



An Analytical Solution to the Kinetics of Breakable Filament Assembly

Tuomas P. J. Knowles *et al.*

Science **326**, 1533 (2009);

DOI: 10.1126/science.1178250

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by [clicking here](#).

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines [here](#).

The following resources related to this article are available online at www.sciencemag.org (this information is current as of September 22, 2012):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/content/326/5959/1533.full.html>

Supporting Online Material can be found at:

<http://www.sciencemag.org/content/suppl/2009/12/10/326.5959.1533.DC1.html>

This article **cites 26 articles**, 6 of which can be accessed free:

<http://www.sciencemag.org/content/326/5959/1533.full.html#ref-list-1>

This article has been **cited by** 4 article(s) on the ISI Web of Science

This article has been **cited by** 15 articles hosted by HighWire Press; see:

<http://www.sciencemag.org/content/326/5959/1533.full.html#related-urls>

This article appears in the following **subject collections**:

Biochemistry

<http://www.sciencemag.org/cgi/collection/biochem>

HQ theropods are separated from each other by branches subtending taxa from other continental faunas, indicating that dispersal between these geographical regions probably occurred during the Camian-Norian. Other contemporaneous theropod assemblages from Europe (22) and South America contain only members of Neotheropoda and do not match the diversity of theropods at the HQ.

Both parsimony (23) and likelihood-based (24) biogeographic methods for ancestral range reconstruction reject scenarios of an endemic North American theropod radiation (10). Analyses differ slightly in support for range reconstructions at individual nodes, but provide high relative support for inferring the South American protocontinent as the ancestral range through the spine of the basal dinosaur tree (10). In most analyses, the distributions of the three HQ theropods are explained by either dispersal to North America from South America or allopatric and/or vicariant speciation from an ancestral widespread range encompassing North and South America (10). This pattern is apparent in many other clades during the Late Triassic, including actosaurs (25), crocodylomorphs (26), shuvosaurids (27), and “traversodont” cynodonts (28). The ubiquity of this phylogenetic pattern in clades encompassing markedly different ecomorphotypes argues against the presence of physiographic barriers isolating the Norian faunas of North America. Thus, the conspicuous absence of sauropodomorphs in the Norian of North America (3, 12) cannot be attributed to their inability to disperse to these areas but rather to their inability to become established in areas sampled by Late Triassic terrestrial sedimentary outcrops. Latitudinal differentiation of Norian faunas attributable to climatic

differences and climatic tolerances remains an intriguing explanation for the global ubiquity of basal theropod taxa such as *Tawa* and the North American absence of sauropodomorphs. Indeed, recent paleoclimate models and proxy data for the Late Triassic reveal a marked dichotomy between low and high paleolatitudes (29). Alternative explanations, including smaller-scale ecological differences, community-level interactions, or facies-dependent sampling biases, cannot be ruled out, nor are these explanations mutually exclusive (12). Explaining these patterns remains an outstanding problem in early dinosaur evolution at the nexus of phylogenetic, geologic, and paleoclimatic studies of the Late Triassic.

References and Notes

- P. C. Sereno, F. E. Novas, *Science* **258**, 1137 (1992).
- P. C. Sereno, *Science* **284**, 2137 (1999).
- S. J. Nesbitt, R. B. Irmis, W. G. Parker, *J. Syst. Palaeontol.* **5**, 209 (2007).
- R. B. Irmis, W. G. Parker, S. J. Nesbitt, J. Liu, *Hist. Biol.* **19**, 3 (2007).
- F. E. Novas, *J. Vertebr. Paleontol.* **16**, 723 (1996).
- R. R. Rogers, C. C. Swisher III, P. C. Sereno, C. A. Forster, A. M. Monetta, *Science* **260**, 794 (1993).
- S. P. Hesselbo, S. A. Robinson, F. Surlyk, S. Piasecki, *Geology* **30**, 251 (2002).
- P. E. Olsen *et al.*, *Science* **296**, 1305 (2002).
- J. C. McElwain, D. J. Beerling, F. I. Woodward, *Science* **285**, 1386 (1999).
- See supporting material on Science Online.
- R. Mundil, G. Gehrels, A. L. Deino, R. B. Irmis, *Eos* **89**, abstract V13A-2108 (2008).
- R. B. Irmis *et al.*, *Science* **317**, 358 (2007).
- N. C. Fraser, K. Padian, G. M. Walkden, A. L. M. Davis, *Palaeontology* **45**, 79 (2002).
- M. C. Langer, M. J. Benton, *J. Syst. Palaeontol.* **4**, 309 (2006).
- M. D. Ezcurra, *Geodiversitas* **28**, 649 (2006).
- R. J. Butler, P. M. Barrett, D. J. Gower, *Biol. Lett.* **5**, 557 (2009).

- M. Kundrát, J. Janáček, *Naturwissenschaften* **94**, 769 (2007).
- N. D. Smith, P. J. Makovicky, W. R. Hammer, P. J. Currie, *Zool. J. Linn. Soc.* **151**, 377 (2007).
- N. D. Smith, A. H. Turner, *Syst. Biol.* **54**, 166 (2005).
- G. T. Lloyd *et al.*, *Proc. R. Soc. London Ser. B* **275**, 2483 (2008).
- R. Irmis, R. Mundil, *J. Vertebr. Paleontol.* **28**, 95A (2008).
- O. W. M. Rauhut, A. Hungerbühler, *Gaia* **15**, 75 (2000).
- F. Ronquist, *Syst. Biol.* **46**, 195 (1997).
- R. H. Ree, S. A. Smith, *Syst. Biol.* **57**, 4 (2008).
- W. G. Parker, *J. Syst. Palaeontol.* **5**, 41 (2007).
- J. M. Clark, X. Xu, C. A. Forster, Y. Wang, *Nature* **430**, 1021 (2004).
- S. J. Nesbitt, *Bull. Am. Mus. Nat. Hist.* **302**, 1 (2007).
- J. A. Hopson, H.-D. Sues, *Paläontol. Z.* **80**, 124 (2006).
- B. W. Sellwood, P. J. Valdes, *Sediment. Geol.* **190**, 269 (2006).
- Fieldwork and research were funded by the National Geographic Society (grant no. 8014-06 to K. Padian), David Clark Inc., the Jean Butz Memorial Fund, the Theodore Roosevelt Memorial Fund (to R.B.I.), the Jurassic Foundation (to S.J.N.), an Explorers Club Exploration Grant (to S.J.N.), an NSF Graduate Fellowship (to S.J.N. and R.B.I.), the Samuel & Doris Welles Research Fund (to R.B.I. and K. Padian), a Bryan Patterson Memorial Grant (to R.B.I.), the University of California-Berkeley Department of Integrative Biology (to R.B.I.), the University of Utah (to R.B.I.), and the Systematics Research Fund (to N.D.S.). We thank the reviewers, R. Ree, and the Field Museum (Chicago) VertMorph discussion group; and the staff, volunteers, and paleontology seminar participants at Ghost Ranch for assistance with fieldwork. Ghost Ranch, the American Museum of Natural History, the Field Museum of Natural History, and the University of California Museum of Paleontology all facilitated the preparation of specimens.

Supporting Online Material

www.sciencemag.org/cgi/content/full/326/5959/1530/DC1
SOM Text
Figs. S1 to S8
Tables S1 to S5
References

10 August 2009; accepted 2 October 2009
10.1126/science.1180350

An Analytical Solution to the Kinetics of Breakable Filament Assembly

Tuomas P. J. Knowles,^{1,2} Christopher A. Waudby,³ Glyn L. Devlin,³ Samuel I. A. Cohen,³ Adriano Aguzzi,⁴ Michele Vendruscolo,³ Eugene M. Terentjev,¹ Mark E. Welland,^{2*} Christopher M. Dobson^{3*}

We present an analytical treatment of a set of coupled kinetic equations that governs the self-assembly of filamentous molecular structures. Application to the case of protein aggregation demonstrates that the kinetics of amyloid growth can often be dominated by secondary rather than by primary nucleation events. Our results further reveal a range of general features of the growth kinetics of fragmenting filamentous structures, including the existence of generic scaling laws that provide mechanistic information in contexts ranging from in vitro amyloid growth to the in vivo development of mammalian prion diseases.

Molecular self-assembly is the basis of phenomena ranging from the construction of materials for nanotechnology (1) to the formation of molecular machineries within living cells (2). The assembly of these frequently complex and highly intricate structures typically depends on a series of individual steps that are inherently simple and are therefore amenable in principle to a quantitative analysis

based on physical principles. An important class of molecular structures that emerges from the self-assembly of simpler components is that of filamentous assemblies of biological macromolecules. Many proteinaceous aggregates of this type, which are increasingly linked with normal and aberrant biological processes (2), form through a nucleation mechanism followed by a self-templated growth where the ends of exist-

ing filaments recruit soluble molecules into aggregates that can themselves multiply through secondary nucleation processes such as fragmentation (Fig. 1A).

One of the key questions in molecular self-assembly phenomena is to determine the relative importance of different microscopic processes and their contribution to the overall reaction (3, 4). Master equation approaches are particularly powerful in this context as they enable the explicit description of microscopic processes and have thus offered a series of insights (5–10) into phenomena including the formation of amyloid fibrils, species that are of increasing interest particularly because of their association with clinical disorders ranging from Alzheimer’s disease to type II diabetes (2). The lack of analytical

¹Cavendish Laboratory, University of Cambridge, J. J. Thomson Avenue, Cambridge CB3 0HE, UK. ²Nanoscience Centre, University of Cambridge, J. J. Thomson Avenue, Cambridge CB3 0FF, UK. ³Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK. ⁴Institute of Neuro-pathology, University Hospital Zurich, Schmelzbergstrasse, 8091 Zurich, Switzerland.

*To whom correspondence should be addressed. E-mail: mew10@cam.ac.uk (M.E.W.); cmd44@cam.ac.uk (C.M.D.)

solutions to such master equations has, however, represented a challenge in the quest to establish general principles and laws governing filamentous growth. Scaling laws, also known in biology as allometric laws, define the relationships between different properties of a given system, and such laws have enabled a variety of fundamental principles to be revealed in fields ranging from condensed matter physics to ecology and from sociology to economics (11, 12). In this work, we illustrate how the availability of an analytical treatment of a master equation (Eq. 1) enables the rationalization of a wide range of experimental results relating to the self-assembly processes of peptides and proteins based on scaling laws and increases the predictive power of the analysis of such phenomena.

The basic processes that enter the master equation describing filamentous protein aggregation have been established previously. Amyloid formation has almost universally been shown to involve a nucleation-dependent polymerization reaction (13), where the formation of growth nuclei from soluble proteins is slower than the elongation of preformed nuclei. In addition, secondary nucleation mechanisms (3) have been identified (5, 7, 14) that lead to the formation of additional nuclei from preexisting filaments; we include here a common type of such mechanism, namely filament fragmentation (5), and other types can also readily be considered within the same scheme [see Supporting Online Material (SOM)]. These molecular processes lead to a master equation describing the time evolution of the concentration $f(t, j)$ of filaments of length j in the system as an infinite set of coupled nonlinear differential equations of the form (6–10)

$$\begin{aligned} \frac{\partial f(t, j)}{\partial t} = & 2m(t)k_+ f(t, j-1) \\ & - 2m(t)k_+ f(t, j) - k_-(j-1)f(t, j) \\ & + 2k_- \sum_{i=j+1}^{\infty} f(t, i) + k_n m(t)^{n_c} \delta_{j, n_c} \quad (1) \end{aligned}$$

where $m(t)$ is the concentration of monomers. The different terms in Eq. 1 represent the elementary microscopic processes. The first term $2m(t)k_+ f(t, j-1)$ accounts for the increase in the number of filaments of length j due to the addition of monomers to either end of a filament of length $j-1$. The term $-2m(t)k_+ f(t, j)$ describes the decrease in their number due to filaments of length j growing further to length $j+1$, the term $-k_-(j-1)f(t, j)$ reflects the possibility of a filament of length j breaking at any of its $j-1$ internal links, and the term $2k_- \sum_{i=j+1}^{\infty} f(t, i)$ accounts for the fact that there are two links in any filament of length $i > j$ where breakage leads to a filament of length j . The last term represents the spontaneous formation of growth nuclei of size n_c as a polynomial form in $m(t)$, which we shall see has the property that it allows the conclusions from the classical theory of homogeneous nucle-

ated growth to be recovered as the limit of our results when setting the fragmentation rate to zero.

Using analytical techniques based on fixed-point mappings (15), we extend linearized early time solutions to describe the full time course of the reaction and reveal [see details in (16)] that the time evolution of the most important experimentally accessible observables, the principal moments of the distribution $f(t, j)$ (the number $P = \sum_{j=n_c}^{\infty} f(t, j)$ and mass $M = \sum_{j=n_c}^{\infty} j \times f(t, j)$ concentrations of filaments) can be written in closed form to a very good approximation (Fig. 1 B) as the integrated rate laws

$$P(t) = \frac{m_{\text{tot}}}{2n_c - 1} - \frac{m_{\text{tot}} k_- e^{-(2n_c - 1)k_- t} \text{Ei}(-C_+ e^{k_- t})}{\kappa} + e^{-(2n_c - 1)k_- t} B_2 \quad (2a)$$

$$M(t) = m_{\text{tot}} [1 - \exp(-C_+ e^{k_- t} + C_- e^{-k_- t}) + k_n m_{\text{tot}}^{n_c - 1} k_-^{-1}] \quad (2b)$$

where $\kappa = \sqrt{2m_{\text{tot}} k_+ k_-}$ represents the rate of multiplication of the filament population, $m_{\text{tot}} = M(t) + m(t)$ is the total protein concentration, $\text{Ei}(x) = -\int_{-x}^{\infty} (e^{-y}/y) dy$ denotes the exponential integral function, $C_{\pm} = k_{\pm} P(0) \kappa \pm M(0) / (2m_{\text{tot}} \pm (k_n m_{\text{tot}}^{n_c - 1}) / (2k_-))$, and the constant B_2 is fixed by the initial condition $P(0)$. We note that the form of the rate of filament multiplication κ —which, as discussed below, is a key parameter in our description of filamentous growth—is consistent with the idea that both the elongation rate $m_{\text{tot}} k_+$ and the fragmentation rate k_- have to double for the overall rate of production of new filaments to double. We point out that we have assumed that the fragmentation rate is independent of the length of the filaments; this assumption was primarily made because of the current difficulty of obtaining reliable experimental estimates of length-dependent fragmentation rates. The approach that we describe, however, can be

in principle modified to take this effect, and indeed others, into consideration when more accurate measurements become available.

In the following, we demonstrate that the theoretical framework provided by Eqs. 2a and 2b has the capacity both to account for and to predict the outcome of protein aggregation reactions of the type associated with protein misfolding diseases. As a first example, we consider the case of a representative protein, insulin; the data in Fig. 2A show that a comprehensive series of reactions performed in vitro with different initial conditions, and resulting in differently shaped kinetic curves ranging from sigmoidal to convex, can be completely accounted for by the solution Eqs. 2a and 2b, with the same values of the two global parameters corresponding to the microscopic elongation and breakage rate constants. According to our treatment, the sigmoidal growth kinetics can be observed as a result of secondary nucleation, here fibril fragmentation, even in the absence of rate-limiting primary nucleation events. This result contrasts with the classical theories of nucleated growth, where homogeneous nucleation is the only source of additional filaments and where such nucleation processes are consequently crucial for determining the duration of the lag phase (13)—the characteristic time of nucleated polymerization reactions during which the initial growth rate is small, and in some cases not measurable using bulk techniques (17).

The filament number concentration $P(t)$ can be reconstructed experimentally (7) from measurements of the free monomer concentration $m(t) = m_{\text{tot}} - M(t)$ and from the measurement of the rate of change $\dot{M}(t)$ in the mass concentration of polymers: $P(t) = \dot{M}(t) / \{2k_+ [m_{\text{tot}} - M(t)]\}$. Comparison of the results of this reconstruction with the analytical expression for $P(t)$ from Eqs. 2a and 2b demonstrates that in this case also the use of the same values of k_+ and k_- yields good agreement with the measurements (Fig. 2B, inset).

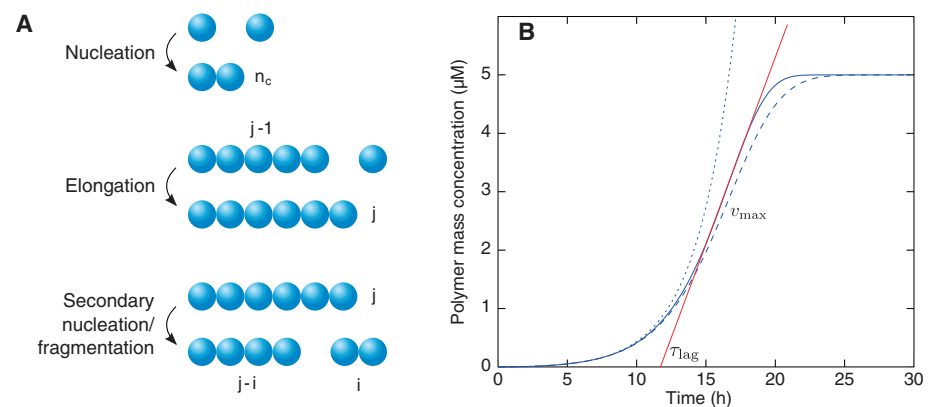


Fig. 1. Kinetics of self-templated aggregation. (A) Growth through nucleation (1), elongation (2), and fragmentation (3) leads to sigmoidal kinetic curves (B) for the mass concentration of fibrils as a function of time. Dashed blue line, first moment $M(t)$ computed from the numerical solution of the master equation, Eq. 1; solid blue curve, analytical solution given in Eqs. 2a and 2b, obtained through fixed-point iteration of the early time limit (dotted curve). The parameters are $k_+ = 5 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$, $k_- = 2 \times 10^{-8} \text{ s}^{-1}$, $m_{\text{tot}} = 5 \times 10^{-6} \text{ M}$, $k_n = 2 \times 10^{-5} \text{ M}^{-1} \text{ s}^{-1}$, $n_c = 2$, and $M(0) = P(0) = 0$. The definitions of the lag phase τ_{lag} and the maximal growth rate v_{max} are shown.

The importance of fibril fragmentation for the phenomenon of amyloid growth can be illustrated with the very high sensitivity of the growth kinetics to mechanical shear. We examine in this context the polymerization of β -lactoglobulin (18) under a series of different controlled rates of shearing that promote breakage. We considered a constant primary nucleation rate, k_n , for the entire data set, and a one-parameter fit to each kinetic curve to allow the breakage rate to vary through different values of κ . As shown in Fig. 2C, the variation of this single parameter can account explicitly for the observed differences in both the lag time and the growth rate under different shear conditions. This result shows that even a moderate increase in fracture rate can have very important consequences for the number of protein molecules incorporated into the aggregates; hence, fragmentation emerges as possessing fundamental importance for determining key observables such as the lag phase. This result is also likely to lie at the heart of the well-known high sensitivity of the kinetics of fibril growth to agitation or sonication, processes known to introduce enhanced fragmentation but likely to leave largely unaffected other rate constants for the system.

Based on knowledge gained from the analysis of the rates of the individual processes underlying the growth of protein fibrils, we can now predict

quantitatively how the course of this reaction is affected by changes in system parameters such as protein concentration. To illustrate this idea, we consider data from (19) for the amyloid growth rates of the WW domain in vitro as the concentration of the monomeric protein is varied over an order of magnitude. We extract the microscopic rates through a fit to one of the kinetic curves (Fig. 2D, blue line), and then quantitatively predict as a function of time, with no free parameters, the behavior of the system for a different set of initial conditions (Fig. 2D, gray line) simply through the scaling of the microscopic rates with concentration according to Eqs. 2a and 2b.

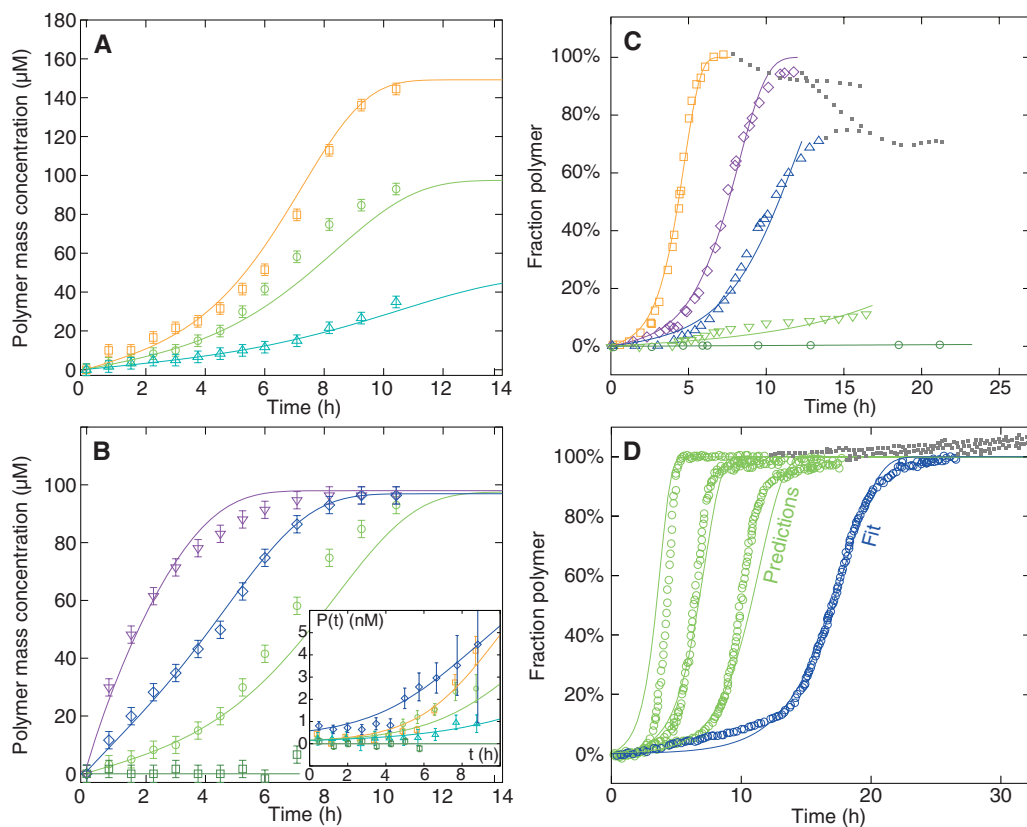
Our analytical theory further reveals general features of the growth of fragmenting filaments. First, considering the normalized maximal growth rate, given as the steepest slope of the sigmoidal kinetic trace (Fig. 1A) divided by the initial soluble protein concentration $v_{\max} = m_{\text{tot}}^{-1} \max_t [M(t)] \approx m_{\text{tot}}^{-1} [M(t)]_{t=t_{\max}} = \kappa/e$, where $t_{\max} = \kappa^{-1} \log(1/C_+)$ and we have kept the leading C_+ term in Eqs. 2a and 2b over constant and decaying terms, our results show that this maximal rate depends on the nature of the aggregates and the environment in which they are formed only through the single parameter κ , and not on other system parameters such as the number of nuclei present initially or on the primary nucleation rate. This re-

sult is in contrast to classical linear growth theories (20), where the growth rate and lag phase have strong dependencies on the specific details of the homogeneous nucleation process.

A closer examination of the role of the parameter κ reveals that it also essentially defines the lag phase. The lag phase exists only if the growth rate $m_{\text{tot}}\kappa/e$ is maximal at the inflection point $t_{\max} = \log(1/C_+)\kappa^{-1}$, or equivalently for $M(0) < M_C = \kappa L_0/(2k_+e)$, where M_C represents a critical seed concentration which, if exceeded, results in a reaction proceeding without a lag phase. This prediction for the existence of such a threshold can be verified experimentally as shown in Fig. 2B. The conditions used result in a threshold of $M_C = 0.9 \mu\text{M}$, and it can be seen that the reaction starting with $M(0) = 0.21 \mu\text{M}$ proceeds with a marked time lag, whereas for $M(0) = 2.4 \mu\text{M}$ the rate profile is convex. If we define the lag time as shown in Fig. 1A by extrapolation from the maximal growth rate, we obtain the expression $\tau_{\text{lag}} = [\log(1/C_+) - e + 1]\kappa^{-1}$.

It is well known that primary nucleation is an essential step in the phenomenon of amyloid growth in the absence of seeds, and a contribution to the duration of the lag phase can indeed result from such a nucleation process. Our results show, however, that the nucleation rate enters only as a logarithmic correction through C_+

Fig. 2. Experimental measurements of the polymer mass concentration $M(t)$ of fibrillar insulin in (A) and (B) are explained using Eqs. 2a and 2b with the two parameters $k_+ = 2.9 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$ and $k_- = 2.1 \times 10^{-9} \text{ s}^{-1}$. In (A), the total monomer concentration was (from top to bottom) $m_{\text{tot}} = 149 \mu\text{M}$, $98 \mu\text{M}$, and $49 \mu\text{M}$, and the seed mass concentration used was $M(0) = 0.21 \mu\text{M}$. In (D), for a constant $m_{\text{tot}} = 98 \mu\text{M}$, the seed concentration was (from left to right) $M(0) = 2.4 \mu\text{M}$, $0.8 \mu\text{M}$, $0.21 \mu\text{M}$, and $0 \mu\text{M}$, and the average polymerization number of the seed fibrils $L_0 = M(0)/P(0) = 1380$ was estimated from AFM measurements (see SOM). The inset in (B) shows the polymer number concentration $P(t)$ determined simultaneously for the data in (A) and (B) for the same values of k_+ and k_- . The effect of fragmentation is shown in (C); the polymerization of β -lactoglobulin was measured by Hill *et al.* (18) for increasing shear rates (right to left: 0, 25, 50, 100, and 200 s^{-1}) and was scaled here between 0% and 100%. The values $n_c = 2$, $k_r/k_- = 22.8$ were held constant in all the data sets, and $\kappa = 1.2 \times 10^{-4} \text{ s}^{-1}$, $4.3 \times 10^{-5} \text{ s}^{-1}$, $2.6 \times 10^{-5} \text{ s}^{-1}$, $3.7 \times 10^{-6} \text{ s}^{-1}$, and 0 s^{-1} . (D) Experimental measurements from Ferguson *et al.* (19) for the polymerization of the WW domain. The two parameters in Eqs. 2a and 2b were derived from the data at $m_{\text{tot}} = 50 \mu\text{M}$ (blue curve); using the resulting values of $k_+, k_- = 1.7 \times 10^{-4} \text{ M}^{-1} \text{ s}^{-1}$, $k_r/k_- = 13.2 \text{ M}^{-1}$ to predict the polymerization time course for $m_{\text{tot}} = 100 \mu\text{M}$, $200 \mu\text{M}$, and $500 \mu\text{M}$ is shown in gray, whereby the parameter $n_c = 2$ is fixed.



The comparison between this prediction (green curves) and the experimental measurements (green circles) demonstrates that both the maximal rate and the time at which this rate occurs are predicted accurately. Data not considered in the fits are shown with filled gray squares.

in the expression for τ_{lag} ; therefore, when secondary nucleation pathways are active, we conclude that the experimental lag time is primarily determined by the exponential growth regime that takes place in the initial phase of the reaction. Under these circumstances, the formation of oligomeric species, a process that has been observed concomitantly with fibril formation and linked

with cellular toxicity (2), is likely to contribute to such logarithmic terms and may therefore have a limited influence on the overall growth kinetics of fibrillar species. These results reveal that, although the kinetic curves for processes dominated by primary nucleation (20) and by secondary nucleation give rise to apparently similar kinetic traces, their origin is fundamentally different.

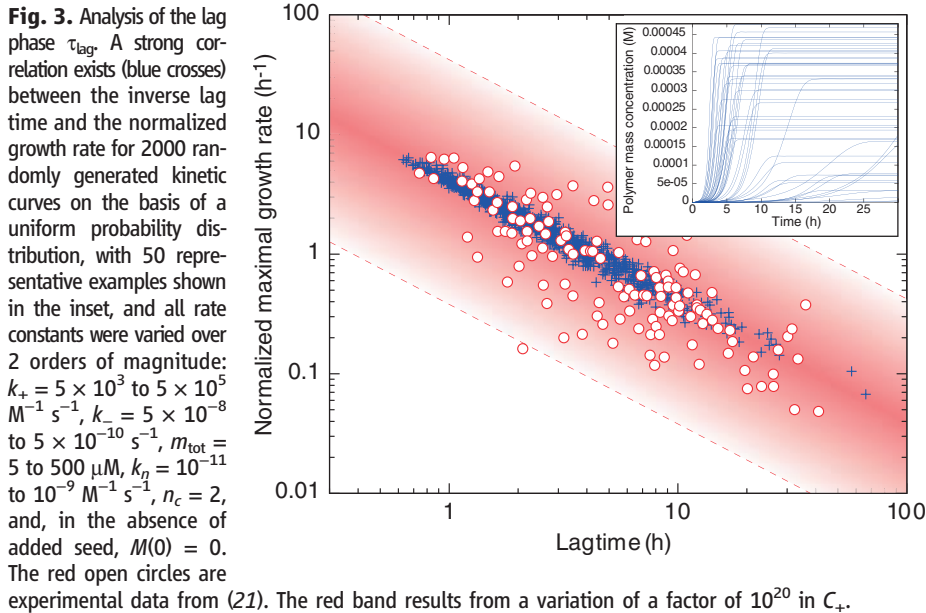


Fig. 3. Analysis of the lag phase τ_{lag} . A strong correlation exists (blue crosses) between the inverse lag time and the normalized growth rate for 2000 randomly generated kinetic curves on the basis of a uniform probability distribution, with 50 representative examples shown in the inset, and all rate constants were varied over 2 orders of magnitude: $k_+ = 5 \times 10^3$ to $5 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$, $k_- = 5 \times 10^{-8}$ to $5 \times 10^{-10} \text{ s}^{-1}$, $m_{\text{tot}} = 5$ to $500 \text{ } \mu\text{M}$, $k_n = 10^{-11}$ to $10^{-9} \text{ M}^{-1} \text{ s}^{-1}$, $n_c = 2$, and, in the absence of added seed, $M(0) = 0$. The red open circles are experimental data from (21). The red band results from a variation of a factor of 10^{20} in C_+ .

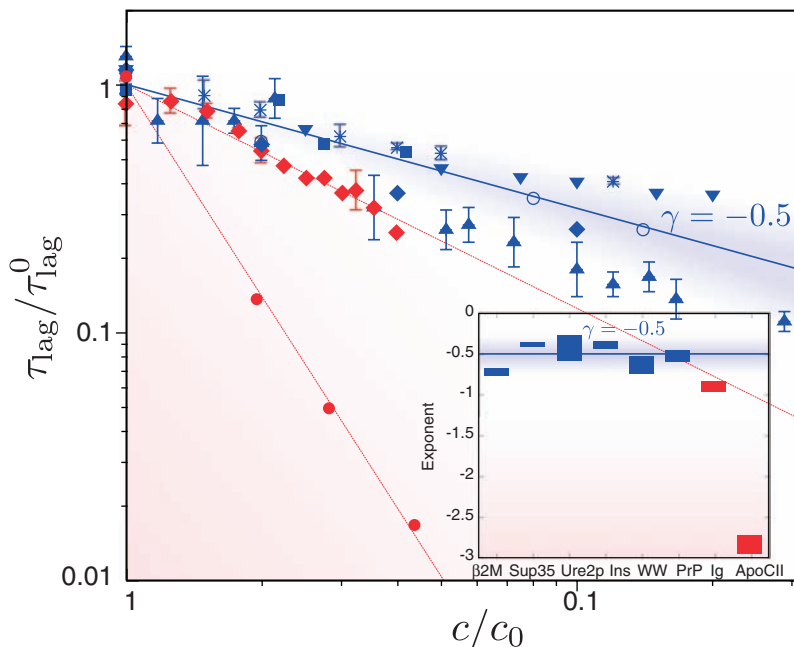


Fig. 4. Scaling of the lag phase with protein concentration. Data are shown for β 2-microglobulin (β 2M) (8) (blue upward triangles); the yeast prion fragment Sup35 (7) (blue crosses); Ure2p yeast prion (22) (blue squares); insulin (27) (blue downward triangles); WW domain (19) (blue diamonds); TI I27 immunoglobulin domain (Ig) (29) (red diamonds); apolipoprotein C-II (ApoCII) (28) (red circles); and times for onset of terminal disease in 50% of mice inoculated with a fixed dose of prions at $t = 0$ (PrP) (26) (blue circles). The slope corresponding to the exponent $\gamma = -0.5$ is shown with a blue solid line, and fits to the data sets with $\gamma < -0.5$ are shown in red. The inset shows the scaling exponents determined from the data in the main figure for the different systems. The box height in the inset shows standard errors from the linear fit.

Our theory also provides a general demonstration of the validity of a conjecture about amyloid fibril growth (21, 22) that the lag time is commonly highly correlated with the inverse maximal growth rate. Indeed, for fragmentation-assisted growth, we have shown that both the lag phase and the maximal growth rate depend primarily on κ . Therefore, the correlation emerges naturally without the requirement for the responses of the processes of nucleation and growth to changes in system parameters to be identical, a situation generally likely to be incompatible with structural and mechanistic information but required to enforce this correlation under the assumption that the lag phase is defined by primary nucleation processes and the growth phase solely by fibril elongation. Although the coefficient $[\log(1/C_+) - e + 1]$ linking the lag time to the maximal growth rate is variable for different proteins and different experimental conditions, the fact that these differences only enter as logarithmic corrections implies that the general inverse correlation between lag time and growth rate remains valid. In Fig. 3, this idea is illustrated with a red band that shows that the overall correlation is maintained, even when the factor C_+ has been artificially varied by a factor of 10^{20} , a value chosen to exceed greatly the range of variability accessible experimentally.

More generally, the parameter κ , which corresponds to the rate of multiplication of the population of filaments, emerges as the most important quantity describing the overall properties of systems that self-assemble by processes that involve elongation and fragmentation. Our results show that in the regime where secondary nucleation through fragmentation of aggregates is a more effective source of additional structures than primary nucleation [$k_-/(k_n m_{\text{tot}}^{n_c-1}) \gg 1$], observables such as the lag time and maximal growth rate depend primarily on just the single parameter κ .

A particularly striking result that follows from the dominance of a single kinetic parameter in this regime of elongation-fragmentation growth is that generic scaling laws emerge for the behavior of the system. For instance, the lag time is predicted to scale weakly with an exponent $\gamma = -0.5$ with respect to the monomer concentration (because $\tau_{\text{lag}} \sim \kappa^{-1} \approx m_{\text{tot}}^{-1}$ with $\gamma = -0.5$). We have analyzed experimental results for eight unrelated systems, including short peptides, proteins such as insulin and β 2-microglobulin, and different types of yeast prions, and find that five of them exhibit this type of weak scaling with high accuracy (Fig. 4). Interestingly, in absolute value, this exponent, 0.5, is crucially smaller than the value of 1.0 or greater required by traditional models (20) in which primary nucleation is a dominant effect determining the duration of the lag phase. In these latter models, the lag time scales with an exponent $\gamma = -(n_c + 1)/2$, which is always $|\gamma| \geq 1$ for a nucleus $n_c \geq 1$ (20), a result that is clearly inconsistent with the experimental observations for the majority of systems we have

evaluated (Fig. 4). Furthermore, in the case of the aggregation of glutamine-rich peptides, nucleus sizes of less than 0 appear to emerge (23) when using classical theories of homogeneous nucleated growth, whereas this type of low concentration dependence emerges naturally from our less-than-linear scaling laws and suggests that breakage is likely to be a key factor in the aggregation of polyglutamine peptides and hence potentially an important contributor to the development of disorders such as Huntington's disease (2).

It is interesting to note that in the limit of a vanishing breakage rate, $k_- \rightarrow 0$, an expansion of the exponentials in Eqs. 2a and 2b up to quadratic order in κ yields a polynomial growth form $M(t) \approx M(0) + k_+ k_n m_{\text{tot}}^{n_c+1} t^2 + 2k_+ m_{\text{tot}} P(0)t + n_c k_+ m_{\text{tot}} t$ at the early stages of the reaction. Our result therefore contains in the appropriate limit the behavior characteristic of classical nucleated growth theories (3, 20). It is likely, therefore, that filamentous growth reactions that deviate significantly from the $\gamma = -0.5$ exponent, such as those shown in red in Fig. 4, are limited by complex primary or secondary nucleation processes rather than by simple breakage. For instance, application of the schema described in this paper to growth phenomena where secondary nucleation processes produce additional seeds at a rate that depends on the presence of both filaments and monomer (14) $dP(t)/dt = k_2 M(t)m(t)^{n_2}$ yields a filament mul-

tiplication rate $\kappa = \sqrt{2k_+ k_2 m_{\text{tot}}^{n_2+1}}$ and therefore a scaling exponent $|\gamma| = (n_2 + 1)/2 \geq 1$ (see SOM), where k_2 represents a rate constant for monomer-dependent secondary nucleation and n_2 is the reaction order of the secondary nucleation.

Finally, we investigate the applicability of the analytical results derived in the present work to in vivo systems. We examine the phenomenon of mammalian prion transmission that is characteristic of disorders such as bovine spongiform encephalopathy and its human analog Creutzfeldt-Jacob disease. The protein-only hypothesis for the propagation of prions describes the conversion of cellular prion proteins (PrP^C) into a pathologically aggregated form (PrP^{Sc}) in a manner that can be transmissible (24). In such cases, PrP^{Sc} aggregates are thought to elongate through the addition of further PrP molecules, and it has been suggested that they multiply rapidly enough through fragmentation of existing structures (5, 25) to allow for transmissibility; these mechanisms are equivalent to the microscopic processes described to the lowest order in Eq. 1. In agreement with this idea, we observe in Fig. 4 that the time for disease onset scales with an exponent $\gamma = -0.52 \pm 0.07$ as a function of the expression level (26) of the prion protein that determines its concentration in vivo; this exponent is fully consistent with the value of -0.5 observed for a range of in vitro systems (Fig. 4) and expected from the analysis developed in this paper for fragmentation-assisted growth, at least when other in vivo limiting factors, such as

cellular clearance mechanisms, have been overwhelmed.

In conclusion, we have provided a unified theoretical framework to address complex biomolecular self-assembly processes that involve primary and secondary nucleation events coupled to linear growth, and demonstrated its value for predicting the kinetics and mechanisms of the proliferation of amyloid fibrils, an example of filamentous growth that is increasingly important in the context of understanding and managing some of the most common and debilitating diseases of the modern era.

References and Notes

- G. M. Whitesides, J. P. Mathias, C. T. Seto, *Science* **254**, 1312 (1991).
- F. Chiti, C. M. Dobson, *Annu. Rev. Biochem.* **75**, 333 (2006).
- F. A. Ferrone, J. Hofrichter, W. A. Eaton, *J. Mol. Biol.* **183**, 611 (1985).
- F. Ferrone, *Methods Enzymol.* **309**, 256 (1999).
- M. Tanaka, S. R. Collins, B. H. Toyama, J. S. Weissman, *Nature* **442**, 585 (2006).
- N. Carulla et al., *Nature* **436**, 554 (2005).
- S. R. Collins, A. Douglass, R. D. Vale, J. S. Weissman, *PLoS Biol.* **2**, e321 (2004).
- W.-F. Xue, S. W. Homans, S. E. Radford, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 8926 (2008).
- T. Pöschel, N. V. Brilliantov, C. Frömmel, *Biophys. J.* **85**, 3460 (2003).
- K. C. Kunes, D. L. Cox, R. R. P. Singh, *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **72**, 051915 (2005).
- A. L. Barabási, R. Albert, *Science* **286**, 509 (1999).
- J. M. Skotheim, L. Mahadevan, *Science* **308**, 1308 (2005).
- J. T. Jarrett, P. T. Lansbury, *Cell* **73**, 1055 (1993).
- A. M. Ruschak, A. D. Miranker, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 12341 (2007).
- A. Granas, J. Dugundji, *Fixed Point Theory* (Springer-Verlag, 2003).
- The kinetic equation for the moment $P(t)$ can be found by taking the sum over j from n_c to ∞ on both sides of Eq. 1. The analogous equation for the first moment $M(t)$ is obtained by multiplying both sides by j and then taking the sum. The telescopic sums vanish; rearranging the order in double sums, we obtain, after some algebra, the coupled system

$$\frac{dP(t)}{dt} = k_- [M(t) - (2n_c - 1)P(t)] + k_n m(t)^{n_c} \quad (3a)$$

$$\frac{dM(t)}{dt} = 2[M(t)k_+ - n_c(n_c - 1)k_-/2]P(t) + n_c k_n m(t)^{n_c} \quad (3b)$$

where $m(t) = m_{\text{tot}} - M(t)$, and we set $f(t, j) = 0$ for all $j < n_c$, where $n_c \geq 2$ is the size of the smallest stable aggregate. Eqs. 3a and 3b make intuitive sense, e.g., the number of polymers at a given time, $P(t)$, can increase if, out of $(M - P)$ total links, the chains break at any location more than $(n_c - 1)$ links distant from each end, thereby forming two stable polymers and thus accounting for the $[M - (2n_c - 1)P]$ factor. The general idea underlying the solution of Eqs. 3a and 3b comes from a fixed-point analysis (25). To reformulate the problem in terms of a fixed-point equation, we first formally solve the differential system

$$P(t) = \int_0^t e^{-(2n_c - 1)k_- (t - \tau)} k_- M(\tau) d\tau + \tilde{B}_2 e^{-(2n_c - 1)k_- t} \quad (4a)$$

$$M(t) = m_{\text{tot}} - n_c(n_c - 1)k_-/2k_+ + B_1 e^{-2k_+ t} \int_0^t P(\tau) d\tau \quad (4b)$$

where B_1 and \tilde{B}_2 are integration constants, and we have neglected monomer consumption by nucleation terms

$O(k_n)$. We introduce an operator A , acting on the space of the two principal moments and defined by the right-hand sides of Eqs. 4a and 4b; by construction, the fixed points $x^*(t) = [M^*(t), P^*(t)]$ of A with $Ax^* = x^*$ are solutions to Eqs. 3a and 3b to within $O(k_n)$. The fixed-point equation is solved iteratively: $x^* = \lim_{n \rightarrow \infty} A^{(n)}(x_0)$; the resulting solution after one [for $M(t)$] and two [for $P(t)$] iterations is already in excellent agreement with the numerical result (Fig. 1) and with experiments (Figs. 2 to 4). Successful iteration requires a good starting value $x_0(t) = [M_0(t), P_0(t)]$, and we chose for this value the early aggregation limit, when $m_{\text{tot}} - M(t) \approx m_{\text{tot}}$ and Eqs. 3a and 3b give

$$P_0(t) = C_1 e^{\kappa t} + C_2 e^{-\kappa t} - \frac{n_c k_n m_{\text{tot}}^{n_c-1}}{2k_+} \quad (5a)$$

$$M_0(t) = \frac{2k_+ m_{\text{tot}} C_1}{\kappa} e^{\kappa t} - \frac{2k_+ m_{\text{tot}} C_2}{\kappa} e^{-\kappa t} - \frac{k_n m_{\text{tot}}^{n_c}}{k_-} \quad (5b)$$

where the effective rate $\kappa = \sqrt{2k_+ k_- m_{\text{tot}}}$, and we have used the condition $m_{\text{tot}} k_+ \gg k_-$, which guarantees the existence of an initially growing filament population. The constants C_1 and C_2 are related to the initial conditions $C_{1,2} = \frac{1}{2}P(0) + n_c k_n m_{\text{tot}}^{n_c-1}/(4k_+) \pm M(0)\kappa/(4m_{\text{tot}}k_+) \pm k_n m_{\text{tot}}^{n_c}\kappa/(4m_{\text{tot}}k_+)$. Using the starting values (5) for the fixed-point iteration, we arrive at the analytical result presented in Eqs. 2a and 2b. Although the agreement between Eqs. 2a and 2b and the numerical solution (Fig. 1) is good (typical differences comparable in magnitude to common experimental errors), systematic differences are present in the intermediate time regime. Qualitatively, these are primarily because the filament number concentration $P(t)$ entering the expression for $M(t)$ in Eqs. 4a and 4b in the first interaction is from the linearized solution and increases more rapidly than the true solution, and therefore the protein conversion rate $dM(t)/dt$ is overestimated. These errors can be minimized by successive iterations of the fixed-point operator beyond the first-order solution discussed here.

- J. S. Pedersen, D. E. Otzen, *Protein Sci.* **17**, 2 (2008).
- E. K. Hill, B. Krebs, D. G. Goodall, G. J. Howlett, D. E. Dunstan, *Biomacromolecules* **7**, 10 (2006).
- N. Ferguson et al., *Proc. Natl. Acad. Sci. U.S.A.* **100**, 9814 (2003).
- F. Oosawa, S. Asakura, *Thermodynamics of the Polymerization of Protein* (Academic Press, New York, 1975).
- M. Fändrich, *J. Mol. Biol.* **365**, 1266 (2007).
- L. Zhu, X.-J. Zhang, L.-Y. Wang, J.-M. Zhou, S. Perrett, *J. Mol. Biol.* **328**, 235 (2003).
- A. K. Thakur et al., *Nat. Struct. Mol. Biol.* **16**, 380 (2009).
- S. B. Prusiner et al., *Ann. Neurol.* **11**, 353 (1982).
- A. Aguzzi, C. Weissmann, *Nature* **389**, 795 (1997).
- M. Fischer et al., *EMBO J.* **15**, 1255 (1996).
- V. Foderà, F. Librizzi, M. Groenning, M. van de Weert, M. Leone, *J. Phys. Chem. B* **112**, 3853 (2008).
- K. J. Binger et al., *J. Mol. Biol.* **376**, 1116 (2008).
- C. F. Wright, S. A. Teichmann, J. Clarke, C. M. Dobson, *Nature* **438**, 878 (2005).
- T.P.J.K. acknowledges support from St John's College, Cambridge through a Research Fellowship. C.M.D. acknowledges support from the Leverhulme and Wellcome Trusts. We are grateful to A. Craig, A. Knowles, D. White, A. Buel, F. Chiti, A. Miranker, and J. Falsig for valuable discussions, and to M. Sandal for preparing a Web server, which is publicly available at www.vendruscolo.ch.cam.ac.uk/cgi-bin/nef/nefcgi.cgi, to enable the use of the equations presented in this work to fit experimental measurements and to predict aggregation kinetics.

Supporting Online Material

www.sciencemag.org/cgi/content/full/326/5959/1533/DC1
Materials and Methods
References

25 June 2009; accepted 14 October 2009
10.1126/science.1178250