P H Y S I C A L   R E V I E W   L E T T E R S

Fluctuations in the Kinetics of Linear Protein Self-Assembly

Thomas C. T. Michaels,[1] Alexander J. Dear,[1] Julius B. Kirkegaard,[1,3] Kadi L. Saar,[1]
David A. Weitz,[2] and Tuomas P. J. Knowles[1,*]

[1]*Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom*
[2]*Department of Physics and School of Engineering and Applied Sciences,*
Harvard University, Cambridge, Massachusetts 02138, USA
[3]*Department of Applied Mathematics and Theoretical Physics, Centre for Mathematical Sciences, University of Cambridge,*
Wilberforce Road, Cambridge CB3 0WA, United Kingdom
(Received 18 April 2015; published 21 June 2016)

Biological systems are characterized by compartmentalization from the subcellular to the tissue level, and thus reactions in small volumes are ubiquitous in living systems. Under such conditions, statistical number fluctuations, which are commonly negligible in bulk reactions, can become dominant and lead to stochastic behavior. We present here a stochastic model of protein filament formation in small volumes. We show that two principal regimes emerge for the system behavior, a small fluctuation regime close to bulk behavior and a large fluctuation regime characterized by single rare events. Our analysis shows that in both regimes the reaction lag-time scales inversely with the system volume, unlike in bulk. Finally, we use our stochastic model to connect data from small-volume microdroplet experiments of amyloid formation to bulk aggregation rates, and show that digital analysis of an ensemble of protein aggregation reactions taking place under microconfinement provides an accurate measure of the rate of primary nucleation of protein aggregates, a process that has been challenging to quantify from conventional bulk experiments.

The formation of protein filaments is a process of central importance for both normal [1,2] and aberrant biology [3,4], as well as for the development of novel materials for nanotechnology [5–8]. The fundamental kinetic equations describing such processes in bulk are well established in the literature and have been studied extensively over the past 50 years [1,2,9–16]. These descriptions rely on the mean-field assumption [9] and therefore neglect statistical number fluctuations. Yet, protein aggregation processes in typical cellular environments (fl–pl) involve significantly smaller numbers of molecules than conventional bulk experiments and, thus, stochastic variability is expected to play an important role [14,15,17–19]. Moreover, recent experimental advances in microdroplet techniques [18,20] allow volumes in the picoliter range (comparable to intracellular volumes) to be probed for synthetic systems, creating the need for a general theoretical framework capable of describing protein filament assembly in small volumes.

Current theoretical descriptions of protein filament formation in small volumes focus on systems characterized by aggregate propagation from a single primary nucleation event [21]. A key question, however, is the nature of the full fluctuation behavior bridging the gap between the limit of classical nucleation theory and bulk behavior. In this Letter, we study stochastic effects in filamentous growth processes with secondary pathways [11–15,22–29] and derive closed-form expressions for the distribution of lag times. Our theoretical framework describes currently available micro-droplet experimental data that are characterized by

aggregate proliferation from multiple nucleation sites. Moreover, our results suggest a powerful method for characterizing the primary nucleation step, which is typically difficult to access from current bulk methods. We demonstrate the power of this approach by obtaining a value for the rate of primary nucleation for bovine insulin aggregation that is significantly better constrained than results obtained from analysis of bulk data.

Stochastic moment equations.—We consider a system of volume V containing a mixture of fibrillar aggregates and monomeric proteins in solution. Since we are interested in the early stages of the assembly process, we assume a constant chemical potential for the available soluble precursor proteins [10–12]. We describe the state of the system by a vector (n, m), where n is the number of fibrils and m is the number of monomers incorporated into aggregates, parameters that relate directly to experimental observables [30]. The probability distribution function (PDF) $P(n, m, t)$ of states (n, m) evolves according to a master equation [21],

$$\frac{\partial P(n, m, t)}{\partial t} = \alpha_1 P(n - 1, m - n_c, t) - \alpha_1 P(n, m, t)$$
$$+ \mu n P(n, m - 1, t) - \mu n P(n, m, t)$$
$$+ \alpha_2 (m - n_2) P(n - 1, m - n_2, t)$$
$$- \alpha_2 m P(n, m, t), \tag{1}$$

where α_1, μ, and α_2 are the transition rates (units s^{-1}) for primary nucleation, filament elongation, and secondary mechanisms, respectively [Fig. 1(a)]. This description

explicitly considers time variations of the PDF in terms of probability fluxes: the positive expressions represent gain terms that account for system transitions into state (n, m), whereas the negative terms describe losses from transitions from (n, m) into other states. The terms on the first line of Eq. (1) describe the initial primary nucleation step as the spontaneous formation of growth-competent aggregates from the interaction of n_c monomers. The increase of aggregate mass through elongation is described by the terms on the second line of Eq. (1). Secondary processes are captured by the third and fourth lines of Eq. (1) and cover several options, including breakage $(n_2 = 0)$ [12,22,24,25], lateral branching $(n_2 = 1)$ [11,26,27], and surface-catalyzed secondary nucleation $(n_2 \geq 2)$ [15,28,29]. Note that in general monomer dissociation from filament ends and rejoining of fibrils are necessary components to ensure microscopic reversibility [31]. The assumption of vanishing rates of monomer dissociation and polymer rejoining employed here, however, is justified as these processes do not significantly affect the early stages of the reaction [32].

The transition rates entering Eq. (1) can be related to the total concentration of proteins, m_{tot}, and the bulk rate parameters k_n, k_+, k_2 for primary nucleation, elongation, and secondary pathways, respectively, by requiring the rate equations for the averages $\langle n \rangle$ and $\langle m \rangle$ to be in agreement with existing early-time deterministic models [9–12] (see the Supplemental Material [34]). This condition yields $\alpha_1 = k_n m_{\text{tot}}^{n_c} N_A V$, $\mu = 2k_+ m_{\text{tot}}$, and $\alpha_2 = k_2 m_{\text{tot}}^{n_2}$, where N_A is the Avogadro number [21]. Importantly, the transition rate for primary nucleation, α_1, explicitly depends on the system size, V, while the parameters μ and α_2 describing autocatalytic growth are determined only by the associated bulk quantities. We expect, therefore, that reducing system size leads to a transition from a situation when the kinetics are controlled by autocatalytic growth to a situation when the fibrillization reaction is limited by primary nucleation. Thus, primary nucleation events becoming infrequent is at the origin of the stochastic behavior of filamentous growth processes in small volumes.

Analytical solution for the PDF.—The master equation (1) yields differential equations for the principal moments of the PDF through summation over system compositions. Solving for moments for times greater than κ^{-1}, with $\kappa = \sqrt{\mu \alpha_2} = \sqrt{2k_+ k_2 m_{\text{tot}}^{n_2+1}}$ being the characteristic time scale for aggregate proliferation [10–12], but still sufficiently short for the constant monomer approximation to be valid, shows that the Pearson correlation coefficient for n and m, $\rho_{n,m} = [\langle nm \rangle - \langle n \rangle \langle m \rangle][(\langle n^2 \rangle - \langle n \rangle^2)(\langle m^2 \rangle - \langle m \rangle^2)]^{-1/2}$, equals 1 in this limit (see the Supplemental Material [34]). This result implies the existence of a linear correlation in this regime between the random variables n and m, whereby the constant of proportionality is $m = (\kappa/\alpha_2)n$, $t \gg \kappa^{-1}$. We can directly test this prediction from numerical

realizations of Eq. (1) generated using the Gillespie algorithm [37] which reveal that n and m are indeed linearly correlated even before aggregation is detected (see the Supplemental Material [34]). The linear correlation between n and m allows recasting the master equation (1) into an equivalent one with a single variable,

$$\frac{\partial P(n,t)}{\partial t} = \alpha_1 P(n-1,t) - \alpha_1 P(n,t) \\ + \kappa(n-1)P(n-1,t) - \kappa n P(n,t). \quad (2)$$

Interestingly, Eq. (2) is analogous to the master equation of bacterial growth [33], whereby bacteria are constantly introduced into the system at rate α_1 and multiply with rate κ. This analogy, first hypothesized by Szabo [33], is a statement of the fact that for times bigger than κ^{-1} the average length of aggregates is constant. To define appropriate initial conditions for Eq. (2), we match the first moments of the PDFs of Eqs. (1) and (2) for times $t \gg \kappa^{-1}$, yielding $\langle n \rangle(t = \log 2/\kappa) = 0$. This condition translates necessarily into $P(n, t = \log 2/\kappa) = \delta_{n,0}$, therefore also ensuring that all higher moments of the PDFs of Eqs. (1) and (2) match for $t \gg \kappa^{-1}$ at leading order. The exact solution of Eq. (2) subject to the above initial conditions is (see the Supplemental Material [34])

$$P(n,t) = \frac{2^{\alpha_1/\kappa} \Gamma(n + \frac{\alpha_1}{\kappa})}{\Gamma(n+1)\Gamma(\frac{\alpha_1}{\kappa})} e^{-(\alpha_1+\kappa n)t}(e^{\kappa t} - 2)^n, \quad (3)$$

where $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$ is the Gamma function. The PDF in terms of the variable m is obtained by implementing
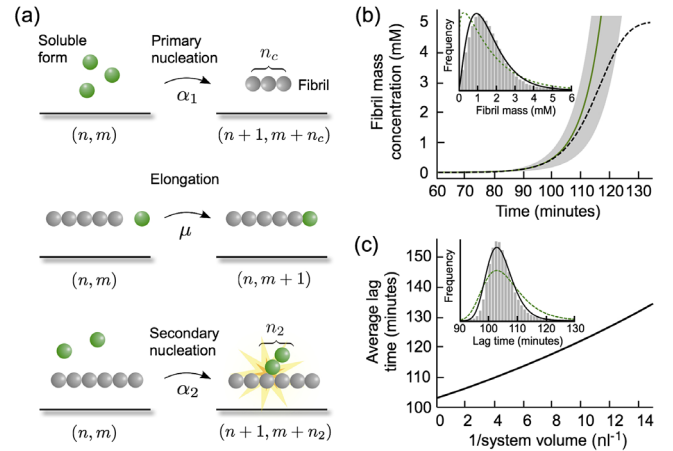

FIG. 1. (a) Different transitions in stochastic protein aggregation. (b) Time evolution of average mass concentration and 68% confidence bands. Inset: PDF for m at $t = 106$ minutes predicted by Eq. (3) (solid line) is compared to numerics. The dashed line is the solution of [33]. (c) The scaling behavior of the average lag time with system volume predicted by Eq. (5). Inset: PDF of lag times for $V = 1$ nl predicted by Eq. (4) (solid line) is compared to numerics. The dashed line is from [33]. Calculation parameters: $k_n = 4 \times 10^{-13}~M^{-1}~s^{-1}$, $n_c = 2$, $n_2 = 0$, $k_2 = 2.5 \times 10^{-8}~s^{-1}$, $k_+ = 2.5 \times 10^4~M^{-1}~s^{-1}$, $m_{\text{tot}} = 5$ mM, and $V = 1$ nl.

the correlation between n and m in Eq. (3). Figure 1(b) shows that Eq. (3) is in agreement with numerical realizations of Eq. (1).

Lag times.—A common qualitative feature of filamentous growth processes is the observation of a lag phase before aggregation can be detected. A commonly used measure of this polymerization delay is the lag time, τ, defined as the time at which the aggregate mass concentration $m(t)/(N_A V)$ reaches an arbitrarily chosen concentration threshold M_{th}. Although it is often the case that a halfway point for the reaction is taken, this may fall outside the realm over which our approximations are valid. Therefore a 10% extent or the experimental limit for aggregate detection [30] are simple choices for M_{th} more in keeping with our solution. Because τ is a random variable, the quantity of interest is the PDF of lag times, i.e., the probability $T(t)$ that τ equals t. According to the theory of first passage times [33], $T(t)$ is computed as $T(t) = -dQ(t)/dt$, where $Q(t)$ is the probability that at time t the process $m(t)/(N_A V)$ has not yet reached M_{th}. Using Eq. (3) we find

$$T(t) = \frac{\kappa 2^{\alpha_1/\kappa} \Gamma(n_{th} + \alpha_1/\kappa)}{\Gamma(n_{th})\Gamma(\alpha_1/\kappa)} e^{-[\alpha_1 + (n_{th}-1)\kappa]t}(e^{\kappa t} - 2)^{n_{th}-1}, \quad (4)$$

where $n_{th} = \alpha_2 N_A V M_{th}/\kappa$. The average lag time is obtained from Eq. (4) in analogy to [33] as:

$$\langle \tau \rangle = \frac{\log(2)}{\kappa} + \sum_{j=0}^{n_{th}-1} \frac{1}{\alpha_1 + j\kappa} \quad (5)$$

and the extent of fluctuations is

$$\sigma^2 = \langle \tau^2 \rangle - \langle \tau \rangle^2 = \sum_{j=0}^{n_{th}-1} \frac{1}{(\alpha_1 + j\kappa)^2}. \quad (6)$$

Limiting behavior of lag time in key regimes.— Inspection of Eqs. (5) and (6) reveal that the level of stochasticity in the system is controlled by the dimensionless parameter $\gamma = \kappa/\alpha_1$. Based on this parameter, we can distinguish three natural regimes of stochastic behavior: bulk ($\gamma = 0$), onset of stochasticity ($\gamma \to 0$), and single-event controlled ($\gamma \to \infty$). We now discuss how Eqs. (5) and (6) can be simplified in these regimes. In bulk ($\gamma = 0$), the sum in Eq. (5) is replaced by an integral and Eq. (5) is determined solely by the propagation time associated with the secondary nucleation chain reaction, τ_{bulk} (see the Supplemental Material [34]). Series approximation of Eq. (5) around $\gamma = 0$ allows us to explore the onset of stochasticity and reveals that $\langle \tau \rangle$ approaches the bulk value τ_{bulk} as $\langle \tau \rangle = \tau_{\text{bulk}} + c_n/(2V)$, where $c_n = 1/(k_n m_{\text{tot}}^{n_c} N_A)$ is the average time of forming nuclei in volume V (see the Supplemental Material [34]). This system size expansion shows, therefore, that in this regime, $\langle \tau \rangle$ approximatively decomposes into a sum of the deterministic lag time and a

term proportional to $1/V$. The extent of fluctuations in this regime is approximatively given by $\sigma^2 = c_n/(\kappa V)$ (see the Supplemental Material [34]). In the opposite limit of very small volumes or slow nucleation ($\gamma \to \infty$), the dominant contribution to Eq. (5) is $\langle \tau \rangle = c_n/V$ [21]. In this regime, V is small enough that eventually only a single nucleation event occurs ahead of the threshold being reached, at which point $\langle \tau \rangle$ is dominated by the waiting time for formation of a single nucleus, which scales inversely proportional to system volume as expected from classical nucleation theory [38]. Furthermore, the extent of fluctuations is now controlled by $\sigma^2 = (c_n/V)^2$ (see the Supplemental Material [34]). Equations (5) and (6) interpolate smoothly between these limiting regimes. In particular, across the entire range of system sizes $\langle \tau \rangle$ approximatively can be written as the sum of the deterministic lag time and a nucleation term proportional to $1/V$, whereby the constant of proportionality satisfies $c_n/2 \le d\langle \tau \rangle/d(1/V) \le c_n$. The transition between these two limiting regimes occurs approximatively at the critical volume $V_c = \kappa c_n$ at which $\gamma = 1$. This critical volume corresponds to the radius of convergence of the system size expansion of $\langle \tau \rangle$ around $\gamma = 0$ and, hence, marks the upper volume limit for the small fluctuation result to be accurate.

Stochastic analysis provides strong constraints for probing primary nucleation events.—Heretofore, rate constants for protein aggregation have been determined by carrying out kinetic experiments in bulk, in which the mass concentration of fibrils over time is measured [30,39], and by fitting such data to rate laws derived from deterministic master equations [39]. Typically, however, the rate constants characterizing the elementary processes of primary and secondary nucleation and growth occur as combinations, and thus it remains challenging to obtain accurate values for the rates of these processes from experimental data. In particular the process of primary nucleation has proven challenging to quantify, in part as the major experimental observables such as the lag-phase display only a weak logarithmic dependence on this parameter. Remarkably, however, our results suggest that a measurement of the volume dependence of the lag time allows the rate of primary nucleation to be determined directly from the slope of a plot of $\langle \tau \rangle$ versus $1/V$. This approach has several advantages over attempting to characterize primary nucleation from bulk polymerization fraction experiments: (i) While fitting of bulk experiments fixes only the combined rate parameter $k_+ k_n$, the analysis of stochastic data allows the rate of primary nucleation to be determined directly, without the necessity of estimating the elongation rate constant or the length of the aggregates from other experimental techniques, factors which are intrinsically sources of significant error. (ii) Microdroplet experiments can be tightly controlled, and lag-time experiments are digital in nature with the exact value of the threshold not entering the gradient and thus not contributing to error.
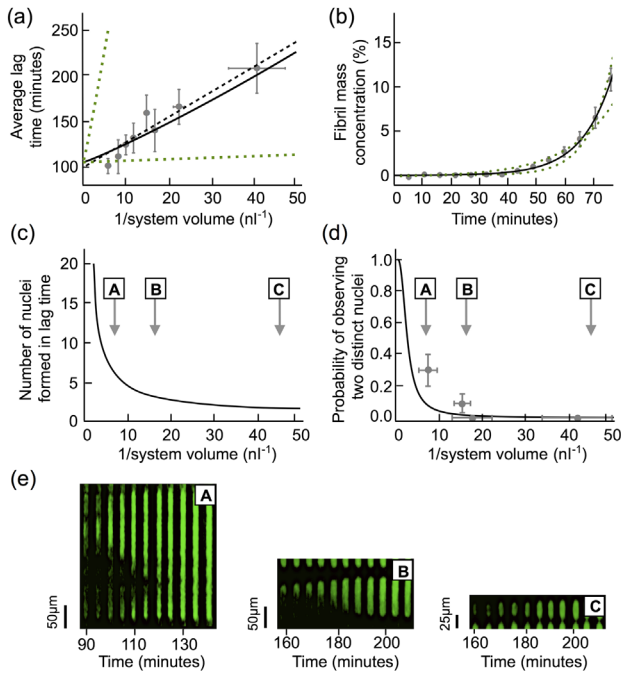
FIG. 2. (a) Analysis of small-volume experiments of bovine insulin fibrillization kinetics from Ref. [18]. Dashed line: best fit to $\langle\tau\rangle \propto c_n/V$ with $c_n^{-1} = 6 \times 10^6$ s^{-1} l^{-1}; solid line: prediction from Eq. (5); dotted lines: k_n is decreased and increased by an order of magnitude. $\langle\tau\rangle$ shows marked volume dependence despite the presence of multiple nucleation sites, as demonstrated by the plot of the number of nuclei formed on average during the mean lag time against $1/V$. (b) Kinetic analysis of insulin aggregation in bulk. Solid line: best-fit curve to initial exponential growth; dotted lines: k_n is decreased and increased by an order of magnitude relative to best fit. (c) Average number of nuclei formed in lag time. (d) Predicted probability to observe more than one nucleation event is compared with measurements from analysis of 80 microscopic droplet images. (e) Fluorescence microscopy images of representative microdroplets of volumes A, B, and C in (c) and (d).

(iii) In the presence of secondary mechanisms, the fitting of bulk data over the full time course is predominantly constrained by the autocatalytic processes rather than the primary nucleation step, contributing to uncertainties of several orders of magnitude for the determined nucleation rate [40]. Even if fitting is limited to the early times, the exponential form of the fitting equation allows for substantial leeway in the values of the fitting parameters, as numerous combinations of such parameters give rise to fairly similar-looking curves. For example, changing the best-fit value for k_n by an order of magnitude would not significantly affect the performance of the bulk analysis fit, but would give rise to a dramatically poorer fit of $\langle\tau\rangle$ versus $1/V$ in a linear relationship [Figs. 2(a)–2(b)].

Connecting small-volume experiments of bovine insulin fibrillization kinetics to bulk experiments.—We have applied this technique to analyze data from microdroplet experiments on bovine insulin [18] and obtained a value for

the rate of primary nucleation of $c_n^{-1} = (6 \pm 1) \times 10^6$ s^{-1} l^{-1} [Fig. 2(a)]. We then carried out bulk experiments of insulin aggregation (see the Supplemental Material [34]) and fitted the data to a standard deterministic models, in conjunction with our calculated value for the rate of primary nucleation in microdroplets [Fig. 2(b)]. The resultant calculated rate of elongation $\mu = 2 \times 10^5$ s^{-1} agrees with those reported in the literature, to within the levels of error expected from the method of calculation [41]; this shows that the stochastic analysis presented in this Letter allows small-volume behavior to be related to conventional bulk experiments. Moreover, the value of the nucleation rate constant is constrained to within better than an order of magnitude, a result that is very challenging to achieve with analysis of bulk data [Fig. 2(b)].

Linking average number of nucleation events with experimental observations.—Finally, we test our model by predicting the number of individual nuclei formed in the mean lag time using the extracted nucleation rate constant [Fig. 2(c)]. Individual nuclei can be counted in microdroplet experiments; however, the fluorescence signal due to fibril growth from previous nuclei is expected to obscure signals from subsequent nucleation events, and thus only some events are observed. In order to compare this prediction with experiments, we therefore devised a probabilistic model capable of quantifying this effect (see the Supplemental Material [34] for details). Applying this methodology to the analysis of 80 droplet images [Figs. 2(d)–2(e)], we see overall good agreement between the predicted and measured probabilities of observing more than one nucleation event, given the limitations of our measurement techniques. Thus, the system behavior results from multiple nucleation sites even though naive visual inspection would suggest single nucleation events.

Conclusions.—We have reported a theoretical study on stochastic effects in nucleated polymerization phenomena in small volumes. We have derived fully analytical results describing the distribution of lag times that allows linking the bulk parameters characterizing large-volume experiments with the statistical properties of polymerization curves in small volumes across the entire range of fluctuation behavior. From the analysis of experimental data, we have shown that small-volume microdroplet experiments of amyloid aggregation are typically characterized by multiple nucleation events. Moreover, our results provide a practical route towards an accurate determination of primary nucleation rates, which represent a key event in the transition of soluble proteins into their aggregated forms.

[*]tpjk2@cam.ac.uk

[1] F. Oosawa and S. Asakura, *Thermodynamics of the Polymerization of Proteins*, (Academic Press, Waltham, MA, 1975).

[2] F. Oosawa and M. Kasai, J. Mol. Biol. **4**, 10 (1962).

[3] C. M. Dobson, Nature (London) **426**, 884 (2003).

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

[5] E. Gazit, Chem. Soc. Rev. **36**, 1263 (2007).

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

[7] T. P. J. Knowles, T. Oppenheim, A. K. Buell, D. Y. Chirgadze, and M. E. Welland, Nat. Nanotechnol. **5**, 204 (2010).

[8] C. Li, S. Bolisetty, and R. Mezzenga, Adv. Mater. **25**, 3694 (2013).

[9] T. C. T. Michaels and T. P. J. Knowles, Am. J. Phys. **82**, 476 (2014).

[10] 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).

[11] M. F. Bishop and F. A. Ferrone, Biophys. J. **46**, 631 (1984).

[12] 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).

[13] 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).

[14] F. A. Ferrone, J. Hofrichter, and W. A. Eaton, Biophys. J. **32**, 361 (1980).

[15] F. A. Ferrone, J. Hofrichter, and W. A. Eaton, J. Mol. Biol. **183**, 591 (1985).

[16] T. C. T. Michaels, S. I. A. Cohen, M. Vendruscolo, C. M. Dobson, and T. P. J. Knowles, Phys. Rev. Lett. **116**, 038101 (2016).

[17] J. Hofrichter, J. Mol. Biol. **189**, 553 (1986).

[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] D. W. Colby, J. P. Cassady, G. C. Lin, V. M. Ingram, and K. D. Wittrup, Nat. Chem. Biol. **2**, 319 (2006).

[20] C. A. Stan, G. F. Schneider, S. S. Shevkoplyas, M. Hashimoto, M. Ibanescu, B. J. Wiley, and G. M. Whitesides, Lab Chip **9**, 2293 (2009).

[21] J. Szavits-Nossan, K. Eden, R. J. Morris, C. E. MacPhee, M. R. Evans, and R. J. Allen, Phys. Rev. Lett. **113**, 098101 (2014).

[22] M. Tanaka, S. R. Collins, B. H. Toyama, and J. S. Weissman, Nature (London) **442**, 585 (2006).

[23] W.-F. Xue, S. W. Homans, and S. E. Radford, Proc. Natl. Acad. Sci. U.S.A. **105**, 8926 (2008).

[24] A. Wegner and P. Savko, Biochemistry **21**, 1909 (1982).

[25] S. R. Collins, A. Douglass, R. D. Vale, and J. S. Weissman, PLoS Biol. **2**, e321 (2004).

[26] C. B. Andersen, H. Yagi, M. Manno, V. Martorana, T. Ban, G. Christiansen, D. E. Otzen, Y. Goto, and C. Rischel, Biophys. J. **96**, 1529 (2009).

[27] K. J. Amann and T. D. Pollard, Nat. Cell Biol. **3**, 306 (2001).

[28] A. M. Ruschak and A. D. Miranker, Proc. Natl. Acad. Sci. U.S.A. **104**, 12341 (2007).

[29] A. Cacciuto, S. Auer, and D. Frenkel, Nature (London) **428**, 404 (2004).

[30] E. Hellstrand, B. Boland, D. M. Walsh, and S. Linse, ACS Chem. Neurosci. **1**, 13 (2010).

[31] P. L. Krapivsky, S. Redner, and E. Ben-Naim, *A Kinetic View of Statistical Physics* (Cambridge University Press, Cambridge, Englandf, 2010).

[32] T. C. T. Michaels and T. P. J. Knowles, J. Chem. Phys. **140**, 214904 (2014).

[33] A. Szabo, J. Mol. Biol. **199**, 539 (1988).

[34] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.116.258103, which includes Refs. [35,36], for details on the mathematical derivations and for further descriptions of the simulation and experimental methods.

[35] D. T. Gillespie, J. Phys. Chem. **115**, 1716 (2001).

[36] S. I. A. Cohen, L. Rajah, C. H. Yoon, A. K. Buell, D. A. White, R. A. Sperling, M. Vendruscolo, E. M. Terentjev, C. M. Dobson, D. A. Weitz, and T. P. J Knowles, Phys. Rev. Lett. **112**, 098101 (2014).

[37] D. T. Gillespie, J. Phys. Chem. **81**, 2340 (1977).

[38] J. L. Barrat and J. P. Hansen, *Basic Concepts for Simple and Complex Liquids* (Cambridge University Press, Cambridge, UK, 2003).

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

[40] K. Eden, R. Morris, J. Gillam, C. E. MacPhee, and R. J. Allen, Biophys. J. **108**, 632 (2015).

[41] Note that the high concentration used for the bulk kinetic experiment is expected to lead to molecular crowding effects, which could cause further error in the calculated rate constants up to a few orders of magnitude [42,43].

[42] F. A. Ferrone and M. A. Rotter, J. Mol. Recognit. **17**, 497 (2004).

[43] F. A. Ferrone, J. Proteins Proteomics **3**, 157 (2012).