Pseudo-one-dimensional nucleation in dilute polymer solutions

Lingyun Zhang and Jeremy D. Schmit

Pseudo-1D nucleation in dilute polymer solutions

Lingyun Zhang[*][1,2] and Jeremy D. Schmit[†][1]

[1]*Department of Physics, Kansas State University, Manhattan, KS 66506, USA*
[2]*Beijing National Laboratory for Condensed Matter Physics,*
Institute of Physics, Chinese Academy of Sciences, Beijing 100190, China

Pathogenic protein fibrils have been shown *in vitro* to have nucleation dependent kinetics despite the fact that one-dimensional structures do not have the size dependent surface energy responsible for the lag time in classical theory. We present a theory showing that the conformational entropy of the peptide chains creates a free energy barrier that is analogous to the translational entropy barrier in higher dimensions. Interestingly, the dynamics of polymer rearrangement make it very unlikely for nucleation to succeed along the lowest free energy trajectory, meaning that most of the nucleation flux avoids the free energy saddle point. We use these results to construct a 3D model for amyloid nucleation that accounts for conformational entropy, backbone H-bonds, and sidechain interactions to compute nucleation rates as a function of concentration.

Amyloids are linear protein aggregates associated with conditions like Alzheimer's, Huntington's, and prion diseases [1]. When studied *in vitro*, amyloidogenic proteins remain soluble through a pronounced lag phase before undergoing a nucleation event that initiates exponential growth of fibrils [2, 3]. Previous theoretical work has shown that the proliferation of fibrils is dominated by secondary nucleation events that follow primary nucleation [4–8]. But these works do not address the underlying molecular mechanism of either primary or secondary nucleation and, therefore, it is not clear if they can be extrapolated to the low concentrations found *in vivo*. In particular, it is a puzzle why linear aggregates have a lag phase in the first place. In bulk phase transitions the nucleation lag time occurs because the surface energy of the new phase, which arises from the translational entropy loss of the bound particles, initially grows faster than the favorable bulk energy. This does not apply to 1D systems where the surface energy is independent of the cluster size. The reason for nucleation kinetics in amyloids is almost certainly because fibrils are only pseudo-1D. Several authors have noted that the steric zipper motif [9, 10] gives the fibrils a thickness in a direction perpendicular to the fibril axis [11–13]. These theories model the fibril as a 2D object, thereby achieving the size-dependent surface energy needed for a lag time.

In this Rapid Communication we investigate the effect of the third dimension: the direction parallel to the peptide backbones. In a growing post-nucleation fibril, each backbone H-bond and steric zipper interaction must overcome a free energy increase f_{CE} from the conformational entropy cost of trapping the backbone in the extended β-sheet conformation. This contrasts with the initial dimerization of two molecules where each intermolecular contact results in the loss of entropy from both chains for a free energy gain of $2f_{CE}$ (the Ramachandran analysis shows that both molecules must adopt β conformation

to maintain the intermolecular bonds). Importantly, the increased entropic penalty for the initial dimerization results in a free energy barrier that will contribute to the nucleation time [14, 15].

We focus on intrinsically disordered molecules at low concentration where monomer assembly pathways are expected to be dominant [16–18]. As a first step we compute the kinetics of a nucleating trimer as a toy model to explore the competition between binding energy and the conformational entropy of the polymers. Using the intermolecular H-bonds as a reaction coordinate [19–21], we map the kinetics to a 2D diffusion problem for which we can obtain an exact solution (albeit in the continuum limit). This model shows that the saddle point state is particularly prone to dissociation so that most of the nucleation flux avoids the saddle point. This implies that self-assembly in this system requires trajectories that are 'far' from equilibrium [22]. Next, we use these results to build a second model that allows us to account for both H-bond and steric zipper interactions as well as the growth of a cluster beyond the trimer stage. This model allows us to treat the nucleus as a 3D object and include the contributions of both conformational and translational entropy to the free energy barrier.

The nucleation rate is the product of a free energy barrier, an attempt rate, and a success probability. We develop our theory by modifying the nucleation rate from classical nucleation theory (CNT)

$$k_{\text{CNT}} = k_{\text{add}} Z e^{-F^{\ddagger}/k_B T}. \qquad (1)$$

The terms in Eq. 1 can be understood as follows. The exponential term is an Arrhenius factor giving the probability of finding a cluster at the free energy maximum $F^{\ddagger}$. Hereafter, we refer to the highest free energy state along a given trajectory as the "critical" state with quantities associated with this state denoted by the ($\ddagger$) superscript. k_{add} is the rate at which a molecule attaches to the cluster making it super-critical. Finally, the Zeldovich factor, Z, describes the probability that a newly super-critical cluster will nucleate the new phase without returning to a sub-critical size. Eq. 1 was derived using the cluster

size as a 1D reaction coordinate and applying a saddle point approximation [23]. We find that a 1D reaction coordinate is not sufficient to describe amyloid nucleation. Rather than re-deriving an analogous expression in 2D, we use Eq. 1 as a heuristic and independently compute each of the three terms.
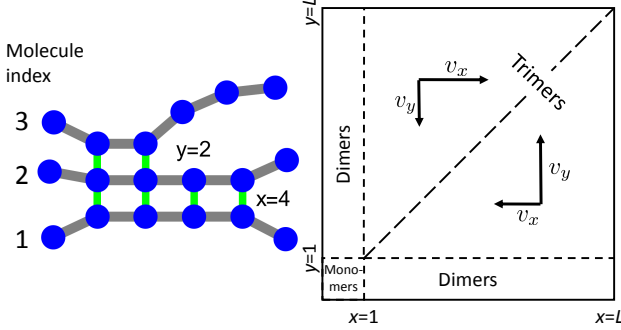

FIG. 1: Cartoon of a nucleating trimer illustrating the two-dimensional reaction coordinate used in the theory. A configuration of the trimer is mapped to the x-y plane where the x coordinate is the number of H-bonds between the middle and lower molecules and y is the number of H-bonds between the top and middle molecules. The formation and breakage of H-bonds results in a random walk in this 2D space. The unequal free energies of H-bond formation mean that the random walk is biased by drift velocities that push the system toward the $x = y$ diagonal.

2D model describes competition between H-bonds and conformational entropy. We first consider a toy model of nucleation consisting of three molecules forming one β-sheet. This model is the simplest one that captures the entropy penalty of initial dimerization and the attractive free energy of a "bulk" binding event. There are two reaction coordinates; x is the number of H-bonds between molecules 1 and 2 and y is the number of H-bonds between molecules 2 and 3 (see Fig. 1). The system can evolve in four ways; it can add or break a H-bond between the first two molecules with rates k_{+x} and k_{-x}, respectively, or between molecules 2 and 3 with rates k_{+y} and k_{-y}. Despite the $x \leftrightarrow y$ symmetry, the rates are not equal. To see this, we use detailed balance to relate the reaction rates to the free energy of H-bond formation

$$\frac{k_{+x}}{k_{-x}} = e^{-f_x/k_BT} \quad \frac{k_{+y}}{k_{-y}} = e^{-f_y/k_BT}. \quad (2)$$

In the configuration shown in Fig. 1, $x > y$ meaning that the transition $x \to x+1$ will result in a loss of conformational entropy from both molecule 1 and 2 and the formation of one favorable H-bond contact. In comparison, the transition $y \to y+1$ only results in a loss of entropy from molecule 3 while providing the same attractive H-bond. Therefore, when $x > y$ we have $f_x = f_{\text{weak}} > 0$ and $f_y = f_{\text{strong}} < 0$. For $x < y$ the situation is reversed and the energies are $f_x = f_{\text{strong}}$ and $f_y = f_{\text{weak}}$. For nucleation to occur within a three molecule system we require

that the fully bound state is lower in free energy than the unbound state so $f_{\text{strong}} + f_{\text{weak}} < 0$. As described below, it is unlikely that peptides satisfy this condition, however, this model captures the effect of conformational entropy on nucleation rate and, therefore, provides useful intuition.

Unstable dimers provide the substrate for nucleation. In analogy to Eq. 1, we write the nucleation rate for a fibril as

$$k_{\text{nuc}} = \sum_{x=1}^{L} k_{\text{diff}} \mathcal{E}_+(x,1) C_{n^{\ddagger}}(x), \quad (3)$$

Here, $C_{n^{\ddagger}}$ is the concentration of a critical cluster containing $n^{\ddagger}$ molecules. In the trimer model, $C_{n^{\ddagger}} = C_2(x)$ [40] where $C_2(x)$ represents the equilibrium concentration of dimers with x H-bonds at the supersaturated monomer concentration C_1

$$C_2(x) = C_1^2 e^{-x f_{\text{weak}}/k_BT}, \quad (4)$$

where all concentrations are scaled by the concentration of pure water (55.5 M) to yield dimensionless number densities and the monomer concentration plays the role of the fugacity, $C_1 = e^{\mu/k_BT}$. The difference between Eq. 1 and Eq. 3 is the explicit summation over trajectories in the 2D reaction coordinate space. These trajectories are indexed by the number of H-bonds in the initial dimer, x, which has a maximum value, L, set by the length of the molecules.

The factor k_{diff} is the diffusion limited rate of monomer-dimer collisions resulting in trimers. Following [20], we approximate k_{diff} by the rate of particles striking an absorbing sphere of radius a, $k_{\text{diff}} = 4\pi a C_1 D_p$, where D_p is the monomer diffusion constant. This rate assumes that all collisions result in binding and, therefore, neglects sequence effects that enforce specific alignments between the molecules. This treatment is certainly appropriate for homopolymers like polyglutamine, but may be more generally valid if small clusters have the flexibility to accommodate sidechain packing errors. The opposite case, where sidechain registry is rigorously enforced, can be described by scaling k_{diff} by L^{-1}.

Finally, like the Zeldovich factor, $\mathcal{E}_+$ describes the probability of successful nucleation. We assume that the initial contact between a dimer and monomer results in a β-sheet H-bond so that a dimer with n H-bonds becomes a trimer with coordinates $(n,1)$ in (x,y) space. This trimer will evolve according to the rate constants in Eq. 2 and perform a random walk in the (x,y) plane. For successful nucleation we require that the walk proceed to (L,L) without striking the boundaries $(x,0)$ or $(0,y)$ which signify dissolution back to the monomer+dimer state. Therefore, $\mathcal{E}_+(x,y)$ is the splitting probability that a random walk starting at (x,y) reaches (L,L) before reaching either the x or y axes.

The nucleation probability is described by 2D diffusion. In the continuum limit the success probability

satisfies the convection-diffusion equation [24]

$$\nabla \cdot (\mathbf{D} \nabla \mathcal{E}_+) + \nabla \cdot (\mathbf{v} \mathcal{E}_+) = 0, \tag{5}$$

where the velocity and diffusion tensor are defined by

$$\mathbf{v} = \begin{pmatrix} v_x \\ v_y \end{pmatrix} = \begin{pmatrix} k_{+x} - k_{-x} \\ k_{+y} - k_{-y} \end{pmatrix} \tag{6}$$

$$\mathbf{D} = \begin{pmatrix} \frac{k_{+x} + k_{-x}}{2} & 0 \\ 0 & \frac{k_{+y} + k_{-y}}{2} \end{pmatrix}. \tag{7}$$

The velocity is constant on either side of the $x = y$ diagonal but switches direction such that the drift always pushes the random walk toward the diagonal (Fig. 1).

The rate constants are related to each other by the free energy of the H-bond being broken/formed (Eqs. 2), however, the values are unknown beyond an expectation that the rates are on the order of ns^{-1} [25]. Therefore, we adopt a convention where the forward and backward rates have a constant average $D = (k_{+x} + k_{-x})/2 = (k_{+y} + k_{-y})/2 \simeq 1$ns. This assumption is motivated by the simplification resulting from isotropic diffusion. However, it also qualitatively captures our expectation that weak bonds form more slowly than strong bonds because both peptide chains must simultaneously adopt β conformation. Similarly, we expect that weak bonds will break faster than strong bonds because there are twice as many degrees of freedom that can lead to bond breakage.

With Eqs. 2, 6, and the assumption of isotropic diffusion, the drift velocities can be expressed as

$$v_{x/y}(f) = 2D \frac{e^{-f/k_B T} - 1}{e^{-f/k_B T} + 1}, \tag{8}$$

where f is either f_{weak} or f_{strong} as appropriate. Eq. 5 can then be solved. See Supplemental Material at [URL will be inserted by publisher] for details of the solution. For $x > y$ we find

$$\mathcal{E}_+(x,y) = e^{-\frac{v_x}{2D}(x-L) - \frac{v_y}{2D}(y-L)}$$

$$\times \frac{\sum_{m=2}^{\infty} B_m \sin(m \arctan(y/x)) I_m(\frac{v\sqrt{x^2+y^2}}{2D})}{\sum_{m=2}^{\infty} B_m \sin(m\pi/4) I_m(\frac{v\sqrt{2}L}{2D})}, \tag{9}$$

where $v = |\mathbf{v}|$, I is the modified Bessel function, and the first four coefficients are $B_2 = 1$, $B_3 = B_5 = 2(v_x - v_y)/v$, and $B_4 = 2(v_x - v_y)^2/v^2$. The solution for $y > x$ can be obtained by exchanging $x \leftrightarrow y$.

Clusters with low free energy have low probability for successful nucleation. Fig. 2 shows the two dominant contributions to the nucleation rate. Fig. 2a shows the probability $\mathcal{E}_+$ that a trimer with (x, y) intermolecular H-bonds proceeds to the fully H-bonded state without dissociating back to monomer+dimer. This is an increasing function of x and y. In particular, $\mathcal{E}_+$ is very small near the origin because these trimer states readily decay by losing either terminal molecule.

Fig. 2b shows the grand free energy of the cluster as a function of x and y. The loss of translational entropy,

$\mu = k_B T \ln C_1$, upon molecular binding results in a discontinuous jump in the free energy when x or y become nonzero. There is a pronounced valley along the $x = y$ diagonal indicating the lowest free energy pathway to nucleation. This is the pathway that would be chosen in an attempt to model nucleation using a 1D reaction coordinate (as in a saddle point approximation). Note that the free energies are repulsive when there is a large mismatch between x and y. Of particular interest are the dimer states, where either x or y are zero, since these states are the starting point for nucleation attempts (Eqs. 3, 4). These states have especially unfavorable free energies due to the entropic cost of straightening both peptides and are strongly suppressed thermodynamically.

These two plots lead us to the primary result of the trimer model. Dimers held together by a small number of H-bonds are plentiful but have a low probability of nucleation because, upon the addition of a third molecule, the ensuing trimer has many trajectories that take it back to the dimer state. Conversely, dimers that have formed many H-bonds are rare but are likely to proceed toward nucleation upon the addition of a third molecule. The result of these competing trends is that the nucleation flux is a non-monotonic function of the starting dimer structure with the peak flux originating from a partially ordered nucleus (Fig. 2c). Note that the suppression of flux at the saddle point is not simply a result of the uniqueness of this trajectory compared to the multitude of non-saddle point trajectories, instead, the system dynamics make is so that successful passage through the saddle point is especially unlikely.

3D model contains both backbone and sidechain interactions. While the trimer model presented above provides useful intuition into the nucleation process, it neglects major features of amyloid nucleation. To address this we seek a model that can handle arbitrary sized clusters and accounts for the contributions of both backbone H-bond and sidechain steric zipper interactions in compensating for the translational and conformational entropy loss upon aggregation. In this model we consider a cluster of N molecules, each of which have x amino acids in the ordered β core and $L - x$ that remain disordered with a reference free energy of zero. Following the usage of Eq. 3, x will be used to index trajectories.

As input for the model, we require estimates of the free energy as a function of x and N. In a mature fibril the free energy gain upon the addition of a new molecule is on the order of $f_{\alpha\alpha} \simeq -0.5 \ k_B T$ per amino acid in the cross-β core [15, 26]. This free energy is the sum of three contributions, $f_{\alpha\alpha} = f_{\text{HB}} + f_{\text{SZ}} + f_{\text{CE}}$, where the terms represent the backbone H-bond, sidechain packing in the steric zipper, and the loss of conformational entropy. Helix-coil models have shown that f_{CE} is on the order of $\sim 2k_B T$ [27] so the two attractive components sum to approximately $-2.5 \ k_B T$. As a crude estimate we will say that $f_{\text{HB}} \simeq -1.5 \ k_B T$ and $f_{\text{SZ}} \simeq -1 \ k_B T$. Using the structures illustrated in Fig. 3a we estimate the energies as $x(2f_{\text{CE}} + f_{\text{HB}})$ for the dimer, $x(3f_{\text{CE}} + f_{\text{HB}} + 2f_{\text{SZ}})$ for
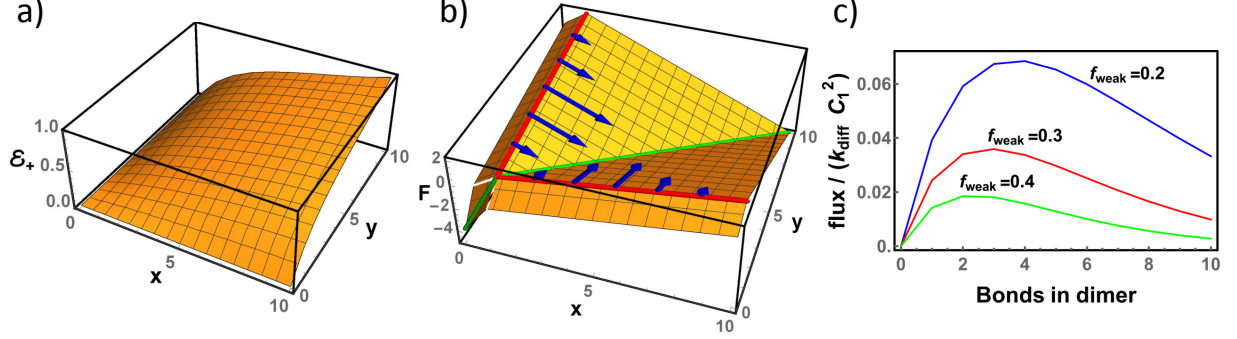
FIG. 2: (a) Probability of a trimer containing x and y intermolecular bonds to proceed to the fully bound state before either terminal molecule unbinds. (b) Grand free energy (in units of $k_B T$) of a trimer as a function of the number of intermolecular bonds. The increase in free energy when x or y increase from 0 to 1 arises from the loss of translational entropy upon molecular binding. This jump, given by $k_B T \ln C_1$, is arbitrarily set to $2 k_B T$ for convenient visualization. There is a pronounced trough along the $x = y$ diagonal (green line), however, due to the low probability of nucleation when x and y are both small (panel a) most nucleation trajectories (blue arrows) avoid this path. (c) The nucleation flux as a function of the number of bonds in the starting dimer (see Eq. 3). This is proportional to the product of the nucleation probability (panel a) along the $y = 1$ line and the dimer Boltzmann weight ($y = 0$ line in panel b). The competition between these terms gives a non-monotonic function that has a peak significantly removed from the lowest free energy pathway ($x = 1$). This is shown schematically by the blue arrows in the middle panel. $L = 10$, $f_{\text{strong}} = -0.5$, $f_{\text{weak}} = 0.3$ unless noted.

the trimer, and $x(4f_{\text{CE}} + 2f_{\text{HB}} + 2f_{\text{SZ}})$ for the tetramer. Therefore, the highest free energy cluster is a trimer. This analysis is valid for molecules that form a single β-strand in the fibril. An important second case to consider is molecules like Aβ and IAPP [28–30] that form a hairpin and contribute β-strands to both layers of the fibril (Fig. 3b). For these molecules the highest free energy state is a hairpin monomer. The formation of a hairpin incurs an entropic cost of $k_B T \ln x_{\text{loop}}$, where x_{loop} is the number of disordered amino acids between the β segments. This term is $\sim$2-3 $k_B T$, which is less than the translational entropy cost of making the equivalent structure from two molecules. The looping penalty can be mitigated by attractive sidechain interactions, which may be responsible for the difference in aggregation propensity between Aβ_{40} and Aβ_{42} [31]. With these calculations, the Arrhenius terms for single strand and hairpin molecules are

$$C_{n^{\ddagger}} = \begin{cases} C_1^3 e^{-x(3f_{\text{CE}} + f_{\text{HB}} + 2f_{\text{SZ}})/k_B T} & \text{single strand} \\ C_1 e^{-2xf_{\text{CE}}/k_B T - \ln(x_{\text{loop}})} & \text{hairpin} \end{cases}$$

(10)

These calculations show that the population of critical clusters is exponentially sensitive to both the number of molecules in the cluster and the size of the ordered core. As the cluster grows the bulk interactions in the core, contributing a free energy $(N - 2)xf_{\alpha\alpha}$, will gradually overwhelm the surface energy $2xf_{\text{CE}} + N \ln C_1$. Adding these terms together and requiring that the resulting free energy is stationary in x, we obtain the minimum number of β-strands, $N^{\star} = 2 - 2f_{\text{CE}}/f_{\alpha\alpha}$, necessary for it to be energetically favorable for x to grow. Thus, when the cluster exceeds $N^{\star} \simeq 10$ β-strands, fibril growth transitions from nucleation to elongation.

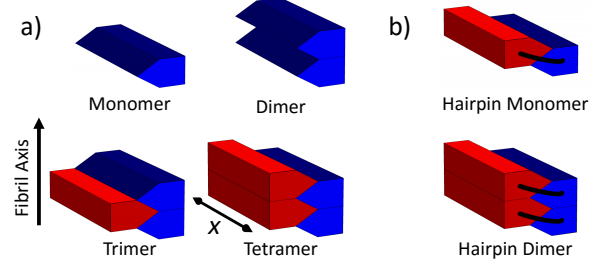Next, we need to evaluate the size evolution of clusters

FIG. 3: Schematic of the assembly process described by the 3D model for (a) peptides that form a single β-sheet and (b) peptides that form two β-sheets separated by a hairpin. Each block represents a β strand consisting of x amino acids. Disordered peptide segments have been omitted for clarity. The two β-sheets of the nascent fibril are color coded so red/blue (light grey/dark grey) interfaces represent sidechain steric zipper interactions while red/red and blue/blue interfaces represent backbone H-bond interactions.

smaller than $N^{\star}$. These clusters will add a molecule with rate k_{diff} and lose a molecule when either terminal strand breaks all x bonds with the neighboring molecules. The rate of such loss events is given by [20]

$$k_{\text{loss}}^{-1} = -\frac{1}{2v_{\text{f}}} + \frac{D}{2v_{\text{f}}^2} e^{v_{\text{f}}x/D}(1 - e^{-v_{\text{f}}/D}), \quad (11)$$

where v_{f} is given by Eq. 8 with $f = f_{\alpha\alpha}$, the factor of two accounts for the two ends of the fibril, and the substitution $x \to 2x$ should be made for hairpin molecules. The probabilities that a cluster gains or loses a molecule are $p_{+\text{mol}} = k_{\text{diff}}(k_{\text{diff}} + k_{\text{loss}})^{-1}$ and $p_{-\text{mol}} = k_{\text{loss}}(k_{\text{diff}} + k_{\text{loss}})^{-1}$. We expect that $in\ vivo$ k_{loss} greatly exceeds k_{diff}, so nucleation depends on the unlikely event

where many molecules add with few detachments [32]. The tendency for backwards steps is prohibitive for nucleation unless the cluster can present a binding surface large enough to prolong the residence time of incoming molecules. Thus, the present model contains the same physics obtained from the trimer model; that successful nucleation events are dominated by highly ordered states that are significantly displaced from the saddle point.

To compute the nucleation rate in our 3D model, we make the approximation that x remains constant until the cluster reaches a stable size and compute the nucleation rate by summing the rates of clusters of all core sizes. This rough approximation is the simplest approach to take without knowing the relative timescales for the evolution of N and x. The reduction of the 2D Markov process (in N and x) to a 1D walk (in N only) assumes that the energetically favorable H-bonds will bias the bonding states of the molecules toward the maximum number of bonds until the actual unbinding event. With this simplification, the required Zeldovich factor in Eq. 3 is the probability of a super-critical cluster with size $n^{\ddagger} + 1$ reaching the stable size $N^{\star}$ without becoming smaller than $n^{\ddagger}$. The probability of such walks is (see Supplemental Material for derivation)

$$\mathcal{E}_{\text{mol}}(1) = \frac{\frac{1}{p_{+\text{mol}}} - 2}{\left(\frac{1}{p_{+\text{mol}}} - 1\right)^{N^{\star} - n^{\ddagger}} - 1}. \quad (12)$$

Theory predicts two power law regimes. The nucleation rate for the 3D model is given by Eq. 3 with Eqs. 10 and 12 for the Arrhenius and Zeldovich terms. At low concentrations, representative of *in vivo* systems, the nucleation rate scales like $C_1^{N^{\star}}$ (Fig. 4). This gives $k_{\text{nuc}} \propto C_1^{10}$ and $k_{\text{nuc}} \propto C_1^5$ for single strand and hairpin molecules, respectively, although this prediction is sensitive to sequence dependent binding energetics. At high concentrations it becomes more likely that molecules bind than unbind and the Zeldovich factor saturates at unity giving a concentration dependence of $C_1^{n^{\ddagger}+1}$. This corresponds to $k_{\text{nuc}} \propto C_1^4$ for single strand molecules and $k_{\text{nuc}} \propto C_1^2$ for hairpin molecules in agreement with previous theory and experiment [33–35].

Measured nucleation rates range from $\sim 10^{-11}$ L^{-1}s^{-1} at μM concentrations (Aβ) to $\sim 10^6$ L^{-1}s^{-1} at mM concentrations (insulin) [3, 8, 34]. Assuming these molecules nucleate by the hairpin mechanism, these rates are considerably slower than the predictions of our theory. While it is well known that nucleation theories often suffer from quantitative discrepancies despite capturing the scaling trends [36–39], there are several features of the theory

that may also be contributing. These include approximations like the projection of the nucleus growth onto a 1D random walk, the absorbing boundary condition at a cluster size of $N^{\star}$, and the neglect of sequence effects. In particular, the latter contribution is particularly large because the requirement of in-register binding is equivalent to reducing the effective concentration by $\sim L^{-1}$.
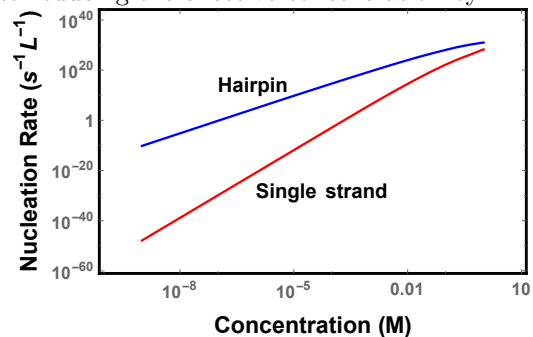

FIG. 4: Nucleation rate of hairpin and single strand molecules. Molecules that form hairpin structures nucleate much faster than molecules that only contribute one strand because of the reduced translational entropy penalty to form the nucleus. The curvature at high concentration is due to the saturation of the Zeldovich factor.

Despite the over-estimate of nucleation rates at *in vitro* concentrations, the predicted rates shown in Fig. 4 lead to the conclusion that homogenous nucleation is too slow to be significant at physiological concentrations. Therefore, nucleation under these conditions will require heterogeneous mechanisms to ameliorate the nucleation barrier. This will be the subject of future work.

With these quantitative limitations, the primary results of our model are the predictions of two different power law regimes and the observation that nucleation is dominated by trajectories that avoid the saddle point. Primary nucleation is a difficult subject to study because it is difficult to disentangle from the background of secondary nucleation events [4–8]. Our hope is that simple models like ours will inspire further studies into the mechanism of amyloid nucleation.

[1] J. Hardy and D. J. Selkoe, Science (New York, N.Y.) **297**, 353 (2002).

[2] W. W.-F. Xue, S. W. S. S. W. Homans, and S. E. Radford, Proc. Nat. Acad. Sci. **105**, 8926 (2008).

[3] T. P. J. Knowles *et al.*, Proc. Nat. Acad. Sci. **108**, 14746 (2011).

[4] T. P. J. Knowles *et al.*, Science (New York, N.Y.) **326**, 1533 (2009).

[5] S. I. A. Cohen *et al.*, J. Chem. Phys. **135**, 065105 (2011).

[6] S. I. A. Cohen, M. Vendruscolo, C. M. M. Dobson, and T. P. J. Knowles, J. Chem. Phys. **135**, 065106 (2011).

[7] S. I. A. Cohen, M. Vendruscolo, C. M. M. Dobson, and T. P. J. Knowles, J. Chem. Phys. **135**, 065107 (2011).

[8] S. I. A. Cohen *et al.*, Proc. Nat. Acad. Sci. **110**, 9758 (2013).

[9] R. Nelson *et al.*, Nature **435**, 773 (2005).

[10] M. R. Sawaya *et al.*, Nature **447**, 453 (2007).

[11] J. Zhang and M. Muthukumar, J. Chem. Phys. **130**, 035102 (2009).

[12] R. Cabriolu, D. Kashchiev, and S. Auer, J. Chem. Phys. **133**, 225101 (2010).

[13] D. Kashchiev and S. Auer, J. Chem. Phys. **132**, 215101 (2010).

[14] S. Auer, C. Dobson, M. Vendruscolo, and A. Maritan, Phys. Rev. Lett. **101**, 17 (2008).

[15] J. D. Schmit, K. Ghosh, and K. A. Dill, Biophys. J. **100**, 450 (2011).

[16] S. E. Hill, T. Miti, T. Richmond, and M. Muschol, PloS One **6**, e18171 (2011).

[17] T. Miti, M. Mulaj, J. D. Schmit, and M. Muschol, Biomacromol. **16**, 326 (2015).

[18] W. Qiang, K. Kelley, and R. Tycko, J. Am. Chem. Soc. **135**, 6860 (2013).

[19] C. F. Lee, J. Loken, L. Jean, and D. J. Vaux, Phys. Rev. E **80**, 041906 (2009).

[20] J. D. Schmit, J. Chem. Phys. **138**, 185102 (2013).

[21] M. Zamparo, A. Trovato, and A. Maritan, Phys. Rev. Lett. **105**, 1 (2010), 1010.4459.

[22] S. Whitelam and R. L. Jack, Ann. Rev. Phys. Chem. **66**, 143 (2015), 1407.2505.

[23] Y. Zeldovich, Acta Physicochim. URSS **18**, 1 (1943).

[24] S. Redner, *A Guide to First-Passage Processes* (Cambridge University Press, 2007).

[25] V. Muñoz, P. A. Thompson, J. Hofrichter, and W. A. Eaton, Nature **390**, 196 (1997).

[26] A. J. Baldwin *et al.*, J. Am. Chem. Soc. **133**, 14160 (2011).

[27] K. Ghosh and K. A. Dill, J. Am. Chem. Soc. **131**, 2306 (2009).

[28] T. Lührs *et al.*, Proc. Nat. Acad. Sci. **102**, 17342 (2005).

[29] A. T. Petkova *et al.*, Proc. Nat. Acad. Sci. **99**, 16742 (2002).

[30] S. Luca, W. M. Yau, R. Leapman, and R. Tycko, Biochemistry **46**, 13505 (2007).

[31] N. G. Sgourakis, Y. Yan, S. A. McCallum, C. Wang, and A. E. Garcia, J. Mol. Biol. **368**, 1448 (2007).

[32] J. T. Jarrett and P. T. Lansbury, Cell **73**, 1055 (1993).

[33] A. M. Ruschak and A. D. Miranker, Proc. Nat. Acad. Sci. **104**, 12341 (2007).

[34] G. Meisl *et al.*, Proc. Nat. Acad. Sci. **111**, 9384 (2014).

[35] F. A. Ferrone, Journal of Molecular Biology **427**, 287 (2015).

[36] R. P. Sear, J. Phys. Cond. Matt. **19**, 033101 (2007).

[37] L. Filion, M. Hermes, R. Ni, and M. Dijkstra, J. Chem. Phys. **133** (2010), 1006.2925.

[38] L. Filion, R. Ni, D. Frenkel, and M. Dijkstra, J. Chem. Phys. **134**, 134901 (2011).

[39] W. Lechner, C. Dellago, and P. G. Bolhuis, J. Chem. Phys. **135**, 154110 (2011).

[40] In the trimer model we are taking the critical state to be the one with the highest free energy and not the one with the highest grand free energy (red line in Fig. 2b). The latter choice would require $C_{n\ddagger} = C_3(x, y = 1)$ with k_{diff} replaced by the rate of forming the first favorable H-bond. Our choice of the dimer as the critical state allows a more facile comparison to CNT and our 3D nucleation model where the attempt rate is diffusion-limited. This somewhat arbitrary choice will affect the quantitative rates, but will not change the primary result of the trimer model, which is the qualitative observation that the saddle point trajectory is uniquely disfavored.