Reaction rate theory for supramolecular kinetics: application to protein aggregation

Michaels, T.C.T.; Liu, L.X.; Curk, S.; Bolhuis, P.G.; Šarić, A.; Knowles, T.P.J.

Taylor & Francis
Taylor & Francis Group

Check for updates

FRENKEL SPECIAL ISSUE

Reaction rate theory for supramolecular kinetics: application to protein aggregation

Thomas C. T. Michaels[a,b], Lucie X. Liu[a*], Samo Curk[c,d*], Peter G. Bolhuis[e], Anđela Šarić[c] and Tuomas P. J. Knowles[a,f]

[a]Department of Chemistry, University of Cambridge, Cambridge, UK; [b]Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA; [c]Department of Physics and Astronomy, Institute for the Physics of Living Systems, University College London, London, UK; [d]Department of Physics, Faculty of Natural Sciences and Mathematics, University of Maribor, Maribor, Slovenia; [e]van't Hoff Institute for Molecular Sciences, University of Amsterdam, Amsterdam, Netherlands; [f]Cavendish Laboratory, Department of Physics, University of Cambridge, Cambridge, UK
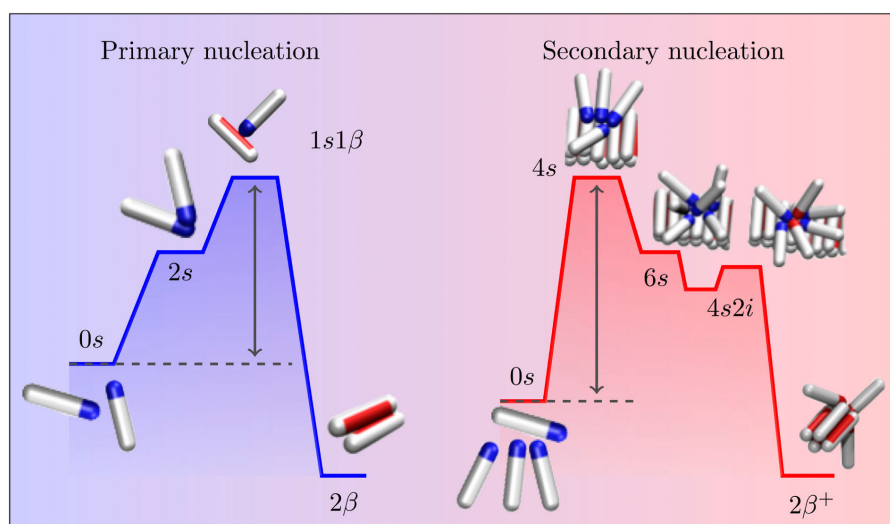
ABSTRACT

Probing reaction mechanisms of supramolecular processes in soft and biological matter, such as protein aggregation, is inherently challenging. This is because these processes involve multiple molecular mechanisms that are associated with the rearrangement of large numbers of weak bonds, resulting in complex free energy landscapes with many kinetic barriers. Reaction rate measurements at different temperatures can offer unprecedented insights into the underlying molecular mechanisms. However, to be able to interpret such measurements, a key challenge is to establish which properties of the complex free energy landscapes are probed by the reaction rate. Here, we present a reaction rate theory for supramolecular kinetics based on Kramers theory of diffusive reactions over multiple kinetic barriers. We find that reaction rates for protein aggregation are of the Arrhenius–Eyring type and that the associated activation energies probe only one relevant barrier along the respective free energy landscapes. We apply this advancement to interpret, in experiments and in coarse-grained computer simulations, reaction rates of amyloid aggregation in terms of molecular mechanisms and associated thermodynamic signatures. These results suggest a practical extension of the concept of rate-determining steps for complex supramolecular processes and establish a general platform for probing the underlying energy landscape using kinetic measurements.

1. Introduction and motivation

The mechanisms of macromolecular reactions in soft and biological matter, such as protein–protein association or protein aggregation, are notoriously difficult to probe in experiments. This difficulty originates from the fact that these complex macromolecular processes involve the concurrent making and breaking of very large numbers of bonds and interactions between the multiple molecular species present. Historically, the key for investigating molecular mechanisms of small molecule reactions has been to probe the underlying free energy landscape by measuring the temperature dependence of the reaction rates. Reaction rate theory then provides the framework for relating these measurements to the thermodynamics of the underlying free energy landscape.

The discipline of rate theory was established when Arrhenius [1] described the temperature dependence of the rate k of a chemical reaction in terms of what is now known as the Arrhenius equation: $k = \nu\, e^{-\beta \Delta G^{\ddagger}}$, where ν is a frequency pre-factor, $\beta = 1/k_B T$ is the inverse temperature and $\Delta G^{\ddagger}$ is the free energy barrier associated with the rate-determining step. The next substantial development came with Eyring [2] in the 1930s, who assumed that the reaction is governed by a rate determining step which corresponds to the breaking of a single quantum mechanical chemical bond. This assumption allows explicit calculation of the frequency pre-factor as $\nu = k_B T/h$, where h is the Planck constant. Eyring's equation has proved very successful in describing the reactions of small molecules, but it is not expected to apply to supramolecular processes involving macromolecules, as these processes require the rearrangement of large numbers of bonds rather than breaking of a single quantum mechanical mode of vibration (Figure 1). Hence, the associated energy landscape in this case involves many kinetic barriers along the reaction coordinate. A more physically realistic model for these systems is offered by Kramers rate theory [3–6]. In this theory, reactions are described as diffusion processes along a complex free energy landscape which is parameterised by just a few important coordinates. The reaction rate corresponds to the inverse of the escape time, and it is found that the reaction rate is of the Arrhenius–Eyring type, whereby the pre-factor depends on key features of the free energy landscape. In particular, it is found that $\nu = \omega_1 \omega_2/(2\pi\gamma)$, where ω_1 and ω_2 denote the curvatures of the potential landscape at the bottom and the top of the free energy barrier, respectively, and γ is the friction coefficient.

Unravelling the molecular mechanisms of macromolecular diffusive processes thus requires solving the inverse problem of characterising the free energy landscape through measurements of the reaction rate. Clearly, the reaction rate will contain the information about the free energy barrier associated with the rate-determining step, $\Delta G^{\ddagger}/k_B T$, via the Kramers equation. Specifically, by measuring the temperature variation of the reaction rate, the information about the free energy barrier becomes directly accessible. A key question therefore is to establish which free energy barrier on a complex multi-barrier landscape is rate-determining and is thus read out by such a measurement.

To address these questions about the interpretation of reaction rate measurements of supramolecular processes, we review and apply Kramers reaction theory [3–5] to model the molecular mechanisms of supramolecular processes governed by diffusive dynamics and, conversely, to establish which information about the free energy landscape can be obtained from the temperature dependence of the associated reaction rate. We find that only one specific barrier from the multi-barrier landscape is probed by such measurements. We then apply this framework to study the energetics of protein aggregation phenomena, a biologically relevant example of multi-step diffusive processes with implications in areas ranging from biomedicine [7] to nanotechnology [8–11]. Specifically, we apply Kramers theory in conjunction with coarse-grained computer simulations and kinetic experiments to determine the thermodynamic characteristics of key steps involved in the aggregation of Alzheimer's Amyloid-β peptide into amyloid fibrils. These results suggest a natural extension of the concept of a rate-determining step in the context of protein aggregation and establish a platform for probing the energetics of complex macromolecular reactions in soft matter.

2. Kramers theory of diffusive reactions with multiple kinetic barriers

Let us consider a diffusive reaction between well-defined initial and final states, x_0 and x_e, taking place over a one-dimensional potential free energy landscape, $G(x)$, with multiple barriers (Figure 1(d)). The following Fokker–Planck equation then describes the time evolution of the probability $p(x, t \mid x_0)$ that, starting at x_0, the system has diffused to position x at time t [4]:

$$\frac{\partial p(x, t \mid x_0)}{\partial t} = -\frac{\partial \mathcal{J}}{\partial x},\qquad (1)$$

where

$$\mathcal{J} = -\frac{1}{\gamma}\frac{\partial G(x)}{\partial x}p(x, t \mid x_0) - D\frac{\partial p(x, t \mid x_0)}{\partial x}.\qquad (2)$$

Here, γ denotes the frictional coefficient and D is the diffusion coefficient. Note that γ and D are related to
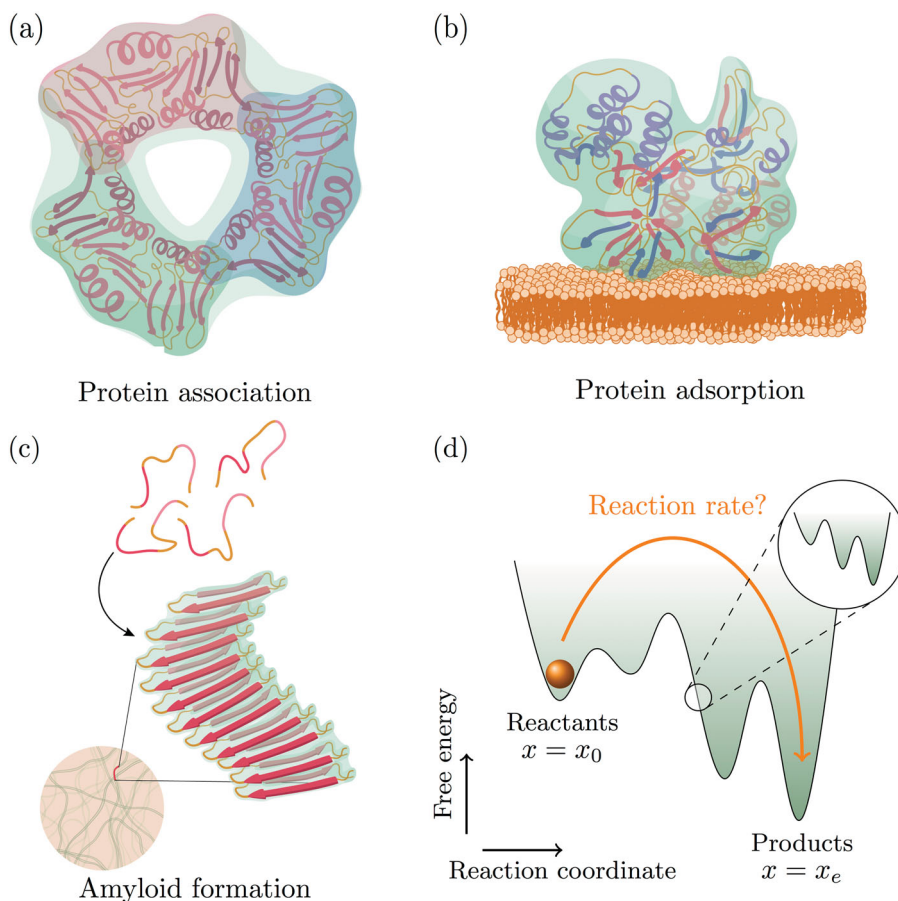
Figure 1. Examples of supramolecular kinetics in soft and biological matter. (a) Formation of protein complexes, (b) protein adsorption and (c) amyloid fibril formation. (d) Supramolecular kinetics are characterised by a complex free energy landscape with multiple kinetic barriers. A key question is to establish which features of the detailed free energy landscape are probed by measurements of the overall reaction rate.

the thermal energy through the Einstein–Smoluchowski relation, $\gamma D = k_B T$.

The key quantity of interest is the average first passage time $\tau(x_0 \to x_e)$, i.e. the time that it takes, on average, for a system described by Equation (1) and starting at x_0 to reach x_e. In fact, the inverse of the average first passage time corresponds to the transition rate from x_0 to x_e [12]

$$k(x_0 \to x_e) = \frac{1}{\tau(x_0 \to x_e)}, \qquad (3)$$

thus providing a link between free energy landscapes and experimental reaction rate measurements. The average first passage time $\tau(x_0 \to x_e)$ is related to the probability $p(x, t \mid x_0)$ through (see Supplementary Material)

$$\tau(x_0 \to x_e) = \int_0^\infty dt \int_{-\infty}^{x_e} dx \, p(x, t \mid x_0). \qquad (4)$$

Integrating the Fokker–Planck Equation (1) using Equation (4), we find that $\tau(x_0 \to x_e)$ satisfies the following differential equation (see Supplementary Material):

$$-\frac{1}{\gamma} \frac{\partial G(x)}{\partial x} \frac{\partial \tau(x \to x_e)}{\partial x} + D \frac{\partial^2 \tau(x \to x_e)}{\partial x^2} = -1.$$
$$(5)$$

The solution to Equation (5) subject to the boundary condition $\tau(x_e \to x_e) = 0$ is (see Supplementary Material) [5]:

$$\tau(x_0 \to x_e) = \beta \gamma \int_{x_e}^{x_0} dy \int_{-\infty}^{y} dz \, e^{\beta[G(y) - G(z)]}. \qquad (6)$$

In the limit when the relevant free energy barrier is much bigger than thermal energy ($\beta \Delta G \gg 1$), the integrals in Equation (6) can be evaluated explicitly using the saddle point approximation [13]. In particular, assuming that the width does not vary significantly between the multiple potential energy barriers, we need to maximise the integrand over the integration range of Equation (6), i.e. we need to find $\max_{z \leq y}[G(y) - G(z)]$. Let $y = x^*$ and $z = x_*$ denote the points in the range of integration

of Equation (6) where the integrand in Equation (6) is maximal. We find (see Supplementary Material):

$$\tau(x_0 \to x_e) \simeq \frac{2\pi\gamma}{\omega_1\omega_2} e^{\beta\Delta G^{\ddagger}}, \qquad (7)$$

where

$$\Delta G^{\ddagger} = \max_{x_0 \leq z \leq y \leq x_e} [G(y) - G(z)] \qquad (8)$$

and ω_1 and ω_2 are the curvatures of the free energy landscape at x_* and x^*, respectively.

2.1. Identifying the rate-determining free energy barrier from kinetic experiments

Using Equation (3) we find that the escape rate is in the form of the Arrhenius–Eyring equation:

$$k(x_0 \to x_e) \simeq A \, e^{-\beta\Delta G^{\ddagger}}, \qquad (9)$$

where $A = \omega_1\omega_2/2\pi\gamma$ is a pre-factor, which depends on the curvatures of the potential landscape at x_* and x^*, respectively. Note that, although the energy landscape includes multiple intermediate kinetic barriers, only one relevant free energy barrier $\Delta G^{\ddagger}$ determines the escape rate $k(x_0 \to x_e)$ and hence can be probed directly by kinetic experiments. This relevant free energy barrier is thus a natural generalisation of the concept of the rate-determining step for small molecule reactions to complex supramolecular processes. The rate-determining free energy barrier is found using Equation (8) and corresponds to the largest possible free energy difference between any local maximum and any local minimum preceding it. Figure 2 illustrates this principle for a series of three examples of energy landscapes.

As Equation (9) predicts that $\tau(x_0 \to x_e)$ has an exponential dependency on the height of the rate-determining energy barrier, using the relationship $\Delta G^{\ddagger} = \Delta H^{\ddagger} - T\Delta S^{\ddagger}$, where $\Delta H^{\ddagger}$ is the enthalpy of activation and $\Delta S^{\ddagger}$ is the entropy of activation, and absorbing the entropy contribution into the pre-factor, we find $k(x_0 \to x_e) \simeq e^{-\beta\Delta H^{\ddagger}}$. Hence, a plot of $\log k(x_0 \to x_e)$ against $\beta = 1/k_BT$ is expected to yield a straight line with the enthalpy of activation corresponding to the rate-determining barrier as the slope:

$$\frac{\Delta H^{\ddagger}}{k_B} = -\frac{\partial \log k(x_0 \to x_e)}{\partial(1/T)}. \qquad (10)$$

This equation provides the key for interpreting reaction rate measurements at varying temperature in terms of the enthalpy of activation of the rate-determining step. Note, however, that replacing $\Delta G^{\ddagger}$ with $\Delta H^{\ddagger}$ in Equation (8) is wrong. This is because, the relevant activation energy barrier is determined by the free energy landscape (Figure 2(a)), while the measured temperature dependence of rate constants only reflects the enthalpic contribution to this barrier, which need not correspond to the highest enthalpy change (Figure 2(b)).

To test numerically the theoretical predictions of Equations (8)–(10), we performed Monte Carlo (MC) and Molecular Dynamics simulations [14] of diffusion of a single particle on a series of examples of one-dimensional potential energy landscapes (see Section 2 of Supplementary Material). We have also tested the theoretical predictions of Equations (8)–(10) by performing an explicit analysis of the spectrum of the rate matrix describing the transitions in the energy landscapes of Figure 2(a) (see Section 3 of Supplementary Material). This analysis establishes all the conditions under which the energy landscapes of Figure 2(a) lead to an effective two-state kinetics and confirms the nature of the rate-determining barrier. This discussion thus provides an alternative view on the results from Kramers theory and allows for an exhaustive analysis of the possible scenarios in the case of a three-state energy landscape. In addition, this analysis illustrates the time-dependence of the population in each of the states (see Section 3.2 of Supplementary Material).

2.2. The frequency pre-factor

It is useful to conclude this section with a remark on the exponential pre-factor. Unlike the enthalpy of activation, $\Delta H^{\ddagger}$, the free energy of activation, $\Delta G^{\ddagger}$, is crucially coordinate dependent. This raises the question of the appropriate choice for the reaction coordinate or, equivalently, of an appropriate frequency pre-factor A. While Kramers theory in principle provides an explicit formula for this pre-factor via Equation (9), this expression depends on parameters such as the curvature of the free energy landscape at the top of the rate-determining barrier, which are commonly inaccessible in experiments. A possible strategy to overcome this limitation consists in partitioning all of the missing information about diffusion along the reaction coordinate into the free energy barrier in the rate equation by re-writing the escape rate as

$$k(x_0 \to x_e) = A^{\text{phys}} \, e^{-\beta\Delta G^{\ddagger,\text{phys}}}, \qquad (11)$$

where $\Delta G^{\ddagger,\text{phys}} = \Delta G^{\ddagger} + T\Delta S^{\text{i}}$ and $\Delta S^{\text{i}} = k_B \log (A^{\text{phys}}/A)$. Here, A^{phys} is a known frequency pre-factor which can be constructed conveniently from the experimentally accessible information about the reactive flux towards the relevant barrier. Hence, ΔS^{i} can be interpreted as an additional entropy term representing the fact that we inevitably do not have complete information
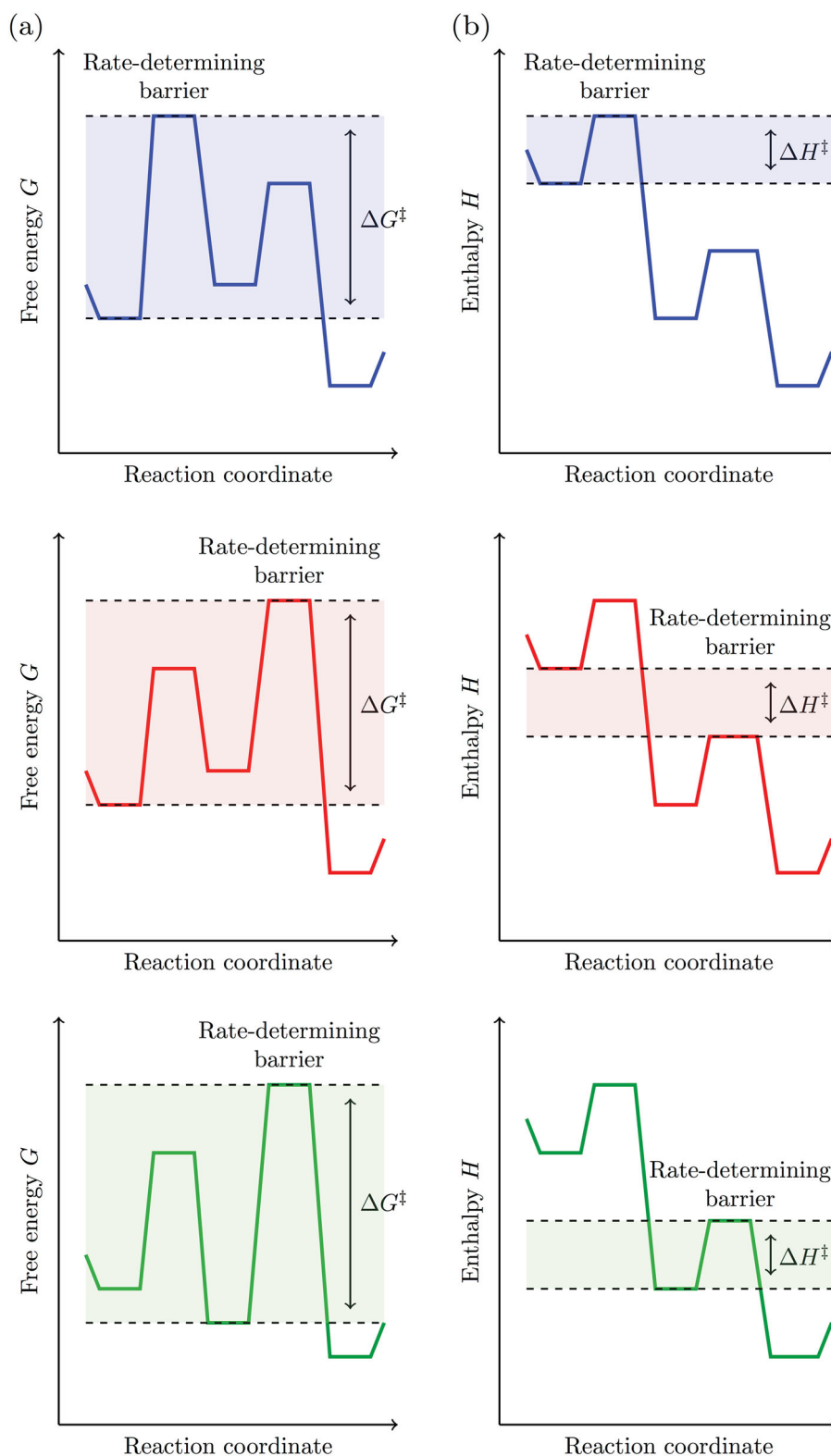
Figure 2. Relating temperature-dependent measurements of reaction rates to the rate-determining barrier along the complex free energy landscape describing macromolecular processes. Equation (8) is used to determine the rate-determining barrier for three examples (top, middle, bottom) of free energy (a) and enthalpy (b) landscapes.

about diffusion along the reaction coordinate. Note that various choices of partitionings are equally possible. Thus, the kinetic pre-factor, A, and the absolute value of the relevant free energy barrier, $\Delta G^{\ddagger}$, are not independent quantities, and the latter is only meaningful if stated together with the corresponding pre-factor.

To illustrate these concepts with a practical example, we consider the processes of primary and secondary nucleation during protein aggregation (see Section 2 for details on the definition of these processes). During primary nucleation, protein molecules come together spontaneously to form fibrillar aggregates. Hence, a choice of the pre-factor, for which the flux towards the relevant barrier can be determined experimentally in a straightforward way, is the rate of diffusional arrival of the protein building blocks into an effective reaction volume r_{eff}, as $A^{\text{phys}} = Dr_{\text{eff}}$, where D is the diffusion coefficient [15]. In the case of secondary nucleation, it is known that fibril nucleation involves the adsorption of monomers onto the surface of existing fibrils. Thus, in addition to the diffusional arrival of the protein building blocks, the pre-factor A^{phys} in this case would take into account also the kinetics of protein adsorption onto the fibril surface. This requires the experimental determination of the probability per unit time that a protein subunit is added to a fibril, which could be achieved, for instance, by measuring the fibril surface coverage as a function of the bulk protein concentration [16]. Note that, depending on whether the kinetics of adsorption is taken into account into the pre-factor for secondary nucleation or not, the resulting value for the rate-determining free energy barrier of secondary nucleation will be different. This highlights that the observed value of the rate-determining free energy barrier is crucially linked to the experimental knowledge about the system. This is in contrast to the enthalpy of activation, which is independent of the choice of reaction coordinate and thus represents a highly robust readout of the key characteristics of the underlying free energy landscape.

3. Application to protein aggregation

In the following, we demonstrate how Kramers rate theory discussed above can be used to study complex multimolecular reactions governed by the diffusive dynamics and, in this manner, extract energetic information about some of its constituent microscopic steps. In particular, we shall focus on protein aggregation kinetics into amyloid fibrils, a process which is attracting great interest due to its connection with over 50 medical conditions, including Alzheimer's and Parkinson's diseases [7,17–19].

Amyloid fibril formation is a process in which soluble proteins spontaneously aggregate into fibrils of a cross-β structure, enriched in β-sheet content [20]. This is a complex phenomenon that typically involves the concomitant action of multiple molecular mechanisms. Recent advances in the available experimental techniques for measuring aggregation kinetics coupled to mathematical analysis of the underlying kinetic equations have allowed

the identification of these mechanisms at a microscopic level [21–23]. In the case of the aggregation of Aβ42 (the 42-residue form of the amyloid-β peptide), a process that is intimately linked to Alzheimer's disease [24], the fundamental steps that underlie amyloid fibril formation involve an initial primary nucleation step, where monomeric proteins spontaneously come together to form new fibrils, coupled to filament elongation. In addition, the aggregation process is accelerated by the fact that fibrils are able to generate copies of themselves through surface catalysis [25,26], a process known as secondary nucleation [27].

Despite the fact that the molecular steps of Aβ42 amyloid formation have been identified, the molecular mechanisms that underlie them have remained challenging to understand [28–30]. Here, we use Kramers theory to study the free energy landscape of two key steps in the formation of Aβ42 amyloid fibrils, namely primary and secondary nucleation. We use a coarse-grained model (Figure 3) which can capture the diffusive motion of proteins on a multi-barrier energy landscape determined by the relevant effective intermolecular interactions, such as hydrophobic forces, hydrogen bonding and screened electrostatic interactions. Our model retains only crucial molecular ingredients needed to reproduce the aggregation behaviour at experimentally relevant scales. The main advantage of this coarse-grained computer simulation approach is that the measurements from simulations can be validated directly against bulk experimental measurements [16], as we also demonstrate here.

In the case of primary nucleation, the free energy landscape along the reaction coordinate in our model involves an initial step whereby monomeric proteins associate, followed by a conformational conversion of proteins from their native soluble states into a β-sheet prone state, and finally the association of the latter state in a β-sheet rich nucleus, which then rapidly grows into an amyloid fibril (Figure 3). Secondary nucleation involves the adsorption of monomers onto the surface of an existing fibril and a subsequent surface-catalysed conformational conversion step. This step leads to aggregate detachment and its conversion into the β-sheet nucleus, which then elongates into a fibril (Figure 3). In both scenarios, we measure the temperature dependence of the overall rates of primary and secondary nucleation and show that, despite the complexity of the underlying processes, these temperature-dependent kinetic measurements probe only one relevant barrier along the free energy landscape. We then compare our simulation results to the equivalent experimental results on the temperature dependence of primary and secondary nucleation during the formation of Aβ42 amyloid fibrils. Just like predicted by Equations (9) and (10), we show
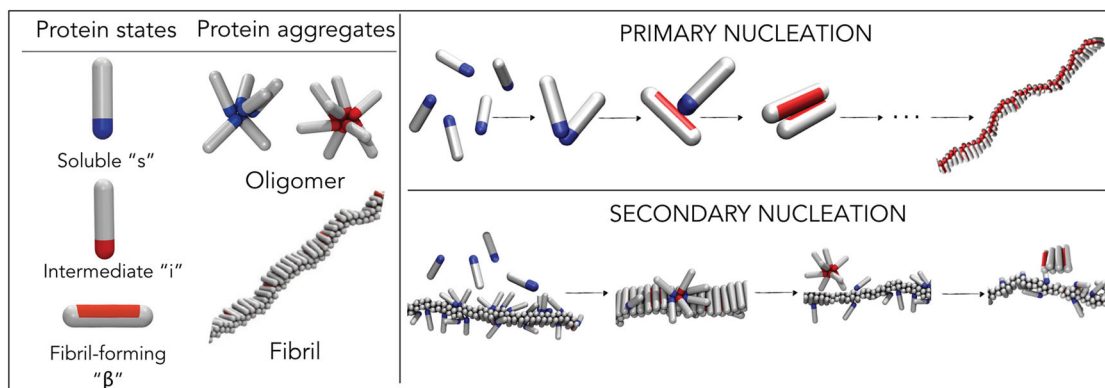
Figure 3. Coarse-grained model for amyloid aggregation. (a) Monomers can exist in two states – the soluble state that forms oligomers and a β-sheet prone state that forms fibrils. When bound to a fibril, monomers can also convert into an intermediate state, which binds stronger to its own kind than to the fibril and hence self-associates into oligomers that detach from the fibril surface. (b) Primary nucleation, over the concentration and temperature regime discussed in this paper ($c = 1.8$ mM, $1.25 < k_B T < 10$), proceeds through protein dimerisation and conversion into a β-sheet dimer which continues growing into a fully elongated fibril. Secondary nucleation ($c = 0.15$ mM, $0.92 < k_B T < 1.03$) proceeds by monomer attachment and oligomerisation on the fibril surface, conversion into an intermediate state, oligomer detachment and finally conversion into the β-sheet rich nucleus in solution.

that the kinetic measurements do not probe the highest enthalpic barrier on the energy landscape, but rather the enthalpic barrier which contributes to the highest free energy barrier on the landscape.

3.1. Computer model

We used a coarse-grained MC model for primary and secondary amyloid nucleation developed in [16,32]. Although minimal in its nature, this model captures many complex features of the aggregation processes. In particular, the model accounts for the fact that an amyloidogenic protein needs to exist in at least two different states: a state in solution ('s') that can occasionally aggregate into small oligomers and a higher free energy state that can form amyloid fibrils ('β') [32,33] (Figure 3). To capture secondary nucleation, a soluble monomer can adsorb onto the fibril surface and can convert into an additional (intermediate 'i') state that lies in between the 's' and 'β' states; the existence of this intermediate state reflects the catalytic role of the fibril in assisting the conformational conversion from the soluble into the fibril-forming state. In this model, a protein is described by a single rod-like particle, decorated with a patch that controls protein aggregation into either non-β-sheet oligomers or fibrils. A protein in an 's' or 'i' state interacts with its own kind via its attractive tip, which mimics non-specific inter-protein interactions. The fibril-forming state interacts with its own kind via an attractive side-patch, which models directional interactions, such as hydrogen bonding, and drives the formation of fibrillar aggregates. The interaction between two proteins in the fibril-forming states is by far the strongest interaction

in the system, and once formed, the fibrils are effectively irreversible. Monomer adsorption onto the fibril is energetically favourable; adsorbed monomers are then able to interact on the fibril surface to form oligomers. Since the protein in the intermediate state interacts with the fibril only weakly, oligomer detachment is favourable only for sufficiently large oligomers. This is because the loss of monomer-fibril interactions is overcome by the energetic gain associated with the interaction between proteins in the oligomer-forming state. Every conversion event from the soluble state into the fibril-forming state is penalised by a change in the excess chemical potential, $\Delta\mu_{s-\beta}$. This property is needed to reflect the fact that amyloidogenic proteins, such as Aβ, are typically not found in the β-sheet prone conformation in solution [34,35]. As in our previous work [16], the conversion from the soluble to the intermediate state on the fibril surface, as well as the conversion from the oligomer-forming state to the fibril-forming state was penalised by $0.5\Delta\mu_{s-\beta}$. Further details on the parameters used in this work are given in the Supplementary Information. To calculate the nucleation rate, we measure the mean first passage time for the particles to diffuse along the energy landscape and create a β-sheet enriched nucleus. The rate of nucleation is then defined as the inverse of such an average first passage time for nucleation [16,36].

3.2. Temperature dependence of primary nucleation

For primary nucleation to take place, proteins need to meet in solution and then convert into the β-sheet prone conformations. The converted monomers interact

strongly and form a β-sheet nucleus which is able to grow into a mature fibril (Figure 3). This process can involve the formation of long-lived pre-nucleation clusters. Such clusters provide a suitable environment for the conformational conversion to take place and hence significantly enhance the rate of nucleation [32,36]. We simulated amyloid nucleation in solution for a range of different temperatures. Interestingly, we find a non-monotonic dependence of the primary nucleation rate on temperature (Figure S5(a)): at low temperatures the nucleation rate decreases with temperature, while at high temperatures the rate increases as the temperature increases. This result can be understood as follows. We find that monomers oligomerise substantially at low temperatures (Figure S5(b)), which increases the rate of fibril nucleation. As the temperature is increased, the pre-nucleation oligomers become smaller and smaller, which in turn decreases the nucleation rate. However, as the temperature is increased further, the high-energy β-sheet prone state becomes more easily accessible by thermal fluctuations, and the conversion rate is enhanced, resulting in an increased overall nucleation rate.

Recently, the temperature dependence of the rate of primary nucleation of the Aβ peptide has been probed in experiments [31], and it has been found that, in the experimentally relevant regime of temperatures for this peptide, the nucleation rate is significantly increased at higher temperatures. Hence, we focus on this regime in our analysis. In this temperature regime, the nucleation process starts by two proteins meeting in solution and forming a dimer (denoted as '2s' in Figure 4). This '2s' dimer usually falls apart many times and reforms elsewhere in solution before one protein in the dimer successfully converts into the β-sheet prone state ('1s1β' in Figure 4). Such a '1s1β' dimer also has a high probability of dissolving back into the solution, before a successful conversion into a β-sheet nucleus ('2β') occurs. However, if the dimer manages to convert successfully into the '2β' state, then it will quickly grow into a fibril. Indeed, by analysing the simulation trajectories, we find that the '2β' nucleus, made of two proteins in the β-sheet prone state, always grows further, without ever dissolving back into the solution. In our simulations, we also detect a significant amount of dimers containing two proteins in the soluble states, while the least probable species in the system is a dimer that contains exactly one protein in the soluble state and one protein in the β-sheet prone state. Thus, we assign the highest free energy barrier in the system to precisely this rare species, as shown in Figure 4. Since in our simulations we know all the interactions in the system, we can explicitly calculate the enthalpic barrier that corresponds to this 'rate-determining' free energy barrier in the simulations. In particular, a measurement of

the relevant enthalpic barrier from the simulation trajectories yields the value of 14.5$k_B T$. This result agrees remarkably well with the variation of the reaction rate with temperature in our simulations that yields a value of 14.8$k_B T$, just like predicted by the reaction rate theory in Equations (9)–(10).

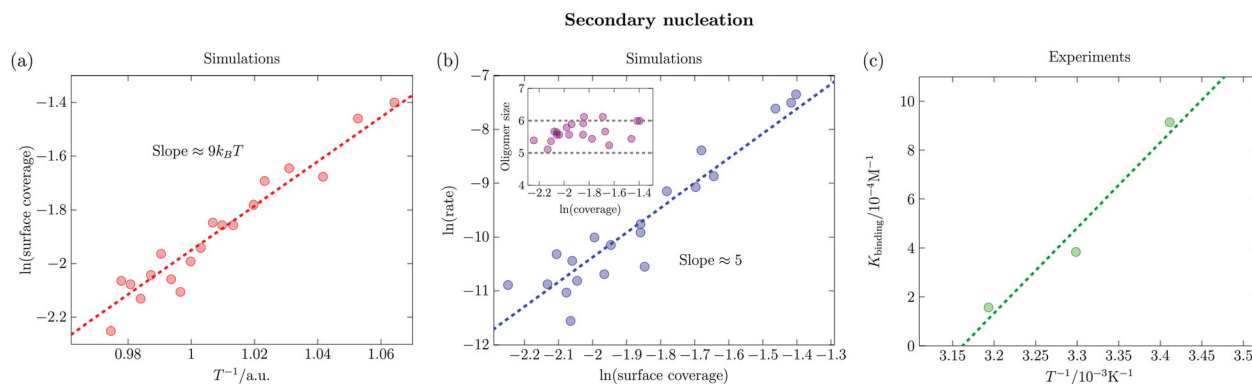### 3.3. Temperature dependence of secondary nucleation

We repeated an analogous set of simulations and rate measurements for the temperature dependence of secondary nucleation. In contrast to primary nucleation, secondary nucleation occurs via protein adsorption and oligomerisation on the fibril surface. These steps are followed by a conformational conversion into an intermediate state, the detachment of the oligomer from the fibril surface and finally the conversion of the detached oligomer into a β-sheet nucleus that further grows into a fibril (Figure 3). From our simulations, at the particular protein concentration we considered, we find that the rarest species in the system is a fibril-bound oligomer which contains four proteins in the 's' state ('4s' in Figure 4). If such a '4s' oligomer survives, it will grow into a hexamer ('6s' in Figure 4), which then partially converts into an intermediate state ('4s2i' in Figure 4), detaches from the fibril surface ('6i') and finally converts into a β-sheet nucleus ('2β^+' in Figure 4). Our rate measurements show that secondary nucleation is strongly hampered at high temperatures. This is exactly the opposite trend compared to the one we observed for primary nucleation. The reason for this difference is that, unlike in primary nucleation, the highest free energy barrier for secondary nucleation corresponds to the protein oligomerisation step on the fibril surface, rather than the oligomer conversion step. At higher temperatures, fewer monomers are adsorbed onto the fibril surface; this leads to a decreased tendency to undergo protein oligomerisation and, hence, to a slower overall nucleation rate. It is important to note that secondary nucleation in our simulations occurs only in a very narrow temperature regime. As previously reported in experiments [26,37] and simulations [16], secondary nucleation is extremely sensitive to environmental conditions. The exact temperature range where secondary nucleation occurs in our simulations is determined by the choice of all the interactions in the system, which is somewhat arbitrary in such a highly coarse-grained model. Hence, one should not try to compare the exact values of the rates or temperatures between primary and secondary nucleation in simulations and experiments. Instead, one should focus on similarities and differences between the trends and qualitative behaviour observed in simulations and experiments.

Primary nucleation

Secondary nucleation
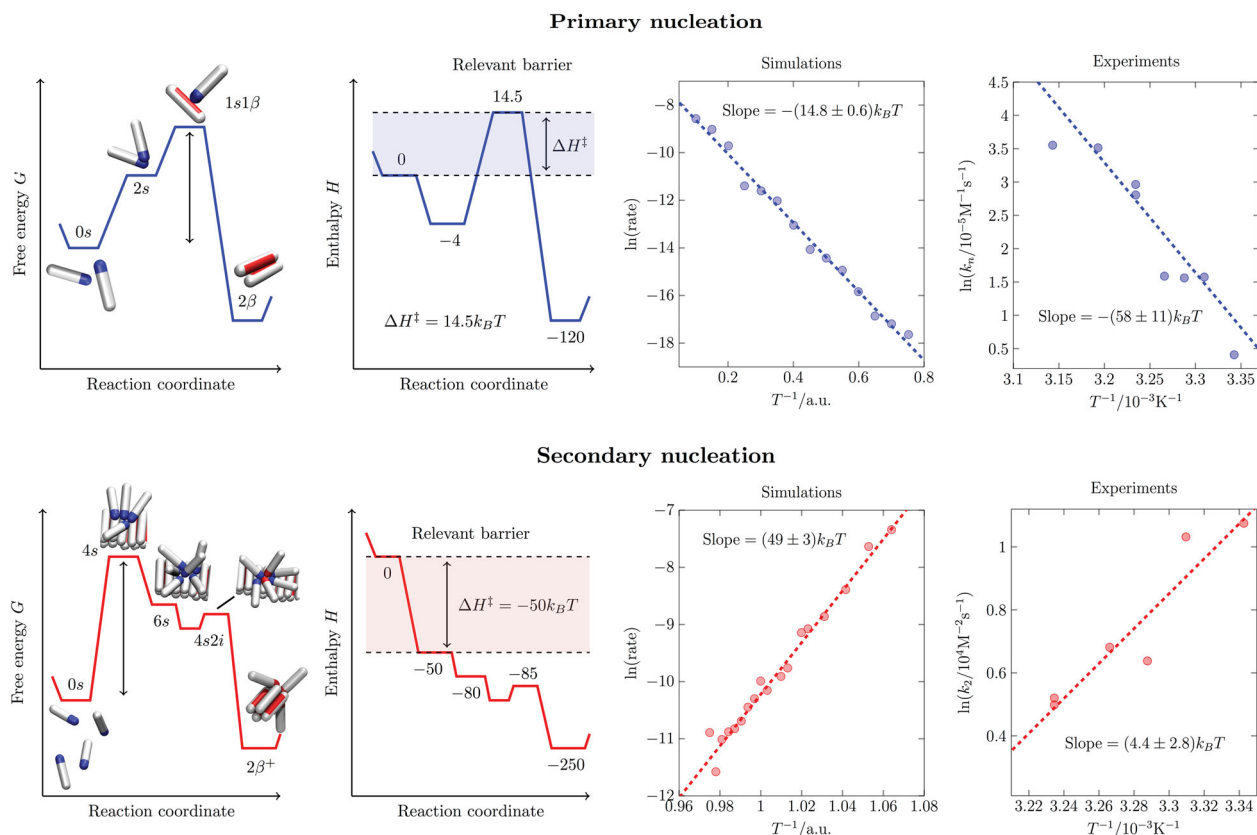
Figure 4. Determining the rate-determining step for primary and secondary nucleation in amyloid fibril formation. Free energy and enthalpy profiles underlying primary (top) and secondary (bottom) nucleation in our computer model and in experiments, and the variation of the respective nucleation rates with temperature. Top panel: The highest free energy barrier for primary nucleation in the simulations considered here corresponds to the conversion of a single monomer within the nucleus to the β-sheet configuration (left). The temperature dependence of the nucleation rate is a readout of the energetic penalty for this conversion event, which, in our model, is imposed by $\Delta\mu_{s-\beta}$ (middle). The experimental measurements of the temperature-variation of the primary nucleation rate for Aβ42 aggregation show qualitatively the same trend as in our simulations (right). In particular, the enthalpy of activation is positive, $\Delta H^{\ddagger}_{\text{exp}} = 58.1 k_B T$. Bottom panel: The highest free energy barrier for secondary nucleation in this model corresponds to the adsorption of monomers onto the fibril surface (left). The variation of the secondary nucleation rate with temperature thus probes the enthalpic barrier for protein adsorption, which is in this case negative (middle). The experimental measurements of the temperature variation of the rate of secondary nucleation for Aβ42 aggregation exhibit qualitatively the same trend as in our simulations [31] (right). Specifically, the secondary nucleation rate is decreased at high temperatures, a fact which is reflected in a negative enthalpy of activation for secondary nucleation, $\Delta H^{\ddagger}_{\text{exp}} = -4.4 k_B T$.

An example demonstrating the power offered by the comparison between simulations and experiments is provided in Figure 5. Figure 5(a) shows the temperature dependence of the surface coverage, while Figure 5(b) plots the dependence of the rate of secondary nucleation on the fibril surface coverage. The two dependencies (Figure 5(a,b)) combined give rise to the temperature-dependent behaviour of the secondary nucleation reaction rate observed in Figure 4. From the measurements of the interactions in our model (Figure 4), we find that the enthalpic barrier corresponding to the formation of the 'critical' oligomer is actually negative, and again, this result aligns very well with the variation of the nucleation rate with temperature. The same trends have

recently been reported experimentally for the kinetics of secondary nucleation of Aβ42 [31]. In particular, secondary nucleation of Aβ42 has been observed to be retarded at high temperatures and appeared to have a negative enthalpic barrier (Figure 4). In this case, the experimentally observed behaviour of the secondary nucleation rate is also caused by the fact that protein-fibril adsorption is reduced at higher temperatures [31] (Figure 5(c)). Secondary amyloid nucleation thus is a clear example of a process in which the kinetic measurements do not read out the highest enthalpic barrier along the energy landscape. The enthalpic barrier probed by secondary nucleation is in fact highly negative, both in simulations and experiments, and contributes to the

Secondary nucleation

Figure 5. Temperature dependence of secondary nucleation. (a) Computer simulations show that the amount of monomers adsorbed on fibrils decreases with increasing temperature. (b) The rate of secondary nucleation in computer simulations depends linearly on the fibril surface coverage, while the size of the nucleating oligomer remains unchanged (Inset). The combination of the dependence in (a) and (b) results in the temperature dependence of secondary nucleation presented in Figure 4. (c) The temperature dependence of the monomer-fibril binding constant measured for the Aβ peptide shows the same trend as the temperature dependence of surface coverage observed in simulations [31].

highest free energy barrier on the free energy landscape, which is clearly dominated by the unfavourable entropic contribution related to protein adsorption.

4. Conclusions

We have discussed a general framework, based on Kramers reaction rate theory, for studying the temperature dependence of complex supramolecular processes with multiple barriers and wells. We find that the enthalpic barrier probed in temperature-dependent kinetic measurements of reaction rates does not necessarily correspond to the highest enthalpic barrier along the reaction coordinate, but rather to the enthalpic barrier corresponding to the highest relative free energy barrier on the free energy landscape. We have then applied Kramers theory to interpret in coarse-grained computer simulations of the fundamental processes underlying the formation of amyloid fibrils – primary and secondary amyloid nucleation. For primary nucleation we find that two regimes can exist for the rate of nucleation – a regime in which the nucleation is faster at low temperatures, and a regime in which nucleation is faster at high temperatures. Guided by recent experimental results, we focus on the latter regime, and find that protein conformational conversion, which is aided at high temperatures, is the rate-determining step that drives primary nucleation in this temperature range. Unlike primary nucleation, we find that the relevant free energy barrier that determines the temperature dependence of the rate of secondary nucleation is the adsorption and oligomerisation on the fibril surface, which is hampered at hightemperatures. The difference in the respective rate-determining steps results in fundamentally distinct thermodynamic

signatures for primary and secondary nucleation, thus highlighting the power of probing free energy landscapes for understanding microscopic processes underlying complex multi-molecular processes.

Acknowledgments

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

References

[1] S. Arrhenius, Z. für Phys. Chem. **4**, 226 (1889).
[2] H. Eyring, J. Chem. Phys. **3**, 107 (1935).
[3] H.A. Kramers, Physica **7**, 284 (1940).
[4] P. Hänggi, P. Talkner and M. Borkovec, Rev. Mod. Phys. **62**, 251 (1990).
[5] R. Zwanzig, Proc. Natl Acad. Sci. USA **85**, 2029 (1988).
[6] A. Zaccone and E.M. Terentjev, Phys. Rev. Lett. **108**, 038302 (2012).
[7] F. Chiti and C.M. Dobson, Annu. Rev. Biochem. **86**, 27 (2017).
[8] E. Gazit, Chem. Soc. Rev. **36**, 1263 (2007).

[9] T.P.J. Knowles and M.J. Buehler, Nat. Nanotechnol. **6**, 469 (2011).

[10] G. Wei, Z. Su, N.P. Reynolds, P. Arosio, I.W. Hamley, E. Gazitf and R. Mezzenga, Chem. Soc. Rev. **46**, 4661 (2017).

[11] R. Mezzenga and P. Fischer, Rep. Prog. Phys. **76**, 046601 (2013).

[12] P. Reimann, G.J. Schmid and P. Hänggi, Phys. Rev. E **60**, R1 (1999).

[13] C.M. Bender, S.A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers* (McGraw-Hill Book Company, New York, 1978).

[14] D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications* (Vol. 1) (Elsevier, San Diego, CA, 2001).

[15] A.K. Buell, J.R. Blundell, C.M. Dobson, M.E. Welland, E.M. Terentjev and T.P.J. Knowles, Phys. Rev. Lett. **104**, 228101 (2010).

[16] A. Šarić, A.K. Buell, G. Meisl, T.C.T. Michaels, C.M. Dobson, S. Linse, T.P.J. Knowles and D. Frenkel, Nat. Phys. **12**, 874 (2016).

[17] F. Chiti and C.M. Dobson, Annu. Rev. Biochem. **75**, 333 (2006).

[18] C.M. Dobson, Cold Spring Harb. Perspect. Biol. **9**, a023648 (2017).

[19] T.P.J. Knowles, M. Vendruscolo and C.M. Dobson, Nat. Rev. Mol. Cell Biol. **15**, 385 (2014).

[20] A.W.P. Fitzpatrick, S.T. Park and A.H. Zewail, Proc. Natl Acad. Sci. USA **110**, 5468 (2013).

[21] T.C.T. Michaels, A. Šarić, J. Habchi, S. Chia, G. Meisl, M. Vendruscolo, C.M. Dobson and T.P.J. Knowles, Annu. Rev. Phys. Chem. **69**, 11 (2018).

[22] S.I.A. Cohen, M. Vendruscolo, C.M. Dobson and T.P.J. Knowles, J. Mol. Biol. **421**, 160 (2012).

[23] G. Meisl, J.B. Kirkegaard, P. Arosio, T.C.T. Michaels, M. Vendruscolo, C.M. Dobson, S. Linse and T.P.J. Knowles, Nat. Prot. **11**, 252 (2016).

[24] D.J. Selkoe and J. Hardy, EMBO Mol. Med. **8**, 595 (2016).

[25] S.I.A. Cohen, S. Linse, L.M. Luheshi, E. Hellstrand, D.A. White, L. Rajah, D.E. Otzen, M. Vendruscolo, C.M. Dobson and T.P.J. Knowles, Proc. Natl Acad. Sci. USA **110**, 9758 (2013).

[26] G. Meisl, X. Yang, E. Hellstrand, B. Frohm, J.B. Kirkegaard, S.I.A. Cohen, C.M. Dobson, S. Linse and T.P.J. Knowles, Proc. Natl Acad. Sci. USA **111**, 9384 (2014).

[27] F.A. Ferrone, J. Hofrichter and W.A. Eaton, J. Mol. Biol. **183**, 611 (1985).

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

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

[30] R. Cabriolu and S. Auer, J. Mol. Biol. **411**, 275 (2011).

[31] S.I.A. Cohen, R. Cukalevski, T.C.T. Michaels, A. Šarić, M. Vendruscolo, C.M. Dobson, A.K. Buell, T.P.J. Knowles and S. Linse, Nat. Chem. **10**, 523 (2018).

[32] A. Šarić, Y.C. Chebaro, T.P.J. Knowles and D. Frenkel, Proc. Natl Acad. Sci. USA **11**, 17869 (2014).

[33] N.S. Bieler, T.P. Knowles, D. Frenkel and R. Vácha, PLoS Comput. Biol. **8**, e1002692 (2012).

[34] M. Fändrich, M.A. Fletcher and C.M. Dobson, Nature **410**, 165 (2001).

[35] J.R. Allison, P. Varnai, C.M. Dobson and M. Vendruscolo, J. Am. Chem. Soc. **131**, 18314 (2009).

[36] A. Šarić, T.C.T. Michaels, A. Zaccone, T.P.J. Knowles and D. Frenkel, J. Chem. Phys **145**, 211926 (2016).

[37] G. Meisl, X. Yang, C.M. Dobson, S. Linse and T.P.J. Knowles, Chem. Sci. **8**, 4352 (2017).