

The attachment of α -synuclein to a fiber: A coarse-grain approach

Ioana M. Ilie, Wouter K. den Otter, and Wim J. Briels

Citation: *The Journal of Chemical Physics* **146**, 115102 (2017); doi: 10.1063/1.4978297

View online: <https://doi.org/10.1063/1.4978297>

View Table of Contents: <http://aip.scitation.org/toc/jcp/146/11>

Published by the [American Institute of Physics](#)

Articles you may be interested in

[A coarse grained protein model with internal degrees of freedom. Application to \$\alpha\$ -synuclein aggregation](#)

The Journal of Chemical Physics **144**, 085103 (2016); 10.1063/1.4942115

[Multiscale simulations of anisotropic particles combining molecular dynamics and Green's function reaction dynamics](#)

The Journal of Chemical Physics **146**, 114106 (2017); 10.1063/1.4977515

[Perspective: Dissipative particle dynamics](#)

The Journal of Chemical Physics **146**, 150901 (2017); 10.1063/1.4979514

[Early stages of clathrin aggregation at a membrane in coarse-grained simulations](#)

The Journal of Chemical Physics **146**, 155102 (2017); 10.1063/1.4979985

[Kinetics of spontaneous filament nucleation via oligomers: Insights from theory and simulation](#)

The Journal of Chemical Physics **145**, 211926 (2016); 10.1063/1.4965040

[Wetting at the nanoscale: A molecular dynamics study](#)

The Journal of Chemical Physics **146**, 114704 (2017); 10.1063/1.4978497

PHYSICS TODAY

WHITEPAPERS

ADVANCED LIGHT CURE ADHESIVES

Take a closer look at what these environmentally friendly adhesive systems can do

READ NOW

PRESENTED BY
 **MASTERBOND**
ADHESIVES | SEALANTS | COATINGS

The attachment of α -synuclein to a fiber: A coarse-grain approach

Ioana M. Ilie,^{1,2,a)} Wouter K. den Otter,^{1,2,3,a)} and Wim J. Briels^{1,2,4,a)}

¹Computational Chemical Physics, Faculty of Science and Technology, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands

²MESA+ Institute for Nanotechnology, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands

³Multi Scale Mechanics, Faculty of Engineering Technology, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands

⁴Forschungszentrum Jülich, ICS-3, D-52425 Jülich, Germany

(Received 25 October 2016; accepted 24 February 2017; published online 20 March 2017)

We present simulations of the amyloidogenic core of α -synuclein, the protein causing Parkinson's disease, as a short chain of coarse-grain patchy particles. Each particle represents a sequence of about a dozen amino acids. The fluctuating secondary structure of this intrinsically disordered protein is modelled by dynamic variations of the shape and interaction characteristics of the patchy particles, ranging from spherical with weak isotropic attractions for the disordered state to spherocylindrical with strong directional interactions for a β -sheet. Flexible linkers between the particles enable sampling of the tertiary structure. This novel model is applied here to study the growth of an amyloid fibril, by calculating the free energy profile of a protein attaching to the end of a fibril. The simulation results suggest that the attaching protein readily becomes trapped in a mis-folded state, thereby inhibiting further growth of the fibril until the protein has readjusted to conform to the fibril structure, in line with experimental findings and previous simulations on small fragments of other proteins. *Published by AIP Publishing.* [<http://dx.doi.org/10.1063/1.4978297>]

I. INTRODUCTION

Intrinsically disordered proteins (IDPs) are proteins that lack a stable three dimensional structure in solution but which are able to fold in suitable environments and perform various functions within the cell. Their misfolding and aggregation are strongly linked to a number of neurodegenerative disorders such as Parkinson's disease (PD) and Alzheimer's disease (AD).¹ The proteins linked to these disorders are α -synuclein (α -syn) and amyloid- β ($A\beta$), respectively. These proteins can change their secondary structure and engage in the formation of various aggregates, ranging from amorphous agglomerates to ordered fibrils, which are known to accumulate in specific brain areas and cause neurodegeneration.²⁻⁶ Depending on the physiological conditions, α -synuclein can aggregate into oligomers of various sizes and geometries, into fibrils of distinct polymorphism, or into higher order aggregates such as Lewy bodies.^{4,7,8} The prevailing ordered structures commonly associated with neurodegradation are amyloids, which are protein aggregates rich in β -sheet content.⁹ Amyloids are characterised by a cross- β architecture^{10,11} in which the β -strands of distinct peptides are running perpendicular and intermolecular hydrogen bonds are running parallel to the long axis of the fibril.⁹ Over the past decades, a profound insight has been gained on how folding occurs in regular, ordered proteins. However, the folding process of intrinsically disordered proteins is still poorly understood, despite being intensively explored by both experiments and simulations.^{4,12} To

understand the folding pathways and the structure of α -synuclein within a fiber, various mechanisms have been proposed and multiple experimental techniques were used. The building blocks of amyloid fibers are the central 60-70 amino acids (AAs) of α -synuclein, roughly coinciding with residues 35 through 100.^{13,14} It has been shown that the central region forms cross- β sheets packed in a highly ordered fashion, parallel, and in-register.¹³⁻¹⁵ This core is formed of five strands, each consisting of about 10-12 residues.¹⁵ The N-terminus of the protein is structurally more heterogeneous than the core, and the C-terminus also remains unfolded.¹³ Experimentally the growth process of a fiber is found to exhibit two phases: growth periods (referred to as "go") during which the fiber is naturally growing by monomer attachment and interrupted by long pauses (referred to as "stop") during which growth is prohibited.¹⁶⁻¹⁹ Moreover, fiber elongation has been proposed to follow a "dock-lock" mechanism during which a free monomer rapidly binds to an existing fiber (dock) followed by slow structural rearrangements towards the ordered state of the template (lock).^{6,9,70}

Computer simulations can provide insight into protein folding at length- and time scales that are difficult to access by experiments. Atomistic simulations provide detailed information on the folding process of peptides, but they are computationally expensive and unable to reach the longer aggregation time scales of entire proteins. Instead, coarse grain simulations reduce the number of configurational degrees of freedom to enable access to longer time scales and to make accurate predictions of slow folding mechanisms.²⁰⁻²² All-atom simulations of fragments of α -synuclein showed that the secondary structure strongly depends on the solvent conditions.^{23,24} Transitions between the conformations

^{a)} Authors to whom correspondence should be addressed. Electronic addresses: i.m.ilie@utwente.nl; w.k.denotter@utwente.nl; and w.j.briels@utwente.nl

of various IDPs have been intensively studied by Molecular Dynamics (MD) simulations.^{25–27} Molecular docking and MD simulations on α -syn showed that β -sheet rich oligomers are less stable than α -helical oligomers.²⁸ Another study showed that β -sheet rich oligomers are more compact than helical oligomers suggesting that they could be the precursors of fibrils.²⁸ Furthermore, the aggregation pathways are highly dependent on the β -sheet propensity of the peptide, i.e., high β -sheet propensity leads to rapid aggregation into fibrillar structures, whereas low β -content leads to the formation of long-lived intermediate oligomers before fiber formation.²⁹

The most intensively used coarse grained models for IDPs are lattice models, in which each residue is represented by a single particle occupying one site on a cubic lattice, while all unoccupied sites are considered as solvent.³⁰ A recent study on small peptides predicts that the self-assembly into fibers occurs either through random aggregation or via a templated-folding process, the latter suggesting strong correlations between folding and aggregation.³¹ Moreover, simulations show that H-bonds are formed both intramolecularly in β -sheet rich folded proteins and in intermolecular cross- β structures.³² To study the aggregation of IDPs into higher order aggregates, several coarser models have been proposed and used. These vary from representing a protein by several beads to single particle descriptions. One of the first among these models represents a protein as a 10-bead-molecule with one internal angular degree of freedom to distinguish between an amyloid-protected and an amyloid-forming state.^{33,34} This model was successful in studying the formation pathways of various oligomeric and fibrillar structures. In another model, a protein is described as a rigid tetrahedron with four beads: two hydrophobic and two hydrophilic.³⁵ This model was used to study the aggregation into amorphous structure and to analyse fiber formation and breakage. A step further on the ladder of highly coarse grain models is that a protein is represented by a single patchy particle. The Alzheimer's A β protein has been coarse grained into a single spherocylinder with surface patches that can spontaneously switch between two distinct binding states, representing the unfolded and folded configurations, respectively.^{36–39} Simulations showed a concentration

dependence of fibrillization: one step nucleation at low concentration and two step nucleation at high concentrations. We observed one-step and two-step nucleations by representing the core of α -synuclein as a single polymorph particle able to change both shape and binding affinities as a function of the environment.⁴⁰ This model also showed that a fiber can act as a catalyst that enhances the transformation of an oligomer into a fiber.

The aforementioned studies employ a variety of coarse-grain models aimed at describing the nucleation and growth mechanisms of amyloid, but, except for the lattice models, they are unsuited to study the folding mechanisms of a protein on the surface of a fiber. One step further towards understanding the complete growth mechanism of amyloidogenic proteins is introducing a model that more closely follows the main features of the protein. These include capturing the conformational dynamics and distribution of binding affinities along the contour of the protein. With this in mind, we introduce a model consisting of a chain of polymorphic patchy particles to describe α -synuclein at the secondary level, see Fig. 1. We capture the agility of the protein to adapt its local secondary structure, e.g., transitions from disordered to β -sheet, by varying both geometry and interaction characteristics of the particles. The disordered state is modelled as a random chain of spherical particles. In the most ordered state, the protein segments can form β -sheets, which are modelled as spherocylinders. The translational, rotational and the internal dynamics of the particles are described by a recently introduced Brownian Dynamics algorithm.^{40,41} The focus of this paper is to study fibrillar growth and the mechanisms of attachment of a free monomer on the surface of a fiber via coarse grained simulations. Since spontaneous attachment is a rare event in coarse-grained simulations, dedicated simulation techniques are required to explore this process. The much rarer event of filament nucleation by several monomers coming together^{42,43} lies outside the scope of this study. This paper first introduces the coarse grain model in Section II. It then continues with a short overview of constrained Brownian Dynamics in Section III. Section IV presents the free energy profiles of a protein attaching to a fiber. The paper ends with conclusions in Section VI.

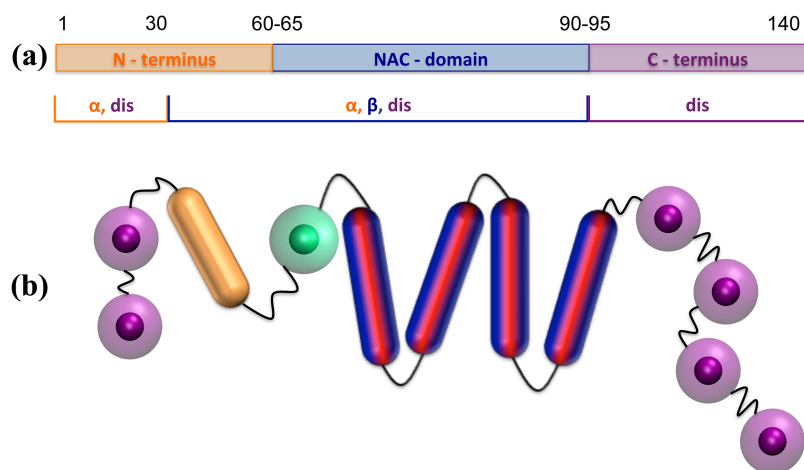


FIG. 1. (a) Schematic representation of the conformations that the α -synuclein segments can adopt. The numbers on top indicate the approximate residues defining these regions. (b) Cartoon of our coarse grained α -synuclein model. The differently coloured and shaped particles indicate different conformations of the segments of the protein.

II. THE SIMULATION MODEL

A. Polymorph particles

α -synuclein is a 14 kDa 140 amino acid (AA) intrinsically disordered protein, characterised by a partial weak secondary and tertiary structure in its native state. The structure and the properties of α -synuclein depend on its amino acid sequence, see Fig. 1(a). Three main regions can be defined: the amphipathic N-terminus consisting of the first 60-65 amino acids, the hydrophobic non-A β component (NAC region) comprising the next 30 AA, and the charged C-terminus composed of the last 50 amino acids. These regions can adapt their conformations in response to the surrounding environment. The protein is mainly disordered in a solvent, its first two regions become α -helical upon binding to a membrane,^{44,45} and the central 60 amino acids form β -sheet rich structures in fibrils and amyloids.^{46,47} We model the protein as a short chain of particles and simulate the conformational flexibility of α -synuclein by enabling every particle to adapt its “conformation” depending on the surrounding environment.⁴⁰ This includes changing shape, binding characteristics, and thermodynamic stability of the particles, referred to as “chameleon” particles. Figure 1(b) shows a cartoon of the model protein illustrating the various conformations its particles can adopt. The dynamic nature of the model is emphasised by combining the particles’ instantaneous shapes with a state dependent colour. We use spheres to describe disordered segments (depicted in purple), long spherocylinders for the components of a β -sheet (in dark blue and red), short spherocylinders for α -helical segments (in orange), and spheres to mark transitory states (in green).

In this paper, we are interested in amyloid growth. The central 60 amino acids (AA 30 to 90) are known to be the building blocks of amyloid fibers.^{13,14,48} Furthermore, both N and C-terminal regions are believed to slow down the aggregation process but not to have any other major influence on the fibrillation.⁴⁹ Given these arguments, we will limit ourselves here to modelling just the core of the protein and postpone extending the model to represent the full α -synuclein in a future study. For ease of presentation, we refer to the central 60 amino acids of α -synuclein, which have the ability to form cross- β -sheets and form the core of the amyloid fiber, as “the protein.”

Starting from the protein’s geometrical features, we coarse grain the core (central 60 amino acids) of α -synuclein into a chain of five “chameleon” particles⁴⁰ connected by four springs. Each particle represents about 12 amino acids which can vary from disordered (sphere) to β -sheet rich (spherocylinders). A disordered protein is represented as a random chain of five spheres. The dimensions of the individual spheres have been chosen such that they match experimental observations. It is known that α -synuclein is more closely packed than a random coil⁵⁰ but less closely packed than a globular protein of the same contour length.⁵¹ We scaled the experimental radius of gyration R_g of about 40 Å of the complete protein to determine that of a 12 amino acids sequence as $R_g = 1.17$ nm, corresponding to a particle diameter $D_S = 2.34$ nm. Upon engaging in fibril formation, α -synuclein adopts a β -sheet rich conformation composed of five β -strands that are able to form intermolecular hydrogen bonds with the neighbouring proteins, i.e., to engage in the cross- β -sheet conformation characteristic for

this protein.⁵² We represent a β -sheet as a folded chain of five spherocylinders closely mimicking the properties of the five β -strands in a β -sheet, see Fig. 2. Experimental measurements on these fragments revealed that the lengths of the β strands vary from 2.0 to 3.5 nm, and the measured inter-strand distance within a fiber is 1.0 nm.¹⁵ With this in mind, we define the tip to tip distance of a spherocylinder as $L_R = 3.5$ nm and the diameter as $D_R = 1.0$ nm.

To mimic the ability of α -synuclein to change its conformation depending on the environment, we employ a model previously introduced by us.⁴⁰ A parameter λ_i is introduced to describe the internal state of the i th particle. Starting from experimental observations that α -synuclein has little structure in solution, we build an internal potential that favours the disordered state over the ordered state. This potential is given by a combination of two third order polynomials

$$\Phi_\lambda(\lambda_i) = \begin{cases} -6.75 \varepsilon_\lambda^S (\lambda_i^2 + \lambda_i^3) & \text{for } \lambda_i < 0 \\ -6.75 \varepsilon_\lambda^R (\lambda_i^2 - \lambda_i^3) & \text{for } \lambda_i \geq 0 \end{cases} \quad (1)$$

with their minima at $\lambda = \pm 2/3$, as shown in Fig. 3(a). Their depths are $-\varepsilon_\lambda^S$ and $-\varepsilon_\lambda^R$ and correspond to the disordered (spherical) and ordered β -sheet rich (spherocylindrical) states, respectively. The geometrical properties of the particles are naturally connected to the internal coordinate. To describe this relation, a smoothed step function $\mu(\lambda_i)$ is introduced,

$$\mu(\lambda_i) = \begin{cases} 0 & \text{for } \lambda_i < -1 \\ \frac{\tanh(A\lambda_i)}{2 \tanh(A)} + \frac{1}{2} & \text{for } -1 \leq \lambda_i \leq 1, \\ 1 & \text{for } \lambda_i > 1 \end{cases} \quad (2)$$

where $A = 10$ determines the width of the transition region. This function smoothly switches from zero, for negative values of λ_i (sphere), to unity, for positive values of λ_i (spherocylinder). Length L and diameter D of particle i are then given by

$$\begin{aligned} L_i &= L_S + \mu(\lambda_i)(L_R - L_S), \\ D_i &= D_S + \mu(\lambda_i)(D_R - D_S), \end{aligned} \quad (3)$$

as shown in Fig. 3(b), where L_R denotes the end-to-end distance of a spherocylinder and L_S equals D_S for a sphere.

B. The repulsive potential

The repulsive component of the potential avoids strong overlap between particles. Naturally, the hardness/softness of the potential depends on the nature of the interacting particles. This property is captured in the potential by introducing

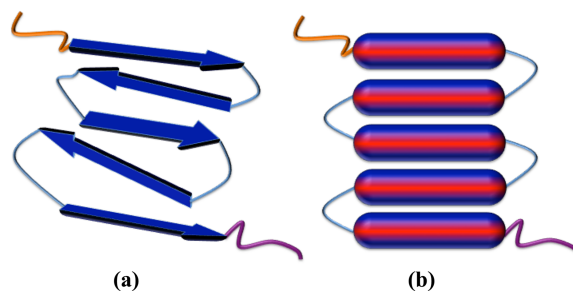


FIG. 2. Cartoon of a single layer of a cross- β -sheet (a) and the matching coarse-grain representation (b). The hydrogen bonds between the sheets are sticking in and out of the plane (red) and the weak van der Waals interactions between the segments belonging to the same protein are highlighted in blue.

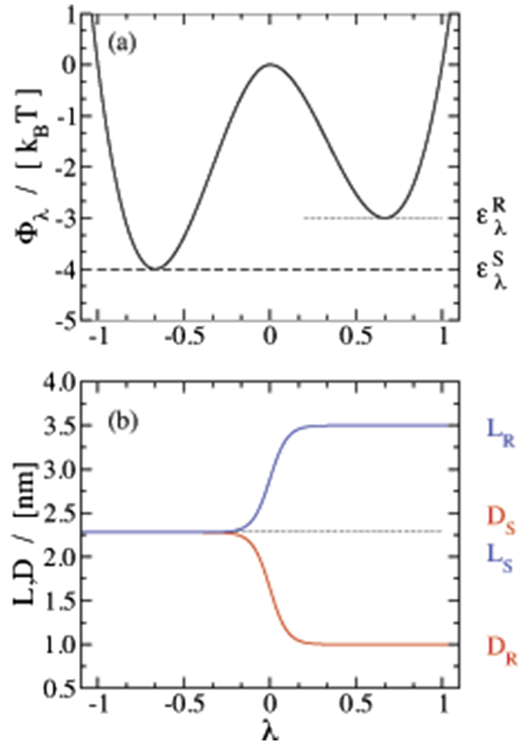


FIG. 3. (a) The internal potential as a function of the internal parameter λ , for $\epsilon_\lambda^S = 4$ and $\epsilon_\lambda^R = 3$. (b) Diameter (red line) and length (blue line) of a “chameleon” particle as functions of λ .

a dependence on the internal states (λ) of the particles. The repulsive potential, as a function of the minimum distance d_{\min} (to be defined below), between particles i and j is given by

$$\frac{\Phi_{\text{rep}}^\lambda(d_{\min})}{k_B T} = \begin{cases} \left(\frac{2\bar{D}}{d_{\min}} - 1\right)^n & \text{for } d_{\min} < 2\bar{D} \\ 0 & \text{for } d_{\min} > 2\bar{D} \end{cases}, \quad (4)$$

and extends up to twice the average diameter $\bar{D} = 0.5(D_i + D_j)$ of the particles. The hardness of the potential is designed to vary with the internal states of the interacting particles, with an exponent $n = 4 + (\mu_i + \mu_j)$ varying from 4, for two interacting spheres, to 6, for two interacting spherocylinders. The adjustment of the hardness is motivated by the physical properties of the protein segments: disordered sections are more loosely packed and hence softer than the ordered β -strands, which are more compact and robust. The minimum distance d_{\min} between the long axes of two particles i and j depends not only on their centres of mass positions \mathbf{r} and orientations $\hat{\mathbf{u}}$ of the long axes⁵³ but also on their internal states λ .⁴⁰ The minimum distance between the long axes reads

$$d_{\min}(\mathbf{r}_i, \mathbf{r}_j, \hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j, \lambda_i, \lambda_j) = \min_{\alpha_i, \alpha_j} |(\mathbf{r}_i + \alpha_i \hat{\mathbf{u}}_i) - (\mathbf{r}_j + \alpha_j \hat{\mathbf{u}}_j)|. \quad (5)$$

This equation accounts for the internal states of the particles by restricting the two minimisation parameters α_i and α_j to $|\alpha_i| \leq \frac{1}{2}(L_i - D_i)$ and $|\alpha_j| \leq \frac{1}{2}(L_j - D_j)$, respectively. For two spherical particles, this distance reduces to the distance between their centres. If two rods are nearly aligned, small changes in their orientations can lead to significant changes in their points of closest approach, and the derivative of Eq. (5) may become singular. Therefore a different calculation of the minimum distance is required in these instances. Defining the

average direction⁴⁰

$$\hat{\mathbf{u}}_{\text{av}} = \begin{cases} \hat{\mathbf{u}}_i + \hat{\mathbf{u}}_j & \text{for } \hat{\mathbf{u}}_i \simeq \hat{\mathbf{u}}_j \\ \frac{|\hat{\mathbf{u}}_i + \hat{\mathbf{u}}_j|}{|\hat{\mathbf{u}}_i - \hat{\mathbf{u}}_j|} & \text{for } \hat{\mathbf{u}}_i \simeq -\hat{\mathbf{u}}_j \end{cases}, \quad (6)$$

the centre to centre vector $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$ is decomposed into projections parallel, r^\parallel , and perpendicular, r^\perp , to this direction. The minimum distance is then calculated as⁵⁴

$$d_{\min} = \begin{cases} r^\perp & \text{for } r^\parallel < \bar{L} - \bar{D} \\ \sqrt{(r^\perp)^2 + (r^\parallel - (\bar{L} - \bar{D}))^2} & \text{for } r^\parallel \geq \bar{L} - \bar{D} \end{cases}, \quad (7)$$

where $\bar{L} = 0.5(L_i + L_j)$ is the average length of the two particles.

C. The attractive potential

The formation of amyloidogenic aggregates is the result of hydrophobic, van der Waals, and hydrogen bonds both within and between proteins.⁵⁵ Their relative strengths are determined by the internal states of the particles, i.e., disordered protein segments are mainly governed by hydrophobic and van der Waals interactions and occasional hydrogen bonds, while in fibrils protein segments adopt a more compact structure and are governed by strong short ranged inter-molecular H-bonds and weak intra-molecular interactions. The potential introduced here captures these features and is given by

$$\frac{\Phi_{\text{attr}}^\lambda}{k_B T} = \eta(d_{\min}) \left\{ C_{\text{SS}} (1 - \mu_i) (1 - \mu_j) + (m_{ij} C_{\text{RR}}^w + g_{ij} C_{\text{RR}}^s) \mu_i \mu_j f_{ij} + C_{\text{SR}} [\mu_i (1 - \mu_j) + \mu_j (1 - \mu_i)] \right\}. \quad (8)$$

It has been designed to convey directionality between spherocylindrical particles, to mimic hydrogen bonding, and to be non-directional between spheres, to resemble anisotropic protein segments. In the following paragraphs, we explain every term on the right hand side of Eq. (8).

We start with introducing the distance dependent function $\eta(d_{\min})$ in the attractive potential,

$$\eta(d_{\min}) = \frac{\tanh(A_{\text{attr}}(d_{\min} - 2\bar{D}))}{\tanh(2A_{\text{attr}}\bar{D})} \quad \text{for } d_{\min} < 2\bar{D}, \quad (9)$$

that smoothly increases from -1 to 0 with increasing distance and remains zero for $d_{\min} \geq 2\bar{D}$. Its steepness is set by the λ dependent parameter $A_{\text{attr}} = a_{\text{attr}}(\mu_i + \mu_j) + b_{\text{attr}}$. Here $a_{\text{attr}} = 1.25 \text{ nm}^{-1}$ and $b_{\text{attr}} = 1.5 \text{ nm}^{-1}$ are two positive parameters chosen such that the steepness of the attractive potential varies between 1.5, for two interacting spherocylinders, and 4 for two interacting spheres.

The sum of the terms between the curly brackets of Eq. (8) reduces to its first term upon the interaction of two spheres, i.e., C_{SS} . The last term between the curly brackets introduces an interaction between a spherical, $\mu(\lambda) \simeq 0$, and a spherocylindrical particle, $\mu(\lambda) \simeq 1$, of interaction strength C_{SR} . Both these terms introduce a non-directional component in the potential, in the sense that the potential depends on the minimum distance only. The second term between the curly

brackets defines the interactions between two spherocylindrical particles and accounts for both weaker, van der Waals and hydrophobic, and strong, H-bonds, interactions. Figure 4 shows the strong intermolecular H-bonds along the direction of fibrillar growth, in red, and the weak intramolecular bonds, in light blue. The strengths of the specific interactions are given by C_{RR}^s for the strong bonds and C_{RR}^w for the weak interactions, respectively. To mimic these interactions one needs, to introduce an orientational dependence in the potential. To describe the asymmetry, we introduce the orientation dependent function f_{ij} in Eq. (8) with

$$f_{ij} = (\hat{\mathbf{u}}_i \cdot \hat{\mathbf{u}}_j)^2 h(\hat{\mathbf{u}}_i \cdot \hat{\mathbf{r}}_{ij}) h(\hat{\mathbf{u}}_j \cdot \hat{\mathbf{r}}_{ij}), \quad (10)$$

which varies from 0 to 1. For maximum contact between the surfaces of two spherocylinders, they must run in parallel, as accounted for in the first factor in Eq. (10). The last two factors of the same equation, where $\hat{\mathbf{r}}_{ij}$ is a unit vector along \mathbf{r}_{ij} and

$$h(x) = \begin{cases} 1 - ax^2 & \text{if } x^2 < 1/a \\ 0 & \text{if } x^2 > 1/a \end{cases}, \quad (11)$$

penalise longitudinal displacement along the spherocylinders' long axes. The parameter $a = 1.5$ sets the cutoff displacement, ensuring an attractive potential covering about 75% of the rod length. The anisotropy of the interactions between two spherocylindrical particles under rotations around their long axes is described by the factor $(g_{ij}C_{RR}^s + m_{ij}C_{RR}^w)$ in Eq. (8), with m_{ij} and g_{ij} defined as

$$\begin{aligned} g_{ij} &= (\hat{\mathbf{r}}_{ij} \cdot \hat{\mathbf{n}}_i)^p (\hat{\mathbf{r}}_{ij} \cdot \hat{\mathbf{n}}_j)^p, \\ m_{ij} &= (\hat{\mathbf{r}}_{ij} \cdot \hat{\mathbf{l}}_i)^p (\hat{\mathbf{r}}_{ij} \cdot \hat{\mathbf{l}}_j)^p. \end{aligned} \quad (12)$$

The vectors $\hat{\mathbf{n}}_i$ and $\hat{\mathbf{l}}_i$ are two perpendicular unit vectors, both perpendicular to $\hat{\mathbf{u}}_i$, see Fig. 4. They are associated with the regions on the surface involved in van der Waals attraction and hydrogen bonding, respectively (Fig. 5). Note that the hydrogen bonds are highly directional, as is accounted for by the power $p = 6$ for the g_{ij} component in Eq. (12). To mimic the flatness of a β -sheet, the same value for p has been chosen in m_{ij} . In addition, this power limits the width of the potential and thereby favours square ordering over hexagonal ordering

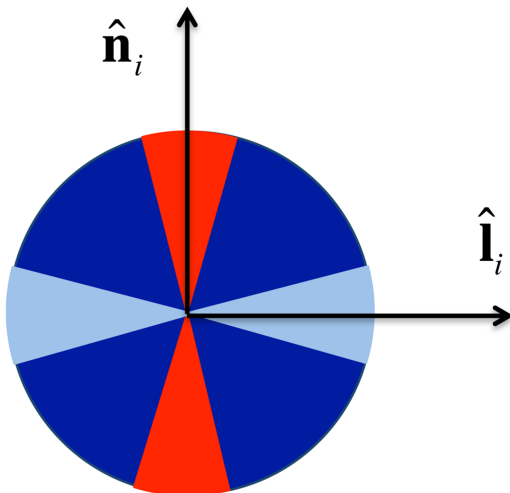


FIG. 4. A schematic representation of the orientation vectors $\hat{\mathbf{l}}_i$ and $\hat{\mathbf{n}}_i$ associated with the van der Waals (light blue surface) and the H-bonds (red surface), respectively, in a cross section of a rod. The long axis $\hat{\mathbf{u}}_i$ is into this plane.

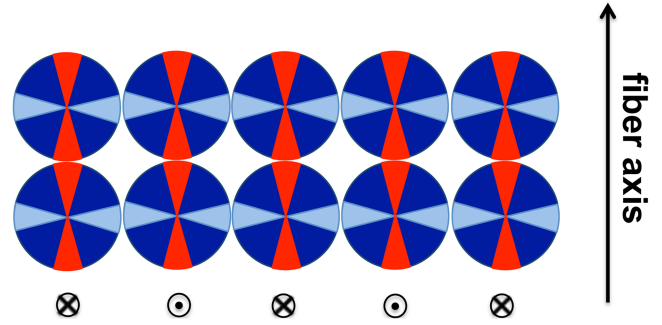


FIG. 5. Front view of two aligned folded proteins showing the strong (red-red) intermolecular hydrogen bonds and the weak (light blue-light blue) intramolecular affinities. The crosses and the dots show the arrangement of the sheets within a fiber, i.e., out of \odot and into \otimes the plane of the paper.

of the rods in a fiber. Summarising, $\eta C_{RR}^s g_{ij} f_{ij}$ represents the strong hydrogen bonds and $\eta C_{RR}^w m_{ij} f_{ij}$ the weak van der Waals and hydrophobic interactions between spherocylinders.

D. The finitely extensible nonlinear elastic (FENE) potential

To model the connecting loops between consecutive beads within the chain, we use the finitely extensible nonlinear elastic (FENE) potential. The expression for the FENE potential reads as

$$\Phi_{\text{FENE}} = -\frac{1}{2} k R_0^2 \log \left[1 - \left(\frac{db}{R_0} \right)^2 \right] \quad \text{for } 0 < db < R_0, \quad (13)$$

where db the distance between the end caps minus a margin of $0.1\bar{D}$; with $b_{i,i+1}$ the end to end distance between the ends of the cylindrical segments of two consecutive particles, see Fig. 6, $db = b_{i,i+1} - 1.1\bar{D}$. The limit to the bond extension, R_0 , is calculated based on the number of amino acids in a connection loop. We start from the assumption that the length of one amino acid is about 0.37 nm.⁵⁶ This implies a maximum distance between the centres of mass of two chameleon particles of 12 amino acids, each of $R'_0 = 12 \times 0.37 = 4.44$ nm. The limit to the bond extension is then obtained by subtracting the dimensions of the particles and their end caps, i.e., $R_0 = R'_0 - (\bar{L} - \bar{D}) - 1.1\bar{D}$. It reduces to $R_0 = 1.86$ nm for the maximum distance between the surfaces of two spheres and to $R_0 = 0.8$ nm for the maximum extensibility between the surfaces of the end caps of the two spherocylinders. To estimate the order of magnitude of the stiffness k of the spring, we approximate the connector as a polymer of length R_0 and a persistence length of $l_p = 0.3$ nm. The latter value matches

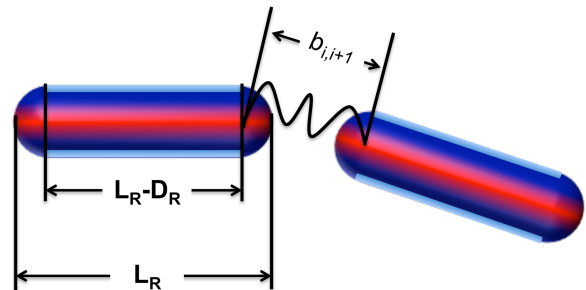


FIG. 6. Cartoon of two aligned folded proteins showing the strong (red-red) intermolecular interactions and the weak (light blue-light blue) intramolecular affinities.

experimental and computational findings for disordered α -synuclein at this temperature.⁵⁷ For a chain of 12 AA, the estimated order of magnitude of the spring constant for disordered proteins⁵⁸ is about $k = 3k_B T / (2R_0 l_p) = 1 k_B T / \text{nm}^2$. For a chain of 4 AA between two spherocylindrical particles, $k = 3 k_B T / \text{nm}^2$. To avoid computational problems, we linearized the FENE potential for distances beyond $0.95R_0$. Since test runs revealed that linkers with $k = 3 k_B T / \text{nm}^2$, in combination with the linearization just mentioned, allow too large distances between the tethered particles, we use $k = 10 k_B T / \text{nm}^2$ in the simulations.

E. The torsion potential

The previously introduced attractive potential accounts for the weak interactions between all spherocylindrical particles in the system. In a β -pleated-sheet, the hairpin, i.e., the connection between the rods, limits the freedom of the particles to rotate around their long axes. This prevents the formation of internal hydrogen bonds between consecutive β -strand sequences. In order to mimic this, we added a torsion potential on the consecutive particles i and $i + 1$ in the same molecule, which reads as

$$\frac{\Phi_{\text{tors}}}{k_B T} = \varepsilon_{\text{tors}} (\hat{\mathbf{n}}_i \cdot \hat{\mathbf{r}}_{i,i+1})^6 \cdot (\hat{\mathbf{n}}_{i+1} \cdot \hat{\mathbf{r}}_{i,i+1})^6, \quad (14)$$

with $\varepsilon_{\text{tors}}$ the strength.

III. CONSTRAINED BROWNIAN DYNAMICS

A. Equation of motion

To describe the dynamics of the particles in the system, we use Brownian Dynamics simulations. This method has been developed to efficiently simulate the motion of colloidal particles in a solvent by including the collisions with the (non-simulated) solvent molecules in the equations of motion. We use Brownian Dynamics equations of motion to describe the translational and rotational motions of anisotropic particles as well as to characterise the evolution of the internal coordinate λ . Here we present only the generalised equation of motion as the expanded equations have been previously introduced.^{40,41} Next we focus on free-energy calculations from the simulations.

The generalised Brownian Dynamics equation of motion^{59,60} reads as

$$\begin{aligned} \mathbf{Q}(t + \Delta t) - \mathbf{Q}(t) = & -\boldsymbol{\mu}^Q \frac{\partial \mathcal{A}}{\partial \mathbf{Q}} \Delta t \\ & + k_B T \frac{\partial}{\partial \mathbf{Q}} \cdot \boldsymbol{\mu}^Q \Delta t \\ & + (\boldsymbol{\mu}^Q)^{1/2} \boldsymbol{\Theta}^Q(t) \sqrt{2k_B T \Delta t}, \end{aligned} \quad (15)$$

with \mathbf{Q} the full set of generalised coordinates. Since translation, rotation and internal motions are sufficiently decoupled, for \mathbf{Q} one may read the position of a particle, \mathbf{R} , the quaternions \mathbf{q} representing the orientation of a particle or the internal coordinate λ , depending on the type of motion Eq. (15) is being applied to. The first term on the r.h.s. describes the displacement over a time step Δt which arises from the balance between a conservative force $\mathcal{F} = -\partial \mathcal{A} / \partial \mathbf{Q}$ and the opposing friction from collisions with the solvent molecules, where

\mathcal{A} represents the free energy as a function of the generalised coordinates. The second term on the r.h.s. represents a drift term accounting for the inhomogeneity of the mobility tensor, which is required in the Itô representation to recover the correct Boltzmann probability distribution. The last contribution on the r.h.s. represents the stochastic displacement that replaces the random kicks from explicit solvent molecules. Here $\boldsymbol{\Theta}^Q$ is a time-dependent Markovian vector, and its elements have zero mean, unit variance, no correlations across them and no memory. The random displacements are related to the mobility matrix $\boldsymbol{\mu}^Q$ via the fluctuation-dissipation theorem, as given by the last term with k_B the Boltzmann constant and T the temperature.

To study the attachment of a protein to a fiber, we simulate a system in which a stable fibril of 10 proteins interacts with the 11th protein binding to the fiber. To analyse the attachment free energy, we fix the distance between the centres of mass of the fiber, $\mathbf{R}_{\mathcal{F}}$, and the 11th protein, $\mathbf{R}_{\mathcal{P}}$, at $d = |\mathbf{R}_{\mathcal{P}} - \mathbf{R}_{\mathcal{F}}|$. Here we define the generic centre of mass calculation as $\mathbf{R} = \sum_{i=1}^N \mathbf{r}_i / N$ with N the number of particles in an aggregate, i.e., 50 for a fiber and 5 for the 11th protein. This distance restriction is realised by a constraint on the translational equation of motion and has no direct effect on the rotational or the internal dynamics of the particles.

The unconstrained positions of the particles after a BD step are given by⁴¹

$$\begin{aligned} \mathbf{R}_i^u(t + \Delta t) = & \mathbf{R}_i^c(t) + \boldsymbol{\mu}_i^{\text{ts}}(t) \mathbf{F}_i(t) \Delta t \\ & + \sqrt{\boldsymbol{\mu}_i^{\text{ts}}(t)} \boldsymbol{\Theta}_i^t(t) \sqrt{2k_B T \Delta t}, \end{aligned} \quad (16)$$

with \mathbf{R}^c the constrained position at the previous time step, $\boldsymbol{\mu}^{\text{ts}}$ the translational mobility matrix in the space frame, $\mathbf{F}_i(t) = -\partial \Phi / \partial \mathbf{R}_i^c(t)$ the conservative force acting on the particle for a force field Φ , and $\boldsymbol{\Theta}_i^t$ the time dependent Markovian vector. Imposing the constraint on the distance, the positions change into

$$\mathbf{R}_i^c(t + \Delta t) = \mathbf{R}_i^u(t + \Delta t) + \Lambda(t) \boldsymbol{\mu}_i^{\text{ts}}(t) \nabla_i d(t), \quad (17)$$

with Λ a Lagrange multiplier and $\nabla_i d$ providing the direction of the constraint force, the latter two being evaluated at the previous time and the former being solved from the condition $|\mathbf{R}_{\mathcal{P}}^c(t + \Delta t) - \mathbf{R}_{\mathcal{F}}^c(t + \Delta t)| = d$. Using $\mathbf{d}^u(t + \Delta t) = \mathbf{R}_{\mathcal{P}}^u(t + \Delta t) - \mathbf{R}_{\mathcal{F}}^u(t + \Delta t)$ as the vector connecting the positions of the unconstrained centres of mass at time $t + \Delta t$, and $\mathbf{d}^c(t) = \mathbf{R}_{\mathcal{P}}^c(t) - \mathbf{R}_{\mathcal{F}}^c(t)$ the corresponding vector after the previous constrained time step, the value of Λ is determined from the quadratic equation

$$|\mathbf{d}^u(t + \Delta t)|^2 + 2\Lambda(t) \mathbf{d}^u(t + \Delta t) \cdot \nabla \mathbf{d}^c(t) + \Lambda^2(t) |\nabla \mathbf{d}^c(t)|^2 = d^2. \quad (18)$$

The constraint term is then introduced in Eq. (17) to determine the updated positions of all particles.

B. Free energy

In simulations of reacting systems, the determination of the equilibrium constant and the reaction rate is usually achieved by calculating the free energy as a function of the reaction coordinate.⁶¹ The free energy along d is defined as

$$\mathcal{A}(\bar{d}) = -k_B T \ln Q(\bar{d}) + \mathcal{A}_0, \quad (19)$$

where \mathcal{A}_0 defines the zero point and the partition function $Q(\bar{d})$ is given by the integral of the Boltzmann factor over a hyperplane defined by the reaction coordinate having a specific value \bar{d} ,⁶²

$$Q(\bar{d}) = \int e^{-\beta\Phi(\mathbf{Q})} \delta[d(\mathbf{Q}) - \bar{d}] d\mathbf{Q}. \quad (20)$$

Evaluating $\mathcal{A}(\bar{d})$ directly is impossible, but one may recover $\mathcal{A}(\bar{d})$, up to an inconsequential constant, by integrating its derivative,⁶³

$$\begin{aligned} \frac{d\mathcal{A}(\bar{d})}{d\bar{d}} = & \left\langle g_d^{-1} \sum_{i=1}^N \nabla_i d \cdot \nabla_i \Phi - k_B T g_d^{-1} \sum_{i=1}^N \nabla_i^2 d \right. \\ & \left. + k_B T g_d^{-2} \sum_{i=1}^N \nabla_i d \cdot \nabla_i g_d \right\rangle_{\bar{d}}, \end{aligned} \quad (21)$$

with $g_d = \sum_{i=1}^N \nabla_i d \cdot \nabla_i d$ an element of the metric tensor, Φ the interaction potential, and $\langle \dots \rangle_{\bar{d}}$ represents the canonical average over configurations with $d(\mathbf{Q}) = \bar{d}$ irrespective of their velocity along the reaction coordinate, i.e., irrespective of

$$\Delta d = \sum \frac{\partial d}{\partial \mathbf{R}_i} \cdot [\mathbf{R}_i(t + \Delta t) - \mathbf{R}_i(t)]. \quad (22)$$

The transition region between reactants and products is often so highly elevated in free energy that it can only be sampled efficiently in biased simulations. To enhance the sampling efficiency, we constrain the coordinate d to the fixed value \bar{d} , in which case the conditional average of a function f is given by^{64,65}

$$\langle f \rangle_{\bar{d}} = \frac{\langle f g_d^{-1/2} \rangle_{\bar{d}}^c}{\langle g_d^{-1/2} \rangle_{\bar{d}}^c}, \quad (23)$$

where $\langle \dots \rangle_{\bar{d}}^c$ denotes the average over a constrained BD simulation and the factor $g_d^{-1/2}$ corrects for the bias introduced by $\Delta d = 0$. As a second route to the free energy, we note that in molecular dynamics simulations, the free-energy can be derived from the average constraint force, including some corrections related to the mass-metric.^{62,63,66} These corrections are typically rather minor, and thereby suggest that free energies can be likewise obtained from constrained Brownian Dynamics simulations,

$$\frac{d\mathcal{A}(\bar{d})}{d\bar{d}} \approx \frac{\langle \Lambda g_d^{-1/2} \rangle_{\bar{d}}^c}{\langle g_d^{-1/2} \rangle_{\bar{d}}^c}. \quad (24)$$

The good numerical agreement with the results obtained through Eq. (21) confirms the validity of this relation.

IV. RESULTS

A. Simulation setup

All simulations were performed with water as the implicit solvent, having a viscosity of $7 \cdot 10^{-4}$ Pa s, at a temperature of 37 °C, corresponding to body temperature. The simulation parameters for the attractive potential were chosen such that fibers consisting of 10 proteins are stable. This was achieved by choosing the constants in the attractive potential according to $C_{RR}^s = 20$, $C_{RR}^w = 10$, and $C_{SR} = 0.5$. Following existing theory on the diffusion coefficients of spherocylindrical particles,

the translational and rotational mobility tensors are easily calculated and then used in the equations of motion.^{40,67} Note that the mobilities reduce for a sphere to the well-known Stokes-Debye-Einstein relations.⁶⁸ The time step was fixed at $\Delta t = 1.7 \cdot 10^{-15}$ s, as this proved to be the maximum time step permitted by the repulsive part of the potential. In this respect the time step limitation is similar to that in atomistic simulations. However, in the present coarse-grained simulations, we need less particles to describe longer length scales. The internal well depths for the spherical and spherocylindrical particles were chosen such that the spherical state is favoured over the spherocylindrical state: $\epsilon_\lambda^S = 5$ and $\epsilon_\lambda^R = 4$. In reality the internal potential is most probably much more asymmetrical in that the difference between the two internal well depths is much larger, with a much deeper well for the disordered state. In that case particles would hardly ever overcome the internal energy barrier, thereby making the use of Brownian Dynamics impossible. We therefore chose the above values to enable efficient Brownian Dynamics simulations, while still capturing the essential physics. The mobility of the internal coordinate was set at $\mu^\lambda = 1.4 \cdot 10^{32}$ (J s)⁻¹ which ensures in combination with Φ_λ that a free particle transforms about once per nanosecond. With these parameters and $\epsilon_{\text{tors}} = 0$, an isolated protein relaxes into a sheet with four sets of internal hydrogen bonds. The formation of these internal hydrogen bond is inhibited when the torsional stiffness of the hairpins is increased to $\epsilon_{\text{tors}} = 10$, in which case an isolated protein adopts a dynamically evolving largely unstructured state. Fibrillar arrangements appear stable for both values of ϵ_{tors} .

B. Results—Free energy profile

In the first set of simulations, systems were initialised by placing a single ordered molecule at a distance d above the center of mass of a fiber consisting of 10 ordered proteins in the centre of the box, see Fig. 7(a). We performed simulations in which the centre of mass of the 11th protein was constrained at distances between 4.3 and 11 nm from the centre of mass of the fiber, both in the presence and absence of a torsion potential. At each distance, the system was equilibrated for about 0.1 μ s followed by a production run of typically 1 μ s.

Fig. 8 shows the free-energy profile as a function of the constrained distance for systems with no torsion potential. This plot shows excellent agreement between the two curves obtained by integrating either Eq. (24), in black, or Eq. (21), in red. The free energy at the minimum near $d = 4.9$ nm has been set equal to zero. This distance is also the equilibrium distance between the protein and the fiber in simulations without distance constraint. In this minimum, all five particles of the protein are in the spherocylindrical state and sequentially arranged with their orientation vectors $\hat{\mathbf{u}}_i$ pointing in alternating directions, to form the 11th layer of the fiber. With decreasing distance, the free energy increases as the protein is effectively being “pushed” into the fiber. A transition occurs for distances below 4.5 nm: in order to obey the distance constraint, while simultaneously keeping as many particles as possible near their preferred locations at the top of the fiber, two terminal particles of the protein dislodge from the top of the fiber and adopt positions along the flank of the fiber, thereby stretching the internal bonds by which these two

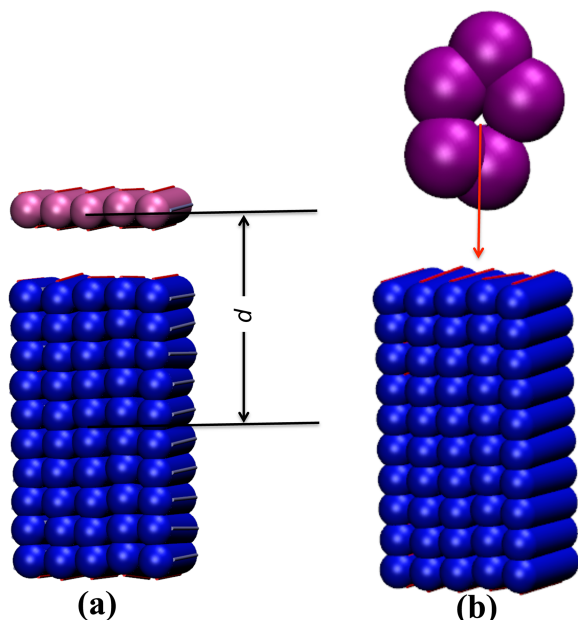


FIG. 7. Representation of the system setup. The free energy profile is calculated as a function of the distance d between the center of mass of a fiber of 10 proteins, in blue, and the center of mass of the free protein, in pink. The first two sets of simulations correspond to an initial state where the isolated protein is neatly folded (a). In the third set of simulations we start from a completely random free protein (b) which is then stepwise moved toward the fiber as indicated by the red arrow.

particles are attached to their respective neighbours, see snapshot in Fig. 8.

For distances exceeding the equilibrium distance, when the protein is pulled away from the fiber, a similar competition arises between the constraint, contacts with the fiber, and intramolecular interactions. With increasing distance, first the outer particles move upwards to form a partial 12th layer of the fiber, so that they can still bind by strong hydrogen bonding interactions to the particles remaining in the 11th layer. To further improve the interaction, the spherocylinders of the protein and those of the top of the fiber are tilted. Next, the particles of the protein are one by one “peeled” off from the fiber, and they subsequently transform into the random state. From the snapshot at $d = 7.5$ nm, see Fig. 8, it is clear that the

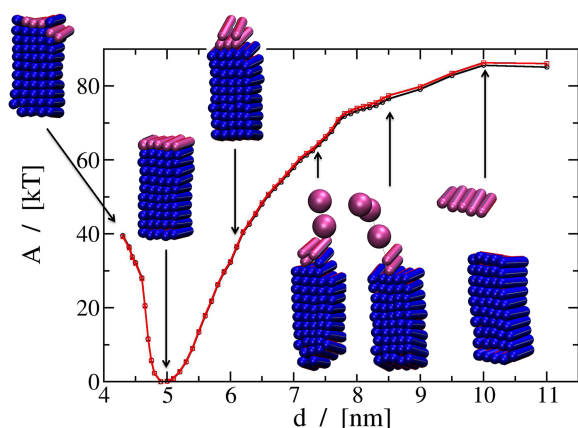


FIG. 8. The free-energy profile calculated from Eq. (21) (red line) and Eq. (24) (black line) for a system with $C_{RR}^s = 20$, $C_{RR}^w = 10$, $C_{SR} = 0.5$, and $\varepsilon_{tors} = 0$. The snapshots show the behaviour of the system at the indicated distances.

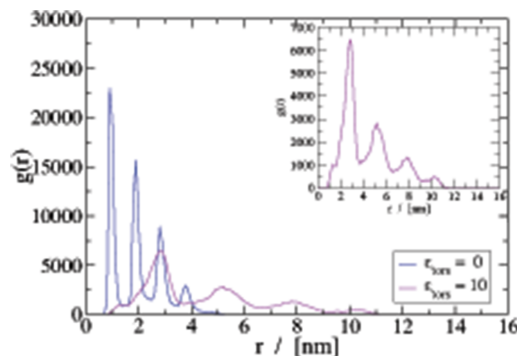


FIG. 9. The pair correlation histogram $g(r)$ between the five particles of a free protein, with the protein folding into a sheet with internal hydrogen bonds for $\varepsilon_{tors} = 0$ (blue line) and evolving through disordered and partially ordered conformations at $\varepsilon_{tors} = 10$ (magenta line).

fiber is being deformed not only by the protein pulling at the top but also by elongation at the bottom to maintain the desired constrained distance between the centers of mass. The smallest distance at which the protein ceases to interact with the fiber, $d_0 \approx 10$ nm, appears as a maximum in the free energy; beyond this distance the free energy is expected to gradually decrease following $\mathcal{A}(d) = \mathcal{A}(d_0) - k_B T \ln(d^2/d_0^2)$. For $\varepsilon_{tors} = 0$, the particles of the unbound protein adopt the spherocylindrical state and rearrange to form intramolecular hydrogen bonds, see the snapshot in Fig. 8 and the regular spacings in the distance in Fig. 9, while the constraint causes the protein to diffuse over a spherical surface around the fibre’s center.

The folding of the free protein is suppressed when the flexibility of the hairpins connecting the β -strands is reduced by the torsion potential of Eq. (14) with $\varepsilon_{tors} = 10$. The free energy profile calculated for this model is shown in Fig. 10. The most stable structure again occurs at $d = 4.9$ nm and is virtually identical to that in the previous case. In particular, the contributions of the torsion potentials are almost zero. The minimum of the free energy curve, again taken as the zero-point of the graph, therefore corresponds to virtually the same absolute free energy as the minima of Fig. 8. The overall picture of the free-energy profile is very similar to that calculated in the absence of this torsion potential, except that the maximum has increased by about $20 k_B T$. The reason for this is

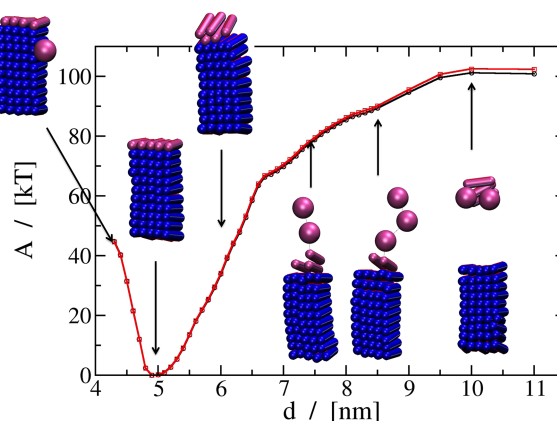


FIG. 10. The free-energy profile calculated from Eq. (21) (magenta line) and Eq. (24) (black line) for a system with $C_{RR}^s = 20$, $C_{RR}^w = 10$, $C_{SR} = 0.5$, and $\varepsilon_{tors} = 10$. The snapshots show the behaviour of the system at the indicated distances.

that the torsion potential has largely removed the possibility to form internal hydrogen bonds in the isolated molecule. In fact, the remaining weak van der Waals bonds between the particles in their folded states are not strong enough to prevent part of the beads from turning spherical, as indicated in Fig. 10 by the snapshot at $d = 10$ nm. The detached protein was observed to form 2 to 3 dynamic internal bonds while the remaining segments adopt dynamical disordered conformations, as is confirmed by the distance histogram in Fig. 9. For constrained distances larger than 5 nm, initially the protein retains its planar structure, until its constituent beads are again peeled off one by one. The added stability provided by the torsion makes the fiber retain its structure much better than in the case with no torsion potential, as no significant structural modifications in the fiber are visible, see Fig. 10. To further test the influence of the torsion potential, we increased its strength to $\epsilon_{\text{tors}} = 20$, but this did not result in any major differences in the free-energy profile (data not shown).

In a third set of simulations, we placed a completely disordered protein at a large distance ($d > 11$ nm) from the center of mass of the fiber and allowed the system to equilibrate, see Fig. 7(b). We weakened the hydrogen bonding to $C_{\text{RR}}^{\text{s}} = 15$, which still ensures the stability of the fiber but is expected to reduce possible trapping of the protein in a mis-folded state on the fiber. After equilibration, the protein is stepwise moved to the fiber, using the final frame of the current time as the initial frame of the next time. As in the previous simulations, at large distances the protein diffuses on a sphere around the fiber. To avoid lateral attaching of the protein to the fiber, in some runs at large distance we returned the protein to a position along the long axis of the fiber before decreasing the distance between the centers of mass. Upon approaching the fiber, the segments of the protein gradually adapt conformation and bind to the fiber. First one segment becomes spherocylindrical and binds followed by a step-by-step transformation and binding of the other segments as the distance is reduced. This behaviour is shared by systems with $\epsilon_{\text{tors}} = 0$ and $\epsilon_{\text{tors}} = 10$ (data not shown). In Fig. 11 the free energy when starting the simulations at every distance with a neatly folded protein (blue line) is compared with that of an initially disordered protein that is stepwise

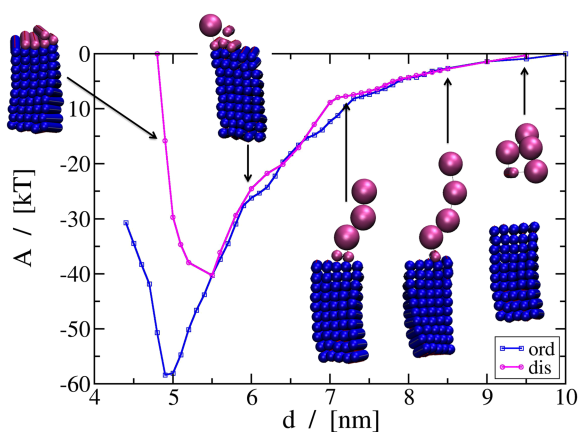


FIG. 11. The free-energy profile for a system with $C_{\text{RR}}^{\text{s}} = 15$, $C_{\text{RR}}^{\text{w}} = 10$, $C_{\text{SR}} = 0.5$, and $\epsilon_{\text{tors}} = 10$ upon stepwise approaching an initially disordered protein to a fiber (magenta line) and restarting every distance with an optimally configured protein (blue line).

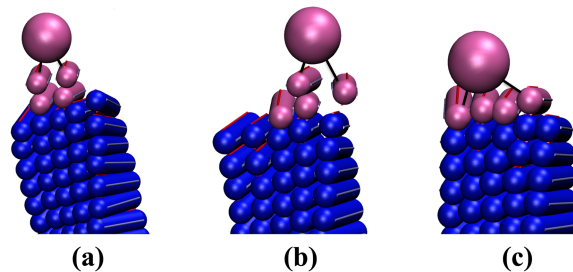


FIG. 12. Examples of proteins trapped in local minima, at constrained distances of 5.8 nm (a), 5.9 nm (b), and 5.4 nm (c), when stepwise brought toward the fiber.

moved to the fiber in a series of simulations (magenta line), for $C_{\text{RR}}^{\text{s}} = 15$ and $C_{\text{RR}}^{\text{w}} = 10$ in the presence of torsion, $\epsilon_{\text{tors}} = 10$. The different minima of the two lines indicate that the initially disordered protein became trapped in a local minimum. We repeated the approach procedure several times, in each case ending in a local minimum. Visual inspection of the system at these minima showed that the 11th protein easily attaches to the fiber in a sub-optimal configuration from where there is little chance to recover. For instance, in some cases the two sequentially neighbouring beads $i \pm 1$ of a bead i attach as two adjacent spherocylinders to a fiber leaving the particle i in a spherical state dangling on top or next to the fiber, see Figs. 12(a) and 12(b). Similar behaviour was also observed for segments that are more than one particle apart, see Fig. 12(c). The simulations indicate that it is easy to attach in a sub-optimal manner, while it is difficult to re-arrange these configurations into the optimal alignment.

V. DISCUSSION

Highly ordered filaments of α -synuclein are known to form intracellular inclusions in the brains of patients with Parkinson's disease. The mechanisms by which these filaments grow and how a protein attaches on the surface of a fiber are not fully understood. Experimental work and small scale simulations on other proteins or fragments thereof describe a generic amyloidogenic “dock-lock” mechanism of fibril elongation.^{69–74} These studies show that monomer attachment to the fibrillar end is a two-step process in which the free monomer rapidly binds to the surface of the fiber (dock) followed by a slow rearrangement towards the structure of the fiber template (lock).^{69,71–73,75} Atomistic and slightly coarse-grained simulations on small fragments of various amyloid forming proteins revealed the small scale details of the restructuring of docked proteins.^{71–76} The free-energy profile of A β 35-40 docking and locking to a fiber were calculated using all-atom simulations in implicit and explicit solvent.⁷² Furthermore, it was shown that a united residue A β 1-40 locks to a fiber by the slow binding of one β -strand followed by the rapid binding of the second β -strand.⁷⁶ Studies on fragments of amyloid- β revealed that the break-up of strand-loop-strand motifs in the attaching monomer plays a key role in a lock phase lasting three to four orders of magnitude longer than the dock phase.⁷³

While simulations just mentioned focus on describing the specific details of the dock-lock attachment mechanism,

we aim at describing the attachment of the much larger α -synuclein protein to a fiber. To this end we introduced a coarse-grain model representing the protein as a chain of polymorphic patchy particles which retain the real protein's binding affinities. The simulations show that this type of model is computationally feasible and captures the transitions in secondary, tertiary, and quaternary structures induced by interactions between proteins. The calculated monotonic decrease of the free-energy profile between the partially ordered protein in solution, $d \geq 10$ nm, and the ordered state in a fibril, $d \approx 5$ nm, would suggest that attachment to the fibril is a diffusion limited process in the current model. This agrees with existing small scale simulations which suggest a diffusion limited locking followed by slow docking.^{72,73} Yet experimental findings imply there may be a rate-limiting barrier.^{43,77} This discrepancy might arise due to the absence of the N- and C-terminal regions in the current coarse-grain model, which only considers the central ≈ 60 amino acids of the protein. The fibril growth rate of the truncated protein is known to be much higher than that of the full protein, in which the highly charged C-terminus is inhibiting the fibrillation process.⁴⁹

More interestingly, and inspired by the dock-lock mechanism discussed above, we suggest an alternative interpretation of the experimental findings, bringing simulations and experiments closer together. While simulations starting with ordered proteins at every intermediate distance between the free protein and the fibril generally lead to a monotonic free energy profile, other simulations starting with a disordered protein placed at a large distance, which is next slowly moved towards the fibril, indicate that the approaching protein easily gets trapped in a sub-optimal configuration. In the latter case, the runs performed at each of the intermediate distances before moving the free protein to a smaller distance, are too short to sample the relevant parts of phase space. The same may happen in experiments when diffusion of the free protein towards the fibril is fast enough compared to its internal dynamics. The protein will get trapped in a metastable minimum of the free energy from which it is difficult to escape, like it occurs in the magenta curve in Fig. 11. Only after the protein has released itself from the metastable state will it be able to properly attach to the fibril, thereby offering a new template for further growth. In the case the time needed to escape from the metastable state is within the time resolution of the experiment, this will be perceived as a “stop-and go” process. Such a stop-and-go mechanism has indeed been observed in various experimental studies,^{16–19,78} in which growth disruptions of amyloid fibrils are observed, with long pauses (“stop”) interrupting growth periods (“go”). Experiments over several hours show a nearly exponential distribution of stop periods, with many short stop periods and only few long ones, see Fig. 4 in Ref. 16. One may easily envision to extrapolate these results to shorter time scales and consider very short stop periods, which with the experimental resolution available have not been recognised as such. An abundance of short stop periods will slow down the growth rate which will be interpreted as a kinetic barrier to the final attached state. The linearisation of Fig. 3 in Ref. 16 at early times in order to get growth rates depends on the time resolution of the experiment. With a coarser resolution, no stop

periods will be discerned and a much smaller growth rate will be attributed to the process, from which much larger kinetic barriers will be deduced.

The above discussion clearly reveals that the distance between the approaching protein and the fibril is not an optimal reaction coordinate by itself. A better choice should include also some structural information about the attaching protein discriminating between correct and incorrect attachment. Together with the distance used in this paper, such a coordinate would provide a two dimensional free energy surface that allows a better description of the attaching process. A very rough example of such a model, reducing the distance coordinate to “attached” or “detached” and the structural information to “correct” or “incorrect,” was recently proposed in an experimental work by Wördehoff *et al.*¹⁶ Further experimental information on activation barriers of amyloidogenic proteins was presented by Buell *et al.*⁴³ Calculating these and similar results from simulations requires a huge expense of CPU time, which we intend to do once improved parameters in our model have been obtained from small scale simulations for the various fragments constituting the protein.

One quantity that can easily be obtained from our data is the detachment free energy of one protein to the fibril. This result will obviously be too large since our internal potential has deliberately been deformed in order to make Brownian Dynamics runs feasible. In the [Appendix](#) to this paper, we describe how this value may be obtained from our free energy profiles. The result is 162 kJ/mol, corresponding to the blue line in Fig. 11, as calculated by Eq. (A7) given in the [Appendix](#). The experimental value is 33 kJ/mol for the full 140 amino acid α -synuclein.⁷⁹

VI. CONCLUSIONS AND PERSPECTIVES

A coarse-grained model of α -synuclein was introduced to study the attachment of a protein to the end of a fiber. The model is highly idealised, but follows the molecular structure and the flexibility of α -synuclein and consequently manages to capture the attachment properties of a protein to a mature fiber. We performed simulations starting with ordered proteins at every intermediate distance between the free protein and the fibril and starting with a disordered protein placed at a large distance, which is slowly moved towards the fibril. Our results indicate that the approaching protein easily gets trapped in a sub-optimal configuration. The simulations thereby suggest a likely explanation for the experimentally observed stop-and-go growth of an amyloid fiber.

The intricate nature of α -synuclein makes it difficult to develop an effective coarse-grain model from scratch, and further developments are required to arrive at a more quantitative understanding of fibril growth. Having established the feasibility of modelling a protein as a chain of polymorphic particles, we can now envisage and explore possibilities to improve the current model. Experiments indicate that proteins bind in a parallel fashion along the growth direction of the fiber, yet the current model does not differentiate between the parallel and anti-parallel binding of a particle to a fiber. In addition, the five β -strands within a sheet have distinct lengths,¹⁵

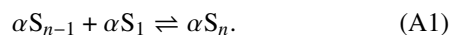
whereas we used equal lengths for all strands; including spherocylinders of distinct lengths within a chain may promote the proper alignment within a fiber. The inclusion of tails at both end of the protein is expected to slow down the growth rate. The internal free energy difference between a particle in the ordered and in the disordered states is smaller than in reality, for computational reasons. Introducing a realistic difference is a necessary improvement which will require the combination of Brownian Dynamics with Monte Carlo moves to overcome the high transition barriers. A systematic route to the force field parameters that will be a topic of future research involves matching the internal free energy profiles and the interaction energies of the coarse-grained particles with free-energy profiles sampled in relatively short fully atomistic Molecular Dynamics simulations of the peptide sequences represented by the coarse-grained particles. These topics are among the features that we plan to include in future applications of the model.

ACKNOWLEDGMENTS

This work is part of the research programme ‘‘A Single Molecule View on Protein Aggregation’’ (Project No. 10SMPA05) of the Foundation for Fundamental Research on Matter (FOM), which is part of the Netherlands Organisation for Scientific Research (NWO).

APPENDIX: FREE ENERGY CHANGE

The attachment of an α -synuclein to a fiber of length $n - 1$ and the release of a protein by a fiber of length n are described by the fibrillation reaction



The equilibrium constant of this reaction is defined as

$$K_n = \frac{(c_n/c_0)}{(c_{n-1}/c_0)(c_1/c_0)} = e^{-\Delta G_n^0/RT}, \quad (\text{A2})$$

where c_n is the concentration of n -mers, c_0 a reference concentration typically taken as 1 molar, ΔG_n^0 is the standard free energy change of the reaction, and R is the gas constant. In statistical mechanics, this equilibrium constant is calculated as^{80–82}

$$K_n = \frac{(q_n/V)}{(q_{n-1}/V)(q_1/V)} c_0, \quad (\text{A3})$$

where q_n denotes the molecular partition of an n -mer in a volume V . In terms of the conditional partition function $Q(\bar{d})$ of Eq. (20), all sampled configurations with a distance \bar{d} below the distance of the transition state, d^\ddagger , represent states with an n -protein fiber, and hence contribute to the partition function of that fiber,

$$q_n = \int_0^{d^\ddagger} Q(\bar{d}) d\bar{d}. \quad (\text{A4})$$

All sampled configurations beyond the transition state, $\bar{d} > d^\ddagger$, represent states in which a monomer is detached from a fiber of $n - 1$ proteins, and hence contribute to the product of their partition functions

$$q_{n-1} q_1 = \int_{d^\ddagger}^{\infty} Q(\bar{d}) d\bar{d}. \quad (\text{A5})$$

For large distances, where the fiber and protein do not interact, the conditional partition function $Q(\bar{d})$ becomes quadratic in \bar{d} for geometric reasons. One may then write

$$Q(\bar{d}) = Q(d_0) \frac{\bar{d}^2}{d_0^2}, \quad (\text{A6})$$

where d_0 is an arbitrary large distance beyond the interaction range. Using Eq. (19) to convert a calculated free energy profile $\mathcal{A}(\bar{d})$ into a conditional partition function $Q(\bar{d})$, where \mathcal{A}_0 may be taken as zero, the above two integrals can be evaluated. The free energy change of the reaction is obtained as

$$\Delta G_n^0 = -RT \ln \left[\frac{4\pi d_0^2}{Q(d_0)} c_0 \int_0^{d^\ddagger} Q(\bar{d}) d\bar{d} \right]. \quad (\text{A7})$$

In our calculations, d_0 is taken as the rightmost point of the free energy profile.

- ¹C. M. Dobson, *Semin. Cell Dev. Biol.* **15**, 3 (2004).
- ²R. L. Redler, D. Shirvanyants, O. Dagliyan, F. Ding, D. N. Kim, P. Kota, E. A. Proctor, S. Ramachandran, A. Tandon, and N. V. Dokholyan, *J. Mol. Cell Biol.* **6**, 104 (2014).
- ³B. Caughey and P. T. Lansbury, *Annu. Rev. Neurosci.* **26**, 267 (2003).
- ⁴H. A. Lashuel, C. R. Overk, A. Oueslati, and E. Masliah, *Nat. Rev. Neurosci.* **14**, 38 (2013).
- ⁵A. Oueslati, K. E. Paleologou, B. L. Schneider, P. Aebischer, and H. A. Lashuel, *J. Neurosci.* **32**, 1536 (2012).
- ⁶A. Recasens, B. Dehay, J. Bové, I. Carballo-Carbajal, S. Dovero, A. Pérez-Villalba, P.-O. Fernagut, J. Blesa, A. Parent, C. Perier, I. Fariñas, J. A. Obeso, E. Bezard, and M. Vila, *Ann. Neurol.* **75**, 351 (2014).
- ⁷J. T. Berryman, S. E. Radford, and S. A. Harris, *Biophys. J.* **100**, 2234 (2011).
- ⁸M. R. Smaoui, F. Poitevin, M. Delarue, P. Koehl, H. Orland, and J. Waldispühl, *Biophys. J.* **104**, 683 (2013).
- ⁹B. H. Toyama and J. S. Weissman, *Annu. Rev. Biochem.* **80**, 557 (2011).
- ¹⁰W. T. Astbury, S. Dickinson, and K. Bailey, *Biochem. J.* **29**, 2351 (1935).
- ¹¹L. C. Serpell, J. Berriman, R. Jakes, M. Goedert, and R. A. Crowther, *Proc. Natl. Acad. Sci. U. S. A.* **97**, 4897 (2000).
- ¹²M. Arai, K. Sugase, H. J. Dyson, and P. E. Wright, *Proc. Natl. Acad. Sci. U. S. A.* **112**, 9614 (2015).
- ¹³A. Der-Sarkissian, C. C. Jao, J. Chen, and R. Langen, *J. Biol. Chem.* **278**, 37530 (2003).
- ¹⁴H. Heise, W. Hoyer, S. Becker, O. C. Andronesi, D. Riedel, and M. Baldus, *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15871 (2005).
- ¹⁵M. Vilar, H.-T. Chou, T. Lührs, S. K. Maji, D. Riek-Loher, R. Verel, G. Manning, H. Stahlberg, and R. Riek, *Proc. Natl. Acad. Sci. U. S. A.* **105**, 8637 (2008).
- ¹⁶M. M. Würdehoff, O. Bannach, H. Shaykhalishahi, A. Kulawik, S. Schiefer, D. Willbold, W. Hoyer, and E. Birkmann, *J. Mol. Biol.* **427**, 1428 (2015).
- ¹⁷T. Ban, M. Hoshino, S. Takahashi, D. Hamada, K. Hasegawa, H. Naiki, and Y. Goto, *J. Mol. Biol.* **344**, 757 (2004).
- ¹⁸J. Ferkinghoff-Borg, J. Fonslet, C. B. Andersen, S. Krishna, S. Pigolotti, H. Yagi, Y. Goto, D. Otzen, and M. H. Jensen, *Phys. Rev. E* **82**, 010901 (2010).
- ¹⁹W. Hoyer, D. Cherny, V. Subramaniam, and T. M. Jovin, *J. Mol. Biol.* **340**, 127 (2004).
- ²⁰M. Fuxreiter, *Computational Approaches to Protein Dynamics: From Quantum to Coarse-Grained Methods*, Series in Computational Biophysics (Taylor & Francis, 2014).
- ²¹C. Clementi, *Curr. Opin. Struct. Biol.* **18**, 10 (2008).
- ²²V. Tozzini, *Curr. Opin. Struct. Biol.* **15**, 144 (2005).
- ²³J. Yoon, J. Park, S. Jang, K. Lee, and S. Shin, *J. Biomol. Struct. Dyn.* **25**, 505 (2008).
- ²⁴A. Irbäck and S. Mohanty, *Computational Methods to Study the Structure and Dynamics of Biomolecules and Biomolecular Processes* (Springer Berlin Heidelberg, 2014), Vol. 1, pp. 433–444.
- ²⁵L. Liu and Z. Cao, *Int. J. Mol. Sci.* **14**, 10896 (2013).
- ²⁶A. Zhmurov, O. Kononova, R. I. Litvinov, R. I. Dima, V. Barsegov, and J. W. Weisel, *J. Am. Chem. Soc.* **134**, 20396 (2012).
- ²⁷R. Qi, Y. Luo, B. Ma, R. Nussinov, and G. Wei, *Biomacromolecules* **15**, 122 (2014).

- ²⁸K. K. Sahu, M. T. Woodside, and J. A. Tuszynski, *Biochimie* **116**, 133 (2015).
- ²⁹G. Bellesia and J.-E. Shea, *J. Chem. Phys.* **131**, 111102 (2009).
- ³⁰S. Abeln and D. Frenkel, *Biophys. J.* **100**, 693 (2011).
- ³¹R. Ni, S. Abeln, M. Schor, M. A. Cohen Stuart, and P. G. Bolhuis, *Phys. Rev. Lett.* **111**, 058101 (2013).
- ³²S. Abeln, M. Vendruscolo, C. M. Dobson, and D. Frenkel, *PLoS One* **9**, e85185 (2014).
- ³³R. Pellarin, E. Guarnera, and A. Caffisch, *J. Mol. Biol.* **374**, 917 (2007).
- ³⁴R. Pellarin and A. Caffisch, *J. Mol. Biol.* **360**, 882 (2006).
- ³⁵B. Barz and B. Urbanc, *J. Phys. Chem. B* **118**, 3761 (2014).
- ³⁶N. S. Bieler, T. P. J. Knowles, D. Frenkel, and R. Vacha, *PLoS Comput. Biol.* **8**, e1002692 (2012).
- ³⁷A. Saric, Y. C. Chebaro, T. P. J. Knowles, and D. Frenkel, *Proc. Natl. Acad. Sci. U. S. A.* **111**, 17869 (2014).
- ³⁸A. Saric, A. K. Buell, G. Meisl, T. C. T. Michaels, C. M. Dobson, S. Linse, T. P. J. Knowles, and D. Frenkel, *Nat. Phys.* **12**, 874 (2016).
- ³⁹R. Vacha, S. Linse, and M. Lund, *J. Am. Chem. Soc.* **136**, 11776 (2014).
- ⁴⁰I. M. Ilie, W. K. den Otter, and W. J. Briels, *J. Chem. Phys.* **144**, 085103 (2016).
- ⁴¹I. M. Ilie, W. J. Briels, and W. K. den Otter, *J. Chem. Phys.* **142**, 114103 (2015).
- ⁴²D. Eisenberg and M. Jucker, *Cell* **148**, 1188 (2012).
- ⁴³A. K. Buell, C. Galvagnion, R. Gaspar, E. Sparr, M. Vendruscolo, T. P. J. Knowles, S. Linse, and C. M. Dobson, *Proc. Natl. Acad. Sci. U. S. A.* **111**, 7671 (2014).
- ⁴⁴N. Lorenzen, L. Lemminger, J. N. Pedersen, S. B. Nielsen, and D. E. Otzen, *FEBS Lett.* **588**, 497 (2014).
- ⁴⁵E. Jo, N. Fuller, R. Rand, P. S. George-Hyslop, and P. E. Fraser, *J. Mol. Biol.* **315**, 799 (2002).
- ⁴⁶N. C. Maiti, M. M. Apetri, M. G. Zagorski, P. R. Carey, and V. E. Anderson, *J. Am. Chem. Soc.* **126**, 2399 (2004).
- ⁴⁷V. N. Uversky, *J. Biomol. Struct. Dyn.* **21**, 211 (2003).
- ⁴⁸B. van Rooijen, K. van Leijenhof-Groener, M. Claessens, and V. Subramaniam, *J. Mol. Biol.* **394**, 826 (2009).
- ⁴⁹D.-P. Hong, W. Xiong, J.-Y. Chang, and C. Jiang, *FEBS Lett.* **585**, 561 (2011).
- ⁵⁰V. N. Uversky, J. Li, and A. L. Fink, *J. Biol. Chem.* **276**, 10737 (2001).
- ⁵¹S. L. Bernstein, D. Liu, T. Wyttenbach, M. T. Bowers, J. C. Lee, H. B. Gray, and J. R. Winkler, *J. Am. Soc. Mass Spectrom.* **15**, 1435 (2004).
- ⁵²J.-E. Suk, S. B. Lokappa, and T. S. Ulmer, *Biochemistry* **49**, 1533 (2010).
- ⁵³C. Vega and S. Lago, *Comput. Chem.* **18**, 55 (1994).
- ⁵⁴M. P. Allen, G. T. Evans, D. Frenkel, and B. M. Mulder, "Hard convex body fluids," in *Advances in Chemical Physics* (John Wiley & Sons, Inc., 2007), pp. 1–166.
- ⁵⁵A. W. P. Fitzpatrick, G. M. Vanacore, and A. H. Zewail, *Proc. Natl. Acad. Sci. U. S. A.* **112**, 3380 (2015).
- ⁵⁶A. Lehninger, D. L. Nelson, and M. M. Cox, *Lehninger Principles of Biochemistry*, 5th ed. (W. H. Freeman, 2008).
- ⁵⁷C. Narayanan, D. S. Weinstock, K.-P. Wu, J. Baum, and R. M. Levy, *J. Chem. Theory Comput.* **8**, 3929 (2012).
- ⁵⁸D. Boal, *Mechanics of the Cell*, 2nd ed. (Cambridge University Press, Cambridge, UK, 2002).
- ⁵⁹H. C. Öttinger, *Stochastic Processes in Polymeric Fluids* (Springer, 1996).
- ⁶⁰C. W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*, Springer Series in Synergetics, 3rd ed. (Springer-Verlag, Berlin, 2004), Vol. 13, p. xviii+415.
- ⁶¹D. Frenkel and B. Smit, *Understanding Molecular Simulation*, 2nd ed. (Academic Press, Inc., Orlando, FL, USA, 2001).
- ⁶²W. K. den Otter and W. J. Briels, *Mol. Phys.* **98**, 773 (2000).
- ⁶³W. K. den Otter, *J. Chem. Phys.* **112**, 7283 (2000).
- ⁶⁴D. C. Morse, *Adv. Chem. Phys.* **128**, 65 (2004).
- ⁶⁵E. J. Hinch, *J. Fluid Mech.* **271**, 219 (1994).
- ⁶⁶W. K. den Otter and W. J. Briels, *J. Chem. Phys.* **109**, 4139 (1998).
- ⁶⁷S. R. Aragon and D. Flamik, *Macromolecules* **42**, 6290 (2009).
- ⁶⁸G. H. Koenderink and A. P. Philipse, *Langmuir* **16**, 5631 (2000).
- ⁶⁹W. P. Esler, E. R. Stimson, J. M. Jennings, H. V. Vinters, J. R. Ghilardi, J. P. Lee, P. W. Mantyh, and J. E. Maggio, *Biochemistry* **39**, 6288 (2000).
- ⁷⁰M. J. Cannon, A. D. Williams, R. Wetzal, and D. G. Myszka, *Anal. Biochem.* **328**, 67 (2004).
- ⁷¹P. H. Nguyen, M. S. Li, G. Stock, J. E. Straub, and D. Thirumalai, *Proc. Natl. Acad. Sci. U. S. A.* **104**, 111 (2007).
- ⁷²E. P. O'Brien, Y. Okamoto, J. E. Straub, B. R. Brooks, and D. Thirumalai, *J. Phys. Chem. B* **113**, 14421 (2009).
- ⁷³W. Han and K. Schulten, *J. Am. Chem. Soc.* **136**, 12450 (2014).
- ⁷⁴M. Schor, A. S. J. S. Mey, F. Noé, and C. E. MacPhee, *J. Phys. Chem. Lett.* **6**, 1076 (2015).
- ⁷⁵G. Reddy, J. E. Straub, and D. Thirumalai, *Proc. Natl. Acad. Sci. U. S. A.* **106**, 11948 (2009).
- ⁷⁶A. Rojas, A. Liwo, D. Browne, and H. A. Scheraga, *J. Mol. Biol.* **404**, 537 (2010).
- ⁷⁷V. V. Shvachak, M. M. A. E. Claessens, and V. Subramaniam, *J. Phys. Chem. B* **119**, 1912 (2015).
- ⁷⁸D. Pinotsi, A. K. Buell, C. Galvagnion, C. M. Dobson, G. S. Kaminski Schierle, and C. F. Kaminski, *Nano Lett.* **14**, 339 (2014).
- ⁷⁹A. J. Baldwin, T. P. J. Knowles, G. G. Tartaglia, A. W. Fitzpatrick, G. L. Devlin, S. L. Shamma, C. A. Waudby, M. F. Mossuto, S. Meehan, S. L. Gras, J. Christodoulou, S. J. Anthony-Cahill, P. D. Barker, M. Vendruscolo, and C. M. Dobson, *J. Am. Chem. Soc.* **133**, 14160 (2011).
- ⁸⁰D. A. McQuarrie, *Statistical Mechanics* (Harper & Row Publishers, New York, NY, USA, 1976).
- ⁸¹P. C. Nelson, *Biological Physics*, updated 1st ed. (W. H. Freeman and Company, New York, NY, USA, 2008).
- ⁸²R. Phillips, J. Kondev, and J. Theriot, *Physical Biology of the Cell* (Garlan Science, New York, NY, USA, 2009).