The Journal of Chemical Physics ARTICLE scitation.org/journal/jcp

The catalytic nature of protein aggregation

Cite as: J. Chem. Phys. 152, 045101 (2020); doi: 10.1063/1.5133635
Submitted: 24 October 2019 • Accepted: 23 December 2019 •
Published Online: 28 January 2020

Alexander J. Dear,[1,2] Georg Meisl,[1,a)] Thomas C. T. Michaels,[1,a)] Manuela R. Zimmermann,[1] Sara Linse,[2] and Tuomas P. J. Knowles[1,3,b)]

AFFILIATIONS

[1]Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom
[2]Department of Biochemistry and Structural Biology, Lund University, SE22100 Lund, Sweden
[3]Cavendish Laboratory, University of Cambridge, J J Thomson Avenue, Cambridge CB3 0HE, United Kingdom

[a)]**Contributions:** G. Meisl and T. C. T. Michaels contributed equally to this work.
[b)]**Author to whom correspondence should be addressed:** tpjk2@cam.ac.uk

ABSTRACT

The formation of amyloid fibrils from soluble peptide is a hallmark of many neurodegenerative diseases such as Alzheimer's and Parkinson's diseases. Characterization of the microscopic reaction processes that underlie these phenomena have yielded insights into the progression of such diseases and may inform rational approaches for the design of drugs to halt them. Experimental evidence suggests that most of these reaction processes are intrinsically catalytic in nature and may display enzymelike saturation effects under conditions typical of biological systems, yet a unified modeling framework accounting for these saturation effects is still lacking. In this paper, we therefore present a universal kinetic model for biofilament formation in which every fundamental process in the reaction network can be catalytic. The single closed-form expression derived is capable of describing with high accuracy a wide range of mechanisms of biofilament formation and providing the first integrated rate law of a system in which multiple reaction processes are saturated. Moreover, its unprecedented mathematical simplicity permits us to very clearly interpret the effects of increasing saturation on the overall kinetics. The effectiveness of the model is illustrated by fitting it to the data of *in vitro* Aβ40 aggregation. Remarkably, we find that primary nucleation becomes saturated, demonstrating that it must be heterogeneous, occurring at interfaces and not in solution.

I. INTRODUCTION

The self-assembly of proteins into amyloid fibrils is a natural biological process that has received increasing attention in recent years due to its close association with a range of widely prevalent, incurable, and fatal human disorders such as Alzheimer's and Parkinson's diseases.[1] Identifying the reaction processes that lead to biofilament proliferation is vital for understanding these diseases. It also holds the key to the rational design of drugs to inhibit these self-assembly phenomena, with a view to halting or preventing the related disorders.

Biofilament assembly from monomeric samples is always initiated by a slow nucleation reaction process in which monomeric protein molecules associate to form new filamentous structures and continued by a fast elongation process, which ensures that the average filament grows to macroscopic lengths.[2,3] Secondary reaction processes that frequently accompany these include filament breakage (often associated with prions[4]), filament branching (seen in, e.g., actin polymerization), and secondary nucleation (nucleation of new filaments on the surface of existing filaments, first observed with sickle hemoglobin polymerization[5–7]). Most of these processes are multistep reactions. Both theoretical and experimental considerations suggest that elongation is properly described as a two-step, catalytic process controlled by a Michaelis-Menten-like rate law.[8,9] It has similarly been shown that secondary nucleation is catalytic and can be well-described by a similar rate law in the case of the aggregation of the Alzheimer's-associated Aβ40 protein.[10] Closed sets of rate laws describing the kinetics of filamentous self-assembly via every possible combination of these reaction processes may be written down by considering only the monomer concentration alongside the number and mass concentrations of filaments, rather than the full size distribution.[11–15] Developing analytical solutions to these

rate laws, and subsequently testing them against experimental data, has permitted the key reaction processes responsible for the proliferation of a range of biofilaments of interest to be determined in the past decade.[16,17] Most notably, it has revealed that secondary nucleation is crucial to the aggregation of the Alzheimer's-associated Aβ42 protein.[18] In these models, it has always been assumed that concentrations are low enough that enzymelike saturation effects are rare and that the sampled range of concentrations is small enough for them to occur in at most one reaction process at a time so that the remaining processes can be accurately approximated by single-step mass-action-like rate laws. Moreover, primary nucleation is often assumed to be homogeneous and occurring in bulk such that no catalytic saturation effects are possible. However, as larger ranges of monomer concentrations are measured with increasingly accurate kinetic experiments, these assumptions no longer hold. Thus, in order to extend our descriptions of amyloid filament formation at all concentrations of interest, there is a need to allow for all processes in the kinetic framework to be explicitly catalytic and have the potential to saturate in a given experiment. Previous integrated rate laws were unable to address such concurrent saturation of several processes.

In this paper, we develop a single, fully general approximate analytical solution to the kinetics of catalytic protein aggregation. We demonstrate its accuracy to be superior to preexisting models for protein aggregation in which at most one reaction process is multistep. We next use the model to explain exactly how the overall kinetics are affected by saturation effects in each reaction process. We finally demonstrate the utility of the model by applying it to the analysis of experimental kinetic data on *in vitro* Aβ40 aggregation, finding that in addition to secondary nucleation, primary nucleation also saturates in this system.

II. RESULTS

A. Aggregation reaction processes can be described by Michaelis-Menten-like equations

To understand how enzymelike saturation effects arise in protein aggregation, we turn to Michaelis-Menten kinetics. This is a well-known model that describes the conversion of a substrate to a product through the catalytic action of an enzyme and admits an analytical solution.[19] The fundamental assumptions behind this model are that the reaction is two-step, with a reversible bimolecular substrate-enzyme binding first step, followed by an irreversible unimolecular second step that regenerates the enzyme and generates the product [Fig. 1(a)]. The rate equations describing this reaction are[19]

$$\frac{d[ES]}{dt} = k_b[E][S] - (k_d + k_{\text{cat}})[ES], \tag{1}$$

$$\frac{d[O]}{dt} = k_{\text{cat}}[ES], \tag{2}$$

$$[ES] + [E] = [E]_0, \tag{3}$$

where $[E]$, $[S]$, $[ES]$, $[O]$, and $[E]_0$ are the concentrations of enzyme, substrate, enzyme-bound substrate, product, and the total concentration of enzyme molecules overall.

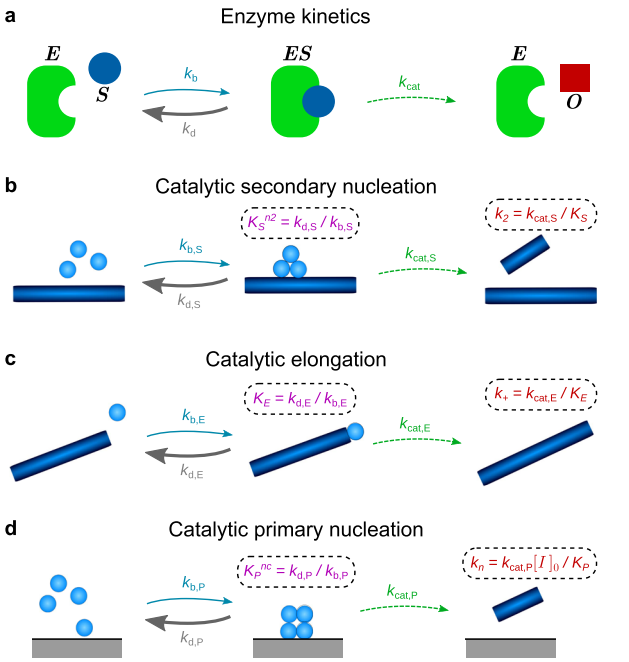

FIG. 1. The catalytic nature of key reaction processes in biofilament assembly. (a) Catalytic conversion of substrate S to product O by an enzyme E, featuring an intermediate enzyme-substrate complex. (b) In secondary nucleation, the fibril surface acts as a catalyst. (c) In elongation, the growing fibril ends act as a catalyst. Although the chemical species (a shorter fibril) is not regenerated, the pseudospecies (the fibril end) is. (d) In heterogeneous primary nucleation, any surface or interface (we denote the total concentration of binding sites on the interface as $[I]_0$) present in the reaction vessel may act as a catalyst. In all cases, where the concentration is high enough, the surface may become completely saturated with monomers; at this point, further increases in concentration do not affect the rate, which is then given simply by $k_{\text{cat}} \cdot [\textit{catalyst}]_{t=0}$. The 50% binding concentration K_x ($x = P, E, S$) is given by setting the intermediate bound state to steady-state, and in the usual case that $k_d \gg k_{\text{cat}}$, $K_x^{n_x}$ is approximately the dissociation constant for the corresponding dissociation reaction. We may thus interpret K_x as the geometric mean of the dissociation constants for each fundamental step in the dissociation reaction.

The key additional assumption that permits an analytical solution to be determined is that the enzyme-substrate complex is either at quasi-steady-state (QSS) or at pre-equilibrium with the free substrate (PE). If this is satisfied, then the rate of reaction r is given by[19]

$$r = \frac{v_{\max}[S]}{K_M + [S]}, \tag{4}$$

where $v_{\max} = k_{\text{cat}}[E]_0$ is the maximum rate achievable by the system and the Michaelis constant K_M is the substrate concentration at which the rate is half of $v_{\max}$.

Almost all of these assumptions hold for the catalytic reaction processes underlying biofilament formation, when the substrate is monomeric protein, and the "enzyme" is a filament surface, a growing end, or another interface in the reaction vessel [see Table I and Figs. 1(b)–1(d)]. The only assumption not generally satisfied is that the first step is bimolecular, since multiple monomers may be

TABLE I. Criteria for Michaelis-Menten type kinetics are only fully satisfied by the elongation reaction. Secondary nucleation does not satisfy assumption 7 and is therefore described by the closely related Hill type kinetics instead. Primary nucleation only satisfies assumptions 4 and 5 when heterogeneous and not homogeneous. Assumption 6 is generally satisfied when the concentration of the catalyst-bound substrate is low. This may not be the case for heterogeneous primary nucleation in general, but it is for Aβ40; if the monomer-bound catalyst was present at a significant concentration relative to monomers, an extended slow increase in the fibril mass concentration toward the end of the reaction would be visible, as remaining monomers detach from the interface and attach to the ends of growing fibrils. This effect is not seen in the aggregation kinetics of Aβ40 suggesting that the concentration of surface-bound monomers is low. Assumption 6 is also satisfied for elongation due to the comparatively low concentration of growing filament ends, and for secondary nucleation because at the early times at which secondary nucleation plays a significant role, the fibril mass concentration is still much lower than the monomer concentration.

Criterion for Michalis-Menten kinetics	1° nucleation	Elongation	2° nucleation
1: Reversible initial step	✓	✓	✓
2: Irreversible final step	✓	✓	✓
3: Unimolecular final step	✓	✓	✓
4: Substrate-catalyst binding in initial step	(✓)	✓	✓
5: Final step regenerates catalyst	(✓)	✓	✓
6: Quasi steady-state bound catalyst	(✓)	✓	✓
7: Substrate-catalyst binding is bimolecular	✗	✓	✗

involved in primary and secondary nucleation. However, the mathematics are easily generalized for higher reaction orders with respect to monomers, yielding instead a Hill equation for r, which resembles the Michaelis-Menten equation (4) but has the monomeric substrate raised to the power of the new reaction order.

For instance, secondary nucleation has been shown to be well-described by the Michaelis-Menten style rate law,[10]

$$r_S = \frac{k_2 m(t)^{n_2}}{1 + \left(\frac{m(t)}{K_S}\right)^{n_2}} M(t), \tag{5}$$

where $m(t)$ is the concentration of monomeric protein at time t, $M(t)$ is the mass concentration of fibrils (proportional to the concentration of fibril surface), K_S is the monomer concentration at which for a given $M(t)$, the rate is equal to half the maximum possible rate $k_2 K_S^{n_2} M(t)$, and k_2 and n_2 are the rate constant and the reaction order of secondary nucleation at low monomer concentrations, respectively. Note that secondary nucleation has alternatively been described as the bimolecular binding of monomers to fibril surfaces, which then are free to diffuse and to react with one another to form fibrils.[20] However, the resultant kinetic curves for biofilament assembly too closely resemble those generated by the Michaelis-Menten style model to be distinguished by data fitting, so we use only the Michaelis-Menten style model here.

Filament elongation has previously been modeled as a dock-lock mechanism[21–23] and as a diffusive barrier-crossing reaction.[8,9] Although there is no chemical species that is regenerated during an elongation reaction, the number of free fibril ends [concentration $2P(t)$] is recovered unchanged after temporarily decreasing, while a rearranging monomer is bound to the end. Therefore, we may consider the fibril ends as a catalytic pseudospecies[24] playing the role of the enzyme E in Eq. (1). By identifying k_b from Eq. (1) as the rate constant for diffusion-limited attachment of the monomer to the fibril end, k_d as the detachment time, and k_{cat} as the rearrangement time for a bound monomer to form a new fibril subunit, a Michaelis-Menten rate law identical in the functional form to those proposed in Refs. 8 and 9 may then be obtained,[24]

$$r_E = \frac{2k_+ m(t)}{1 + \frac{m(t)}{K_E}} P(t), \tag{6}$$

where k_+ is the rate of elongation at low monomer concentrations and K_E is the monomer concentration at which the rate is 50% of the theoretical maximum for a given concentration $P(t)$. Note that both K_E and K_S can also be interpreted as approximate equilibrium dissociation constants of monomers from fibril ends and fibril surfaces, respectively.

Finally, if primary nucleation is heterogeneous, then it too can be expected to be catalytic and to obey a Michaelis-Menten-style law of the form

$$r_P = \frac{k_n m(t)^{n_c}}{1 + \left(\frac{m(t)}{K_P}\right)^{n_c}}, \tag{7}$$

where k_n and n_c are the rate constant and the reaction order of primary nucleation at low monomer concentrations, respectively, and K_P is the equilibrium dissociation constant of the monomer to nucleation surface and also the monomer concentration at which primary nucleation is 50% saturated.

B. Rate equations for catalytic filamentous self-assembly

The classical single-step moment equations describing the kinetics of filament number and mass concentrations[11,12,16] can therefore be generalized for catalytic aggregation to read

$$\frac{dP}{dt} = \frac{k_n m(t)^{n_c}}{1 + \left(\frac{m(t)}{K_P}\right)^{n_c}} + \frac{k_2 m(t)^{n_2}}{1 + \left(\frac{m(t)}{K_S}\right)^{n_2}} M(t), \tag{8a}$$

$$\frac{dM}{dt} = \frac{2k_+ m(t)}{1 + \frac{m(t)}{K_E}} P(t). \tag{8b}$$

As usual, the small contributions of nucleation processes to the rate of increase in filament mass concentration have been neglected,[26] as have contributions from processes such as filament annealing.[27] Also note that fragmentation, instead of secondary nucleation, may be captured in these equations by setting $n_2 = 0$. These equations can therefore describe a wide variety of filament self-assembly reactions.

C. Analytical solutions to the kinetics of catalytic filamentous self-assembly

To solve Eqs. (8), we first nondimensionalize them and combine them into a single equation (Appendix A), yielding

$$-\left[\frac{1}{\mathcal{K}_E}\frac{d^2\mu(\tau)}{d\tau^2} + \frac{d^2 \log \mu(\tau)}{d\tau^2}\right] = \frac{\lambda^2}{\kappa^2}\frac{\mu(\tau)^{n_c}}{1 + \mu(\tau)^{n_c}/\mathcal{K}_P^{n_c}} + \frac{\mu(\tau)^{n_2}\left[1 - \mu(\tau)\right]}{1 + \mu(\tau)^{n_2}/\mathcal{K}_S^{n_2}}, \tag{9}$$

where $\lambda = \sqrt{2k_+ k_n m_{\text{tot}}^{n_c}}$ and $\kappa = \sqrt{2k_+ k_2 m_{\text{tot}}^{n_2+1}}$ are the effective non-catalytic fibril proliferation rates through primary and secondary processes, respectively;[28] $\mu(\tau) = m(\tau)/m_{\text{tot}}$ is the nondimensionalized monomer concentration; and $\mathcal{K}_P$, $\mathcal{K}_E$, and $\mathcal{K}_S$ are the nondimensional dissociation constants K_P/m_{tot}, K_E/m_{tot}, and K_S/m_{tot}, respectively. Finally, the rate constant k_+ now affects the kinetics solely via the nondimensionalized time $\tau = \kappa t$.

In most systems, the formation of new filaments is dominated by secondary processes, suggesting that a productive approach might be to seek a perturbation series solution in $\varepsilon = \lambda^2/2\kappa^2$. The resultant perturbative solution describes the early-time dynamics of Eqs. (8) as exponential growth. However, this formula is not accurate at long times as it does not account for monomer depletion and therefore does not converge. To remove this divergence and develop a global solution, we employ a perturbative renormalization group methodology similar to that used in Ref. 29 (see Appendix B for details). The resulting closed-form solution is

$$\frac{M(t)}{m_{\text{tot}}} = 1 - \left[1 + \frac{\varepsilon'}{c'}\left(e^{\kappa' t} + e^{-\kappa' t} - 2\right)\right]^{-c'}, \text{ where} \tag{10a}$$

$$\varepsilon' = \frac{k_n' m_{\text{tot}}^{n_c}}{2k_2' m_{\text{tot}}^{n_2+1}}, \tag{10b}$$

$$\kappa' = \sqrt{2k_+' k_2' m_{\text{tot}}^{n_2+1}}, \tag{10c}$$

$$c' = \frac{3}{2n_2' + 1}, \tag{10d}$$

and ε' can be interpreted as the relative importance of primary vs secondary nucleation as a source of new fibrils. k_n', k_+', and k_2' are the perturbed rate constants for primary nucleation, elongation, and secondary nucleation, respectively, given by

$$k_n' = k_n \frac{\mathcal{K}_P^{n_c}}{1 + \mathcal{K}_P^{n_c}}, \tag{11a}$$

$$k_+' = k_+ \frac{\mathcal{K}_E}{1 + \mathcal{K}_E}, \tag{11b}$$

$$k_2' = k_2 \frac{\mathcal{K}_S^{n_2}}{1 + \mathcal{K}_S^{n_2}}. \tag{11c}$$

Note that these are simply the effective rate constants at $t = 0$, as shown by comparison with Eqs. (5)–(7). Finally, $n_2' + 1$ is the effective reaction order of filament proliferation, which also depends on the monomer concentration, given by

$$n_2' = n_2 \frac{\mathcal{K}_S^{n_2}}{1 + \mathcal{K}_S^{n_2}} - \frac{2}{1 + \mathcal{K}_E}. \tag{12}$$

In the limit that $\mathcal{K}_P$, $\mathcal{K}_E$, and $\mathcal{K}_S$ tend to infinity (i.e., when initial monomer concentrations are far below the saturation concentrations), $k_n' \to k_n$, $k_+' \to k_+$, $k_2' \to k_2$, and $n_2' \to n_2$, and single-step kinetics are recovered as required.

Previously, a range of different approximate solutions have been developed for the kinetics of biofilament formation in which at most one of the processes displays saturation effects. Comparison of these legacy models to the solution given by Eqs. (10) in Appendix C demonstrates that in almost all cases when reactions from initially monomeric samples are considered, our new solution describes the kinetics with improved accuracy. Moreover, it has a far simpler mathematical form than most of these earlier models (see Table II). In Appendix D, we further generalize the solution so that it accurately describes the kinetics of biofilament formation even when secondary processes no longer dominate over primary processes.

D. The effect of saturation on filamentous growth kinetics

The simple structure of the analytical model derived here, and its general nature, makes it particularly easy to determine the effect of saturation in different processes on the overall kinetics. From Eqs. (10), we see that the most significant effect of saturation in any process is to reduce its effective rate constant according to Eqs. (11). Saturation in elongation or in secondary nucleation also has the subsidiary effect of reducing the effective monomer dependence of the self-assembly process, by reducing n_2'. In addition to affecting the relative difference between reactions performed at different initial monomer concentrations, this also governs how the filament mass concentration approaches its maximal value toward the end of the aggregation reaction, as the monomer is heavily depleted. Primary nucleation saturation does not affect this, as in typical small-ε' systems, secondary nucleation dominates and primary nucleation ceases to affect the kinetics relatively early in the reaction.[30]

It has been shown elsewhere[31] that κ may be interpreted as the effective first order rate constant for the proliferation of mature fibrils via secondary processes under constant-monomer conditions; κ' inherits this interpretation. The kinetics are most sensitive to κ' due to the exponential dependence of $M(t)$ on κ'; saturation of secondary nucleation or of elongation therefore has a stronger effect on the kinetics than the saturation of primary nucleation [Figs. 2(b)–2(d)], which does not enter κ'. Increasing saturation in these processes dilates the reaction time, or stretches the kinetic curve on the t-axis, by reducing κ'.

TABLE II. Performance compared to pre-existing models for biofilament assembly featuring primary nucleation, elongation, and secondary processes. ✓ indicates improvement and ≈ indicates little difference. No models have previously been derived for biofilament assembly via a catalytic primary nucleation process, but earlier models do exist in which either elongation or secondary nucleation is catalytic. In every case, it is found that the general model derived in this paper is superior to these previous specialized models, either in accuracy or in mathematical simplicity (or in both). The most notable improvement is in our description of biofilament formation with secondary nucleation and catalytic elongation for which the previous model gave a comparatively poor description. See Appendix C for full details. Note that many other combinations of reaction processes have never previously been modeled but can also be accurately described by our new approach.

Earlier model	Reaction steps	Improvement in accuracy?	Improvement in model simplicity?
Reference 11	1° nucleation; elongation; fragmentation	✓	≈
Reference 13	1° nucleation; elongation; 2° nucleation	≈	✓
Reference 25	1° nucleation; elongation; 2° nucleation	✓	≈
Reference 16	1° nucleation; catalytic elongation; fragmentation	≈	✓
Reference 16	1° nucleation; catalytic elongation; 2° nucleation	✓ ✓	✓
Reference 10	1° nucleation; elongation; catalytic 2° nucleation	✓	✓

Saturation of elongation does not affect the parameter ε', whereas saturation of secondary nucleation increases ε', which counteracts the reduction in κ' to a certain extent. However, the kinetics are typically more sensitive to saturation in secondary nucleation overall [Figs. 2(b) and 2(c)], due to the higher reaction order of secondary nucleation compared to elongation [Fig. 2(a)]. ε' gives the relative rate of primary nucleation compared to secondary processes and is typically expected to be small in most amyloid-forming systems. It may also be interpreted as the logarithm of an effective starting time; the only effect of increasing saturation in primary nucleation on the kinetics is therefore to shift the concentration curve to the right [Fig. 2(d)].

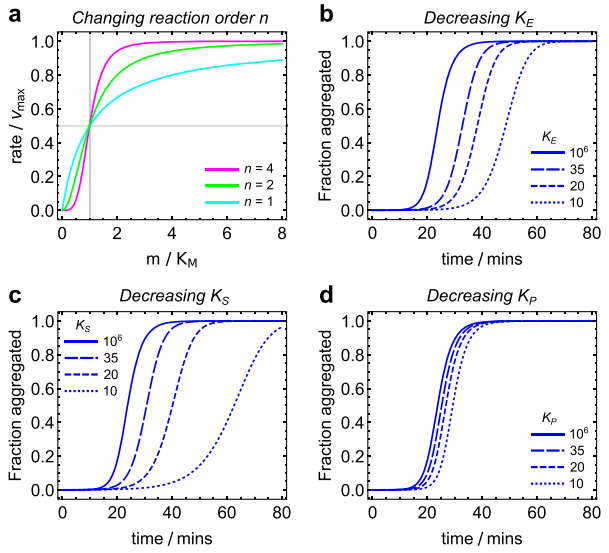

FIG. 2. (a) Plot of the rate, scaled by the maximal rate v_{max}, vs the monomer concentration m, scaled by the half-saturation concentration K_M, for an elongation reaction (cyan) and for nucleation reactions with orders 2 (green) and 4 (magenta). Elongation obeys Michaelis-Menten kinetics precisely, with a sublinear dependence of the rate on monomer concentration, whereas the higher-order nucleation reactions obey Hill kinetics with the rate exhibiting a sigmoidal monomer dependence. (b)–(d) investigating the effect of saturation in elongation, secondary nucleation, and primary nucleation, respectively, on aggregation curves. Aβ40 rate constants employed with $m(0)$ = 35 μM. Solid lines: K_M = 1M, i.e., no saturation. Dashed lines: K_M = 35, 20, or 10 μM. Shorter dashed lines correspond to lower saturation concentrations. Saturation in elongation and secondary nucleation mainly reduces the aggregation rate, whereas the sole effect of saturation in primary nucleation is to increase the lag time. Due to the logarithmic dependence of the half time on primary nucleation,[26] saturation in the latter has the smallest effect on aggregation kinetics. Saturation in secondary nucleation has the largest overall impact, despite increasing ε, due to the higher reaction order of secondary nucleation compared to elongation.

At concentrations far below the saturation concentration for a given process, its effective rate constant is unchanged and loses its dependence on the saturation concentration, which can therefore not be determined with any accuracy through kinetic model fitting. Far above the saturation concentration, by contrast, the effective constant for process ($x = P, E, S$) becomes $k_x K_x^{n_x}$, and neither the rate constant nor the saturation concentration may be determined with any accuracy but only their product, the conversion rate constant $k_{cat,x}$ (multiplied by the nucleation site concentration $[I]_0$ in the case of primary nucleation). This follows since the effective rate constants enter the model as $k_n' m_{tot}^{n_c}$, $k_+' m_{tot}$, and $k_2' m_{tot}^{n_2}$, which reduce in this limit to $k_n K_P^{n_c} = k_{cat,P}[I]_0$, $k_+ K_E = k_{cat,E}$, and $k_2 K_S^{n_2} = k_{cat,S}$. Modulo factors of $M(t)$ and $P(t)$ for secondary nucleation and elongation, respectively, these are also the maximum rates $v_{max,x}$ attainable by each reaction process. Thus, depending on the monomer concentrations spanned by the data to which the model is fitted, for a given reaction process, it may be appropriate to report both the saturation concentration and the rate constant (if the concentrations span the saturating region), or just the rate constant (if the concentrations measured are well below the saturation concentration), or just v_{max} (if the concentrations measured are well above the saturation concentration).

E. Aβ40 undergoes heterogeneous primary nucleation

We now apply our model to study the aggregation of Aβ40 peptides into amyloid fibrils, which is believed to be a key upstream event in the development of Alzheimer's disease. In addition to the obligatory primary nucleation and elongation processes, the aggregation reaction has been shown to depend critically on secondary nucleation.[10] In the case of Aβ40, this process is known to undergo saturation at a monomer concentration of c. 6 μM, at

pH 7.4, and 37 °C. It was also suspected to undergo saturation in filament elongation at a higher concentration,[10] since catalytic secondary nucleation alone could not describe the lack of concentration dependence of the aggregation rate at the highest concentrations studied [Fig. 3(a)]. However, this could not be verified, since the kinetic models available at the time could describe either catalytic secondary nucleation or catalytic elongation, but not both simultaneously. Equation (10a), however, permits this kind of analysis and allows simultaneous determination of K_E, K_P, and K_S from kinetic curves for filament formation from a wide range of initial monomer concentrations. Although catalytic elongation and secondary nucleation together offer a reasonable description of the kinetics [Fig. 3(b)], the data are in fact better described by catalytic primary and secondary nucleation, without saturation in elongation [Fig. 3(c)]. Additional support for this conclusion comes from a re-examining of the highly seeded aggregation experiments from Ref. 10. The initial rates of filament formation under highly seeded conditions are dominated by the elongation process.[16] The linear dependence of these rates on monomer concentration indicates the absence of significant saturation effects in the elongation process at the concentrations studied (Fig. 4). Therefore, all data indicate that a saturation of primary nucleation, in addition to a saturation of secondary nucleation, occurs in the aggregation of Aβ40 at high micromolar concentrations (Fig. 5). This finding provides further insight into the molecular details of the nucleation mechanism responsible for the formation of the first aggregates directly from the monomer; since homogeneous primary nucleation cannot display saturation effects, the primary nucleation process undergone by Aβ40 *in vitro* must be a heterogeneous process in order to saturate. Heterogeneous nucleation is indeed much more common than its homogeneous counterpart for a wide range of nucleation phenomena, from simulations of hard sphere interactions to water condensation on dust particles.[32–36] In the context of the highly purified Aβ40 samples studied in this work, nucleation most likely occurs at an interface such as the air-water interface or the surface of the reaction vessel. For some amyloid forming proteins, the effect of surface nucleation is so significant that aggregation can be completely inhibited in the absence of an air-water interface.[37–39] In the context of the study of disease-related amyloids, our findings highlight the importance that surfaces may play a role in determining

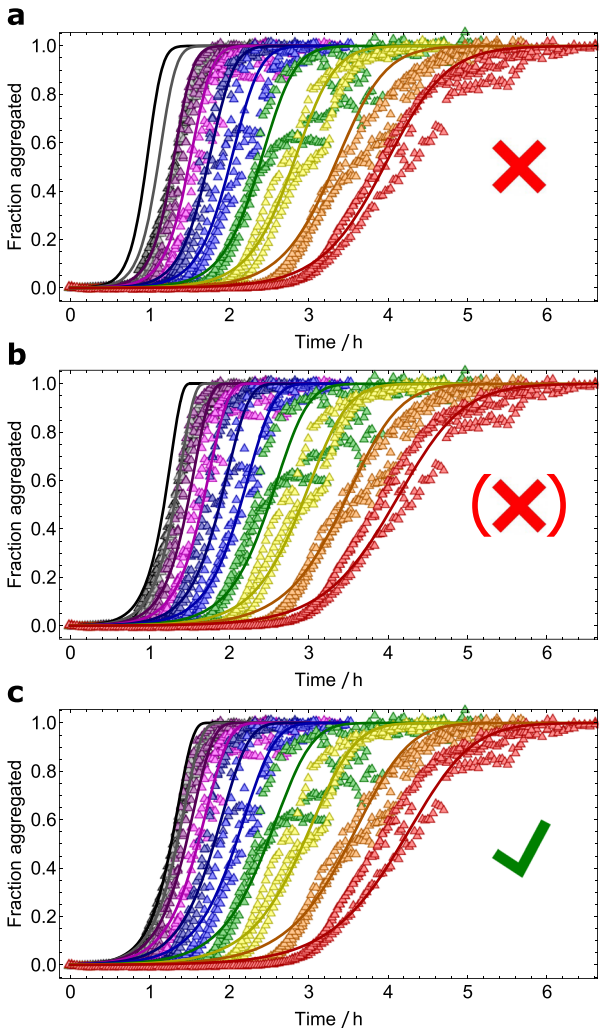

FIG. 3. Aggregation of Aβ40 monomers into filaments; fractional aggregation monitored by ThT fluorescence vs time. Initial monomer concentrations ranging from 3.5 μM to 70 μM (only 9.1–70 μM shown here). Data taken from Ref. 10. (a) The fit to the model by Meisl *et al.*[10] is only able to reproduce aggregation curves below 35 μM. (b) Equation (10a) with K_P set to an arbitrarily large value fits the data significantly better, yielding K_E = 103 μM. (c) When fully unconstrained, Eq. (10a) fits the data better still, yielding a larger value of K_E and a moderate value of K_P, suggesting that saturation in primary nucleation is more important than in elongation for this system. However, the improvement in fit quality is too small to reach a firm conclusion and more analysis is needed.

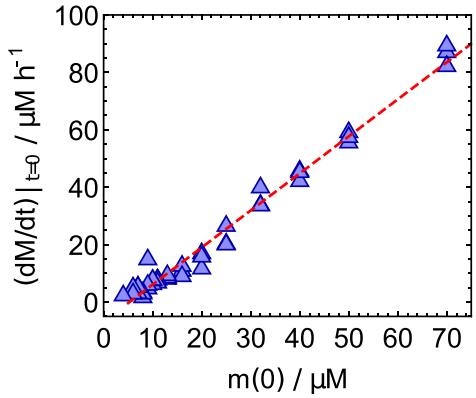

FIG. 4. Initial rates of aggregation of Aβ40 monomers into filaments in the presence of 21 μM seeds; initial monomer concentrations ranging from 3.5 μM to 70 μM. Data taken from Ref. 10. The initial rates in the presence of such high seeds should depend only on the elongation rate; if saturation effects are present in the elongation reaction, the initial aggregation should have a sublinear dependence on initial monomer concentration [see the cyan curve in Fig. 2(a)]. Instead, an approximately linear dependence on the monomer concentration is observed, demonstrating the absence of significant saturation in elongation at the monomer concentrations studied. This supports the tentative conclusion from Fig. 3 that the saturation effects additional to those from secondary nucleation, visible at the upper end of the concentration range studied, originate from primary nucleation and not from elongation.

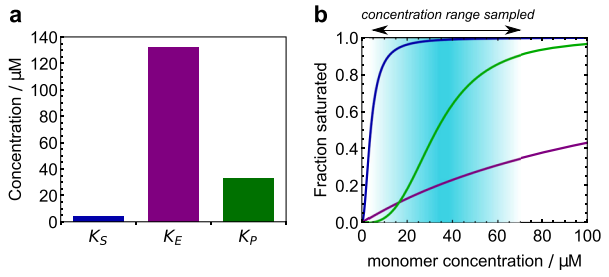

FIG. 5. The degree of saturation of the reaction processes of Aβ40 kinetics: blue for secondary nucleation; purple for elongation, and green for primary nucleation. (a) The concentrations at which the different reaction processes are 50% saturated. At these concentrations, saturation effects reduce the overall rates by 50%. (b) The fractional occupation of the catalytic surface for each reaction process over the range of monomer concentrations studied here. This makes a suitable definition for the degree of saturation of each process. Secondary nucleation is essentially fully saturated over all concentrations of interest, whereas elongation is largely unsaturated. Primary nucleation is fully saturated at the higher end of the range of concentrations studied. Where the fractional occupation is small, dissociation dominates over binding; where it is large, binding dominates over dissociation.

the rates of primary nucleation and that this effect should be taken into account when drawing disease-related conclusions from *in vitro* measurements.

III. CONCLUSIONS

We have derived a single unified analytical solution to the kinetics of amyloid filament formation in which all reaction processes may be multistep and catalytic in nature. This model encompasses a wide range of filamentous growth kinetics and is superior to all prior analytical solutions in terms of accuracy and generality for the description of unseeded aggregation kinetics.

It has been suspected for some time that primary nucleation in amyloid formation is heterogeneous. This study presents some of the first direct evidence that this is indeed the case. The use of our model to fit kinetic data on amyloid aggregation at high monomeric protein concentrations provides a route to confirming the heterogeneous nature of amyloid nucleation for other proteins in the future.

To date, reaction conditions have often been chosen carefully to avoid multiple saturation effects, as they were previously difficult to interpret. By providing the ability to analyze such multiple saturated data, our results open the field up to wider studies of the effects of temperature, pH, and other important reaction conditions on filamentous growth kinetics. This may permit a move toward studying amyloid formation in body fluids, where reaction conditions may promote saturation effects, and toward studying other aggregating systems previously deemed pathological from a modeling perspective. Finally, our model will facilitate thermodynamic studies of activation energies, which require data to be collected over a range of temperatures, often including those at which multiple saturation effects occur. Overall, this new model has the potential to be transformative in the field of amyloid aggregation kinetics.

ACKNOWLEDGMENTS

We acknowledge support from the Swiss National Science Foundation (T.C.T.M.), Peterhouse, Cambridge (T.C.T.M.), the Schiff Foundation (A.J.D.), Ramon Jenkins Fellowship, Sidney Sussex College Cambridge (G.M.), the Swedish Research Council (S.L.), the Wellcome Trust (T.P.J.K.), the Cambridge Centre for Misfolding Diseases (T.P.J.K.), the BBSRC (T.P.J.K.), and the Frances and Augustus Newman foundation (T.P.J.K.) for financial support. The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007–2013) through the ERC grants PhysProt (Agreement No. 337969) and MAMBA (Agreement No. 340890).

APPENDIX A: REDUCING EQS. (8) TO A SINGLE DIMENSIONLESS EQUATION

Making use of the conservation-of-mass relation $m(t) + M(t) = m_{\text{tot}}$, Eq. (8b) can be written as

$$\frac{1+\frac{m(t)}{K_E}}{m(t)}\frac{dM(t)}{dt} = -\left[\frac{1}{m(t)}+\frac{1}{K_E}\right]\frac{dm(t)}{dt} = 2k_+P(t). \tag{A1}$$

Using $\tau = \kappa t$, this may be written as

$$-\left[\frac{d\log m(\tau)}{d\tau}+\frac{1}{K_E}\frac{dm(\tau)}{d\tau}\right] = \frac{2k_+}{\kappa}P(\tau). \tag{A2}$$

Differentiating with respect to τ and noting that $\kappa dP/d\tau = dP/dt$, we obtain

$$-\left[\frac{d^2\log m(\tau)}{d\tau^2}+\frac{1}{K_E}\frac{d^2m(\tau)}{d\tau^2}\right] = \frac{2k_+}{\kappa^2}\left(\frac{k_n m(\tau)^{n_c}}{1+\left(\frac{m(\tau)}{K_P}\right)^{n_c}}+\frac{k_2\, m(\tau)^{n_2}}{1+\left(\frac{m(\tau)}{K_S}\right)^{n_2}}M(\tau)\right), \tag{A3}$$

and Eq. (9) is recovered by the substitution of $m(t) = m_{\text{tot}}\mu(t)$ and $M(t) = m_{\text{tot}}(1-\mu(t))$.

APPENDIX B: DERIVATION OF GENERAL ANALYTICAL SOLUTION

To develop a perturbative solution to Eq. (9), we first define $\varepsilon = \lambda^2/2\kappa^2$ as our small parameter,

$$\mu(\tau) = \mu^{(0)}(\tau) + \varepsilon\,\mu^{(1)}(\tau) + \varepsilon^2\,\mu^{(2)}(\tau) + \mathcal{O}(\varepsilon^3), \tag{B1}$$

where $\mathcal{O}(\varepsilon^3)$ denotes a remainder term of the same order of magnitude as ε^3. After solving the resulting equations separately for each order of ε and applying the initial conditions $\mu(0) = 1$ and $\mu'(0) = 0$, we find the following perturbation expansion:

$$\mu(\tau) = 1 - \varepsilon a e^{b\tau} + \frac{2 + n_2'}{3}\varepsilon^2 a^2 e^{2b\tau} + \mathcal{R}, \tag{B2}$$

$$a = \frac{1 + \mathcal{K}_S}{\mathcal{K}_S}\frac{\mathcal{K}_P}{1 + \mathcal{K}_P}, \tag{B3}$$

$$b = \sqrt{\frac{\mathcal{K}_E}{1 + \mathcal{K}_E}\frac{\mathcal{K}_S}{1 + \mathcal{K}_S}}, \tag{B4}$$

$$n_2' = n_2\frac{\mathcal{K}_S^{n_2}}{1 + \mathcal{K}_S^{n_2}} - \frac{2}{1 + \mathcal{K}_E}, \tag{B5}$$

where $\mathcal{R}$ denotes either terms of $\mathcal{O}(\varepsilon^3)$ or terms that vanish in comparison to the dominant terms at each order in the limit $e^\tau \gg 1$. Since outside of this limit all terms of $\mathcal{O}(\varepsilon)$ and above can be neglected anyway, we may ignore $\mathcal{R}$ for the time being.

We see that the perturbative solution describes the early-time dynamics of Eqs. (8) as exponential growth. However, this formula is not accurate at long times, as it does not account for monomer depletion and therefore does not converge. To proceed, we note that this perturbation expansion has the same form as that given by the kinetics of noncatalytic biofilament self-assembly in Ref. 29 but with $\varepsilon \to \varepsilon a$, $\tau \to b\tau$, and $n_2 \to n_2'$. In that paper, the divergence was tamed by the use of dynamical renormalization group methods, yielding a remarkably accurate solution in the form of a Richards curve $y(x) = 1 - (1 + x/c)^{-c}$, with $x = \varepsilon e^\tau$, $y = \mu$, and $c = 3/(1 + 2n_2)$. Given the identical functional form of the perturbative expansions for catalytic and noncatalytic biofilament self-assembly, the Richards solution may be immediately adopted here with appropriate replacements for x and c.

For small but nonvanishing values of εa, i.e., when most new filaments are formed through secondary processes, an expression of the above Richards form is expected to lose some accuracy at very early times, since the boundary conditions are met only in the limit $\varepsilon a \to 0$. However, since our perturbative expansion only corresponds to the Richards solution asymptotically in the limit $e^{b\tau} \gg 1$, we are free to modify our solutions so that they satisfy the boundary conditions $M(0) = 0$, as long as they recover their original form in this limit. The obvious choice is to replace $\varepsilon a e^{b\tau}$ with $\varepsilon a(e^{b\tau} + e^{-b\tau} - 2)$, which additionally ensures that the solution matches the first-order perturbative solution exactly. This yields Eq. (10) directly.

APPENDIX C: COMPARISON TO PRIOR ANALYTICAL SOLUTIONS

1. The analytical solution is almost exact in the low-concentration limit

For initial monomer concentrations sufficiently far below the saturation concentrations and in the limit $e^\tau \gg 1$, Eq. (10a) recovers the universal single-step model for filamentous growth discovered in Ref. 29. We may investigate the solution error in this limit explicitly by comparing the ratio of the 3rd order terms of the perturbation expansions of the exact kinetics, $\mu^{(3)}$, and of the Richards curve, $L^{(3)}$,

$$\frac{L^{(3)}}{\mu^{(3)}} = \frac{80 + 64n_2}{81 + 63n_2}. \tag{C1}$$

Not only is this remarkably close to 1, it is also exact for $n_2 = 1$. We therefore expect this solution to be highly accurate. Indeed, in Fig. 6(a), we see that it is even more accurate for secondary nucleation than the highly accurate Hamiltonian solution given in Ref. 25.

To date,[11,25] the kinetics of breakable filament assembly have been modeled with fairly high accuracy with a Gompertz curve, which is equal to the Richards solution presented here in the limit $c \to \infty$. However, our identification of $c = 3/(1 + 2n_2)$ suggests that this is not the best approximation possible. Instead, setting $n_2 = 0$ gives $c = 3$. We discover that this Richards curve is indeed a superior approximation, converging on the exact kinetics almost precisely [Fig. 6(b)].

The Richards curve is traditionally used to model biological growth kinetics. More generally, the Richards curve describes autocatalytic growth kinetics for a system with a defined maximum product concentration. The parameter $1/c$ identifies the substrate dependence of the growth rate (here, the substrate is identified as monomer). The Gompertz solution, given by $1/c = 0$, describes a system in which the substrate concentration does not limit the growth rate, and other factors, e.g., physical space limitations, give rise to a maximum product concentration. Viewed in this way, it is eminently reasonable that we do not use the Gompertz solution to describe fragmentation kinetics. Although the secondary process itself does not have substrate dependence (unlike secondary nucleation), the elongation process does, and aggregation should therefore still slow down as the monomer is depleted. The value given by our new approximation, $1/c = 1/3$, may therefore be justified on conceptual and mathematical grounds.

2. Singly catalytic kinetics are captured with greater accuracy than earlier models

We now investigate the limit in which one of secondary nucleation or elongation may saturate but not both. Models have been previously derived for such singly saturating kinetics; we also compare our new solution to these models in this limit.

When only elongation may saturate, i.e., $\mathcal{K}_S, \mathcal{K}_P \to \infty$, Eq. (10a) is usually highly accurate and found to be generally more accurate than the existing analytical solution based on the Lambert-W function (Fig. 7).

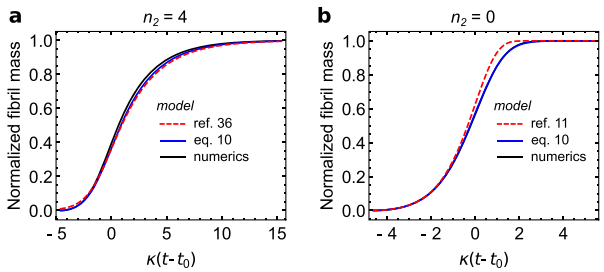

FIG. 6. (a) The Richards solution (in blue) captures the dynamics of a system with secondary nucleation ($n_2 = 4$) even more accurately than the Hamiltonian solution (in red). (b) The kinetics of fragmenting filament assembly are almost exactly captured by the new $c = 3$ Richards solution, significantly more accurately than the current standard; the Gompertz solution (in red) and the numerical solution (in black) to the moment equation for fragmenting filament assembly. $\varepsilon = 0.01$.

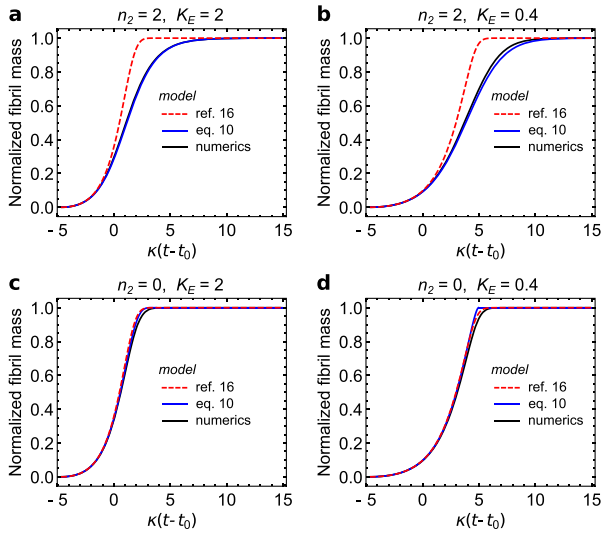

FIG. 7. The kinetics of filament assembly with secondary nucleation and saturating elongation are almost exactly captured by Eq. (10a), which is a significant improvement upon an earlier analytical solution based on the Lambert-W function. However, the Richards solution offers little advantage over the Lambert solution for fragmenting systems. Black: numerical solution to the moment equations. Red: old (Lambert) analytical approximate solution. Blue: new (Richards) analytical approximate solution. $\varepsilon = 0.01$. (a) $n_2 = 2,\ \mathcal{K}_E = 2$. (b) $n_2 = 2,\ \mathcal{K}_E = 0.4$. (c) $n_2 = 0,\ \mathcal{K}_E = 2$. (d) $n_2 = 0,\ \mathcal{K}_E = 0.4$.

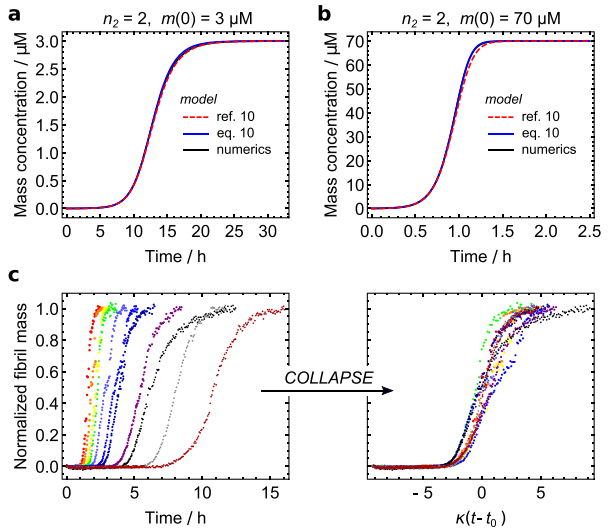

FIG. 8. (a) and (b): using rate constants for Aβ40 aggregation taken from Ref. 10, we calculate kinetic curves for initial monomer concentrations of 3 μM (a) and 70 μM (b). The curve calculated according to the solution developed in this paper [blue, Eq. (10a)] matches the numerical solutions (black) almost precisely; even better than the already highly accurate analytical solution developed in Ref. 10 (red). (c) Raw Ab40 data from Ref. 10 at a range of concentrations 35–3.5 μM (d) Scaling by κ', and addition of $t_0 = \ln(\varepsilon')/\kappa'$, effectively collapses the data onto a universal curve, since experimental errors in these data are a greater source of variation from universality than the small differences in n_2' for each individual curve.

When only secondary nucleation may saturate, i.e., $\mathcal{K}_E,\ \mathcal{K}_P \to \infty$, Eq. (10a) is almost an exact match to the numerical solutions of Eq. (8). The existing analytical solutions are highly accurate; however, Eq. (10a) still offers a small improvement in accuracy [Figs. 8(a) and 8(b)]. Moreover, the scaling of the solution is demonstrably correct since by scaling time by κ' and offsetting it by $t_0 = \ln(\varepsilon')/\kappa'$, the aggregation curves for Aβ40 at different starting concentrations almost fully collapse onto a single curve. A perfect collapse is prevented only by small variations in experimental conditions, for instance initial monomer concentration or sites for primary nucleation, and by each curve's different value for c' [Figs. 8(c) and 8(d)].

APPENDIX D: EXTENSION TO THE CASE OF SLOW SECONDARY PROCESSES

For larger values of ε such that it can no longer be said that secondary processes dominate over primary nucleation for the production of new filaments, Eq. (10a) starts to break down at later times, since the secondary reaction order n_2 no longer should control the late-time kinetics. To improve our solution, we can first investigate what form it should take in the limit $\tau \to 0$, i.e., when secondary processes are absent. We drop the final term in Eq. (9) and solve perturbatively, yielding (to second order in ε)

$$\mu(\tau) = 1 - \varepsilon\tau^2 \frac{\mathcal{K}_E}{1+\mathcal{K}_E}\frac{\mathcal{K}_P^{n_c}}{1+\mathcal{K}_P^{n_c}} + \varepsilon^2\tau^4 \frac{\mathcal{K}_E^2\mathcal{K}_P^{2n_c}(3\mathcal{K}_E + 3\mathcal{K}_E\mathcal{K}_P^{n_c} + \mathcal{K}_P^{n_c}n_c + \mathcal{K}_E\mathcal{K}_P^{n_c}n_c)}{6(1+\mathcal{K}_E)^3(1+\mathcal{K}_P^{n_c})^3}. \quad \text{(D1)}$$

We find that this matches our solution Eq. (10a) to the first order in $\varepsilon\tau^2$ already; moreover, if we replace c' with

$$c' = \frac{3}{n_c'}, \text{ where } n_c' = \frac{\mathcal{K}_P^{n_c}}{1+\mathcal{K}_P^{n_c}}n_c - \frac{3}{1+\mathcal{K}_E}, \quad \text{(D2)}$$

it is matched to the second order in $\varepsilon\tau^2$ as well. We may compute the error in this approximation by comparing the third-order term of its expansion in $\varepsilon\tau^2$ with that of the exact Oosawa solution in the unsaturated limit. The ratio of these terms, r_{err}, is

$$r_{\text{err}} = \frac{15(1+n_c) + 10n_c^2/3}{15(1+n_c) + 4n_c^2}. \quad \text{(D3)}$$

We see that this is exact in the limit $n_c = 0$ and loses accuracy as n_c increases (with a limiting value of 5/6 as $n_c \to \infty$). We therefore expect that such a representation is accurate for values of n_c that are not too large, e.g., $n_c < 4$.

Of course, this is insufficient for describing intermediate values of ε' and is not a single solution. By appropriate manipulation to obtain an effective n_2' in the limit of no secondary processes, n_{2p}', we can instead use a switching function for the secondary processes reaction order,

$$n_2^\dagger = n_{2p}' + (n_2' - n_{2p}')\exp(-2\varepsilon'), \quad \text{(D4)}$$

where

$$n'_{2p} = \frac{1}{2}\left[n_c \frac{\mathcal{K}_P^{n_c}}{1+\mathcal{K}_P^{n_c}} - 1 - \frac{3}{1+\mathcal{K}_E}\right]. \quad \text{(D5)}$$

Using $n_2^{\dagger}$ instead of n'_2 in Eq. (10a) yields a fully universal solution valid for any possible kind of filamentous growth kinetics, single-step or catalytic multistep, and with or without secondary processes. Its only caveat is that it may not accurately describe filamentous growth under the rather specific circumstances of slow significant secondary processes, a high value of $n_c > 3$, and no saturation in primary nucleation.

APPENDIX E: EXTENSION TO THE CASE OF SIMULTANEOUS FRAGMENTATION AND SECONDARY NUCLEATION

When a protein aggregation reaction that features secondary nucleation occurs under shaking conditions, both fragmentation and secondary nucleation may occur simultaneously. The same methods as employed in the rest of the paper may be employed to capture this behavior, yielding a modified expression for k'_2 and n'_2,

$$k'_2 = \frac{\mathcal{K}_S^{n_2} + f(1+\mathcal{K}_S^{n_2})}{1+\mathcal{K}_S^{n_2}}, \quad \text{(E1)}$$

$$n'_2 = n_2 \frac{\mathcal{K}_S^{n_2}}{1+\mathcal{K}_S^{n_2}} \frac{\mathcal{K}_S^{n_2}}{\mathcal{K}_S^{n_2} + f(1+\mathcal{K}_S^{n_2})} - \frac{2}{1+\mathcal{K}_E}, \quad \text{(E2)}$$

$$f = \frac{k_-}{k_2 m_{\mathrm{tot}}^{n_2}}, \quad \text{(E3)}$$

where $k\_$ is the rate constant for fibril fragmentation [and $k\_M(t)$ is the corresponding rate].

This approximation is reasonably accurate but less so than the other cases considered in this paper.

REFERENCES

[1] F. Chiti and C. M. Dobson, Annu. Rev. Biochem. **75**, 333 (2006).
[2] F. Oosawa and M. Kasai, J. Mol. Biol. **4**, 10 (1962).
[3] F. Oosawa and S. Asakura, *Thermodynamics of the Polymerization of Protein* (Academic Press, 1975).
[4] S. R. Collins, A. Douglass, R. D. Vale, and J. S. Weissman, PLoS Biol. **2**, e321 (2004).
[5] F. A. Ferrone, J. Hofrichter, H. R. Sunshine, and W. A. Eaton, Biophys. J. **32**, 361 (1980).
[6] F. A. Ferrone, J. Hofrichter, and W. A. Eaton, J. Mol. Biol. **183**, 591 (1985).
[7] F. A. Ferrone, J. Hofrichter, and W. A. Eaton, J. Mol. Biol. **183**, 611 (1985).
[8] 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).
[9] J. D. Schmit, J. Chem. Phys. **138**, 185102 (2013).
[10] 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. U. S. A. **111**, 9384 (2014).
[11] T. P. J. 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).
[12] S. I. A. Cohen, M. Vendruscolo, M. E. Welland, C. M. Dobson, E. M. Terentjev, and T. P. J. Knowles, J. Chem. Phys. **135**, 065105 (2011).
[13] S. I. A. Cohen, M. Vendruscolo, C. M. Dobson, and T. P. J. Knowles, J. Chem. Phys. **135**, 065106 (2011).
[14] S. I. A. Cohen, M. Vendruscolo, C. M. Dobson, and T. P. J. Knowles, J. Chem. Phys. **135**, 065107 (2011).
[15] S. I. A. Cohen, M. Vendruscolo, C. M. Dobson, and T. P. J. Knowles, Int. J. Mol. Sci. **12**, 5844 (2011).
[16] G. Meisl, J. B. Kirkegaard, P. Arosio, T. T. C. Michaels, M. Vendruscolo, C. M. Dobson, S. Linse, and T. P. J. Knowles, Nat. Protoc. **11**, 252 (2016).
[17] 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**, 273 (2018).
[18] 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. U. S. A. **110**, 9758 (2013).
[19] L. Michaelis and M. Menten, Biochem. Zeitschrift **49**, 333 (1913).
[20] A. Šarić, A. Buell, G. Meisl, T. C. T. Michaels, C. M. Dobson, S. Linse, T. P. J. Knowles, and D. Frenkel, Nat. Phys. **12**, 874 (2016).
[21] W. 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).
[22] T. Scheibel, J. Bloom, and S. L. Lindquist, Proc. Natl. Acad. Sci. U. S. A. **101**, 2287 (2004).
[23] E. Chatani, R. Ohnishi, T. Konuma, K. Sakurai, H. Naiki, and Y. Goto, J. Mol. Biol. **400**, 1057 (2010).
[24] G. Meisl, L. Rajah, S. A. I. Cohen, M. Pfammatter, A. Šarić, E. Hellstrand, A. K. Buell, A. Aguzzi, S. Linse, M. Vendruscolo, C. M. Dobson, and T. P. J. Knowles, Chem. Sci. **8**, 7087 (2017).
[25] T. C. T. Michaels, S. I. A. Cohen, M. Vendruscolo, C. M. Dobson, and T. P. J. Knowles, Phys. Rev. Lett. **116**, 038101 (2016).
[26] T. C. T. Michaels and T. P. J. Knowles, Am. J. Phys **82**, 476 (2014).
[27] T. C. T. Michaels and T. P. J. Knowles, J. Chem. Phys. **140**, 214904 (2014).
[28] T. C. Michaels, A. J. Dear, and T. P. Knowles, Int. Rev. Phys. Chem. **35**, 679 (2016).
[29] T. C. T. Michaels, A. J. Dear, and T. P. J. Knowles, Phys. Rev. E **99**, 062415 (2019).
[30] P. Arosio, T. P. J. Knowles, and S. Linse, Phys. Chem. Chem. Phys. **17**, 7606 (2015).
[31] T. C. T. Michaels, A. J. Dear, and T. P. J. Knowles, New J. Phys. **20**, 055007 (2018).
[32] R. P. Sear, CrystEngComm **16**, 6506 (2014).
[33] V. Marghussian, in *Nano-Glass Ceramics*, edited by V. Marghussian (William Andrew Publishing, Oxford, 2015), pp. 1–62.
[34] J. R. Espinosa, C. Vega, C. Valeriani, D. Frenkel, and E. Sanz, Soft Matter **15**, 9625 (2019).
[35] R. Smallman and A. Ngan, in *Modern Physical Metallurgy*, 8th ed., edited by R. Smallman and A. Ngan (Butterworth-Heinemann, Oxford, 2014), pp. 93–119.
[36] H. Pruppacher and J. Klett, "Homogeneous nucleation," *Microphysics of Clouds and Precipitation* (Springer Netherlands, Dordrecht, 2010), pp. 191–215.
[37] S. Campioni, G. Carret, S. Jordens, L. Nicoud, R. Mezzenga, and R. Riek, J. Am. Chem. Soc. **136**, 2866 (2014).
[38] M. A. Rubio, D. E. Schlamadinger, E. M. White, and A. D. Miranker, Biochemistry **54**, 987 (2015).
[39] C. L. L. Pham, A. Rey, V. Lo, M. Soulès, Q. Ren, G. Meisl, T. P. J. Knowles, A. H. Kwan, and M. Sunde, Sci. Rep. **6**, 25288 (2016).