Inherent Variability in the Kinetics of Autocatalytic Protein Self-Assembly

Juraj Szavits-Nossan,* Kym Eden,† Ryan J. Morris, Cait E. MacPhee, Martin R. Evans, and Rosalind J. Allen‡
SUPA, School of Physics and Astronomy, University of Edinburgh, Mayfield Road, Edinburgh EH9 3JZ, United Kingdom
(Received 17 February 2014; published 26 August 2014)

In small volumes, the kinetics of filamentous protein self-assembly is expected to show significant variability, arising from intrinsic molecular noise. This is not accounted for in existing deterministic models. We introduce a simple stochastic model including nucleation and autocatalytic growth via elongation and fragmentation, which allows us to predict the effects of molecular noise on the kinetics of autocatalytic self-assembly. We derive an analytic expression for the lag-time distribution, which agrees well with experimental results for the fibrillation of bovine insulin. Our expression decomposes the lag-time variability into contributions from primary nucleation and autocatalytic growth and reveals how each of these scales with the key kinetic parameters. Our analysis shows that significant lag-time variability can arise from both primary nucleation and from autocatalytic growth and should provide a way to extract mechanistic information on early-stage aggregation from small-volume experiments.

DOI: 10.1103/PhysRevLett.113.098101 PACS numbers: 87.14.em, 05.10.Gg, 87.15.nr, 87.18.Tt

The self-assembly of protein molecules into amyloid fibrils is associated with many degenerative diseases [1] but also presents potential opportunities for the development of new materials [2]. In both cases, it is of outstanding importance to identify the specific microscopic steps responsible for amyloid aggregation, especially in its early stages. An important success of recent biophysical work has been to show that *in vitro* kinetic data for amyloid fibril self-assembly can often be described by deterministic mechanistic models [3–11]. However, it is unclear how far the results of these large-volume experiments can be translated to clinically relevant intracellular aggregation phenomena, which occur in far smaller volumes.

In large-volume *in vitro* experiments (typically 100–1000 μl), measurements of the total mass of aggregated (fibrillar) protein as a function of time typically produce sigmoidal curves, as in Fig. 1(a) [2]. These data show an initial lag phase in which no aggregated protein is detectable, followed by a rapid growth phase, terminating in a plateau once all the protein is in the aggregated form. In large volumes, these characteristic sigmoidal growth curves can often be well fitted by deterministic kinetic models involving homogeneous primary nucleation [Fig. 1(c), I], filament elongation by monomer addition [Fig. 1(c), II], and autocatalysis via filament fragmentation [Fig. 1(c), III] [2,7,12–14]—although the contributions of primary nucleation and autocatalytic growth in the early stages of aggregation are often poorly distinguished [15]. Importantly, these models lead to analytical predictions for scaling behavior; for example, if autocatalysis is dominant, the mean lag time scales as the inverse square root of the product of the protein concentration, elongation, and fragmentation rates [7].

In a clinical context, however, fibril formation happens in much smaller volumes, on the scale of a human cell (typically 500–3000 fl). In small volumes, the stochastic nature of the underlying chemical reactions ("intrinsic molecular noise") is expected to become important, leading to inherent variability in the aggregation kinetics. Figure 1(b) shows the results of replicate kinetic Monte Carlo simulations of a stochastic version of the autocatalytic growth model [17] in a volume of 830 fl. These simulations predict significant variability in the lag time. Lag-time variability has

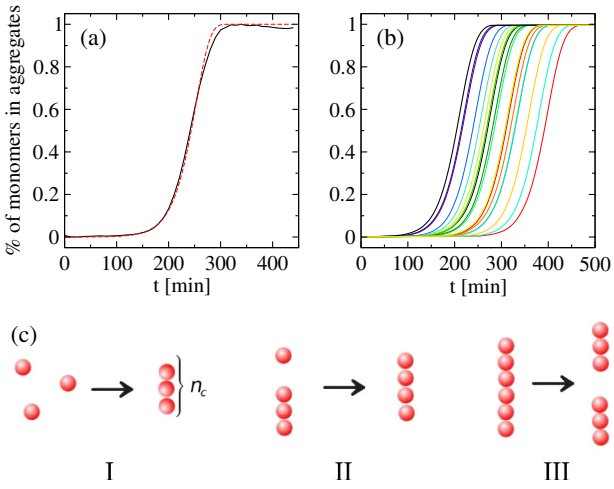

FIG. 1 (color online). (a) Experimental kinetic curve (black line) for the aggregation of bovine insulin in a volume of 100 μl from our own experiments, fitted to the theoretical prediction of a model [7] involving primary nucleation, elongation, and fragmentation (dashed red line); for full experimental details, see the Supplemental Material [16]. (b) Kinetic curves obtained from kinetic Monte Carlo simulations of a stochastic version of the same model [7,17] and the fit parameters extracted from (a), but for a much smaller volume of 830 fl. (c) Schematic illustration of (I) primary nucleation, (II) elongation via polymerization, and (III) fragmentation. The critical nucleus size for primary nucleation is denoted by n_c.

also been observed in recent pioneering experiments on bovine insulin fibril formation in microdroplets [18], as well as in classic small-volume experiments on the polymerization of sickle cell hemoglobin [19–21]. Theoretical models which explain such data should provide a powerful tool for probing the mechanisms involved in early-stage aggregation. In particular, an important question concerns the relative roles of primary nucleation [Fig. 1(c), I] and autocatalytic growth [Fig. 1(c), II and III] in determining the lag-time distribution for amyloid fibril formation. So far, however, analytic predictions for lag-time distributions have been achieved only for models that do not fully take into account autocatalytic growth [22,23].

In this Letter, we present an analytic prediction for the lag-time distribution, for a stochastic model of filamentous protein self-assembly that includes primary nucleation, irreversible filament elongation, and autocatalysis via fragmentation. We show that our prediction is in agreement with recent data for bovine insulin fibril formation in microdroplets [18]. This analytical solution allows a decomposition of the lag-time variability into contributions from primary nucleation and autocatalytic growth and reveals how each of these scales with the key kinetic parameters.

A coarse-grained model for autocatalytic protein self-assembly.—Deterministic kinetic models for amyloid fibril self-assembly usually consist of dynamical equations for the mean number of fibrils $\langle n_i \rangle$ of a given length $i \geq n_c$, where n_c is the size of the smallest growth-competent fibril (the "nucleus") [2,7,12–14,24]. For a model including homogeneous nucleation, irreversible elongation and fibril fragmentation [Fig. 1(c), I–III], these equations are nonlinear, but various approximations have been successfully employed to obtain their full time-dependent solution [7,12–14]. The stochastic version of this model, where the number of each species n_i is allowed to fluctuate, is, however, analytically intractable, although it can be simulated numerically as we have done in Fig. 1(b).

To obtain an analytic prediction for the lag-time distribution, we coarse-grain the model, while retaining the key processes of nucleation, elongation, and fragmentation. Rather than tracking the full distribution of fibril lengths, we track only the total number of fibrils $n = \sum_i n_i$ and the number of monomers in aggregates $m = \sum_i i n_i$ but treat them as discrete random variables, which can fluctuate due to intrinsic noise. This coarse-graining, which amounts essentially to summing over fibril lengths in the full model [25], results in the following set of possible transitions between states n, m of the system:

$$n, m \to n+1, m+n_c \quad \text{at rate } \alpha[c(t)]/\epsilon, \tag{1a}$$

$$n, m \to n, m+1 \quad \text{at rate } 2k_+ c(t) n, \tag{1b}$$

$$n, m \to n+1, m \quad \text{at rate } k_f m. \tag{1c}$$

Primary nucleation is modeled by Eq. (1a) as a one-step process in which a new filament (called a nucleus) is created instantaneously from n_c free monomers at rate $\alpha[c(t)]/\epsilon$. The rate α is assumed to depend on the molar concentration of free monomers $c(t)$ and $\epsilon = 1/(VN_A)$, where V is the volume and N_A is Avogadro's constant. Transition (1b) represents filament growth by monomer addition at rate $2k_+c(t)$; the factor of 2 accounts for the fact that filaments can grow at both ends. Transition (1c) represents fragmentation; this amounts to an autocatalytic creation of new fibrils from existing ones at rate k_f; the probability that any given fibril breaks is assumed to be proportional to its length. Although this assumption is somewhat simplistic [26], we have also studied a model where fibrils break more frequently at their ends [27]. This latter model, which is presented in the Supplemental Material [16], also obeys detailed balance by including backward reactions such as rejoining of fragmented fibrils and loss of monomers at fibril ends [28,29]; however, none of these changes were found to affect the early-stage aggregation phenomena studied here. In the rest of the Letter, we will further simplify the model by neglecting monomer depletion, which amounts to approximating the free monomer concentration $c(t)$ by c_{tot}; this has little effect on the lag phase.

The probability distribution $P_{n,m}(t)$ for a given n and m obeys the following master equation:

$$\frac{d}{dt}P_{n,m} = (\alpha/\epsilon)P_{n-1,m-n_c} + \mu n P_{n,m-1} + \lambda m P_{n-1,m} - (\alpha/\epsilon + \mu n + \lambda m)P_{n,m}, \tag{2}$$

where $\alpha \equiv \alpha(c_{\text{tot}})$, $\mu \equiv 2k_+ c_{\text{tot}}$, and $\lambda \equiv k_f$. Starting with an initial condition $P_{n,m}(0) = \delta_{n,n_0}\delta_{m,m_0}$, we aim to solve for $P_{n,m}(t)$ and then to find the probability distribution for the lag time, i.e., for the time needed for the number of aggregated monomers m to reach some predefined threshold m_T, which we define as 10% of the total number of monomers (which is given by c_{tot}/ϵ, assuming that c_{tot} is measured in moles per unit volume).

Analytic solution for the probability distribution $P_{n,m}$.—In order to obtain an analytic solution, we replace the master equation (2) with a corresponding Fokker-Planck equation via the linear noise approximation (LNA), also known as van Kampen's system size expansion [30,31]. The LNA assumes that n and m can be decomposed into deterministic and fluctuating parts:

$$n = N_A V \phi(t) + \sqrt{N_A V} x_1, \tag{3a}$$

$$m = N_A V \psi(t) + \sqrt{N_A V} x_2, \tag{3b}$$

where the fluctuating parts x_1 and x_2 are scaled by $\sqrt{N_A V}$ and are assumed to be small compared to the deterministic terms. The deterministic parts $\phi(t)$ and $\psi(t)$, expressed in units of concentration (here moles per unit volume), solve the following differential equations:

$$\frac{d\phi}{dt} = \lambda\psi + \alpha, \qquad \phi(0) = \epsilon n_0 \equiv \phi_0, \tag{4a}$$

$$\frac{d\psi}{dt} = \mu\phi + \alpha n_c, \qquad \psi(0) = \epsilon m_0 \equiv \psi_0. \tag{4b}$$

Equations (4a) and (4b) may be solved to yield

$$\phi(t) = \sqrt{\frac{\lambda}{\mu}}\Psi_0 \sinh(\tau) + \Phi_0 \cosh(\tau) - \frac{\alpha n_c}{\mu}, \tag{5a}$$

$$\psi(t) = \sqrt{\frac{\mu}{\lambda}}\Phi_0 \sinh(\tau) + \Psi_0 \cosh(\tau) - \frac{\alpha}{\lambda}, \tag{5b}$$

where we have adopted the following notation: $\tau = \sqrt{\mu\lambda}t$, $\Phi_0 = \phi_0 + \alpha n_c/\mu$, and $\Psi_0 = \psi_0 + \alpha/\lambda$. Equations (5a) and (5b) describe the time evolution of the mean concentrations of fibrils and aggregated protein, respectively, at early times. Solving $\psi(T) = m_T\epsilon$, where m_T is the threshold concentration, yields the mean lag time T:

$$T = \frac{1}{\sqrt{\mu\lambda}} \ln \frac{D + \sqrt{D^2 - \Psi_0^2 + (\mu/\lambda)\Phi_0^2}}{\Psi_0 + \sqrt{\mu/\lambda}\Phi_0}, \tag{6}$$

where $D = \alpha/\lambda + m_T\epsilon$. Equation (6) is a good approximation to the lag time reported in Ref. [7] and predicts the same $T \propto (k_+ c_{\mathrm{tot}} k_f)^{-1/2}$ scaling.

To determine the effects of intrinsic noise, we now turn to the fluctuating parts x_1 and x_2, which are governed by the following Fokker-Planck equation for the probability density $P(x_1, x_2, t)$:

$$\frac{\partial P}{\partial t} = -\sum_i \frac{\partial}{\partial x_i}(A_i P) + \frac{1}{2}\sum_{i,j}\frac{\partial^2}{\partial x_i x_j}(B_{ij}P), \tag{7}$$

where we assumed that $P(x_1, x_2, 0) = \delta(x_1)\delta(x_2)$. The drift vector $\vec{A}$ and the diffusion matrix B are given, respectively, by

$$\vec{A} = \begin{pmatrix} \lambda x_2 \\ \mu x_1 \end{pmatrix}, \qquad B = \begin{pmatrix} \lambda\psi + \alpha & \alpha n_c \\ \alpha n_c & \mu\phi + \alpha n_c^2 \end{pmatrix}. \tag{8}$$

Equation (7) describes a two-variable (time-dependent) Ornstein-Uhlenbeck process which can be solved by standard techniques [31] and yields a bivariate Gaussian distribution with zero mean and time-dependent covariance matrix $\Sigma_{ij} = \langle x_i x_j \rangle$. To calculate the lag-time distribution we only need to know $\Sigma_{22} = \langle x_2(t)^2 \rangle$; the time dependence of the other matrix elements can be found in the Supplemental Material [16].

Lag-time distribution.—Building on these results, we now obtain an analytic expression for the lag-time distribution $L(t)$. This is essentially a first-passage time problem; to calculate $L(t)$, we look for all events such that m has *just* exceeded m_T at a time t, given that it will exceed m_T eventually:

$$L(t) = \frac{\frac{d}{dt}\mathrm{Prob}[m > m_T, t]}{\mathrm{Prob}[m > m_T, t \to \infty]}. \tag{9}$$

The probability $\mathrm{Prob}[m > m_T, t]$ can easily be calculated by integrating $P(x_1, x_2, t)$ and reads

$$\mathrm{Prob}[m > m_T, t] = \frac{1}{2}\mathrm{erfc}\left(\frac{m_T\epsilon - \psi(t)}{\sqrt{2\epsilon\langle x_2(t)^2\rangle}}\right). \tag{10}$$

A lengthy but straightforward calculation for $\langle x_2(t)^2 \rangle$ gives

$$\begin{aligned}\langle x_2(t)^2\rangle = {} & \cosh(2\tau)\left[\frac{1}{6}\left(\Phi_0\frac{\mu}{\lambda} + \Psi_0\right) + \frac{\alpha n_c}{2\lambda}\right] \\ & + \sinh(2\tau)\sqrt{\frac{\mu}{\lambda}}\left[\frac{\Phi_0 + \Psi_0}{3} + \frac{\alpha n_c(n_c - 1)}{4\mu}\right] \\ & + \frac{\cosh\tau}{3\lambda}(\lambda\Psi_0 - 2\mu\Phi_0) + \frac{\sinh\tau}{3}\sqrt{\frac{\mu}{\lambda}}(\Phi_0 - 2\Psi_0) \\ & - \frac{\alpha n_c(n_c - 1)t}{2} + \frac{1}{2}\left(\Phi_0\sqrt{\frac{\mu}{\lambda}} - \Psi_0 - \frac{\alpha n_c}{\lambda}\right).\end{aligned} \tag{11}$$

It now proves useful to introduce a new variable $r(t)$:

$$r(t) = \frac{\psi(t) - m_T\epsilon}{\sqrt{\epsilon\langle x_2(t)^2\rangle}}, \tag{12}$$

which measures the deviation of the mean fibril concentration $[\psi(t)]$ from the threshold $(m_T\epsilon)$, scaled by the root mean square of $m\epsilon - \psi(t)$. Using this variable, we combine expressions (9)–(11) to give our central result: an analytical expression for the lag-time distribution in the linear noise approximation of the master equation (2), which takes the form of a Gaussian in r in the range $-\infty < r < r(\infty)$ [32]:

$$L(t)dt = \frac{dr/dt}{\sqrt{2\pi}Z}e^{-[r(t)^2/2]}dt = \frac{1}{\sqrt{2\pi}Z}e^{-(r^2/2)}dr, \tag{13}$$

where $Z = \mathrm{erfc}[-r(\infty)/2]$. Importantly, Eq. (13) allows us to easily calculate moments of the lag-time distribution. For example, to calculate the mean lag time $\langle t \rangle$ and its standard deviation σ, we express t and t^2 as functions of r and perform a Taylor expansion around $r = 0$ (see the Supplemental Material [16] for details). This gives

$$\langle t \rangle \approx T \quad \text{and} \quad \sigma \approx \frac{\sqrt{\epsilon\langle x_2^2(T)\rangle}}{\mu\phi(T) + \alpha n_c}. \tag{14}$$

For most proteins, the fragmentation rate $\lambda \equiv k_f$ is much smaller than the net fibril elongation rate $\mu \equiv 2k_+ c_{\mathrm{tot}}$; i.e., $\lambda \ll \mu$. If we also assume that no fibrils are present at time $t = 0$ ($\phi_0 = \psi_0 = 0$), we can write a simpler expression for the standard deviation of the lag time:

$$\sigma = \frac{(2/3)^{1/2}}{(\mu\lambda)^{1/4}(\alpha N_A V)^{1/2}}. \tag{15}$$

Remarkably, Eq. (15) implies that the lag-time variance scales in a simple way with the model parameters. Like the mean lag time, the variance is predicted to scale as $\sqrt{\mu\lambda} \sim \sqrt{k_f k_+}$. Interestingly, however, the mean and

variance of the lag time may show different dependencies on the protein concentration $c_{\rm tot}$; while the mean scales as $c_{\rm tot}^{-1/2}$, in expression (15) for the variance this factor (which arises from μ) is multiplied by an additional factor due to the $c_{\rm tot}$-dependent nucleation rate α; the scaling of this factor depends on the nucleus size n_c.

It is important to note that results (13)–(15) hold only in the regime dominated by growth, where fluctuations in n and m are much smaller than their averages, *for all times*. In contrast, for slow nucleation rates, a significant portion of the lag time is spent waiting for the first nucleus to be spontaneously created, which is a fluctuation-driven process. We take this into account by convolving $L(t-t')$ with the waiting time distribution $(\alpha/\epsilon)\exp(-\alpha t'/\epsilon)$ for the primary nucleation event, to give

$$L_1(t) = (\alpha/\epsilon)\int_0^t dt' e^{-(\alpha/\epsilon)t'} L(t-t'), \qquad (16)$$

where in the expression for $L(t-t')$ we set $\phi(t') = \epsilon$ and $\psi(t') = n_c\epsilon$ (i.e., assume one fibril of size n_c at time t'). Figure 2 shows that the lag-time distributions predicted by Eqs. (13) and (16) are in good agreement with the results of stochastic simulations of the full model (which takes into account fibril lengths), for several values of the primary nucleation rate α. For relatively fast nucleation rates, our "bare" LNA prediction $L(t)$ [Eq. (13)] is sufficient (main plots in Fig. 2); for slower nucleation rates (inset in Fig. 2), Eq. (16) should be used instead (inset in Fig. 2).

For slow nucleation rates, we can separate the contributions of primary nucleation and autocatalytic growth to the lag-time variance in a simple way. Assuming $L(t)$ can be replaced by a Gaussian in t, we can use (14) to compute the integral in (16) in a closed form which reveals that $L_1(t)$ has mean T_1 and standard deviation σ_1 given by

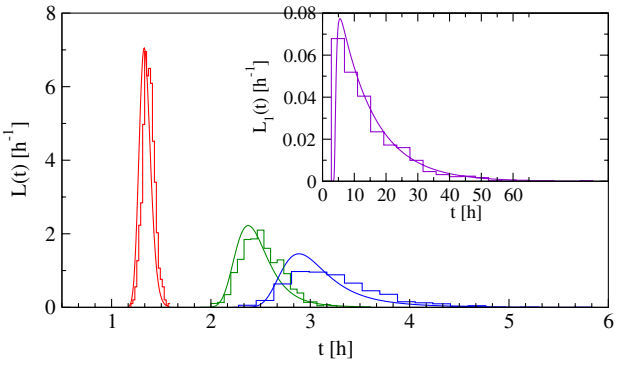

FIG. 2 (color online). The lag-time distribution $L(t)$ for several values of α, compared to that obtained by running 1000 independent kinetic Monte Carlo simulations of the full stochastic model (in which individual fibril lengths are resolved) [17]. From left to right: $\alpha = 50$ (full line), 5 (dashed line), and 1.5 (dot-dashed line), all in units of 10^{-15} mol/(ls). Inset: $L_1(t)$ (dashed line) compared to simulations for $\alpha = 5\times10^{-17}$ mol/(ls). The other parameters are $V = 830$ fl, $M_T = 10\%$ of $c_{\rm tot}$, $c_{\rm tot} = 100$ μmol/l, $n_c = 2$, $k_+ = 5\times10^4$ 1/(mol s), and $k_f = 3\times10^{-8}$ s^{-1}.

$$T_1 = \frac{\epsilon}{\alpha} + T, \qquad \sigma_1 = \sqrt{\left(\frac{\epsilon}{\alpha}\right)^2 + \sigma^2}. \qquad (17)$$

Thus the lag-time variance is given by a simple sum of the variance of the exponential waiting time distribution for the primary nucleation event and the contribution from autocatalytic growth, given by Eq. (15).

Comparison with experimental results for bovine insulin.—So far, the only available experimental data on amyloid fibril nucleation in small volumes is that of Knowles *et al.*, who tracked the fibrillation of bovine insulin in 52 microdroplets of volumes in the range 10–300 pl, using thioflavin T fluorescence [18]. Figure 3(a) shows the resulting lag times (red dots) as a function of droplet volume, compared to our theoretical prediction; the green line shows the mean lag time T_1, and the error bars show the standard deviation σ_1, from Eq. (17). No fitting parameters were used in this plot; rather the parameters k_+, k_f, and α were taken directly from

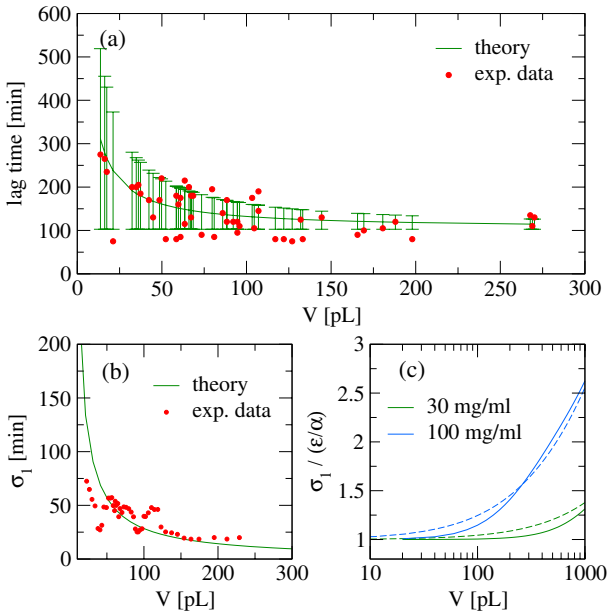

FIG. 3 (color online). (a) Volume dependence of the lag time for the aggregation of bovine insulin in microdroplets of varying volume (red dots) [18], compared to the mean (solid line) and the standard deviation (error bars) from Eq. (17), using the following values, which were obtained from Ref. [18]: $T = 104$ min, $\alpha = 1/(1.7\times10^{-7}N_A)$ mol/(l s), $k_+ = 8.9\times10^4$ 1/(mol s), $k_f = 2\times10^{-8}$ s^{-1}, and assuming $n_c = 2$. (b) Corresponding volume dependence of the standard deviation; the red dots are six-point moving standard deviations from the experimental data, and the solid green line is from Eq. (17). (c) Theoretical predictions for the standard deviation σ_1 as a function of volume, relative to ϵ/α. The green (lower) lines correspond to the protein concentration 30 mg/ml used in Ref. [18], while the blue (upper) lines are for a higher protein concentration, 100 mg/ml, assuming that $\alpha(c_{\rm tot}) \propto c_{\rm tot}^{n_c}$ [7]. In both cases, the dashed lines correspond to (17), while the solid lines are calculated numerically from (16).

the measurements of Ref. [18] (see also [33]). While there are not enough experimental data points to plot a lag-time distribution for any given volume, Fig. 3(a) shows that the variability observed in the experiments is consistent with our theory. This is further evidenced in Fig. 3(b), where we plot directly the volume dependence of the standard deviation.

We can also use our result, Eq. (17), to explore the relative contributions of primary nucleation and autocatalytic growth to the lag-time variability. Figure 3(c) shows our theoretical prediction for the standard deviation σ_1, relative to that for primary nucleation only, ϵ/α. The relative contribution of autocatalytic growth increases strongly as the volume increases (although the total variability decreases with V). For the protein concentration of 30 mg/ml used in Ref. [18], primary nucleation is the main contributor. However, for higher protein concentrations, we predict that autocatalytic variability becomes significant even at smaller volumes, on the scale of a human cell.

Conclusion.—We have presented an analytic expression for the lag-time distribution, for a stochastic model of autocatalytic protein self-assembly which includes nucleation, elongation, and fragmentation. Our solution provides simple scaling relations for the contributions to lag-time variability due to primary nucleation and autocatalysis, both of which can be significant under realistic conditions. The implications of molecular noise for variability in clinical outcomes between individuals, as well as the possible connection to variability between replicates in large-volume experiments [34], present interesting and important directions for future work.

We thank Tuomas Knowles for kindly providing the data from Ref. [18] which we have used in Fig. 3 and Line Jourdain for discussions in the early stages of this work. K. Eden and J. Szavits-Nossan contributed equally to this work. This research was supported by EPSRC under Grant No. EP/J007404/1. K. E. was supported by an EPSRC DTA studentship, and R. J. A. was supported by a Royal Society University Research Fellowship.

*jszavits@staffmail.ed.ac.uk
†k.eden@ed.ac.uk
‡rallen2@staffmail.ed.ac.uk

[1] F. Chiti and C. M. Dobson, Annu. Rev. Biochem. **75**, 333 (2006).
[2] J. E. Gillam and C. E. MacPhee, J. Phys. Condens. Matter **25**, 373101 (2013).
[3] A. Wegner, Nature (London) **296**, 266 (1982).
[4] A. Wegner and P. Savko, Biochemistry **21**, 1909 (1982).
[5] F. Ferrone, Methods Enzymol. **309**, 256 (1999).
[6] M. Tanaka, S. R. Collins, B. H. Toyama, and J. S. Weissman, Nature (London) **442**, 585 (2006).
[7] 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).
[8] F. A. Ferrone, J. Hofrichter, H. R. Sunshine, and W. A. Eaton, Biophys. J. **32**, 361 (1980).
[9] M. F. Bishop and F. A. Ferrone, Biophys. J. **46**, 631 (1984).
[10] A. M. Ruschak and A. D. Miranker, Proc. Natl. Acad. Sci. U.S.A. **104**, 12341 (2007).
[11] L. Jean, C. F. Lee, C. Lee, M. Shaw, and D. J. Vaux, FASEB J. **24**, 309 (2010).
[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] J. P. Bernacki and R. M. Murphy, Biophys. J. **96**, 2871 (2009).
[16] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.113.098101 for experimental setup and details of the full stochastic model.
[17] R. J. Morris, K. Eden, R. Yarwood, L. Jourdain, R. J. Allen, and C. E. MacPhee, Nat. Commun. **4**, 1891 (2013).
[18] T. P. J. Knowles, D. A. White, A. R. Abate, J. J. Agresti, S. I. A. Cohen, R. A. Sperling, E. J. De Genst, C. M. Dobson, and D. A. Weitz, Proc. Natl. Acad. Sci. U.S.A. **108**, 14746 (2011).
[19] F. A. Ferrone, J. Hofrichter, and W. A. Eaton, J. Mol. Biol. **183**, 591 (1985).
[20] F. A. Ferrone, J. Hofrichter, and W. A. Eaton, J. Mol. Biol. **183**, 611 (1985).
[21] J. Hofrichter, J. Mol. Biol. **189**, 553 (1986).
[22] A. Szabo, J. Mol. Biol. **199**, 539 (1988).
[23] M. R. D'Orsogna, G. Lakatos, and T. Chou, J. Chem. Phys. **136**, 084110 (2012).
[24] Note that both the deterministic models to which we refer in this Letter, and our stochastic model, assume that the system is "well mixed," i.e., spatially homogeneous.
[25] To derive the coarse-grained model from the full one including fragmentation, we have to ignore the fact that any fibril created by fragmentation that is smaller than n_c is considered to be unstable and thus dissolves, adding n_c free monomers to the solution.
[26] C. F. Lee, Phys. Rev. E **80**, 031134 (2009).
[27] J. Paturej, A. Milchev, V. G. Rostiashvili, and T. A. Vilgis, J. Chem. Phys. **134**, 224901 (2011).
[28] L. Hong and W.-A. Yong, Biophys. J. **104**, 533 (2013).
[29] J. S. Schreck and J.-M. Yuan, J. Phys. Chem. B **117**, 6574 (2013).
[30] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry* (Elsevier, Amsterdam, 2007).
[31] C. Gardiner, *Stochastic Methods* (Springer-Verlag, Berlin, 2009).
[32] The variable $r(t)$ has the properties that $r(0) = -\infty$ and $r(T) = 0$; also, $r(\infty)$ is finite and can be calculated by noting that for large times $\psi(t) \approx A \exp(\sqrt{\mu\lambda}t)$ and $\langle x_2^2(t)\rangle \approx C \exp(2\sqrt{\mu\lambda}t)$ yielding $r(\infty) = A/\sqrt{\varepsilon C}$.
[33] In Ref. [18], k_+ and k_f were calculated from the speed and width of the aggregation front in the droplets; we also assumed, as in previous work, that $n_c = 2$ and $\phi(0) = \psi(0) = 0$ [7,17].
[34] W.-F. Xue, S. W. Homans, and S. E. Radford, Proc. Natl. Acad. Sci. U.S.A. **105**, 8926 (2008).