Self-assembly of protein amyloids: A competition between amorphous and ordered aggregation

Chiu Fan Lee (李超帆)*

Physics Department, Clarendon Laboratory, Oxford University, Parks Road, Oxford OX1 3PU, United Kingdom

(Received 9 October 2008; revised manuscript received 12 August 2009; published 30 September 2009)

Protein aggregation in the form of amyloid fibrils has important biological and technological implications. Although the self-assembly process is highly efficient, aggregates not in the fibrillar form would also occur and it is important to include these disordered species when discussing the thermodynamic equilibrium behavior of the system. Here, we initiate such a task by considering a mixture of monomeric proteins and the corresponding aggregates in the disordered form (micelles) and in the fibrillar form (amyloid fibrils). Starting with a model on the respective binding free energies for these species, we calculate their concentrations at thermal equilibrium. We then discuss how the incorporation of the disordered structure furthers our understanding on the various amyloid promoting factors observed empirically, and on the kinetics of fibrilization.

DOI: 10.1103/PhysRevE.80.031922 PACS number(s): 87.14.ef, 82.35.Pq, 46.25.Cc, 87.19.xh

I. INTRODUCTION

Amyloids are insoluble fibrous protein aggregations stabilized by a network of hydrogen bonds and hydrophobic interactions [1–4]. They are intimately related to many neurodegenerative diseases such as Alzheimer's disease, Parkinson's disease, and other prion diseases [5]. Better characterization of the various properties of amyloid fibrils is therefore of high importance for the understanding of the associated pathogenesis. More recently, viewing protein amyloid formation as a highly efficient self-assembly process, possible applications have also been proposed. For instance, amyloid fibrils were shown to possess great tensile strength [6,7] and complex phase behavior similar to liquid crystals [8,9], and have been employed as nanowire templates [10,11]. Given the high importance of protein amyloid in biology and potentially in technology, it is being studied intensively. In particular, much effort has been spent on investigating the amino-acid dependency on amyloid propensity [10,12–17]; the possibility of primary-sequence-based amyloid propensity predictions [18–23]; the mechanical properties of protein amyloid [7,24,25]; the kinetics of amyloid formation [26–34]; as well as the thermodynamical behaviors of the aggregation process [35–38].

Although the protein amyloid self-assembly process is highly efficient, aggregates not in the fibrillar form would also occur and it is important to include these disordered species when discussing the thermodynamic equilibrium behavior of the system. This motivates us to consider here a system consisting of a mixture of monomers, aggregates with a linearly ordered structure (fibrils) and aggregates with a disordered structure (a micellelike aggregate) (c.f. Fig. 1). Starting with a discussion on their respective binding free energies, we deduce the concentrations for the various species at thermal equilibrium, and consider the experimental implications of our investigation. In particular, we study the effect of temperature and pressure variations in the average fibrillar length. We then discuss how our work relates to the empirically observed variation in amyloid propensity with respect to the primary sequences of the proteins. Finally, we employ the formalism developed to study the kinetic process of aggregation.

The plan of the paper is as follows: In Sec. II, we introduce our model of a amyloid-forming self-assembly system. In Sec. III, we discuss the experimentally relevant predictions from our model. In Sec. IV, we consider how our findings relate to empirical observations on amyloid propensity. In Sec. V, we investigate the kinetic process of self-assembly from the perspective of our free-energy picture.

II. MODEL

In this work, we are primarily concerned with amyloid fibrilization of short peptides. Peptides interact via an array of interactions, such as hydrophobic interactions, hydrogen bonding, electrostatic interactions, etc. (for a review, see, e.g., [40,41]). Due to these interactions, aggregation may occur and we consider here two different types of aggregates: (i) linearly structured aggregates (amyloid fibrils) and (ii) disordered aggregates (micelles) (c.f. Fig. 1). For the micellar species, we assume that there is an optimal configuration consisting of M proteins, where M is in the order of tens [42]. For the fibrillar species, we assume that the only ordered structure is a two-tape structure, i.e., each fibril consists of stacking two cross-beta structures [c.f. Fig. 1(c)]. We note that amyloid fibrils can exhibit structural variations even when prepared under the same condition, and the precise structural details will be highly primary sequence dependent (see, e.g., [43]).

Now a note on terminology: we will call a free protein in solution a monomer, and a fibril consisting of i proteins a i-mer fibril. We will also denote from now on the numbers of monomers, micelles, and i-mer fibril in a solution of volume V by $N^{(a)}$, $N^{(b)}$, and $N_i^{(c)}$, respectively. In particular, if N denotes the total number of monomers, we have

$$N^{(a)} + MN^{(b)} + \sum_i iN_i^{(c)} = N. \tag{1}$$

Given the three different species: monomers, micelles and fibrils, we are interested in determining their respective vol-

*c.lee1@physics.ox.ac.uk

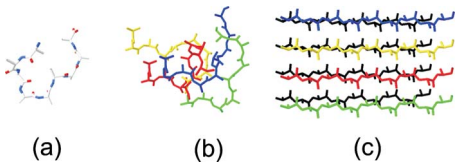

FIG. 1. (Color online) Schematic diagrams of the three species considered in this paper: (a) a monomeric protein in solution; (b) a micelle, or an amorphous aggregate; (c) an eight-monomer segment of an amyloid fibril consisting of two cross-beta structures (one cross-beta structure is colored, the other is black). The hydrogen bonds stabilize the beta sheets in the vertical direction (not shown in this figure). (Drawn with DEEPVIEW [39].)

ume fractions given an initial volume fraction C. In this work, we will always set the unit of volume to be the volume of one monomer. With this convention, $C=N/V$ where V is the volume of the system.

To calculate the relative abundances for the various species, we first need to obtain their respective species-specific binding free energy (BFE). Without loss of generality, we will set the monomeric BFE to zero, and denote the micellar BFE by γ, where

$$\gamma = -T\Delta s_b + \Delta\epsilon_b + p\Delta v_b. \tag{2}$$

In the above equation, Δs_a, $\Delta\epsilon_b$, and Δv_b are the entropic, binding energy, and volumic differences between the monomers and the micelles. In other words, Δs_b quantifies the free-energy contribution from the loss in configurational freedom due to the rigidity of the aggregates, $\Delta\epsilon_b$ quantifies the change in internal energy resulting from the various interprotein and intraprotein interactions, and Δv_b denotes the change in volume for a monomer as a result of being part of a larger aggregate.

For the fibrillar species, we will denote the BFE of an infinitely long fibril as

$$f_\infty = -T\Delta s_c + \Delta\epsilon_c + p\Delta v_c. \tag{3}$$

The terms in the right-hand side above have similar definitions as in the micellar case aforementioned. For a finite-size i-mer fibril, we will make the following assumption typically made in the study of linearly aggregating systems (e.g., see [44]):

$$f_i = \left(\frac{i-\xi}{i}\right) f_\infty, \tag{4}$$

where ξ ($\xi>0$) accounts for the boundary effect at the fibril's ends and is of order one [44]. For instance, it accounts for the loss of hydrophobic interactions at both ends of a fibril.

With the BFE defined, we can calculate the concentrations of the various species by finding the minimum of the total free energy of the system. We will start by writing the total partition function as [45]

$$Q_{\text{tot}} = \prod_i{}' \frac{(AV)^{N^{(a)}}(BV)^{N^{(b)}}(C_iV)^{N_i^{(c)}}}{N^{(a)}!\,N^{(b)}!\,N_i^{(c)}!}, \tag{5}$$

where, by the previous discussion on the BFE [46],

$$A = 1, \tag{6}$$

$$B = \exp(-M\gamma/k_BT), \tag{7}$$

$$C_j = \exp[-(i-\xi)f_\infty/k_BT]. \tag{8}$$

Note that the prime in the product in Eq. (5) denotes the restriction that the total number of peptides is conserved [c.f. Eq. (1)]. The distribution of the various species can now be obtained by determining the minimum of the total free-energy density,

$$\mathcal{F}_{\text{tot}} = -\frac{k_BT\ln Q_{\text{tot}}}{V}, \tag{9}$$

subject to the constraint shown in Eq. (1). This optimization problem can be solved by the Lagrange multiplier method and the results are (see, e.g., [47])

$$n^{(b)} = [n^{(a)}e^{-\gamma/k_BT}]^M \tag{10}$$

$$n_i^{(c)} = [n^{(a)}e^{-(i-\xi)f_\infty/ik_BT}]^i. \tag{11}$$

The lower case n denotes the volume fraction of the corresponding species, i.e., $n^{(.)}\equiv N^{(.)}/V$.

For the micellar species, due to the magnitude of M ($M\geq 10$ [42]), if $C<\exp(\gamma/k_BT)$, the micellar volume fraction will be negligible in comparison to the monomer volume fraction; conversely, if $C>\exp(\gamma/k_BT)$, then the monomeric volume fraction will be $\exp(\gamma/k_BT)$ and all the excess monomers will be in the micellar form, i.e., $n^{(b)}\simeq C-\exp(\gamma/k_BT)$ [44]. It is therefore legitimate to define a critical concentration at $C_{\text{crit}}=\exp(\gamma/k_BT)$. We will call this the critical micellar concentration (CMC). For the fibrillar species, a similar reasoning indicates that the critical concentration for fibrilization is at $C_{\text{crit}}=\exp(f_\infty/k_BT)$. We will call this the critical fibrillar concentration (CFC).

If CMC$<$CFC and $C\gg$CMC, CFC, almost all monomers would be in the micellar form and the concentrations of monomers and fibrils are negligible by comparison. On the other hand, if $C\gg$CFC, CMC and CFC$<$CMC, then the concentrations of monomers and micelles will be negligible while the concentration of the fibrillar species will be abundant. In this fibril-dominant regime, $n_i^{(c)}$ follows the following distribution [44]:

$$n_i^{(c)} = \exp[-i/L + \xi f_\infty/k_BT], \tag{12}$$

where L is the average number of monomers in a fibril such that

$$L \equiv \langle iN_i^{(c)}\rangle = \sqrt{Ce^{-\xi f_\infty/k_BT}}. \tag{13}$$

Since a fibril is a linear structure, the average fibrillar length is thus proportional to L. According to Eq. (13), the average fibril length scales with $\sqrt{C}$. This fact is observed in other linear aggregating systems and is a manifestation of the one-dimensional nature of the aggregates [44,48]. The profile of $i\times n_i^{(c)}$ versus i is depicted in the inset plot in Fig. 2. This analytical result is qualitatively confirmed by the experimental observations on β-lactoglobulin amyloid fibrils [49,50].

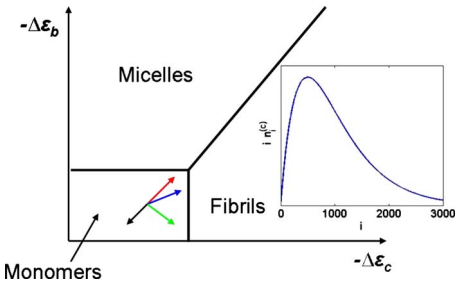

FIG. 2. (Color online) Dominance diagram of the three-species system at concentration higher than the critical concentrations: CMC and CFC. The colored arrows depict how the dominance may shift under increase in hydrophobicity (red), increase in the number of aromatic side chains (blue), increase in alternating hydrophobic-hydrophilic amino-acid sequence (green), and increase in unpaired charges in the side chains (black). *Inset plot:* The volume fractions of i-mer fibrils versus i in the fibrillar phase (i.e., $C \gg$ CFC and CMC $>$ CFC), where $L \equiv \langle iN_i^{(c)} \rangle$ is set to be 500.

We will now try to estimate the magnitudes of the terms appearing in Eq. (3). For the first term, let us assume that a protein in monomeric form is in the denatured state, and a fibrilized protein corresponds to the folded state. It has been experimentally and theoretically estimated that, by going from the denatured state to the folded state, a protein loses on average around $k_B \ln 10$ per amino acid in entropy [40,51,52]. We will therefore estimate Δs_c as $-Rk_B \ln 10 = -2.3 \times Rk_B$, where R is the number of amino acids in the protein. For the third term in Eq. (3), it has been demonstrated that the change in the protein's volume upon folding is very small [53]. Indeed, it is found that the change in volume per amino acid upon folding is in the order of 0.01 nm^3 [53], which suggests that $p\Delta v_c \sim 0.07 \times Rk_BT$ at atmospheric pressure. It is therefore negligible in comparison to the entropic contribution. The second term in Eq. (3) involves a combination of interactions, such as hydrogen bonding, hydrophobic interactions, electrostatic interactions, etc., among which hydrophobic interactions, which are of the order of a few k_BT per amino acid, are believed to be dominant [41,54]. Since hydrophobic interactions involves effective burying of hydrophobic side chains inside the protein structure, it indicates the need for a multilayered fibrillar structure (such as our two-tape model employed here), as universally observed in amyloid fibrils formed from different proteins [4].

III. EXPERIMENTAL IMPLICATIONS

We will now focus on the fibrillar phase, i.e., we are in the scenario where $C >$ CFC and CFC $<$ CMC. According to Eq. (13)

$$\ln L = -\frac{\xi f_\infty}{2k_BT} + \frac{1}{2}\ln C. \tag{14}$$

If we equate a monomeric protein to a denatured protein, and a fibrilized protein to a folded protein, then experimental work indicates that f_∞/T is a concave up function with respect to T such that the minimum occurs at around 20 °C [55,56]. This suggests that in an isobaric experiment, the average fibril length would first increase and then decrease as temperature increases.

The situation for pressure variation is more complicated due to the fact that the compressibility differs for different amino acids. Nevertheless, it has been found generally that at low pressure ($\sim$1 atm), the change in volume upon folding is small while the change is positive at very high pressure (7500 atm [57]) due to the fact that denatured protein has greater compressibility [58,53]. In other words, if we again equate a monomeric protein to a denatured protein, and a fibrilized protein to a folded protein, we would expect that, in the very high-pressure regime, an increase in pressure would lead to an exponential decrease in the average fibrillar length in an isothermal experiments.

IV. RELEVANCE TO PREVIOUS EMPIRICAL FINDINGS

As discussed in Sec. II, if $C \gg$ CMC, CFC, the dominant species in the system will be the one with a lower critical concentration. Namely, in terms of BFE, the dominant species will be fibrillar if $f_\infty < \gamma$, and vice versa. We will now discuss how the primary sequence may affect amyloid propensity in terms of the BFE. To simplify the discussion, we will assume that substituting an amino-acid affects predominantly the binding-energy term, $\Delta\epsilon$, in the BFE (c.f. Fig. 2).

As a result of empirical observations [18–23], it is generally agreed that the following factors promote amyloid formation: (i) an increase in hydrophobicity, and (ii) an increase in length of an alternating hydrophobic-hydrophilic amino-acid sequence; while it is found that an increase in the number of charged amino acids decreases amyloid propensity. In terms of the binding energies, an increase in hydrophobicity would decrease both $\Delta\epsilon_b$ and $\Delta\epsilon_c$ and as such would on average increase the fibrilization probability if the protein's parameters are already close to the monomer-fibril boundary in the dominance diagram (the red arrow in Fig. 2). For our two-tape model for the amyloid fibril [c.f. Fig. 1(c)], an increase in alternating hydrophobic-hydrophilic amino-acid sequence would allow for packing the hydrophobic side chains inside the cross-beta sheet structure, while having the hydrophilic side chains outside, this would decrease $\Delta\epsilon_c$. On the other hand, having such a pattern would conceivably decrease the average energy gained inside a micellar structure given the amorphous structural nature, i.e., $\Delta\epsilon_b$ will be increased. Such a modification would therefore increase amyloid propensity (the green arrow in Fig. 2). If there is an increase in paired charges in the protein, i.e., charges that are not accompanied by ionic bonds, electrostatic interaction would deter aggregation and as such both $\Delta\epsilon_b$ and $\Delta\epsilon_c$ will be increased (the black arrow in Fig. 2).

Another insight we can gain from the above consideration concerns the importance of aromatic residues in amyloid propensity [10,12–17]. Beside the heightened hydrophobicity in aromatic residues, the offsetted π-stacking interaction is directional along the fibrillar axis [59,60]; hence ϵ_c may be decreased more than ϵ_b (the blue arrow in Fig. 2). This suggests that aromatic interaction, or any interactions directional

along the fibrillar axis, contributes to amyloid stability in a way different from ordinary hydrophobic interactions.

V. KINETICS

We discuss now how the picture developed in this paper helps to describe the kinetics of the protein amyloid self-assembly process investigated experimentally. According to the model proposed in [61], the series of events leading up to the fibrilization of amyloid-β proteins is depicted in Fig. 3. In this scenario, the direct pathway from monomers to stable nucleus (depicted by the broken arrow in Fig. 3) is in a time scale too long to be probed experimentally. Therefore, the only possible fibrilization pathway is for the monomers to first formed micelles (a fast process, depicted by the thick black arrow), stable nuclei are then formed out of the micelles (a slow process, depicted by the thin black arrow). Based on this model, within the temporal constraint of experiments, fibrilization is only possible if $C > \mathrm{CMC}$ [c.f. Fig. 3(b)]. In the case of the amyloid-β protein, the CMC has been measured to be in the order of 10 μM [62]. This is substantially higher than the concentration of amyloid-beta in the cerebral spinal fluid, which is in the subnanomolar concentration range [63]. It therefore poses the question are current experimental methods only probing the fast pathway—monomers to micelles to nucleus (depicted by the two solid arrows), while the physiologically relevant pathway is the slow pathway—monomers to nucleus (depicted by the broken arrow).

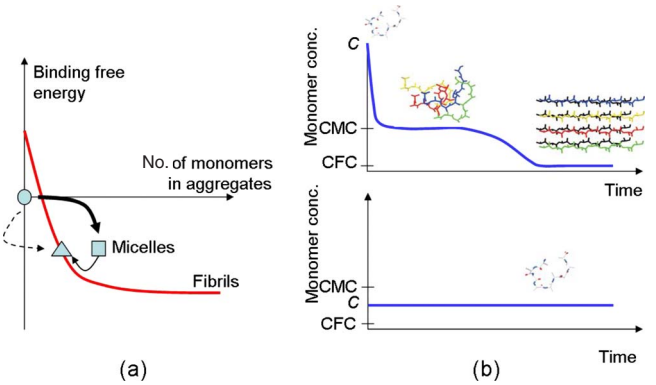

FIG. 3. (Color online) (a) A schematic diagram depicting the amyloid-beta self-assembly process proposed in [61]. The circle denotes the free monomeric state, the square denotes the typical micellar state (the micellar size, M, is estimated to be 25 [61]), and the triangle denotes a stable nucleus (the nucleus size is estimated to be 10 [61]). The thick arrow depicts the fast pathway from free monomer to micelles and the thin arrow depicts the slow process of nucleation from micelles. The broken arrow depicts that very slow process of nucleation from free monomers, which is out of the range of experimental time scale, but may play an important role in actual pathogenesis under physiological time scale. (b) The temporal evolution of monomer concentration. Upper plot: When $C > \mathrm{CMC}$, the monomers are quickly converted into micelles and then slowly into fibrils. The figures above the curve depict the dominant species in the solution as time progresses. Lower plot: When $\mathrm{CFC} < C < \mathrm{CMC}$, the proteins remain in monomeric form for a time longer than can be probed experimentally. These two plots show the curious phenomenon of the possibility of ending up with a lower monomeric concentration when the initial concentration is higher.

VI. CONCLUSION

In this work, we have considered the thermodynamic equilibrium behavior of a system with a mixture of monomeric proteins, the corresponding micellar aggregates, and fibrillar aggregates. We have deduced the concentrations of these species at thermal equilibrium and we have found that the average fibrillar length is very sensitive to temperature or pressure variation. We have also discussed the relevance of our investigation to previous empirical findings and to the understanding of the kinetical process of fibrilization.

ACKNOWLEDGMENTS

The author thanks Catherine Davison, Létitia Jean, and David Vaux at the Dunn School of Pathology (Oxford) for many stimulating discussions, and the Glasstone Trust (Oxford) and Jesus College (Oxford) for financial support.

[1] M. Sunde, L. C. Serpella, M. Bartlama, P. E. Frasera, M. B. Pepysa, and C. C. F. Blakea, J. Mol. Biol. **273**, 729 (1997).
[2] C. M. Dobson, Nature (London) **426**, 884 (2003).
[3] S. Radford, Trends Biochem. Sci. **25**, 611 (2000).
[4] M. R. Sawaya *et al.*, Nature (London) **447**, 453 (2007).
[5] J. D. Harper and P. T. Lansbury, Annu. Rev. Biochem. **66**, 385 (1997).
[6] T. P. Knowles, A. W. Fitzpatrick, S. Meehan, H. R. Mott, M. Vendruscolo, C. M. Dobson, and M. E. Welland, Science **318**, 1900 (2007).
[7] J. F. Smith, T. P. Knowles, C. M. Dobson, C. E. Macphee, and M. E. Welland, Proc. Natl. Acad. Sci. U.S.A. **103**, 15806 (2006).
[8] A. M. Corrigan, C. Muller, and M. R. Krebs, J. Am. Chem. Soc. **128**, 14740 (2006).
[9] C. F. Lee, Phys. Rev. E **80**, 031902 (2009).
[10] M. Reches and E. Gazit, Science **300**, 625 (2003).
[11] T. Scheibel, R. Parthasarathy, G. Sawicki, X.-M. Lin, H. Jaeger, and S. L. Lindquist, Proc. Natl. Acad. Sci. U.S.A. **100**, 4527 (2003).
[12] M. Reches, Y. Porat, and E. Gazit, J. Biol. Chem. **277**, 35475 (2002).
[13] S. M. Tracz, A. Abedini, M. Driscoll, and D. P. Raleigh, Biochemistry **43**, 15901 (2004).
[14] B. Ma and R. Nussinov, Curr. Opin. Chem. Biol. **10**, 445 (2006).
[15] F. Bemporad, N. Taddei, M. Stefani, and F. Chiti, Protein Sci. **15**, 862 (2006).

[16] P. Marek, A. Abedini, B. Song, M. Kanungo, M. E. Johnson, R. Gupta, W. Zaman, S. S. Wong, and D. P. Raleigh, Biochemistry **46**, 3255 (2007).
[17] L. Jean, C. F. Lee, M. Shaw, and D. J. Vaux, PLoS ONE **3**, e1834 (2008).
[18] S. Yoon and W. J. Welsh, Protein Sci. **13**, 2149 (2004).
[19] A. M. Fernandez-Escamilla, F. Rousseau, J. Schymkowitz, and L. Serrano, Nat. Biotechnol. **22**, 1302 (2004).
[20] G. G. Tartaglia, A. Cavalli, R. Pellarin, and A. Caflisch, Protein Sci. **14**, 2723 (2005).
[21] O. V. Galzitskaya, S. O. Garbuzynskiy, and M. Y. Lobanov, PLoS Comput. Biol. **2**, e177 (2006).
[22] K. F. Dubay, A. P. Pawar, F. Chiti, J. Zurdo, C. M. Dobson, and M. Vendruscolo, J. Mol. Biol. **341**, 1317 (2004).
[23] A. Pawar, K. Dubay, J. Zurdo, F. Chiti, M. Vendruscolo, and C. Dobson, J. Mol. Biol. **350**, 379 (2005).
[24] T. P. J. Knowles, J. F. Smith, A. Craig, C. M. Dobson, and M. E. Welland, Phys. Rev. Lett. **96**, 238301 (2006).
[25] T. P. Knowles, J. F. Smith, G. L. Devlin, C. M. Dobson, and M. E. Welland, Nanotechnology **18**, 044031 (2007).
[26] T. R. Serio, A. G. Cashikar, A. S. Kowal, G. J. Sawicki, J. J. Moslehi, L. Serpell, M. F. Arnsdorf, and S. L. Lindquist, Science **289**, 1317 (2000).
[27] W. Yong, A. Lomakin, M. D. Kirkitadze, D. B. Teplow, S.-H. Chen, and G. B. Benedek, Proc. Natl. Acad. Sci. U.S.A. **99**, 150 (2002).
[28] H. D. Nguyen and C. K. Hall, Proc. Natl. Acad. Sci. U.S.A. **101**, 16180 (2004).
[29] R. Pellarin and A. Caflisch, J. Mol. Biol. **360**, 882 (2006).
[30] P. H. Nguyen, M. S. Li, G. Stock, J. E. Straub, and D. Thirumalai, Proc. Natl. Acad. Sci. U.S.A. **104**, 111 (2007).
[31] R. Pellarin, E. Guarnera, and A. Caflisch, J. Mol. Biol. **374**, 917 (2007).
[32] M. Cheon, I. Chang, S. Mohanty, L. M. Luheshi, C. M. Dobson, M. Vendruscolo, and G. Favrin, PLoS Comput. Biol. **3**, e173 (2007).
[33] W.-F. Xue, S. W. Homans, and S. E. Radford, Proc. Natl. Acad. Sci. U.S.A. **105**, 8926 (2008).
[34] J. Zhang and M. Muthukumar, J. Chem. Phys. **130**, 035102 (2009).
[35] J. van Gestel and S. W. de Leeuw, Biophys. J. **90**, 3134 (2006).
[36] A. Aggeli, I. A. Nyrkova, M. Bell, R. Harding, L. Carrick, T. C. B. Mcleish, A. N. Semenov, and N. Boden, Proc. Natl. Acad. Sci. U.S.A. **98**, 11857 (2001).
[37] I. A. Nyrkova, A. N. Semenov, A. Aggeli, and N. Boden, Eur. Phys. J. B **17**, 481 (2000).
[38] G. Tiana, F. Simona, R. A. Broglia, and G. Colombo, J. Chem. Phys. **120**, 8307 (2004).
[39] N. Guex and M. C. Peitsch, Electrophoresis **18**, 2714 (1997).
[40] K. Sneppen and G. Zocchi, *Physics in Molecular Biology*, 1st ed. (Cambridge University Press, Cambridge, 2005).
[41] M. B. Jackson, *Molecular and Cellular Biophysics* (Cambridge University Press, Cambridge, 2006).
[42] For instance, M is in the order of 30 for amyloid-β as demonstrated experimentally in [27].
[43] J. L. Jiménez, E. J. Nettleton, M. Bouchard, C. V. Robinson, C. M. Dobson, and H. R. Saibil, Proc. Natl. Acad. Sci. U.S.A. **99**, 9196 (2002).
[44] J. N. Israelachvili, *Intermolecular and Surface Forces*, 2nd ed. (Academic Press, London, 1991).
[45] B. Mutaftschiev, *The Atomistic Nature of Crystal Growth*, 1st ed. (Springer, Berlin, 2001).
[46] Note that A, B, and C_i in Eq. (5) correspond to the partition functions for a monomer, a micelle, and a i-mer fibril, respectively. For instance, $B=\int_{\Lambda_b} d[q]e^{-U(\{q\})/k_BT}$, where Λ_b denotes the volume of the coordinate space defining a micelle, the set $\{q\}$ denotes the set of degrees of freedom of the aggregates, i.e., the coordinates of all the atoms making up the micelle, and $U(\{q\})$ is the internal energy of the system.
[47] P. van der Schoot and M. E. Cates, Langmuir **10**, 670 (1994).
[48] J. N. Israelachvili, D. J. Mitchell, and B. W. Ninham, J. Chem. Soc., Faraday Trans. **72**, 1525 (1976).
[49] S. S. Rogers, P. Venema, L. M. Sagis, E. van der Linden, and A. M. Donald, Macromolecules **38**, 2948 (2005).
[50] S. Rogers, P. Venema, J. van der Ploeg, L. Sagis, A. Donald, and E. van der Linden, Eur. Phys. J. E **18**, 207 (2005).
[51] J. F. Brandts, J. Am. Chem. Soc. **86**, 4291 (1964).
[52] J. D. Bryngelson and P. G. Wolynes, Proc. Natl. Acad. Sci. U.S.A. **84**, 7524 (1987).
[53] Y. Harpaz, M. Gerstein, and C. Chothia, Structure **2**, 641 (1994).
[54] K. A. Dill, Biochemistry **29**, 7133 (1990).
[55] J. A. Schellman, Biophys. J. **73**, 2960 (1997).
[56] R. L. Baldwin, Proc. Natl. Acad. Sci. U.S.A. **83**, 8069 (1986).
[57] S. A. Hawley, Biochemistry **10**, 2436 (1971).
[58] C. A. Royer, Biochim. Biophys. Acta **1595**, 201 (2002).
[59] C. A. Hunter, J. Singh, and J. M. Thornton, J. Mol. Biol. **218**, 837 (1991).
[60] E. A. Meyer, R. K. Castellano, and F. Diederich, Angew. Chem. Int. Ed. **42**, 1210 (2003).
[61] A. Lomakin, D. B. Teplow, D. A. Kirschner, and G. B. Benedek, Proc. Natl. Acad. Sci. U.S.A. **94**, 7942 (1997).
[62] E. Terzi, G. Holzemann, and J. Seelig, Biochemistry **36**, 14845 (1997).
[63] P. Seubert *et al.*, Nature (London) **359**, 325 (1992).