time t_{wait} before the incorrectly bound molecule releases and there is a new opportunity for in-register binding. This waiting time

is related to the residence time t_{bond} calculated above, but the functional form depends on the templating mechanism as described in section II D.

In the rare event that the polypeptide correctly binds to the fibril there are two possible outcomes. First, with a probability P_+ the polypeptide may form H-bonds across the fibril end and thereby become incorporated in the fibril. Alternatively, the polypeptide may dissociate from the fibril with a probability $P_- = 1 - P_+$. The average times for these two events are the previously calculated t_{attach} and t_{rel} , respectively. The time required to sample all $2LN_\beta$ binding conformations is

$$\tau_{\text{all}} = (2LN_\beta - N_\beta)(t_{\text{wait}} + t_{\text{diff}}) + P_+ N_\beta(t_{\text{attach}} + t_{\text{diff}}) + P_- N_\beta(t_{\text{rel}} + t_{\text{diff}}). \quad (9)$$

In Eq. (9) the first term is the time for a polypeptide to diffuse to the fibril end, attach incorrectly, and release $2LN_\beta - N_\beta$ times. The second and third terms are the weighted times for the polypeptide to bind correctly and either fully attach or release.

After a time τ_{all} the fibril will have added, on average, $P_+ N_{\text{bond}}$ new molecules. Therefore, assuming the residence time $t_{\text{bond}}(\epsilon_0)$ for correctly bound molecules is very long, the growth rate is

$$k_{\text{grow}} = \frac{P_+ N_\beta}{\tau_{\text{all}}}. \quad (10)$$

Since $t_{\text{wait}} \simeq t_{\text{bond}}$ and $t_{\text{bond}} \gg t_{\text{attach}}, t_{\text{rel}}$, Eq. (9) is dominated by the first term. Thus, it is reasonable to approximate the growth rate as

$$k_{\text{grow}} \simeq \frac{P_+}{(2L - 1)(t_{\text{wait}} + t_{\text{diff}})}, \quad (11)$$

which has the same functional form proposed in previous work.³⁴ However, Eq. (11) predicts that the time required to add a molecule to the fibril is much longer than either of the two timescales appearing in the rate equation due to the low probability of in-register binding.

The key timescales in the growth process can be summarized $k_{\text{grow}}^{-1} \gg t_{\text{wait}}, t_{\text{diff}} \gg k_+^{-1}$, with an impressive 8 to 9 orders of magnitude separating the macroscopic growth rate from the microscopic timescale of H-bond formation.^{26,27} In addition, Eq. (11) shows that the timescale governing growth in the reaction limited regime, $t_{\text{wait}} \gg t_{\text{diff}}$, is not the annealing of newly bound molecules, but the clearance of off-pathway intermediates. Finally, Eq. (11) can be immediately generalized to allow for cases where multiple alignments can be tolerated within the fibril. For example, if n different alignments can be accommodated within the steric zipper, then the prefactor is modified $(2L - 1)^{-1} \rightarrow n/(2L - n)$.

The final step in the calculation is to compute t_{wait} , which depends on the mechanism of templating.

D. Successful templating occurs when incorrectly bound molecules arrest growth

I will consider two possibilities for the templating mechanism. In the equilibrium formalism described by Eq. (1), these mechanisms correspond to the case where the energy penalty

$\epsilon_1 - \epsilon_0$ acts on the bonds preceding a mis-aligned molecule and the case where the penalty is applied to the bonds following the mis-aligned molecule.

1. Growth by direct templating

One possibility for the templating mechanism is that specificity is encoded by the direct sidechain-sidechain interactions between the incoming molecule and the terminal molecule on the fibril end. This mechanism, which I refer to as “direct templating” is a two-body interaction between the fibril end and the incoming molecule. This is to be distinguished from the “steric templating” mechanism to be discussed shortly that involves a three-body interaction. In direct templating the binding affinity between two molecules in the optimal alignment is described by ϵ_0 , while the binding affinity between any other set of sidechain pairings is ϵ_1 .

The growth behavior of the system can be controlled by using the protein concentration to vary t_{diff} . There are four concentration regimes to consider, which are labeled by the numerals I–IV in Fig. 2(a). Region I encompasses the very lowest concentrations, where $t_{\text{diff}} > t_{\text{bond}}(\epsilon_0)$, meaning that the fibrils are thermodynamically unstable and will have a negative growth rate. At somewhat higher concentrations, $t_{\text{bond}}(\epsilon_0) > t_{\text{diff}} > t_{\text{bond}}(\epsilon_1)$, only bonds between in-register molecules are thermodynamically stable (regions II and III). However, between in-register binding events there will be numerous out-of-register binding attempts with a lifetime $t_{\text{bond}}(\epsilon_1)$. If an additional binding event occurs (either in- or out-of-register) while an out-of-register molecule is attached, the second molecule will also be unstable and have the same lifetime as the out-of-register molecules. Therefore, in Eq. (11) the waiting time t_{wait} is the average time for all out-of-register molecules to detach so that the fibril presents a clean surface for in-register binding. This waiting time can also be modeled as a random walk with forward steps representing out-of-register binding events and backward steps representing detachment events. t_{wait} is, therefore, the first passage time for a walk starting at one bound molecule to reach zero bound molecules. This time is given by Eq. (5) with $N_\beta \rightarrow \infty$ and a negative velocity $v = t_{\text{diff}}^{-1} - t_{\text{bond}}^{-1}(\epsilon_1)$. The resulting growth rate from Eqs. (5) and (11) is

$$k_{\text{grow}} = \frac{t_{\text{diff}} - t_{\text{bond}}(\epsilon_1)}{2Lt_{\text{diff}}^2}. \quad (12)$$

This expression predicts a maximum growth rate when $t_{\text{diff}} = 2t_{\text{bond}}(\epsilon_1)$. When t_{diff} exceeds this value the growth rate increases with concentration (region II). However, when the concentration becomes high enough that $t_{\text{diff}} < 2t_{\text{bond}}(\epsilon_1)$ the increased rate of binding attempts is offset by the greater probability of mis-bound molecules occupying the fibril end. This results in a declining growth rate as the concentration increases (region III).

When $t_{\text{diff}} = t_{\text{bond}}(\epsilon_1)$, Eq. (12) erroneously predicts that the growth rate reaches zero. In fact, at concentrations above this point, out-of-register binding becomes thermodynamically stable and the system will rapidly form disordered precipitates (region IV). In this regime, the aggregation rate be-

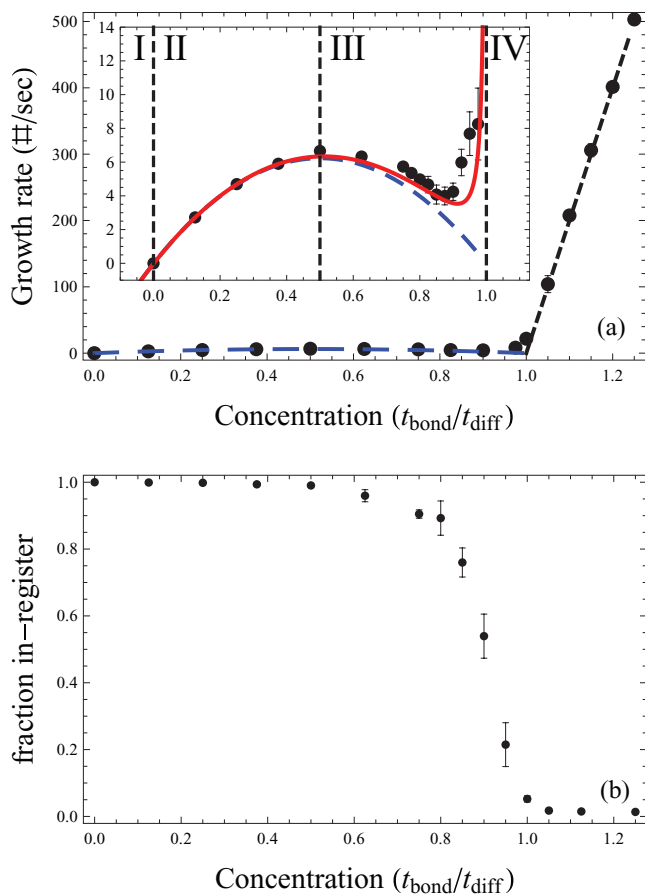


FIG. 2. Numerical simulation of growth master equations for molecules with $L = 40$, $t_{\text{bond}}(\epsilon_1) = 0.5$ ms, and $t_{\text{bond}}(\epsilon_0) = 50$ s. (a) Simulated growth rates as a function of polypeptide concentration. Polypeptide concentration is expressed as a ratio of the diffusion and residence times. In these units $c = 1$ corresponds to approximately $100 \mu\text{M}$. Theoretical curves are shown for Eqs. (12) (blue, long dashes), (15) (black, short dashes), and (16) (red, solid). (inset) Close-up of ordered growth region ($c < 1$). Roman numerals I–IV and the vertical dotted lines denote the four concentration regimes discussed in the text. (b) Fraction of molecules in the simulated fibrils that are bound in the favored, in-register alignment. A rapid decrease in fibril order occurs in region III, defined by $0.5 < t_{\text{bond}}/t_{\text{diff}} < 1$.

comes

$$k_{\text{grow}} = t_{\text{diff}}^{-1} - t_{\text{bond}}^{-1}(\epsilon_1), \quad (13)$$

which increases linearly with the concentration.

Region III, defined by $2t_{\text{bond}}(\epsilon_1) > t_{\text{diff}} > t_{\text{bond}}(\epsilon_1)$, is the transition from completely ordered growth to disordered aggregation. Within this range, molecules that bind out-of-register are thermodynamically unstable but may become trapped within the fibril if they become “capped” by stable ordered regions. The presence of a stable cap will prevent the mis-aligned molecules from regaining conformational entropy upon the breakage of their backbone H-bonds. This will dramatically increase the stability of these bonds causing the mis-aligned molecule to become kinetically trapped within the fibril.³⁵

The additional growth due to these trapped, disordered molecules can be estimated as follows. The formation of an in-register cap requires a single ordered molecule at the end of a disordered segment, which occurs with a probability $(2L)^{-1}$,

followed by an additional in-register binding event, which occurs at a rate $(2Lt_{\text{diff}})^{-1}$. This leads to a fibril growth of d molecules, where d is the average number of bound disordered molecules. d can be estimated from the average lifetime of an uncapped disordered region t_{wait} . In this time approximately d molecules will attach and subsequently detach

$$d(t_{\text{diff}} + t_{\text{bond}}(\epsilon_1)) = t_{\text{wait}} \quad (14)$$

$$= \frac{t_{\text{diff}}t_{\text{bond}}(\epsilon_1)}{t_{\text{diff}} - t_{\text{bond}}(\epsilon_1)}, \quad (15)$$

where the second step follows from Eq. (5). Therefore, Eq. (12) can be modified as follows to account for the trapping of disordered segments within the fibril

$$k_{\text{grow}} = \frac{t_{\text{diff}} - t_{\text{bond}}(\epsilon_1)}{2Lt_{\text{diff}}^2} + \frac{1}{(2L)^2} \frac{t_{\text{bond}}(\epsilon_1)}{t_{\text{diff}}^2 - t_{\text{bond}}^2(\epsilon_1)}, \quad (16)$$

where the two terms account for ordered growth and disordered growth, respectively.

The approximate formulas Eqs. (12), (15), and (16) are plotted in Fig. 2(a). Also shown are numerical growth rates from a stochastic simulation of the growth process using the Gillespie algorithm³⁶ for molecules of $L = 40$. The simulations show the predicted reduced growth rate in region III $2t_{\text{bond}}(\epsilon_1) > t_{\text{diff}} > t_{\text{bond}}(\epsilon_1)$ that coincides with a rapid increase in the fraction of out-of-register molecules that are incorporated in the fibril (Fig. 2(b)).

2. Growth by steric templating

Here I assume that direct sidechain-sidechain interactions account for only minor perturbations in the fibril stability and the sequence specificity arises, instead, from steric complementarity of the sidechains. In this case, mis-aligned molecules would result in a fibril core with either voids or steric clashes between adjacent sidechains. Such clashes would be easily resolved for a single mis-aligned molecule at the fibril end where the sidechains would have the freedom to adopt rotamers that extend the sidechains parallel to the fibril axis. However, these sidechains would then occupy the volume needed for the next molecule to bind. Therefore, in this model the first mis-bound molecule has a residence time that is minimally perturbed from the optimal alignment $t_{\text{bond}}(\epsilon_0)$, however, subsequent molecules have a much shorter lifetime $t_{\text{bond}}(\epsilon_1)$ due to the propagation of steric clashes.

Within this model a single mis-bound molecule will effectively arrest growth and we can approximate $t_{\text{wait}} \simeq t_{\text{bond}}(\epsilon_0)$. Therefore, the growth rate becomes

$$k_{\text{grow}} \simeq \frac{P_+}{(2L-1)(t_{\text{bond}}(\epsilon_0) + t_{\text{diff}})} + \frac{1}{(2L)^2} \frac{t_{\text{bond}}(\epsilon_1)}{t_{\text{diff}}^2 - t_{\text{bond}}^2(\epsilon_1)}, \quad (17)$$

where the ordered growth term comes from Eq. (11) and the disordered term is unchanged from Eq. (16). The primary difference between this expression and Eq. (16) is that the initial growth rate plateau and the disorder transition are governed by two different timescales ($t_{\text{bond}}(\epsilon_0)$ and $t_{\text{bond}}(\epsilon_1)$) whereas in Eq. (16) both features are a function of $t_{\text{bond}}(\epsilon_1)$.

III. COMPARISON TO EXPERIMENT

A. Growth rates are limited by the low probability of in-register binding

The growth rate of $A\beta$ fibrils was estimated by total internal reflection fluorescence (TIRF) microscopy to be about 10 molecules/s.²⁷ To compare this value to the theory I approximate t_{diff} (Eq. (3)) using the Stokes-Einstein relation for the diffusion constant

$$D_p = \frac{k_B T}{6\pi\eta(R_0 L^{0.6})}, \quad (18)$$

where η is the viscosity of the solvent and the radius of the molecule is assumed to follow Flory scaling with $R_0 = 0.133$ nm.³⁷ With these parameters $t_{\text{diff}} = 0.9$ ms at the 50 μM experimental concentration. If this is in the diffusion limited regime ($t_{\text{diff}} \gg t_{\text{wait}}$) with $L = 40$ the result would be a growth rate of about 14 molecules per second, in good agreement with the experiment. At the opposite extreme, an upper bound for t_{wait} can be obtained from $t_{\text{bond}}(\epsilon_0)$. The stability of $A\beta_{40}$ fibrils is $N_\beta \epsilon_b = -13.1 k_B T^{21}$ or $\epsilon_b \simeq 1.3$ kJ/mol for $N_\beta = 24$.²⁹ The calculated parameters for $A\beta_{40}$ are $t_{\text{bond}} = 6.6 \times 10^{-4}$ s, $t_{\text{attach}} = 4.8 \times 10^{-8}$ s, $t_{\text{rel}} = 2.4 \times 10^{-9}$ s, and $P_+ = 0.41$ for a final growth rate of $k_{\text{grow}} = 3.4$ molecules/s. This factor of three error is also within the expected range for such a coarse model, but it would seem that the major limitation on the growth rate is diffusion and the $(2L)^{-1}$ probability of in-register binding.

More extensive kinetic data are available for insulin fibrils. Insulin is somewhat challenging to model because of the unknown fibril structure and the conformational constraints imposed by the disulfide bonds. Also, while the stability of insulin fibrils is known³⁸ this value is a combination of both intermolecular bonds and intramolecular contacts between the two peptide strands. Rather than attempt a blind deconvolution of these factors in the absence of a structure, I will neglect the attractive contribution from sidechain packing interactions and the repulsive contribution from conformational entropy and assume that the intermolecular H-bonds have the same thermodynamic properties as the bonds studied in detail by Ross and Rekharsky,³⁹

$$\epsilon(T) = \left(-1650 \frac{\text{cal}}{\text{mol}} + T \frac{4.2 \text{ cal}}{\text{mol K}} \right). \quad (19)$$

While these parameters were obtained in a carbohydrate system, the overall binding strength is comparable to that found in amyloids³⁸ and the explicit enthalpy and entropy contributions will permit the exploration of temperature effects. I take $N_\beta = 21$ as suggested by Jiménez *et al.*²⁸ and use the total number of amino acids on both peptide strands for the length $L = 51$.

In order to aggregate folded proteins such as insulin it is usually necessary to destabilize the native state. Knowles *et al.* used both acidic conditions (pH 2.0) and guanidinium to denature the proteins.²⁶ The fraction of unfolded proteins as a function of guanidinium concentration is

$$c/c_{\text{tot}} = (1 + e^{-(\Delta G^0 + mc_d)/RT})^{-1}, \quad (20)$$

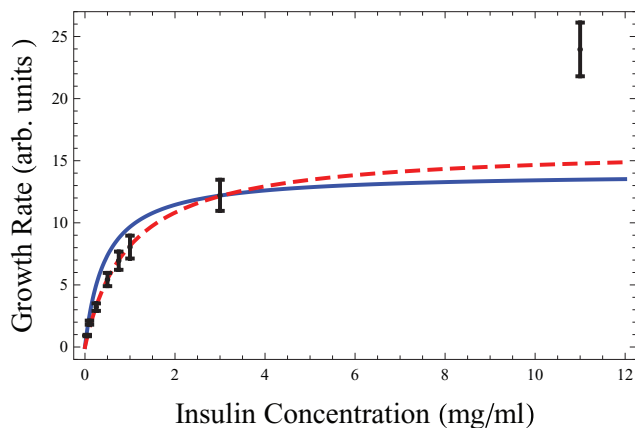


FIG. 3. Comparison of Eq. (10) with $t_{\text{wait}} = t_{\text{bond}}(\epsilon_0)$ (solid) to the growth of insulin fibrils.²⁶ Due to the uncertainty in the experimental fibril density, the theory has been scaled to match the experimental range. Also shown (dashed) is a two parameter fit to $k_{\text{grow}} = (\tau_r + \tau_d/c)^{-1}$ which suggests that the 11 mg/ml data point is either an outlier, the onset of disordered growth, or shows the transition to growth by a different mechanism.

where c_{tot} is the total concentration of proteins, c is the unfolded (aggregation prone) concentration, $\Delta G^0 = -2.5$ kcal/mol is the free energy of folding in the absence of denaturant, c_d is the denaturant concentration, and $m = 0.55$ (kcal/mol)/mol is a constant describing the effect of the denaturant.²⁶

With these assumptions the growth rate in the steric templating model is shown in Fig. 3 as a function of concentration (the direct templating model will prove to be inconsistent with the temperature and denaturant effects discussed below). The theory captures the qualitative trend of a saturating growth rate, although it underestimates the saturation concentration. The time per attachment at 1 mg/ml is calculated to be 0.4 s. This is somewhat less than the 3.1 s estimated in the experiments, however, there is considerable uncertainty in both the experimental fibril density and the theoretical parameters k_+ , N_β , and ϵ_b . As a result the calculated rate in Fig. 3 has been scaled to match the experiments.

B. Temperature and denaturants accelerate fibril growth

Weakening peptide-peptide interactions with denaturants or temperature changes will have multiple effects on fibril growth kinetics. First, weaker interactions will destabilize the native fold, increasing the concentration of aggregation competent molecules and, therefore, reducing t_{diff} (Eq. (20)). Second, weaker H-bonds will reduce t_{bond} allowing the protein to more rapidly sample binding alignments. I assume that the H-bond formation rate k_+ is primarily determined by the diffusion dynamics of the peptide backbone, so weaker bonds will only increase k_- . The temperature dependence of the H-bonds is described by Eq. (19) while denaturants are assumed to have a linear effect

$$k_+/k_- = e^{-(\epsilon_b - m_H c_d)/RT}, \quad (21)$$

where m_H is a constant. Finally, increasing the temperature will increase the diffusion rates. However, including the tem-

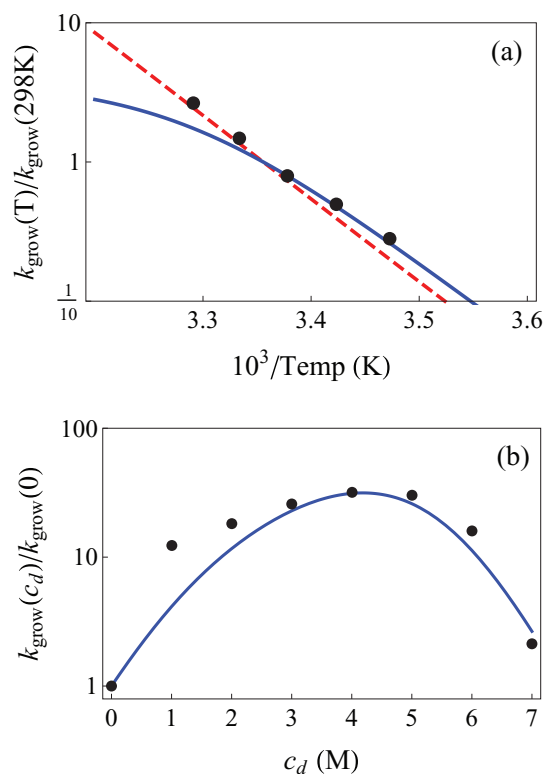


FIG. 4. (a) Effect of temperature on growth rates. The theory incorrectly predicts a deviation from the Arrhenius trend at high temperature (solid blue line). This error vanishes if t_{diff} diminishes sufficiently fast that the reaction remains limited by the binding/unbinding kinetics (red dashes). (b) Effect of denaturant on fibril growth rates. A similar error in t_{diff} is responsible for the discrepancy near 1M. Experiments from Ref. 26.

perature dependence of Eq. (18) and a Rouse-like $k_+ \propto e^T$ dependence⁴⁰ has a negligible effect on calculated growth rates.

Temperature and denaturants are predicted to have very different effects on the two proposed templating mechanisms. In Fig. 3 it is readily apparent that the onset of fully disordered aggregation must occur at a concentration greater than ~ 11 mg/ml. This means that at the 1 mg/ml concentration at which the temperature and denaturation experiments were performed, the direct templating model would require that the ratio $t_{\text{bond}}(\epsilon_1)/t_{\text{diff}}$ is less than 0.1. In this case, Eq. (12) would predict that changes in the temperature or denaturant concentration too small to promote significant unfolding would affect the growth rate by less than 10%. In contrast, the experiments show order of magnitude increases in the growth rate that occur below room temperature and at denaturant concentrations below 1M where unfolding effects are minimal. Therefore, the analysis below employs the steric templating model exclusively.

Figure 4(a) shows the effect of temperature. The experiments show an Arrhenius behavior over the measured temperature range. The theory captures this at low T but deviates at higher temperature. This is because diffusion becomes limiting in the theory, suggesting that either Eq. (20) underestimates the denaturation of the insulin at elevated temperatures, or there is an energetic barrier to binding that is missing from Eq. (3). If either of these is true, the growth rate would

remain reaction limited. This case is shown by the dashed line in Fig. 4(a), which is in good agreement with the experiments.

Arrhenius behavior is common in fibril growth kinetics.^{32,41} This may be derived in the steric model from Eqs. (5) and (11) in the limits of high concentration $t_{\text{diff}} \ll t_{\text{bond}}$ and weak bias $v \ll k_+$

$$k_{\text{grow}} \propto t_{\text{bond}}^{-1} \sim e^{-vN_{\beta}/D} \simeq (1 - v/k_+)^{N_{\beta}} \\ = (k_-/k_+)^{N_{\beta}} = e^{N_{\beta}\epsilon_b/RT}. \quad (22)$$

Importantly, the energy barrier that must be surmounted does not correspond to the transition state. Rather, it describes the recovery from off-pathway kinetic traps.

The effect of denaturants on the growth rate is shown in Fig. 4(b). The single fitting parameter is $m_H = 0.29$ (kJ/mol)/mol, which is between the 0.10 (kJ/mol)/mol found for globular proteins⁴² and 0.42 (kJ/mol)/mol for fibrils.⁴³ At the 1 mg/ml experimental concentration $t_{\text{bond}} > t_{\text{diff}}$, so the initial effect is from the increased bond breakage rate, while the effects of the increased diffusion rate set in at denaturant concentrations greater than $\sim 2\text{M}$. At even greater concentrations the growth rate starts to decline when the denaturant concentration becomes high enough to melt the fibrils.

IV. EFFECT OF POLYPEPTIDE LENGTH ON GROWTH RATES

The growth rate in the steric templating model is extremely sensitive to the length of the polypeptide through the variables L and N_{β} . There are two cases to consider. The first of these is when the entire polypeptide is incorporated into the cross- β core of the fibril such that $L = N_{\beta}$. This is applicable in model systems where the amyloidogenic core of the protein has been isolated. Here the dominant effect of the length arises from the exponential factor in t_{bond} (Fig. 5(a), dashed line). At lower concentrations the diffusion time will become limiting (Fig. 5(a), solid line), which is most dramatic for short polypeptides; although small molecules diffuse faster, this effect is small compared to the exponential dependence of t_{bond} on N_{β} .

Alternatively, a portion of the polypeptide may remain outside the core of the fibril such that $L > N_{\beta}$. For example, in $A\beta$ fibrils the N-terminal 10–15 residues remain disordered.^{22,29} In this case increasing L (with N_{bond} fixed) leads to a linear increase in the attachment time due to the increased probability for incorrect alignments. Figure 5(b) shows a very different behavior where the diffusion becomes limiting for long polypeptides. This is because t_{bond} remains constant as L increases. Since the protein diffusion constant scales according to $D_p \sim L^{-0.6}$ and the probability for correct binding is proportional to L^{-1} , the overall growth rate scales such as $L^{-1.6}$. The result is that assembly times become prohibitive for very long polypeptides—an effect that is exacerbated by the attainable concentrations for large mass proteins. This may explain why large proteins are rarely associated with amyloid diseases.⁴⁴ The kinetic limitation on the fibrilization of large proteins may be alleviated if the non-amyloidogenic regions retain a native fold. A partially folded protein would influence the growth rate in two ways. First,

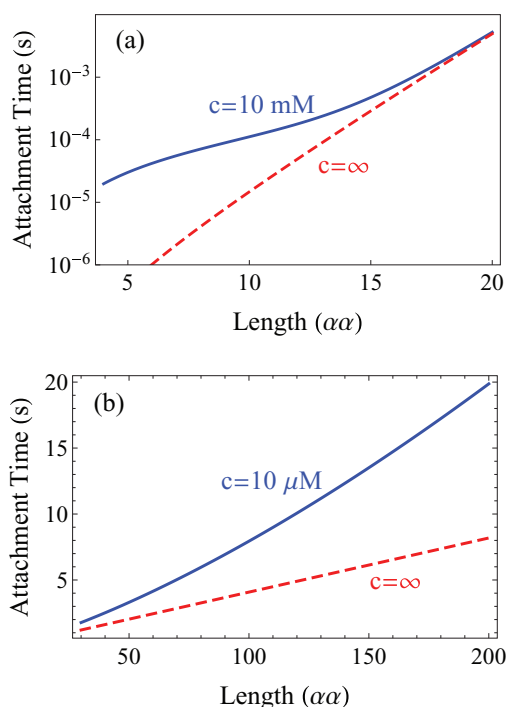


FIG. 5. Length dependence of the attachment time for (a) polypeptides that are completely incorporated into the core of the fibril, and (b) polypeptides with a fixed core component of $N_{\beta} = 30$ and variable length disordered region. At high concentrations the rate is limited by the detachment rate for incorrectly bound polypeptides (red dashes). At lower concentrations, however, the diffusion time is limiting (solid blue). The diffusion limited regime may be more pronounced for either short or long polypeptides depending on whether the added length is incorporated into the cross- β fibril core.

the more compact hydrodynamic radius would reduce t_{diff} , with the diffusion constant scaling such as $L^{-1/3}$. Second, the residues in the folded core would be prevented from making incorrect H-bonds with the fibril end. However, the exposed residues would still only have a N_{β}^{-1} probability of making correct contacts. This would modify the prefactor of Eq. (11) from $(2L - 1)^{-1}$ to N_{β}^{-1} , which does not depend on the size of folded domain.

V. DISCUSSION

The aggregation process involves events occurring over a wide range of timescales, ranging from nanoseconds for the formation of H-bonds (k_+^{-1}), milliseconds for diffusion (t_{diff}) and the release of mis-aligned molecules (t_{bond}), to seconds for the overall growth rate (k_{grow}^{-1}). This fact complicates efforts to observe aggregation *in silico*. All-atom simulations of amyloid fibrils have shown that H-bond formation and breakage dominates the accessible simulation times.^{30,45} According to the model presented here, the timescale characterizing this process is $t_{\text{bond}} \simeq 1$ ms for cross- β fibril cores consisting of 20–25 amino acids, which is at the upper limit of what is computationally tractable.

Amyloid growth has been described in terms of a two state “Dock and Lock” process based on distinct populations observed in dissociation experiments.⁴⁶ In these experiments the defining feature of the docked state is that it can reversibly unbind from the fibril. Therefore, in the present theory the

docked state can be identified with mis-aligned polypeptides and the locked state with the correctly bound molecules. This is in agreement with simulations showing a very fast locking step when the register between the fibril and incoming molecule is correct.⁴⁷ On the other hand, it is also possible to envision a model that does allow a direct conversion from the docked state to the locked state. This would require a polypeptide to bind to the fibril end at multiple positions along its length separated by unbound coil regions. Each of these bound segments would then perform random walks similar to those described in Fig. 1. Such direct interconversion could be tested by rapidly exposing fibrils to a pulse of labeled monomers and checking if the ratio of final incorporation to initial binding exceeds the $(2L)^{-1}$ predicted by theory.

While the model assumes that the initial contact between the soluble polypeptide and the fibril is a random event, in some cases binding between the fibril and the polypeptide may be non-random. Electrostatic interactions will always disfavor parallel, in-register binding, while enhanced binding probabilities may arise from complementary electrostatic interactions in anti-parallel fibrils⁴⁸ or residual structure in the incoming polypeptide. For example, structural differences in the disordered state ensembles of $A\beta_{40}$ and $A\beta_{42}$ may be partly responsible for the differences in the aggregation of these two species.⁴⁹ However, perturbations such as non-random binding or folded domains in the soluble molecules will not alter the qualitative result that ordered fibril growth requires a conformational search over a large number of energetically favorable binding states. This explains why the growth rates for the partially folded insulin system show the same qualitative trends that would be expected for a disordered polypeptide.

In summary, the coarse-grained model of amyloid growth presented here explains many features of the aggregation process including the non-monotonic effect of denaturants, the Arrhenius temperature dependence, and the long timescales characterizing aggregation.

ACKNOWLEDGMENTS

I would like to thank T. Knowles for providing experimental data, A. Buell, P. Tessier, and M. Muschol for valuable discussions, and S. Whitelam and L. Nemzer for critical reading of the paper. This work was supported by KSU startup funds. Publication of this article was funded in part by the Kansas State University Open Access Publishing Fund.

APPENDIX: PASSAGE TIMES FOR THE H-BOND RANDOM WALK

1. Incorrectly bound dwell time

In the language of the random walk described in Fig. 1, t_{bond} corresponds to the mean first passage time for a particle that starts at $x = 1$ to reach an absorbing boundary at $x = 0$. The mean time $t(x)$ required for a particle starting at x to reach an absorbing boundary satisfies the following recursion relationship:³¹

$$t(x) = p_+ t(x+1) + p_- t(x-1) + \tau, \quad (\text{A1})$$

where $p_{\pm} = k_{\pm}/(k_+ + k_-)$ are the probabilities of left/right steps and the average waiting time between steps is $\tau = (k_+ + k_-)^{-1}$. Equation (A1) reflects the fact that after waiting one time step τ the particle that started at x begins a new random walk at $x+1$ with probability p_+ , or at $x-1$ with probability p_- . In the continuum limit this can be written as the differential equation,³¹

$$-1 = Dt''(x) + vt'(x), \quad (\text{A2})$$

where the diffusion constant and velocity are $D = (k_+ + k_-)/2$ and $v = k_+ - k_-$. The boundary conditions are $t(0) = 0$ and $t'(N_{\beta}) = 0$ representing the absorbing boundary on the left and the reflecting boundary on the right. With these boundary conditions the solution to Eq. (A2) is

$$t(x) = -\frac{x}{v} + \frac{D}{v^2} e^{vN_{\beta}/D} (1 - e^{-vx/D}). \quad (\text{A3})$$

So, the average dwell time for a polypeptide that makes an incorrect initial H-bond is

$$t_{\text{bond}} = t(1) = -\frac{1}{v} + \frac{D}{v^2} e^{vN_{\beta}/D} (1 - e^{-v/D}). \quad (\text{A4})$$

2. Correctly bound proteins

a. Splitting probabilities

To describe events where the polypeptide binds with the correct alignment and orientation Eq. (10) requires both the probability P_+ that the polypeptide fully binds before detaching and the average times for full binding and detachment. In the random walker mapping, the attachment probability corresponds to $P_+ = \mathcal{E}_+(1)$ where $\mathcal{E}_{\pm}(x)$ is the probability that a walker starting at x reaches the right/left boundary first. In analogy with Eq. (A2), these splitting probabilities obey the differential equation,³¹

$$D\mathcal{E}_{\pm}'' + v\mathcal{E}_{\pm}' = 0. \quad (\text{A5})$$

With the boundary conditions $\mathcal{E}_+(0) = 0$, $\mathcal{E}_+(N_{\beta}) = 1$, $\mathcal{E}_-(0) = 1$, and $\mathcal{E}_-(N_{\beta}) = 0$ the solutions are

$$\begin{aligned} \mathcal{E}_+(x) &= \frac{e^{-vx/D} - 1}{e^{-vN_{\beta}/D} - 1}, \\ \mathcal{E}_-(x) &= \frac{-e^{-vx/D} + e^{-vN_{\beta}/D}}{e^{-vN_{\beta}/D} - 1}. \end{aligned} \quad (\text{A6})$$

b. Binding and release times

The next step is to calculate the average times for correctly bound random walks to reach the left or right boundaries. First, define the quantity $t_+(x)$ as the average time for a particle that starts at x to reach the right boundary considering only paths that do not touch the left boundary. Similarly, $t_-(x)$ describes the average time to reach the left absorbing boundary without touching the right. These times can be determined from the recursion relationship for the combined function $\mathcal{E}_{\pm}t_{\pm}$,³¹

$$\begin{aligned} \mathcal{E}_{\pm}(x)t_{\pm}(x) &= p_+ \mathcal{E}_{\pm}(x+1)t_{\pm}(x+1) \\ &+ p_- \mathcal{E}_{\pm}(x-1)t_{\pm}(x-1) + \tau \mathcal{E}_{\pm}(x). \end{aligned} \quad (\text{A7})$$

In the continuum limit this becomes

$$D \frac{d^2(\mathcal{E}_{\pm}(x)t_{\pm}(x))}{dx^2} + v \frac{d(\mathcal{E}_{\pm}(x)t_{\pm}(x))}{dx} = -\mathcal{E}_{\pm}(x). \quad (\text{A8})$$

Equation (A8) can be integrated to solve for $\mathcal{E}_{\pm}t_{\pm}$. Applying Eq. (A6), the solutions for the average times are³¹

$$t_{+}(x) = \frac{N_{\beta} e^{-vN_{\beta}/D} + 1}{v} \frac{x e^{-vx/D} + 1}{1 - e^{-vN_{\beta}/D}} + \frac{x e^{-vx/D} + 1}{v e^{-vx/D} - 1}, \quad (\text{A9})$$

$$t_{-}(x) = \frac{N_{\beta}}{v} \frac{2e^{-vN_{\beta}/D}}{e^{-vN_{\beta}/D} - 1} \frac{1 - e^{-vx/D}}{e^{-vx/D} - e^{-vN_{\beta}/D}} + \frac{x e^{-vx/D} + e^{-vN_{\beta}/D}}{v e^{-vx/D} - e^{-vN_{\beta}/D}}. \quad (\text{A10})$$

Setting $t_{\text{attach}} = t_{+}(1)$ and $t_{\text{rel}} = t_{-}(1)$ yields Eqs. (6) and (7).

- ¹I. Cherny and E. Gazit, *Angew. Chem.* **47**, 4062 (2008).
- ²G. G. Tartaglia, A. P. Pawar, S. Campioni, C. Dobson, F. Chiti, and M. Vendruscolo, *J. Mol. Biol.* **380**, 425 (2008).
- ³D. J. Selkoe, *Nature (London)* **426**, 900 (2003).
- ⁴M. M. Pallitto and R. M. Murphy, *Biophys. J.* **81**, 1805 (2001).
- ⁵D. Hall, N. Hirota, and C. Dobson, *J. Mol. Biol.* **351**, 195 (2005).
- ⁶C.-C. Lee, A. Nayak, A. Sethuraman, G. Belfort, and G. J. McRae, *Biophys. J.* **92**, 3448 (2007).
- ⁷E. T. Powers and D. L. Powers, *Biophys. J.* **94**, 379 (2008).
- ⁸T. Knowles, C. A. Waudby, G. L. Devlin, S. I. A. Cohen, A. Aguzzi, M. Vendruscolo, E. M. Terentjev, M. E. Welland, and C. M. Dobson, *Science* **326**, 1533 (2009).
- ⁹D. Kashchiev and S. Auer, *J. Chem. Phys.* **132**, 215101 (2010).
- ¹⁰J. Hardy and D. J. Selkoe, *Science (N.Y.)* **297**, 353 (2002).
- ¹¹S. Auer, F. Meersman, C. Dobson, and M. Vendruscolo, *PLOS Comput. Biol.* **4**, e1000222 (2008).
- ¹²H. D. Nguyen and C. K. Hall, *J. Am. Chem. Soc.* **128**, 1890 (2006).
- ¹³B. Urbanc, L. Cruz, S. Yun, S. V. Buldyrev, G. Bitan, D. B. Teplow, and H. E. Stanley, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 17345 (2004).
- ¹⁴C. Wu and J.-E. Shea, *Curr. Opin. Struct. Biol.* **21**, 209 (2011).
- ¹⁵T. R. Serio, A. G. Cashikar, A. S. Kowal, G. J. Sawicki, J. J. Moslehi, L. Serpell, M. F. Arnsdorf, and S. L. Lindquist, *Science* **289**, 1317 (2000).
- ¹⁶S. R. Collins, A. Douglass, R. D. Vale, and J. S. Weissman, *PLOS Biol.* **2**, e321 (2004).
- ¹⁷S. E. Hill, J. Robinson, G. Matthews, and M. Muschol, *Biophys. J.* **96**, 3781 (2009).
- ¹⁸S. E. Hill, T. Miti, T. Richmond, and M. Muschol, *PLoS ONE* **6**, e18171 (2011).
- ¹⁹P. Seubert, C. Vigo-Pelfrey, F. Esch, M. Lee, H. Dovey, D. Davis, S. Sinha, M. Schlossmacher, J. Whaley, and C. Swindlehurst, *Nature (London)* **359**, 325 (1992).
- ²⁰P. Sengupta, K. Garai, B. Sahoo, Y. Shi, D. J. E. Callaway, and S. Maiti, *Biochemistry* **42**, 10506 (2003).
- ²¹J. D. Schmit, K. Ghosh, and K. A. Dill, *Biophys. J.* **100**, 450 (2011).
- ²²T. Lührs, C. Ritter, M. Adrian, D. Riek-Loher, B. Bohrmann, H. Döbeli, D. Schubert, and R. Riek, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 17342 (2005).
- ²³C. Wasmer, A. Lange, H. Van Melckebeke, A. B. Siemer, R. Riek, and B. H. Meier, *Science* **319**, 1523 (2008).
- ²⁴R. Nelson, M. R. Sawaya, M. Balbirnie, A. O. Madsen, C. Riek, R. Grothe, and D. Eisenberg, *Nature (London)* **435**, 773 (2005).
- ²⁵S. Whitelam, R. Schulman, and L. Hedges, *Phys. Rev. Lett.* **109**, 265506 (2012).
- ²⁶T. Knowles, W. Shu, G. L. Devlin, S. Meehan, S. Auer, C. Dobson, and M. E. Welland, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 10016 (2007).
- ²⁷T. Ban, M. Hoshino, S. Takahashi, D. Hamada, K. Hasegawa, H. Naiki, and Y. Goto, *J. Mol. Biol.* **344**, 757 (2004).
- ²⁸J. L. Jiménez, E. J. Nettleton, M. Bouchard, C. V. Robinson, C. Dobson, and H. R. Saibil, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 9196 (2002).
- ²⁹A. T. Petkova, Y. Ishii, J. J. Balbach, O. N. Antzutkin, R. D. Leapman, F. Delaglio, and R. Tycko, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 16742 (2002).
- ³⁰C. F. Lee, L. Jean, and D. J. Vaux, *Phys. Rev. E* **80**, 041906 (2009).
- ³¹S. Redner, *A Guide to First-Passage Processes* (Cambridge University Press, 2007), p. 328.
- ³²A. K. Buell, A. Dhulesia, D. A. White, T. Knowles, C. M. Dobson, and M. E. Welland, *Angew. Chem.* **51**, 5247 (2012).
- ³³V. Muñoz, P. A. Thompson, J. Hofrichter, and W. A. Eaton, *Nature (London)* **390**, 196 (1997).
- ³⁴A. K. Buell, J. R. Blundell, C. Dobson, M. E. Welland, E. M. Terentjev, and T. Knowles, *Phys. Rev. Lett.* **104**, 228101 (2010).
- ³⁵W. Hwang, S. Zhang, R. D. Kamm, and M. Karplus, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 12916 (2004).
- ³⁶D. T. Gillespie, *J. Phys. Chem.* **81**, 2340 (1977).
- ³⁷J. E. Kohn, I. S. Millett, J. Jacob, B. Zagrovic, T. M. Dillon, N. Cingel, R. S. Dotherager, S. Seifert, P. Thiyagarajan, T. R. Sosnick, M. Z. Hasan, V. S. Pande, I. Ruczinski, S. Doniach, and K. W. Plaxco, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 12491 (2004).
- ³⁸A. J. Baldwin, T. Knowles, G. G. Tartaglia, A. W. Fitzpatrick, G. L. Devlin, S. L. Shammass, C. A. Waudby, M. F. Mossuto, S. Meehan, S. L. Gras, J. Christodoulou, S. J. Anthony-Cahill, P. D. Barker, M. Vendruscolo, and C. Dobson, *J. Am. Chem. Soc.* **133**, 14160 (2011).
- ³⁹P. D. Ross and M. V. Rekharsky, *Biophys. J.* **71**, 2144 (1996).
- ⁴⁰M. Doi and S. F. Edwards, *The Theory of Polymer Dynamics, International Series of Monographs on Physics* (Oxford University Press, USA, 1988), p. 408.
- ⁴¹R. Carrotta, M. Manno, D. Bulone, V. Martorana, and P. L. San Biagio, *J. Biol. Chem.* **280**, 30001 (2005).
- ⁴²K. Ghosh and K. A. Dill, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 10649 (2009).
- ⁴³S. S. L. Shammass, T. Knowles, A. J. Baldwin, C. C. E. Macphee, M. M. E. Welland, C. Dobson, G. L. Devlin, E. Cait, and L. Glyn, *Biophys. J.* **100**, 2783 (2011). Using Eq. (1) of Ref. 38 the midpoint of the melting curve corresponds to $\Delta G = -23$ kJ/mol, so $m_H = (-23 + 56)(\text{kJ/mol})/3.75M/N_{\beta}$.
- ⁴⁴H. Ramshini, C. Parrini, A. Relini, M. Zampagni, B. Mannini, A. Pesce, A. A. Saboury, M. Nemat-Gorgani, and F. Chiti, *PLoS ONE* **6**, e16075 (2011).
- ⁴⁵N. L. Fawzi, Y. Okabe, E.-H. Yap, and T. Head-Gordon, *J. Mol. Biol.* **365**, 535 (2007).
- ⁴⁶W. P. Esler, E. R. Stimson, J. M. Jennings, H. V. Vinters, J. R. Ghilardi, J. P. Lee, P. W. Mantyh, and J. E. Maggio, *Biochemistry* **39**, 6288 (2000).
- ⁴⁷G. Reddy, J. E. Straub, and D. Thirumalai, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 11948 (2009).
- ⁴⁸M. R. Sawaya, S. Sambashivan, R. Nelson, M. I. Ivanova, S. A. Sievers, M. I. Apostol, M. J. Thompson, M. Balbirnie, J. J. W. Wiltzius, H. T. McFarlane, A. O. Madsen, C. Riek, and D. Eisenberg, *Nature (London)* **447**, 453 (2007).
- ⁴⁹N. G. Sgourakis, Y. Yan, S. A. McCallum, C. Wang, and A. E. Garcia, *J. Mol. Biol.* **368**, 1448 (2007).