AIP | The Journal of Chemical Physics

A coarse grained protein model with internal degrees of freedom. Application to α-synuclein aggregation

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

Citation: The Journal of Chemical Physics **144**, 085103 (2016); doi: 10.1063/1.4942115
View online: http://dx.doi.org/10.1063/1.4942115
View Table of Contents: http://scitation.aip.org/content/aip/journal/jcp/144/8?ver=pdfcov
Published by the AIP Publishing

THE JOURNAL OF CHEMICAL PHYSICS **144**, 085103 (2016)

A coarse grained protein model with internal degrees of freedom. Application to α-synuclein aggregation

Ioana M. Ilie,[1,2,a)] Wouter K. den Otter,[1,2,3,a)] and Wim J. Briels[1,2,4,a)]
[1]*Computational Chemical Physics, Faculty of Science and Technology, University of Twente, P.O. Box 217, 7500 AE, Enschede, The Netherlands*
[2]*MESA+ Institute for Nanotechnology, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands*
[3]*Multi Scale Mechanics, Faculty of Engineering Technology, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands*
[4]*Forschungszentrum Jülich, ICS, D-52425 Jülich, Germany*

(Received 6 November 2015; accepted 4 February 2016; published online 23 February 2016)

Particles in simulations are traditionally endowed with fixed interactions. While this is appropriate for particles representing atoms or molecules, objects with significant internal dynamics—like sequences of amino acids or even an entire protein—are poorly modelled by invariable particles. We develop a highly coarse grained polymorph patchy particle with the ultimate aim of simulating proteins as chains of particles at the secondary structure level. Conformational changes, e.g., a transition between disordered and β-sheet states, are accommodated by internal coordinates that determine the shape and interaction characteristics of the particles. The internal coordinates, as well as the particle positions and orientations, are propagated by Brownian Dynamics in response to their local environment. As an example of the potential offered by polymorph particles, we model the amyloidogenic intrinsically disordered protein α-synuclein, involved in Parkinson's disease, as a single particle with two internal states. The simulations yield oligomers of particles in the disordered state and fibrils of particles in the "misfolded" cross-β-sheet state. The aggregation dynamics is complex, as aggregates can form by a direct nucleation-and-growth mechanism and by two-step-nucleation through conversions between the two cluster types. The aggregation dynamics is complex, with fibrils formed by direct nucleation-and-growth, by two-step-nucleation through the conversion of an oligomer and by auto-catalysis of this conversion. © *2016 AIP Publishing LLC.* [http://dx.doi.org/10.1063/1.4942115]

I. INTRODUCTION

Proteins are long chains of amino acids that regulate many processes in living cells. In order to perform their biological functions, proteins undergo a folding process to form a three dimensional structure stabilised by numerous internal interactions, hydrophobicity, and hydrophilicity. Frequently recurring localised folding patterns include the α-helix and the β-sheet, which collectively are referred to as the secondary structure, the primary structure being the sequence of the amino acids. Computer simulations present a powerful tool to investigate the structure of proteins in aggregates.[1] All-atom simulations provide insight into the precise aggregation mechanism for small peptides.[2–5] Aggregation pathways involving the collective refolding of larger proteins are difficult to study by all-atom simulations due to excessive computational demands. By sacrificing the finer details of the system, so-called coarse-grained (CG) models aim to reach the larger time and length scales at which aggregation occurs.[6] Many experimental techniques are also restricted to resolving the overall structure of proteins and protein aggregates in terms of secondary structure elements and their relative positions and orientations, i.e., the tertiary and quaternary structure. This is by no means a drawback, since a conceptual understanding of the protein structure and functioning is often most easily obtained in terms of secondary, tertiary, and quaternary structure while detailed information at the atomic level is only typically required to study binding sites and catalytic sites.

Against this background, it is therefore just natural and attractive to simulate proteins in terms of their secondary structure rather than at a more detailed level. We here present our first results on the development of a highly coarse-grained protein simulation model at the secondary level. Our ultimate goal is to simulate a protein as a chain of particles, which each represent about a dozen amino acids. Changes in the secondary structure of the sequences represented by the coarse particles, e.g., a transition from disordered to β-sheet, are modelled by varying the shape and interaction characteristics of the particles. The chain flexibility allows the tertiary and quaternary structure to develop. In this paper we do not go into the full complexity of the chain model, but restrict ourselves to developing polymorph patchy particles that smoothly change their internal states—and thereby their external appearance—in response to their environment, while the linking together of several polymorph particles into a protein will be the topic of a forthcoming paper.[7] As an illustration, the polymorph particle is applied as the idealisation of a protein with a very dynamic secondary and tertiary structure, namely, the

---

[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 intrinsically disordered amyloidogenic protein α-synuclein (α syn) that aggregates into both disordered oligomers and ordered fibrils.

The self-assembly of proteins into amyloids has been strongly linked to a number of neurodegenerative disorders such as Parkinson's disease (PD)[8–13] and Alzheimer's disease (AD).[14–16] The proteins associated with these diseases are α syn and amyloid-β (Aβ), respectively.[17] Both proteins are intrinsically disordered, meaning that they have no prevalent secondary structure in their native solvated state, but they can adapt their internal state in response to the surrounding environment. Upon changing their secondary structure they can interact with other proteins or membranes and exhibit biological functions.[18,19] Both proteins are known to fold into α-helices against membranes[20–22] and to "misfold" into cross-β-sheets in fibrils or amyloids.[23,24] Aggregation of these proteins gives rise to numerous structures such as oligomers,[25–27] fibrils,[28] and Lewy bodies,[29] localised mainly in brain cells. Some of these aggregates are believed to be toxic[24,26,30,31] and lead to neuronal death.[29,32–34] Several other proteins such as insulin, amylin, and huntingtin[17,35] show similar characteristics, suggesting generic behaviour.

The "protein chameleon" α-synuclein is a small 14 kDa intrinsically disordered protein composed of 140 amino acids (AA). The structures of its various aggregates and their formation pathways have been studied intensively over the past decades.[18,19,25,36–39] The presence of oligomer accumulations during fibrillation suggests that the oligomers are the direct precursor of fibrils (on-pathway mechanism).[40–42] Other studies suggest that protofibrils serve as transient assembly intermediates on the fibril formation pathway.[39,43,44] A wide range of techniques, including atomic force microscopy (AFM), nuclear magnetic resonance (NMR), and small angle X-ray scattering (SAXS), have been used to study the structure of fibrils revealing that the core of the 4-10 nm wide fibers contains a characteristic cross-β secondary structure.[24,28,45–47]

Various coarse grained simulation models have been developed to gain insight into the aggregation process of amyloidogenic proteins. Particles with two diametrically opposed binding patches readily self-assemble into fibrils.[5,48,49] Pellarin and Caflisch model an amphipathic polypeptide as a super-molecule of ten beads with one internal dihedral degree of freedom to convert between amyloid-forming and amyloid-protected states.[50] The model produces a number of distinct oligomer and fiber morphologies, by various thermodynamically or kinetically selected pathways.[50,51] Barz and Urbanc represent a protein as a rigid tetrahedron with two hydrophobic and two hydrophilic beads.[52] Their simulations show elongated aggregates restructuring to form a denser packing, and emphasise the role of breakage in the overall aggregation process. Reducing the number of particles per protein even further, Bieler *et al.* model the Aβ protein as a single spherocylinder with surface patches that instantaneously switch between two binding states, predicting a critical nucleus to fibril formation of four monomers.[53] In a recent study using a modified version of this model, Saric *et al.* observe a protein concentration dependence of the aggregation pathway, which are referred to as one-step nucleation (1SN) for direct growth of fibers and two-step nucleation for fiber formation via an oligomer intermediate.[54,55] The cited studies show that widely disparate highly coarse-grained models share the ability to assemble into oligomers and fibrils, thereby providing valuable insights into the intricate nucleation and interconversion pathways of the amyloid formation processes. Yet, understanding the aggregation mechanism of amyloidogenic proteins involves more than reproducing the aggregation characteristics—the aggregation should result from the protein model faithfully representing the main features of the actual protein, like the distribution of binding affinities along its contour and the agility of its secondary structure. The envisaged model of a protein as a chain of polymorph patchy particles, as discussed earlier, appears an attractive route to attaining this objective. As a first step in this direction, we here introduce a polymorph CG model that represents the α-synuclein protein as a single particle able to adjust its internal state, reflecting the common ability of intrinsically disordered proteins (IDPs) to adapt their folding structure to the surrounding environment. The disordered state is represented by a single soft sphere and the folded β-sheet state by a hard spherocylinder with attractive patches on its surface, see Fig. 1. The state, and hence the shape of the particle, continuously evolves in time in response to internal and inter-particle interactions. To investigate the dynamics of the system, a novel Brownian Dynamics algorithm[56] is used to simulate the translational and rotational motions of the particles as well as their internal dynamics.

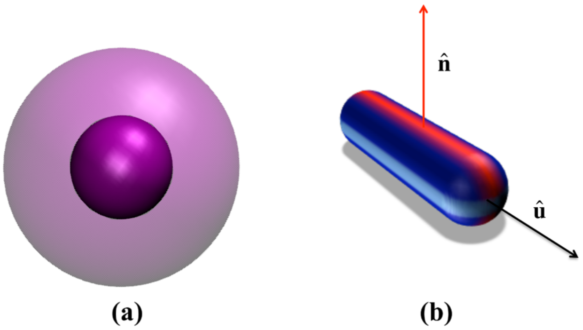

FIG. 1. Cartoons of the coarse grained representation of the amyloid protein in (a) the disordered state and (b) in the compactly folded state. The colours indicate surfaces with distinct interaction properties. The unit vector $\hat{\mathbf{u}}$ along the long axis of the particle, in combination with the perpendicular unit vector $\hat{\mathbf{n}}$, defines the orientation of the particle relative to the laboratory coordinate system.

This paper is structured as follows: Section II introduces the coarse grained model, the interaction potentials, and describes the Brownian Dynamics algorithm. The simulation results on the self-assembly of the particles into oligomers and fibrils is discussed in Section III and the paper ends with a summary and discussion of the main results in Section IV.

II. MODEL AND METHODS

The model introduced here is a single particle highly coarse grained generic representation of amyloid forming proteins. In principle, the force field can be deduced from fully atomistic simulations of these proteins, but we here opt for the pragmatic approach of selecting potentials based on a combination of experimentally established interaction properties and computational simplicity. We will refer to α-synuclein but emphasise that since its folding and aggregation properties are shared by a number of amyloid forming intrinsically disordered proteins, a model capturing these generic properties can be parametrised to model any specific member of this family. The number of internal states of the particle will be limited to two, representing a disordered and a β-rich conformation.

A. Polymorphism

The radius of gyration of α-synuclein in the disordered conformation, $R_g \approx 40$ Å, is a bit smaller than that of a 140 AA random coil, with $R_g \approx 52$ Å,[19] but significantly bigger than that of a 140 AA folded globular protein, $R_g \approx 15$ Å,[57] suggesting that α syn is slightly more compact than the classic random coil.[38] In fibrils, the approximately 60 residues between positions ∼30 and ∼90 adopt a condensed structure that enables this segment to bind strongly to identically folded neighbouring segments by numerous hydrogen bonds.[17,24] The N and C-terminal tails flanking this segment are solvent exposed and appear not involved in the hydrogen bonding in the core of the fiber.[24,58,59] Experiments show that fibrillation proceeds quicker after shorting the N-terminal[58] or neutralising the large negative charge on the C-terminal,[60] which suggests that the two tails may play a role in preventing amyloid accumulation. The structural differences between the core and the tails are best treated by a chain of multiple particles with differing properties. At this early stage in the model development, however, we focus on creating a coarse-grained patchy particle representing the central ∼60 amino acids of α-synuclein.

An essential feature of the protein, and hence of the model particle, is its ability to switch between two internal states, i.e., a disordered state and a folded β-sheet-rich state. This is realised in the simulations by introducing a parameter λ_i to describe the internal state of the ith particle. In line with experimental findings, the probability distribution of this coordinate must be constructed in such a way that for unbound particles the disordered state is favoured over the ordered state. Hence, the internal potential $\Phi_\lambda(\lambda)$ should be of the form shown in Fig. 2(a), where an activation barrier separates the disordered state ($\lambda < 0$) from the ordered state ($\lambda > 0$). For mathematical convenience, we construct this potential by combining two third-order polynomials,

$$\Phi_\lambda(\lambda) = \begin{cases} -\varepsilon^{\mathrm{S}}_\lambda \dfrac{27}{4}\left(\lambda^2+\lambda^3\right) & \text{for } \lambda < 0 \\ -\varepsilon^{\mathrm{R}}_\lambda \dfrac{27}{4}\left(\lambda^2-\lambda^3\right) & \text{for } \lambda \geq 0 \end{cases}, \tag{1}$$

that reach local minima of $-\varepsilon^{\mathrm{S}}_\lambda$ and $-\varepsilon^{\mathrm{R}}_\lambda$ for $\lambda = \pm\frac{2}{3}$, and both have a local maximum for $\lambda = 0$. Although the folded state is unfavourable in solution, it is the stable state for proteins in a fiber. This shift indicates that the stability gained in fibrils arises from attractive interactions between proteins, which are modelled by the attractive potential discussed below.

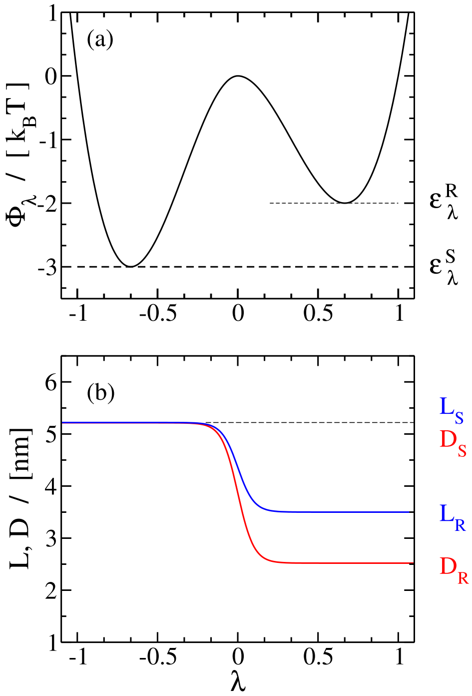

FIG. 2. (a) The internal potential plotted against the internal coordinate for $\varepsilon^{\mathrm{S}}_\lambda = 3$ and $\varepsilon^{\mathrm{R}}_\lambda = 2$, see Eq. (1). (b) Length and diameter as functions of λ, see Eq. (3), for the parameter values employed in the simulation. Note that both dimensions decrease as the protein evolves from the low density disordered state to the compactly folded ordered state.

The shape of the protein is strongly linked to its internal state, hence the shape of the particle should depend on the parameter λ. We idealize the disordered state as a sphere with soft isotropic interactions and the folded state as a sphero-cylinder with hard anisotropic interactions, see Fig. 1. This selection of shapes allows for continuous morphing as a function of the internal coordinate. To restrict the shape changes to the barrier region, we introduce the smoothed step-function

$$\mu(\lambda) = \begin{cases} 0 & \text{for} \quad \lambda < -1 \\ \dfrac{\tanh(A_\mu \lambda)}{2\tanh(A_\mu)} + \dfrac{1}{2} & \text{for} \quad -1 \leq \lambda \leq 1, \\ 1 & \text{for} \quad \lambda > 1 \end{cases} \tag{2}$$

where A_μ determines the width of the transition region. The tip-to-tip length and the diameter of the particle are then defined following

$$\begin{aligned} L(\lambda) &= L_{\mathrm{S}} + \mu(\lambda)(L_{\mathrm{R}} - L_{\mathrm{S}}), \\ D(\lambda) &= D_{\mathrm{S}} + \mu(\lambda)(D_{\mathrm{R}} - D_{\mathrm{S}}), \end{aligned} \tag{3}$$

where the spherical and sphero-cylindrical extremes of the morphing process are marked by subscripts S and R, respectively, and where $D_{\mathrm{S}} = L_{\mathrm{S}}$ for a sphere.

B. The repulsive potential

Having defined the shape and size of the particle, we next focus on the interactions between the particles. Because of the polymorphism of the particles, the potential energy is coupled to the internal states of the interacting particles. Overlap of particles i and j is avoided by a repulsive potential based on the minimum distance $d_{\min}$ between their long axes,[61]

$$d_{\min}(\mathbf{r}_i,\mathbf{r}_j,\hat{\mathbf{u}}_i,\hat{\mathbf{u}}_j,\lambda_i,\lambda_j) = \min_{\alpha_i,\alpha_j}\left|(\mathbf{r}_i+\alpha_i\hat{\mathbf{u}}_i)-(\mathbf{r}_j+\alpha_j\hat{\mathbf{u}}_j)\right|, \tag{4}$$

with $\mathbf{r}_i$ and $\mathbf{r}_j$ the centre of mass positions of the two particles, and $\hat{\mathbf{u}}_i$ and $\hat{\mathbf{u}}_j$ denoting unit vectors pointing along the long axes of the particles. For two infinitely long cylinders, the analytic expressions for the minimisation parameters α_i and α_j are readily solved from two coupled linear equations. By restricting their values to the λ-dependent ranges $|\alpha_i| \le \frac{1}{2}(L_i - D_i)$ and $|\alpha_j| \le \frac{1}{2}(L_j - D_j)$, the points of closest approximation are guaranteed to lie either within the cylindrical segment of a spherocylindrical particle or at the centre of a spherical particle. Substitution of these restricted values into Eq. (4) yields the requested minimum distance. When two rod-like particles are nearly parallel, a small change of their relative orientations can cause a major change in the points of closest approach; these points may even jump from one end of the cylindrical segment to the opposite end. Hence, for $(\hat{\mathbf{u}}_i \cdot \hat{\mathbf{u}}_j)^2 \ge 0.99$ the centre to centre vector $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$ is decomposed into components parallel, $\mathbf{r}_\parallel = (\mathbf{r}_{ij}\cdot\hat{\mathbf{u}}_{\text{av}})\hat{\mathbf{u}}_{\text{av}}$, and perpendicular, $\mathbf{r}_\perp = \mathbf{r}_{ij} - \mathbf{r}_\parallel$ to the average direction $\hat{\mathbf{u}}_{\text{av}} = (\hat{\mathbf{u}}_i + \gamma\hat{\mathbf{u}}_j)/|\hat{\mathbf{u}}_i + \gamma\hat{\mathbf{u}}_j|$, where $\gamma = (\hat{\mathbf{u}}_i\cdot\hat{\mathbf{u}}_j)/|\hat{\mathbf{u}}_i\cdot\hat{\mathbf{u}}_j|$ equals $+1$ (-1) when the orientation vectors point in the same (opposite) direction. The minimum distance is then calculated as

$$d_{\min} = \begin{cases} r_\perp & \text{for } |r_\parallel| \le (\overline{L}-\overline{D}) \\ \sqrt{r_\perp^2 + \left[r_\parallel - (\overline{L}-\overline{D})\right]^2} & \text{for } |r_\parallel| > (\overline{L}-\overline{D}) \end{cases}, \tag{5}$$

with the first line the perpendicular distance between two cylindrical segments and the second line the distance between two end caps, as illustrated in Fig. 3.

The repulsive potential should become relevant, i.e., increase above the thermal energy k_BT with k_B Boltzmann's constant and T the temperature, when the minimum distance between two particles reduces below the average of their radii, $\overline{D} = \frac{1}{2}(D_i + D_j)$. Due to the marked difference in amino acid packing densities, the repulsion between two disordered proteins will be much softer than that between two compactly folded proteins. This is accounted for by the potential

$$\frac{\Phi^\lambda_{\text{rep}}}{k_BT} = \begin{cases} \left(\dfrac{2\overline{D}}{d_{\min}} - 1\right)^n & \text{for } d_{\min} < 2\overline{D} \\ 0 & \text{for } d_{\min} > 2\overline{D} \end{cases}, \tag{6}$$

where the exponent, i.e., the hardness of the potential, was selected to vary from 4 to 6 following $n = 4 + (\mu_i + \mu_i)$.

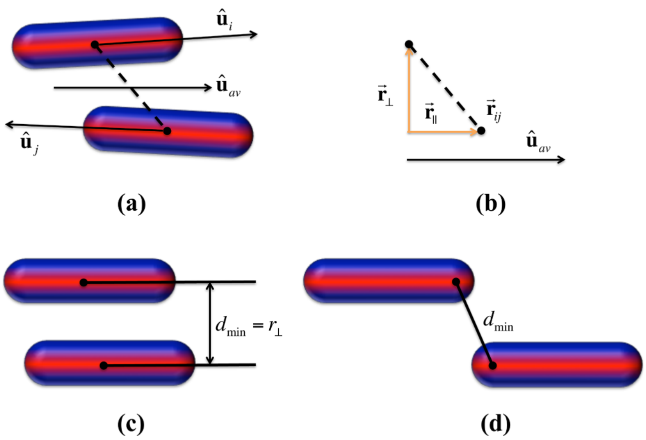

FIG. 3. (a) Cartoon of two nearly parallel rods pointing in opposite directions, $\gamma = -1$, (b) the vector decomposition of the centre to centre vector $\mathbf{r}_{ij}$ parallel and perpendicular to the average direction $\hat{\mathbf{u}}_{\text{av}}$, (c) and (d) the minimum distance between two nearly parallel rods as calculated in Eq. (5).

C. The attractive potential

The aggregation of amyloid proteins into oligomers and fibrils is driven by attractive interactions between the amino acids: van der Waals interactions, hydrophobic interactions, and hydrogen bonding. The number of those interactions and their overall strength depend on the internal state of the protein: disordered proteins have larger radii of gyration and hence larger interaction ranges than densely folded proteins; the interactions between two random coils gradually change with distance, while hydrogen bonds between two compacted proteins are short-ranged. Furthermore, the disordered state is in good approximation isotropic, while the folded state will experience anisotropic interactions. For two particles i and j, we propose the attractive potential

$$\frac{\Phi^\lambda_{\text{attr}}}{k_BT} = \eta(d_{\min})\Big\{C_{\text{SS}}\,(1-\mu_i)\,(1-\mu_j) + \mathcal{F}_{\text{RR}}\mu_i\mu_j + C_{\text{SR}}\big[\mu_i(1-\mu_j) + \mu_j(1-\mu_i)\big]\Big\}, \tag{7}$$

with $d_{\min}$ the previously introduced minimum distance. The distance dependence of the potential is given by

$$\eta(d_{\min}) = \begin{cases} \dfrac{\tanh(A_{\text{attr}}(d_{\min} - 2\overline{D}))}{\tanh(2A_{\text{attr}}\overline{D})} & \text{for } d_{\min} < 2\overline{D} \\ 0 & \text{for } d_{\min} \ge 2\overline{D} \end{cases}, \tag{8}$$

where $A_{\text{attr}} = a_{\text{attr}}(\mu_i + \mu_j) + b_{\text{attr}}$, with $a_{\text{attr}} > 0$ and $b_{\text{attr}} > 0$, determines the steepness of the transition between $\eta(0) = -1$ and $\eta(2\overline{D}) = 0$. The resulting binding energy between two particles in the spherical state, $\mu_i \simeq \mu_j \simeq 0$, depends on their center-to-center distance $d_{\min}$ only, and decays smoothly with decreasing distance, from zero at the cutoff distance to a minimum of $-C_{\text{SS}}$. For two spherocylindrical particles, $\mu_i \simeq \mu_j \simeq 1$, the function $\eta(d_{\min})$ is both shorter-ranged due to the smaller $\overline{D}$ and its decay with decreasing distance to a minimum of $-\mathcal{F}_{\text{RR}}$ is steeper, i.e., A_{attr} is larger, reflecting a tightly packed protein. The interaction between a spherical particle, $\mu_i \simeq 0$, and a spherocylindrical particle, $\mu_j \simeq 1$, is accounted for by the last term in the r.h.s. of Eq. (7), with a decay rate intermediate between those of the two combinations mentioned before and a minimum at $-C_{\text{SR}}$. All three cases are illustrated in Fig. 4.

The interaction between two non-isotropic proteins, i.e., the sum of the contacts between atoms on their surfaces, varies with the relative orientations of the proteins. This complex dependence is qualitatively accounted for in Eq. (7) by the rod-rod contribution $\mathcal{F}_{\text{RR}}\mu_i\mu_j$, with

$$\mathcal{F}_{\text{RR}} = \left[C^{\text{w}}_{\text{RR}} + g_{ij}C^{\text{s}}_{\text{RR}}\right] f_{ij}. \tag{9}$$

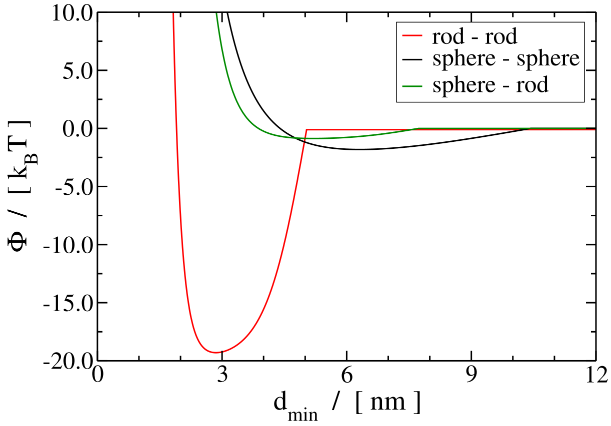

FIG. 4. Plot of the distance dependence of the total interaction energy of two particles, see Eq. (7). Two particles in the disordered state (black) have a longer-ranged attraction and a softer repulsion than two aligned particles in the folded state (red), while the interaction in a mixed pair is intermediate (green). Here $C_{SS}=3$, $\mathcal{F}_{RR}=20$, and $C_{SR}=1$; the employed particle dimensions are discussed in Section III A.

We postpone the presentation of g_{ij} until Eq. (12) and start here with the discussion of

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

with $0 \le f_{ij} \le 1$. The first factor promotes, at given $d_{\min}$ and $\mathbf{r}_{ij}$, the alignment of the long axes $\hat{\mathbf{u}}$ of the spherocylinders. The second and third factors, with

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

reflect energy variations when one rod is transported along its long axis, at fixed orientation and $d_{\min}$, with the lowest interaction energy reached when the connecting vector $\mathbf{r}_{ij}$ is perpendicular to both rods, i.e., for $\alpha_i = \alpha_j = 0$ and $|\mathbf{r}_{ij}| = d_{\min}$. The parameter a sets the parallel displacement of two parallel rods beyond which they do not interact anymore. The parameter $C^{\mathrm{w}}_{\mathrm{RR}}$ in Eq. (9) represents the interaction strength of the weak van der Waals and hydrophobic interactions between two proteins. Folded proteins in fibers are known to form an array of hydrogen bonds with both neighbours along the direction of the fiber, creating a cross β-sheet arrangement.[17,24] This indicates that the interactions of spherocylindrical particles should be non-uniform under rotations along their long axes, as illustrated in Fig. 5. Every particle is therefore endowed with a second orientational body-fixed unit vector $\hat{\mathbf{n}}_i$, perpendicular to the long axis $\hat{\mathbf{u}}_i$, see Fig. 1(b), and $\mathcal{F}_{\mathrm{RR}}$ contains a contribution proportional to

$$g_{ij} = (\hat{\mathbf{r}}_{ij}\cdot\hat{\mathbf{n}}_i)^p(\hat{\mathbf{r}}_{ij}\cdot\hat{\mathbf{n}}_j)^p \tag{12}$$

that favours alignment of the orientation vectors $\hat{\mathbf{n}}$ parallel to the center-to-center direction $\hat{\mathbf{r}}_{ij}$, see Fig. 5. Here p is an even integer, $0 \le g_{ij} \le 1$, and the parameter $C^{\mathrm{s}}_{\mathrm{RR}}$ in Eq. (9) represents the strength of the hydrogen bonds.

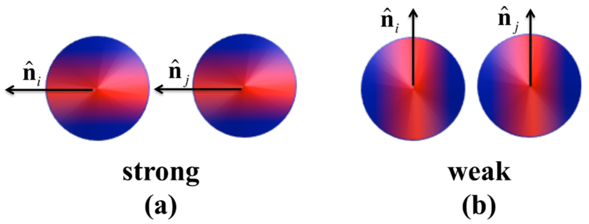

FIG. 5. Cartoons of two aligned folded proteins, with their long axes perpendicular to the projection plane. (a) The strongest binding is achieved when the surfaces capable of forming hydrogen bonds, coloured in red, are facing each other. (b) Weaker affinities apply between the remaining surface areas, in blue, and between unlike surfaces.

The symmetries of the model introduced here inhibit the development of an intrinsic twist in the fiber, which for α-synuclein has been measured as one revolution per ~100 proteins,[62,63] nor does it form multi-stranded fibrils sharing a common hydrophobic core. Since our objective here is to illustrate the feasibility of polymorph-patchy-particle simulations, we opted to leave these features out in order not to overburden the model.

D. Brownian Dynamics algorithm

The evolution of the collection of polymorph particles is simulated by Brownian Dynamics. In this section, we provide the overall equation of motion of the system, which is subsequently split into equations of motion for translation, rotation, and internal dynamics.

Brownian Dynamics algorithms have been developed to simulate the motion of colloidal particles in a solvent, where the particles interact with each other as well as continuously experience multiple collisions with the solvent molecules. The generalised equation of motion[64,65] is expressed as

$$\mathbf{Q}(t+\Delta t)-\mathbf{Q}(t) = -\boldsymbol{\mu}^Q\frac{\partial\mathcal{A}}{\partial\mathbf{Q}}\Delta t + k_BT\frac{\partial}{\partial\mathbf{Q}}\cdot\boldsymbol{\mu}^Q\Delta t + \left(\boldsymbol{\mu}^Q\right)^{1/2}\boldsymbol{\Theta}^Q(t)\sqrt{2k_BT\Delta t}, \tag{13}$$

where $\mathbf{Q}$ represents the full set of generalised coordinates, i.e., positions $\mathbf{r}$, orientations $\mathbf{q}$, and internal coordinates λ for all particles. The first contribution on the r.h.s. gives the displacements over a time step Δt due to the balance between the thermodynamic force $\mathcal{F} = -\partial\mathcal{A}/\partial\mathbf{Q}$, with $\mathcal{A}$ the free energy as a function of $\mathbf{Q}$, and the opposing solvent friction, whose inverse is the mobility matrix $\boldsymbol{\mu}^Q$. The second contribution arises in the Itô interpretation, i.e., all terms on the r.h.s. are evaluated at time t, due to the inhomogeneity of the mobility tensor $\boldsymbol{\mu}^Q$. The last contribution represents the erratic Brownian displacements of the coordinates, where the components of the time-dependent vector $\boldsymbol{\Theta}^Q$ have zero mean, unit variance, and are not correlated to each other nor to previous values (Markovian). The size of these random displacements is connected to the mobility through the fluctuation-dissipation theorem, which is incorporated in the last term of Eq. (13), where the square root $\sqrt{\mathbf{M}}$ of a matrix $\mathbf{M}$ has the property $\sum_{\beta=1}^{3}\sqrt{\mathbf{M}}_{\alpha\beta}\sqrt{\mathbf{M}}_{\gamma\beta} = M_{\alpha\gamma}$.

The simplest application of the Brownian Dynamics algorithm is the translational motion of a particle in a solvent. In this case, the coordinates denote the centre of mass position of the particle, $\mathbf{R}$, which combined with the translational mobility tensor in the space-frame $\boldsymbol{\mu}^{\mathrm{t,s}}$ yields the equation of motion

$$\mathbf{R}(t+\Delta t)-\mathbf{R}(t) = \boldsymbol{\mu}^{\mathrm{t,s}}\mathbf{F}\Delta t + \sqrt{\boldsymbol{\mu}^{\mathrm{t,s}}}\,\boldsymbol{\Theta}^{\mathrm{t}}\sqrt{2k_BT\Delta t}, \tag{14}$$

where the conservative force is given by $\mathbf{F} = -\partial\Phi/\partial\mathbf{R}$, and $\mathbf{\Theta}^{\mathrm{t}}$ is a three-dimensional Markovian vector. For non-spherical rotating particles, the translational mobility in the space-frame varies with the orientation and is given by

$$\boldsymbol{\mu}^{\mathrm{t,s}} = \mathbf{A}\boldsymbol{\mu}^{\mathrm{t,b}}\mathbf{A}^{\mathrm{T}}, \tag{15}$$

where $\boldsymbol{\mu}^{\mathrm{t,b}}$ is the constant body-fixed translational mobility tensor and $\mathbf{A}$ represents the rotation matrix from the body frame to the laboratory frame, see Appendix A. The square root of the laboratory mobility tensor is readily calculated from its constant body-fixed counterpart by

$$\sqrt{\boldsymbol{\mu}^{\mathrm{t,s}}} = \mathbf{A}\sqrt{\boldsymbol{\mu}^{\mathrm{t,b}}}. \tag{16}$$

For a body-fixed coordinate system that diagonalises $\boldsymbol{\mu}^{\mathrm{t,b}}$, the square root of $\boldsymbol{\mu}^{\mathrm{t,b}}$ is simply obtained by taking the square roots of its diagonal elements.

The rotational Brownian motion is simulated by a recently introduced algorithm[56] using quaternions $\mathbf{q}$ as the rotational coordinates. This approach proved to be very accessible—it avoids the singularities, metric tensor correction, and Itô correction that complicate BD simulations in other rotational coordinates[66–70]—making it the ideal candidate for the current study. The rotational equation of motion reads as

$$\mathbf{q}(t+\Delta t) - \mathbf{q}(t) = \mathbf{B}\boldsymbol{\mu}^{\mathrm{r,b}}\mathbf{A}^{\mathrm{T}}\mathbf{T}\Delta t + \mathbf{B}\sqrt{\boldsymbol{\mu}^{\mathrm{r,b}}}\mathbf{\Theta}^{\mathrm{r}}\sqrt{2k_BT\Delta t} + \Lambda\mathbf{q}, \tag{17}$$

where $\boldsymbol{\mu}^{\mathrm{r,b}}$ represents the rotational mobility matrix in the body-fixed frame. The first term on the r.h.s. represents the rotational displacement that arises from the balance between the torque $\mathbf{T} = -\partial\Phi/\partial\boldsymbol{\phi}^{\mathrm{s}}$ and the rotational friction, where $\boldsymbol{\phi}^{\mathrm{s}}$ denotes infinitesimal rotations around the laboratory coordinate axes. The 4×3 transformation matrix $\mathbf{B}$ converts angular displacements in the body-frame into quaternion increments; the employed expressions for $\mathbf{A}$ and $\mathbf{B}$ in terms of $\mathbf{q}$ are given in Appendix A. The second term on the r.h.s. represents the Brownian contribution, where the components of the random vector $\mathbf{\Theta}^{\mathrm{r}}$ are again chosen such that they have zero mean, unit variance, no memory, and are unrelated to the translational random contributions. The last term introduces a constraint on the length of $\mathbf{q}$, as for finitely sized time steps the length of the quaternion vector—which must be unity in order for $\mathbf{A}(\mathbf{q})$ to be a rotation matrix—is only approximately conserved. A simple rescaling, $\mathbf{q} \to \mathbf{q}/|\mathbf{q}|$ will change the sampled phase space distribution. The constraint calculation applied here follows the method of undetermined Lagrange multipliers, as also used in the seminal SHAKE algorithm.[71] The strength of the constraint force is determined by the Lagrange multiplier Λ, which is readily solved from the quadratic equation

$$|\mathbf{q}(t+\Delta t)|^2 = \left|\mathbf{q}^{\mathrm{uncons}}(t+\Delta t) + \Lambda\mathbf{q}(t)\right|^2 = 1, \tag{18}$$

where $\mathbf{q}^{\mathrm{uncons}}(t+\Delta t)$ denotes the unconstrained propagated coordinates in the absence of the constraint force.

The evolution of the internal coordinate λ, characterising the internal state of a particle, is also simulated by Brownian Dynamics. The equation of motion reads as

$$\lambda(t+\Delta t) - \lambda(t) = \mu^{\lambda}F^{\lambda}\Delta t + \sqrt{\mu^{\lambda}}\,\Theta^{\lambda}\sqrt{2k_BT\Delta t}, \tag{19}$$

where μ^{λ} denotes the mobility of the internal coordinate, $F^{\lambda} = -\partial\Phi/\partial\lambda$ the force on the internal coordinate arising from internal and external potentials, and Θ^{λ} a Markovian random number of zero average and unit variance.

III. RESULTS

The parameter values entering the model were based on experimental data wherever possible, as described below. The only parameters that could not be established, the internal well depths $\varepsilon_{\lambda}^{S}$ and $\varepsilon_{\lambda}^{R}$, were systematically varied to construct a phase diagram.

A. Model parameters

The geometrical parameters of the particles were chosen to match experimental data. Small angle X-ray scattering on monomeric α-synuclein in solution yielded a radius of gyration, R_g, of 4.0 nm.[38] For the 60 AA segment, assuming the scaling behaviour of a Gaussian chain, this converts into $R_g^{60\mathrm{AA}} = 2.61$ nm. The diameter of the particle in the disordered state was selected by demanding that the repulsive potential between two particles reaches 1 k_BT at a distance of $2R_g$, hence $D_{\mathrm{S}} = L_{\mathrm{S}} = 5.2$ nm. For the folded state, we based the dimensions on the structural data on fibrils from the solid state hydrogen/deuterium (H/D) exchange NMR measurements by Vilar *et al.*[24] The length of the rod-like particle, $L_{\mathrm{R}} = 3.5$ nm, matches the average length of the five β-strands per protein. The diameter was set at $D_{\mathrm{R}} = 2.5$ nm, which is the average of the collective width of ~4 nm of the five β-strands and the distance ~1 nm between two proteins along the fibril direction. A geometrically more realistic platelet-like description of the protein's core region will be achieved in the future model by representing the five β-strands as a chain of five (thinner) spherocylindrical particles. In the conversion from λ to μ, see Eq. (2), the transition is confined to a small region by $A_{\mu} = 10$. For the steepness of the attractive potential, see Eq. (8), we selected $A_{\mathrm{attr}} = 0.4\left(\mu_i + \mu_j\right) + 0.2$, with A_{attr} in units of nm^{-1}.

To obtain appropriate binding energies, we first studied the aggregation behaviour of the disordered and folded particles separately. The stability limits of oligomers were established by simulating systems containing pre-constructed oligomers of various sizes, surrounded by disordered particles, while freezing all internal coordinates at $\lambda = -2/3$. The concentration was fixed at one particle per $3\cdot 10^{-5}$ nm^3, corresponding to an experimental concentration of 50 μM, which lies within the typical range used in experiments on wild-type α-synuclein. All oligomers were observed to disintegrate when $C_{\mathrm{SS}} < 2.4$. For $C_{\mathrm{SS}} = 3$, oligomers of less than about six particles tended to dissolve while larger oligomers showed a propensity to grow through monomer attachment and occasional fusions with other aggregates, suggesting a critical aggregate size of roughly 6 particles. Because particles in the disordered state have isotropic interactions, there is no inherent restriction on the maximum size of the agglomerates. The stability limits of fibrils were assessed in a similar fashion, from the fate of pre-constructed fibrils in simulations with all internal coordinates fixed at

$\lambda = 2/3$. A series of simulations at various values of the directional interaction parameter yielded a critical binding strength $C^{\rm s}_{\rm RR} \approx 17$ at the aforementioned concentration. Fibrils were observed to stick together, forming multi-stranded bundles, when the weak non-directional interactions were turned on, $C^{\rm w}_{\rm RR} > 0$, and/or when the ranges of the functions g and h were too wide. For simplicity the weak interactions have been turned off, $C^{\rm w}_{\rm RR} = 0$, for the simulations presented here. To ensure the narrowness specific to intra-fiber hydrogen bonds, the power p in Eq. (12) was set to a value of 6 and the parameter a in Eq. (11) to a value of 1.5, thus effectively suppressing inter-fiber hydrogen bonding. In the production runs, the attractive interactions between like particles were set slightly above their respective critical values for aggregation: the sphere-sphere interaction was set at $C_{\rm SS} = 3$, and the directional rod-rod interaction at $C^{\rm s}_{\rm RR} = 20$; these were also the values used in Fig. 4. A small attraction was introduced between unlike particles, by setting the sphere-rod interaction at $C_{\rm SR} = 1$. The depths of the two wells in the internal potential were systematically varied, with $\varepsilon^{\rm S}_\lambda \geq \varepsilon^{\rm R}_\lambda$. Note that the internal potential favours the disordered state over the ordered state, whereas the external interactions promote the ordered state.

Simulation systems were generated by placing 200 particles, with random orientations and $\lambda = -2/3$, at random positions in a cubic simulation box with periodic boundary conditions at the aforementioned concentration. The temperature was set at body temperature, 310 K, and the viscosity equalled that of water, $\eta = 7 \cdot 10^{-4}$ Pa s. The translational and rotational mobility tensors were calculated using the equations presented in Appendix B. By simulating a pre-assembled fibril at a range of time steps, we established that the system permits a time step $\Delta t = 1.7 \times 10^{-13}$ s. The mobility of the internal parameter μ^λ, represents the "internal friction" of the protein supplemented with a solvent contribution. We settled on a value of $\mu^\lambda = 1.4 \times 10^{30}$ (J s)$^{-1}$, which made internal transitions rare events yet numerous transitions occurred during the entire length of a simulation. For instance, for a non-interacting particle with the well depths $\varepsilon^{\rm S}_\lambda = 7$ and $\varepsilon^{\rm R}_\lambda = 5$, the rate constants of the forward and reverse reactions were measured as once per 1.4×10^5 steps and once per 1.9×10^4 steps, respectively. To test whether the proposed combination of Brownian propagators correctly samples the internal coordinate, the sampled probability distribution along λ was compared against the theoretical prediction $P(\lambda) \propto \exp(-\beta\Phi_\lambda)$ in the absence of interactions. Good agreement was obtained, demonstrating that the sampling of λ was not influenced by the shape-dependent mobilities of the particles.

B. Oligomers and fibers

The polymorph particles spontaneously form both oligomers of particles in the disordered state and fibers of particles in the folded state, as shown in Fig. 6. Representative growth curves are shown in Figs. 7–9. For each stored configuration, the system was analysed using a clustering routine[72] assigning a particle to a cluster if its distance $d_{\rm min}$ to any particle within that cluster lies below a threshold distance of $2\overline{D}$. Figure 7 shows the cluster size evolution in time for a system that forms oligomers only. The plot shows the simultaneous growth, predominantly by monomer addition, of five large oligomers, as well as numerous transient aggregates that never crossed the critical nucleus size. The oligomers readily fuse when they come into contact, e.g., two oligomers of about 30 monomers each fuse into a bigger aggregate at $t \approx 65$ μs. The fusion is marked by arrows pointing at the discontinuations of the lines for the original oligomers and the commencement of a line for the "new" oligomer. The latter oligomer continues growing via monomer attachment.

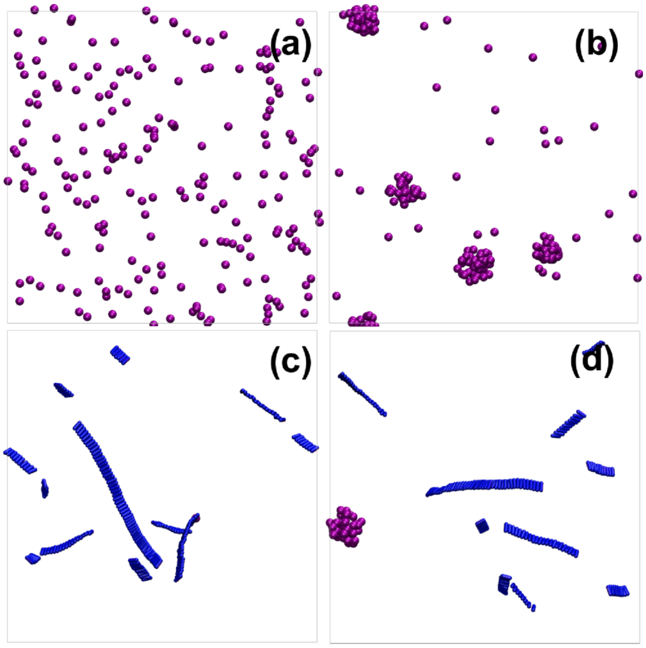

FIG. 6. Snapshots of aggregates that formed spontaneously in simulations started from (a) random configurations of particles in the spherical state. All runs share identical interaction parameters, as specified in the main text. (b) Oligomers are formed for $\varepsilon^{\rm S}_\lambda = 7$ and $\varepsilon^{\rm R}_\lambda = 3$, (c) fibrils at $\varepsilon^{\rm S}_\lambda = 7$ and $\varepsilon^{\rm R}_\lambda = 6$, and (d) both are present at $\varepsilon^{\rm S}_\lambda = 5$ and $\varepsilon^{\rm R}_\lambda = 3$.

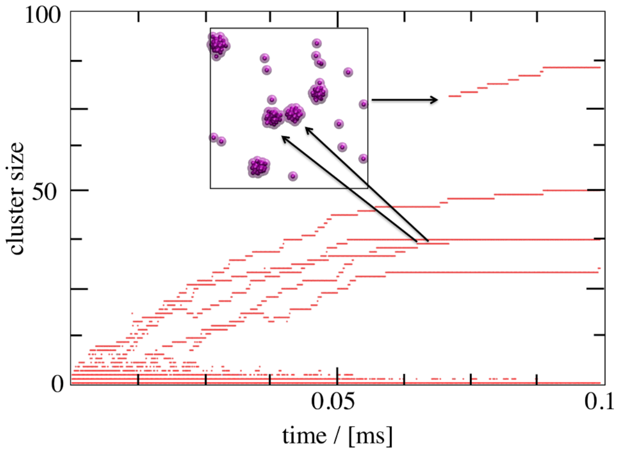

FIG. 7. Time evolution of cluster sizes for a system growing only oligomers. Colours represent the cluster-averaged internal state, $\langle\lambda\rangle$, running from red for disordered ($\langle\lambda\rangle \leq -2/3$), to blue for ordered ($\langle\lambda\rangle \geq 2/3$). The inset shows the system just before two oligomers merge; the two lines representing these aggregates terminate and a new line starts. Here $\varepsilon^{\rm S}_\lambda = 6$ and $\varepsilon^{\rm R}_\lambda = 2$.

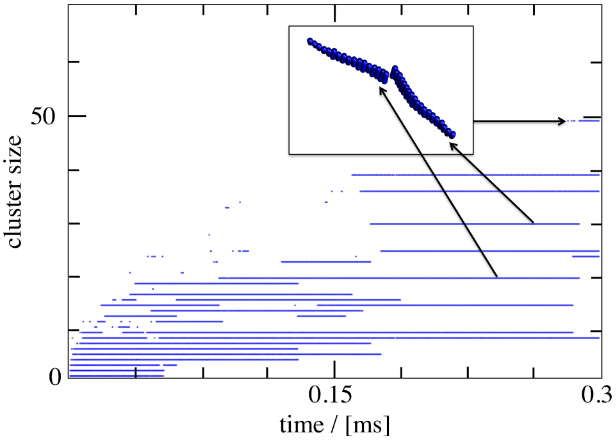

FIG. 8. Time evolution of cluster sizes for a system growing only fibers. The inset shows the system just before two fibers merge. Here $\varepsilon_\lambda^S = 2$ and $\varepsilon_\lambda^R = 1$. Color scheme as in Fig. 7.

Figure 8 shows a system growing fibers only. The abundance of small long-lived aggregates, relative to Fig. 7, reflects the absence of a critical nucleus size in fibril formation. Again the aggregates grow predominantly by monomer addition. Fibers only merge when their ends meet in properly aligned orientations, as in the example shown in the inset to Fig. 8.

At times, a particle in the interior of a fiber spontaneously transforms into a disordered state. Such events, which are more frequent if the folded state is not highly favoured over the disordered state, result in several cases in the break-up of the fiber into two shorter filaments. The spontaneous simulations assembly of oligomers and fibrils is illustrated by Figs. 6(d) and 9. In encounters of oligomers with fibrils, large oligomers tend to break down the fibril and absorb the released particles. An example hereof occurs in Fig. 9 at about 0.1 ms; the red curve representing an oligomer ends and a brown line, for a cluster with $\langle\lambda\rangle \approx 0$, starts. The latter line subsequently turns red, as the particles in the original fiber convert into oligomers. Small oligomers coming into contact with the end of a fibril are typically absorbed by the fibril; the resulting reduction in the number of sphere-sphere interactions makes it relatively easy for particles to escape from the oligomer, hence the fibril often grows by less than the original oligomer size.

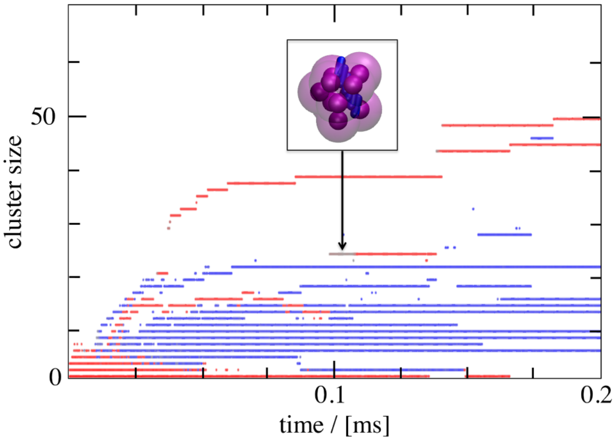

FIG. 9. Time evolution of cluster sizes for a system growing both oligomers and fibers. The inset shows the system just before a fiber transforms and is absorbed into an oligomer. Here $\varepsilon_\lambda^S = 6$ and $\varepsilon_\lambda^R = 4$. Color scheme as in Fig. 7.

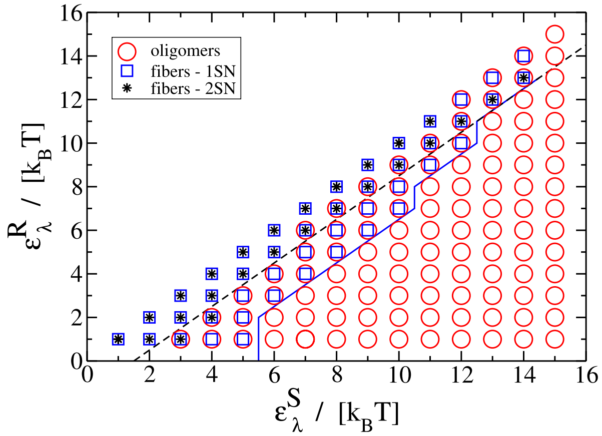

FIG. 10. Phase diagram of the aggregation behaviour of the polymorph particles, plotted against the internal stabilities ε_λ^S and ε_λ^R of the spherical and rodlike states of the particle, respectively. Oligomers are represented by red circles, blue squares and black stars mark the formation of fibrils via one step nucleation and two step nucleation, respectively. See also Figs. 6 and 11. The lines highlight approximate locations of phase boundaries. The inter-particle interactions are $C_{SS} = 3$, $C_{SR} = 1$, $C_{RR}^{s} = 20$, and $C_{RR}^{w} = 0$.

Figure 10 shows a phase diagram of the aggregation behaviour as a function of the two well depths ε_λ^S and ε_λ^R of the internal potential, see also Fig. 2(a). Each point represents three to five independent simulations, initiated with all particles in the disordered state, with most of the runs lasting up to the order of $2 \cdot 10^9$ steps, i.e., over 0.3 ms. During this time interval, a non-interacting spherical particle diffuses over $\sim 5 \times 10^2$ nm, equivalent to almost three box lengths, while the smaller rod-shaped particle travels even further. The aggregate species present at the ends of the simulations, as determined by visual inspection in VMD,[73] see Fig. 6, are marked in the phase diagram with squares and stars for fibrils and circles for oligomers. In the bottom-right corner, below the blue drawn line, only oligomers are formed because the internal potential strongly favours the disordered state. Both oligomers and fibrils are observed between this stability boundary and the black dashed line. Because there typically are only a few aggregates present at the end of a run, some boxes contain aggregates of one type only while others contain both types. Above the second phase boundary, mainly fibrillar aggregates are present at the end of the simulations, indicating that the directional intermolecular interactions between rod-like particles are dominant in determining the final state.

C. Fiber nucleation mechanisms

The cluster growth curves and visualisation of the simulations in VMD indicate that most aggregates arise and evolve by the nucleation-and-growth mechanism. A typical sequence for a fibril is presented in Fig. 11(a) (Multimedia view). Randomly diffusing particles meet each other and—depending on their internal states and relative orientations—can stick together. For two rod-like particles the result will be a fibril nucleus, for two spherical particles an oligomer results. Note that the latter dimer is smaller than

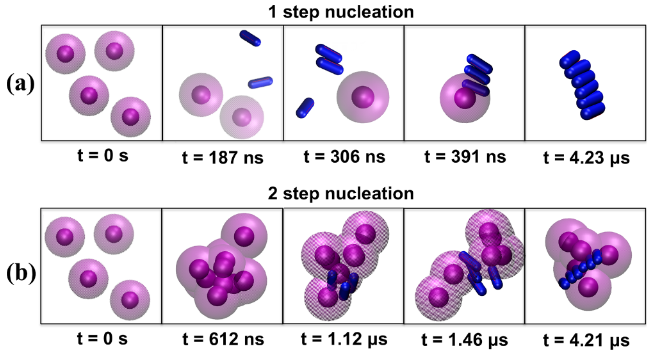

FIG. 11. Simulation snapshots illustrating the major pathways to fibril formation: direct aggregation (top) of particles in the folded state and two step nucleation (bottom) via a transient oligomer gradually evolving into a fiber. Here $\varepsilon_\lambda^{\mathrm{S}} = 3$ and $\varepsilon_\lambda^{\mathrm{R}} = 2$. (Multimedia view) [URL: http://dx.doi.org/10.1063/1.4942115.1]

the critical oligomer cluster, estimated at roughly 6 particles under the prevailing conditions, while there is no critical nucleus for fibrillar aggregates. We regularly observed that the two particles in a dimer transit, in rapid succession, from one internal state to the other. Eventually, the dimers either fall apart due to the perpetual Brownian motion or grow into long-lived aggregates by binding additional particles. Particles attaching to an aggregate are observed to quickly conform to the dominant internal state in that aggregate.

An exception to the above nucleation-and-growth mechanism is found for the combinations of well depths marked by stars in the phase diagram of Fig. 10. In these systems, small oligomers form as usual. Next, the particles in the oligomer change in rapid succession from the disordered state to the folded state, forming a fibril in the process. A typical example of the two step nucleation mechanism (2SN) is highlighted in Fig. 11(b) (Multimedia view) and a movie of a second example is available in Multimedia view. A quantitative analysis of a third example is presented in Fig. 12, by tracking a cluster and the states of its constituting particles during the conversion process. Above the dashed black line in the phase diagram, this mechanism converts almost all spontaneously formed oligomers into fibrils, explaining the absence of oligomers at the end of those simulations. Two step nucleation is suppressed upon reducing the stability of rods over spheres, i.e., by enlarging $\varepsilon_\lambda^{\mathrm{S}} - \varepsilon_\lambda^{\mathrm{R}}$, as is clear from the phase diagram. Two step nucleation is also suppressed by the growth of the oligomer: because the interactions between particles in the disordered state are soft and wide-ranged, the potential energy of a small oligomer increases supra-linearly with the number of particles, while for a fibril the energy increases linearly. Whether or not there is a critical size to the oligomeric intermediate in two step nucleation will depend on the details of the interactions between and within the proteins. For the simulated particles, we estimate a critical size of roughly 10 particles. Two step nucleation is not limited to the formation of fibrils: in several simulations we observed the assembly of rod-shaped particles into a fibrillar intermediate that subsequently disordered into an oligomer.

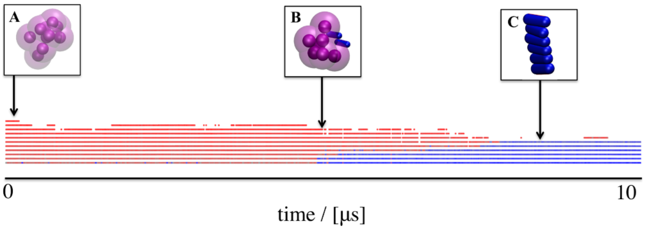

FIG. 12. Evolution of the number of particles (height) and their internal states (red for disordered and blue for ordered) of an aggregate transforming from an oligomer to a fiber. Here $\varepsilon_\lambda^{\mathrm{S}} = 5$ and $\varepsilon_\lambda^{\mathrm{R}} = 4$.

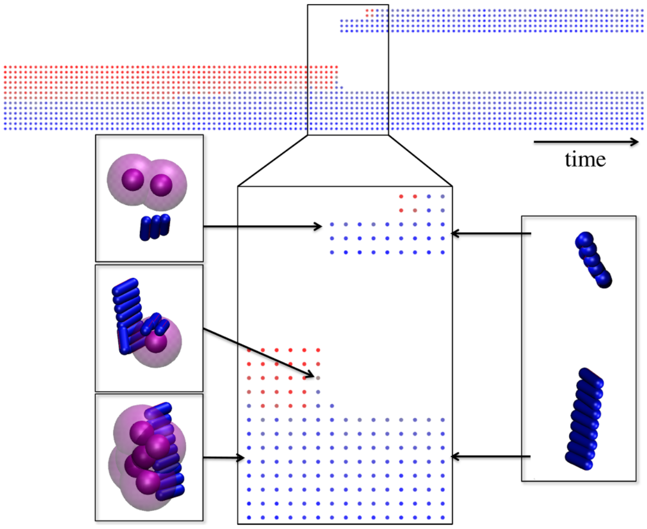

FIG. 13. (Top) Evolution of the number of particles (height) and their internal states (red for disordered and blue for ordered) in an aggregate consisting of an oligomer attached to a fiber. The break-up of the band reflects the break-up of the aggregate. (Bottom) Enlargement of the transition process, with snapshots of the aggregate at selected times. The interval between analysed frames is 0.17 μs and $\varepsilon_\lambda^{\mathrm{S}} = 4$ and $\varepsilon_\lambda^{\mathrm{R}} = 2$. In deviation from all previous figures, this figure was produced at $C_{\mathrm{RR}}^{\mathrm{s}} = 30$.

Interestingly, oligomers were also observed to transform into short fibrils when attached to the central region of a large stable fibril. An example hereof is depicted in Fig. 13. The near-simultaneous conversion of several particles in the oligomer from the disordered to the ordered state weakens their binding to the primary fiber, and consequently a three-particle secondary fiber is released, as well as a pair of disordered particles. Subsequently, the Brownian motion causes the pair to attach to the short fiber and both particles transform to the ordered state, creating a secondary fiber of five particles. This process, observed in a number of simulations, suggests that the primary fibril acts as a catalyst in the conversion of oligomers into secondary fibrils.

IV. DISCUSSION

A polymorph coarse grained particle has been developed and applied successfully to simulate the rich aggregation behaviour of the ~60 AA core of the chameleon protein α-synuclein. This was achieved by creating a particle that smoothly morphs between a spherical shape with weak isotropic interactions, representing the disordered state of the protein in solution and in oligomers, and a sphero-cylindrical shape with strong directional interactions, representing the folded state of the protein in a fibril. The particles aggregate into oligomers of the disordered state and fibrils of the ordered state. Aggregation proceeds by a nucleation-and-growth mechanism, as well as by a two-step mechanism in which an aggregate transforms from one type into the other. These observations confirm experimental and theoretical studies on the formation of fibrils.[74–78] In the simulations, the conversion of oligomers into fibrils appears to be enhanced by the presence of fibrils, which has been proposed as a mechanism to explain the exponential growth phase of the amyloid fraction in aggregation experiments.[17]

Developing a highly coarse grained particle representing a protein involves judicious choices on the protein features to be included in the model and the less relevant properties to be omitted. By sacrificing detail, it becomes possible to study the complex behaviour of the "simple" model—a study that is computationally excessively demanding for a model with fine-grained details on the level of the atoms or residues. A property of wild-type α-synuclein that was intentionally omitted in this study, by focussing on the central ~60 residues that form the core of oligomers and fibrils, is the flexible N and C-terminal tails. Without the tails, especially the hydrophilic C-terminal tail, the protein is unlikely to assemble into the micellar-like oligomers that have been reported in a number of studies.[41,79–82] Instead, the truncated protein is more expected to form droplet-like oligomers of unspecified sizes, as observed for our model. It is well-conceivable that the size and number of oligomers are important in determining the rates of fibril nucleation and growth. We note that the highly coarse-grained model by Saric *et al.*,[54] who employed a particle of fixed sphero-cylindrical shape with Monte Carlo moves to instantaneously swap the single attractive patch from the endcap (the disordered state) to the cylindrical segment (the ordered state), by construction has the ability to form oligomers of limited maximum size. The oligomers that transform into fibrils are about the same size in both their and our simulations.

As discussed in the Introduction, the development of a smoothly morphing patchy particle model constitutes a major step towards modelling a protein as a short chain of polymorph particles.[7] We envisage a model where each particle represents a sequence of about a dozen amino acids, with the internal state of the particle describing the secondary structure and interaction characteristics of this protein segment and the relative positions and orientations of the particles accounting for tertiary and quaternary structure. The internal and external interactions of the particles are to be extracted from all atom simulations of the individual segments and their pair combinations, respectively. The much faster simulations at the coarse-grained level then provide insights into protein aggregation processes that are too slow to study by conventional Molecular Dynamics simulations. Of course, the coarse-grained model can also be used as an efficient method to sample configuration phase, by reinstating atomic details in selected coarse-grained structures.

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 A: MATRICES

The position and orientation of a rigid body in the laboratory frame can be fully described by the coordinates of its centre of mass **R** and by its rotation matrix **A**. The coordinate in the space frame of any point on this particle, given its body-fixed coordinates **x**, are obtained by

$$\mathbf{r} = \mathbf{A}\mathbf{x} + \mathbf{R}. \tag{A1}$$

The rotation matrix is expressed in terms of quaternions **q** of unit length, $q = |\mathbf{q}| = 1$, by the following expression:

$$\mathbf{A} = \begin{pmatrix} q_0^2 + q_1^2 - 2(q_2^2 + q_3^2) & 2(q_1q_2 - q_0q_3) & 2(q_1q_3 + q_0q_2) \\ 2(q_1q_2 + q_0q_3) & q_0^2 + q_2^2 - 2(q_1^2 + q_3^2) & 2(q_2q_3 - q_0q_1) \\ 2(q_1q_3 - q_0q_2) & 2(q_2q_3 + q_0q_1) & q_0^2 + q_3^2 - 2(q_1^2 + q_2^2) \end{pmatrix}. \tag{A2}$$

Differentiation yields that the angular velocity ω^{b} in the body-frame is related to the quaternion velocity by $\dot{\mathbf{q}} = \mathbf{B}\omega^{\mathrm{b}}$, with

$$\mathbf{B} = \frac{1}{2}\begin{pmatrix} -q_1 & -q_2 & -q_3 \\ q_0 & -q_3 & q_2 \\ q_3 & q_0 & -q_1 \\ -q_2 & q_1 & q_0 \end{pmatrix}. \tag{A3}$$

This transformation matrix **B** features in the equation of motion, Eq. (17).

APPENDIX B: MOBILITY TENSORS

The translational and rotational mobility tensors in Eqs. (14) and (17) are determined by the geometry of the particles, which is coupled to the particle's internal state. The mobility tensors are updated every time step, following the expressions of Aragon and Flamik[83] for the translational and rotational diffusion matrices for spherocylindrical particles in the body-fixed frame. The translational mobilities perpendicular and parallel to the long axis, at an aspect ratio $p = L/D$, are given by

$$\begin{aligned}\mu_{\perp}^{\mathrm{t,b}} &= \frac{1}{4\pi\eta L}\left[\ln(p) + 0.902\,952 - \frac{0.311\,044}{\sqrt[4]{p}} + \frac{0.180\,85}{\sqrt{p}} - \frac{0.344\,085}{p^2} - \frac{0.190\,732\ln(p)}{p^2}\right],\\ \mu_{\parallel}^{\mathrm{t,b}} &= \frac{1}{4\pi\eta L}\left[2\ln(p) - 0.113\,192 - \frac{1.304\,29}{\sqrt[4]{p}} + \frac{1.190\,32}{\sqrt{p}} + \frac{3.127\,56}{p} - \frac{1.566\,99}{p^2} - \frac{0.930\,791\ln(p)}{p^2}\right]\end{aligned} \tag{B1}$$

with η the viscosity of the solvent. These mobilities appear twice and once, respectively, along the diagonal of the body-fixed mobility matrix $\boldsymbol{\mu}^{\mathrm{t,b}}$ in Eqs. (15) and (16). Similarly, the perpendicular and parallel rotational mobility components are calculated from

$$\begin{aligned}\mu_{\perp}^{\mathrm{r,b}} &= \frac{3}{\pi\eta L^3}\left[\ln(p) - 0.372\,093 - \frac{0.956\,22}{\sqrt[4]{p}} + \frac{1.247\,92}{\sqrt{p}}\right.\\ &\quad\left. + \frac{1.230\,85}{p} - \frac{1.994\,98}{p^2} + \frac{1.842\,01}{p^3} - \frac{0.664\,147}{p^4}\right],\\ \mu_{\parallel}^{\mathrm{r,b}} &= \frac{p^2}{\pi\eta L^3}\left[1 - 0.001\,586\,88\ln(p)\right.\\ &\quad + \frac{0.036\,951\,8\ln(p)}{\sqrt{p}} - \frac{0.239\,276\ln(p)}{p}\\ &\quad\left. + \frac{0.610\,134\ln(p)}{p^2} - \frac{0.808\,254\ln(p)^3}{p^3}\right],\end{aligned} \tag{B2}$$

and together yield the rotational mobility matrix $\boldsymbol{\mu}^{\mathrm{r,b}}$ in Eq. (17). Equations (B1) and (B2) reduce to the well known Stokes drag expressions[84] $\mu^{\mathrm{t,b}} = (3\pi\eta D)^{-1}$ and $\mu^{\mathrm{r,b}} = (\pi\eta D^3)^{-1}$ for spherical particles ($p = 1$).

[1] R. 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).

[2] K. Lindorff-Larsen, S. Piana, R. O. Dror, and D. E. Shaw, Science **334**, 517 (2011).

[3] J. E. Straub and D. Thirumalai, Curr. Opin. Struct. Biol. **20**, 187 (2010).

[4] K. A. Ball, D. E. Wemmer, and T. Head-Gordon, J. Phys. Chem. B **118**, 6405 (2014).

[5] J. A. Luiken and P. G. Bolhuis, Phys. Chem. Chem. Phys. **17**, 10556 (2015).

[6] C. Wu and J.-E. Shea, Curr. Opin. Struct. Biol. **21**, 209 (2011).

[7] I. M. Ilie, W. K. den Otter, and W. J. Briels, "The attachment of α-synuclein to a fiber: A coarse grained approach" (unpublished).

[8] B. Silva, L. Breydo, A. Fink, and V. Uversky, Mol. Neurobiol. **47**, 598 (2013).

[9] C. W. Olanow and P. Brundin, Mov. Disord. **28**, 31 (2013).

[10] N. A. Haelterman, W. H. Yoon, H. Sandoval, M. Jaiswal, J. M. Shulman, and H. J. Bellen, Annu. Rev. Neurosci. **37**, 137 (2014).

[11] O. W. Wan and K. K. K. Chung, PLoS ONE **7**, e38545 (2012).

[12] E. Maries, B. Dass, T. J. Collier, J. H. Kordower, and K. Steece-Collier, Nat. Rev. Neurosci. **4**, 727 (2003).

[13] H. Chaudhary, A. N. D. Stefanovic, V. Subramaniam, and M. M. A. E. Claessens, FEBS Lett. **588**, 4457 (2015).

[14] M. A. Chabrier, D. Cheng, N. A. Castello, K. N. Green, and F. M. LaFerla, Neurobiol. Dis. **64**, 107 (2014).

[15] A. Griciuc, A. Serrano-Pozo, A. R. Parrado, A. N. Lesinski, C. N. Asselin, K. Mullin, B. Hooli, S. H. Choi, B. T. Hyman, and R. E. Tanzi, Neuron **78**, 631 (2013).

[16] J. L. Crimins, A. Pooler, M. Polydoro, J. I. Luebke, and T. L. Spires-Jones, Ageing Res. Rev. **12**, 757 (2013).

[17] T. P. J. Knowles, M. Vendruscolo, and C. M. Dobson, Nat. Rev. Mol. Cell Biol. **15**, 384 (2014).

[18] V. N. Uversky, C. J. Oldfield, and A. K. Dunker, Annu. Rev. Biophys. **37**, 215 (2008).

[19] P. Tompa and A. Fershtr, *Structure and Function of Intrinsically Disordered Proteins* (Chapman and Hall/CRC, 2009).

[20] S. Tomaselli, V. Esposito, P. Vangone, N. A. J. van Nuland, A. M. J. J. Bonvin, R. Guerrini, T. Tancredi, P. A. Temussi, and D. Picone, ChemBioChem **7**, 257 (2006).

[21] W. S. Woods, J. M. Boettcher, D. H. Zhou, K. D. Kloepper, K. L. Hartman, D. T. Ladror, Z. Qi, C. M. Rienstra, and J. M. George, J. Biol. Chem. **282**, 34555 (2007).

[22] I. Dikiy and D. Eliezer, Biochim. Biophys. Acta, Biomembr. **1818**, 1013 (2012).

[23] N. C. Maiti, M. M. Apetri, M. G. Zagorski, P. R. Carey, and V. E. Anderson, J. Am. Chem. Sci. **126**, 2399 (2004).

[24] M. 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).

[25] V. N. Uversky, J. Neurochem. **103**, 17 (2007).

[26] I. Benilova, E. Karran, and B. De Strooper, Nat. Neurosci. **15**, 349 (2012).

[27] A. Stefanovic, M. Stockl, M. Claessens, and V. Subramaniam, FEBS J. **281**, 2838 (2014).

[28] M. E. van Raaij, I. M. Segers-Nolten, and V. Subramaniam, Biophys. J. **91**, L96 (2006).

[29] H. A. Lashuel, O. C. R. , A. Oueslati, and E. Masliah, Nat. Rev. **14**, 38 (2013).

[30] F. Bemporad and F. Chiti, Chem. Biol. **19**, 315 (2012).

[31] S. I. A. Cohen, P. Arosio, J. Presto, F. R. Kurudenkandy, H. Biverstål, L. Dolfe, C. Dunning, X. Yang, B. Frohm, M. Vendruscolo, J. Johansson, C. M. Dobson, A. Fisahn, T. P. J. Knowles, and S. Linse, Nat. Struct. Mol. Biol. **22**, 207 (2015).

[32] D. Scott, I. Tabarean, Y. Tang, A. Cartier, E. Masliah, and S. Roy, J. Neurosci. **30**, 8083 (2010).

[33] A. Oueslati, K. E. Paleologou, B. L. Schneider, P. Aebischer, and H. A. Lashuel, J. Neurosci. **32**, 1536 (2012).

[34] O. Marques and T. F. Outeiro, Cell Death Dis. **3**, e350 (2012).

[35] M. Stefani and C. M. Dobson, J. Mol. Med. **81**, 678 (2003).

[36] P. Tompa, Trends Biochem. Sci. **37**, 509 (2012).

[37] P. Tompa, E. Schad, A. Tantos, and L. Kalmar, Curr. Opin. Struct. Biol. **35**, 49 (2015).

[38] V. N. Uversky, J. Li, and A. L. Fink, J. Biol. Chem. **276**, 10737 (2001).

[39] B. Caughey, J. Peter, and T. Lansbury, Annu. Rev. Neurosci. **26**, 267 (2003).

[40] L. Giehm, D. I. Svergun, D. E. Otzen, and B. Vestergaard, Proc. Natl. Acad. Sci. U. S. A. **108**, 3246 (2011).

[41] N. Zijlstra, C. Blum, I. M. J. Segers-Nolten, M. M. A. E. Claessens, and V. Subramaniam, Angew. Chem. **124**, 8951 (2012).

[42] K. Ono, M. M. Condron, and D. B. Teplow, Proc. Natl. Acad. Sci. U. S. A. **106**, 14745 (2009).

[43] H. A. Lashuel, B. M. Petre, J. Wall, M. Simon, R. J. Nowak, T. Walz, and P. T. L. , Jr., J. Mol. Biol. **322**, 1089 (2002).

[44] H. A. Lashuel, D. Hartley, B. M. Petre, T. Walz, and P. T. Lansbury, Nature **418**, 291 (2002).

[45] M. E. van Raaij, J. van Gestel, I. M. Segers-Nolten, S. W. de Leeuw, and V. Subramaniam, Biophys. J. **95**, 4871 (2008).

[46] S. A. Semerdzhiev, D. R. Dekker, V. Subramaniam, and M. M. A. E. Claessens, ACS Nano **8**, 5543 (2014).

[47] J. D. Schmit, K. Ghosh, and K. Dill, Biophys. J. **100**, 450 (2011).

[48] J. Zhang and M. Muthukumar, J. Chem. Phys. **130**, 035102 (2009).

[49] F. Sciortino, E. Bianchi, J. F. Douglas, and P. Tartaglia, J. Chem. Phys. **126**, 194903 (2007).

[50] R. Pellarin, E. Guarnera, and A. Caflisch, J. Mol. Biol. **374**, 917 (2007).

[51] R. Pellarin, P. Schuetz, E. Guarnera, and A. Caflisch, J. Am. Chem. Sci. **132**, 14960 (2010).

[52] B. Barz and B. Urbanc, J. Phys. Chem. B **118**, 3761 (2014).

[53] N. S. Bieler, T. P. J. Knowles, D. Frenkel, and R. Vacha, PLoS Comput. Biol. **8**, e1002692 (2012).

[54] A. Saric, Y. C. Chebaro, T. P. J. Knowles, and D. Frenkel, Proc. Natl. Acad. Sci. U. S. A. **111**, 17869 (2014).

[55] A. Saric, A. K. Buell, T. C. Michaels, S. Linse, T. P. J. Knowles, and D. Frenkel, "Physical determinants for the self-replication of protein fibrils" (unpublished).

[56] I. M. Ilie, W. J. Briels, and W. K. den Otter, J. Chem. Phys. **142**, 114103 (2015).

[57] S. 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).

[58] J. C. Kessler, J.-C. Rochet, and P. T. Lansbury, Biochemistry **42**, 672 (2003).
[59] N. Lorenzen, L. Lemminger, J. N. Pedersen, S. B. Nielsen, and D. E. Otzen, FEBS Lett. **588**, 497 (2014).
[60] D.-P. Hong, W. Xiong, J.-Y. Chang, and C. Jiang, FEBS Lett. **585**, 561 (2011).
[61] C. Vega and S. Lago, Comput. Chem. **18**, 55 (1994).
[62] A. Sidhu, I. Segers-Nolten, and V. Subramaniam, Biochim. Biophys. Acta, Proteins Proteomics **1844**, 2127 (2014).
[63] L. Bousset, L. Pieri, G. Ruiz-Arlandis, J. Gath, P. H. Jensen, B. Habenstein, K. Madiona, V. Olieric, A. Böckmann, B. H. Meier, and R. Melki, Nat. Commun. **4**, 2575 (2013).
[64] H. C. Öttinger, *Stochastic Processed in Polymeric Fluids* (Springer, 1996).
[65] C. W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*, 3rd ed. Springer Series in Synergetics Vol. 13 (Springer-Verlag, Berlin, 2004), p. XVIII+415.
[66] S. N. Naess and A. Elgsaeter, Macromol. Theory Simul. **13**, 419 (2004).
[67] T. R. Evensen, S. N. Naess, and A. Elgsaeter, Macromol. Theory Simul. **17**, 121 (2008).
[68] T. R. Evensen, A. Elgsaeter, and S. N. Naess, Colloids Surf., B **56**, 80 (2007).
[69] T. R. Evensen, S. N. Naess, and A. Elgsaeter, Macromol. Theory Simul. **18**, 50 (2009).
[70] T. R. Evensen, S. N. Naess, and A. Elgsaeter, Macromol. Theory Simul. **17**, 403 (2008).
[71] J.-P. Ryckaert, G. Ciccotti, and H. J. Berendsen, J. Comput. Phys. **23**, 327 (1977).
[72] S. D. Stoddard, J. Comput. Phys. **27**, 291 (1978).
[73] W. Humphrey, A. Dalke, and K. Schulten, J. Mol. Graphics **14**, 33 (1996).
[74] S. E. Blondelle, B. Forood, R. A. Houghten, and E. Pérez-Payá, Biochemistry **36**, 8393 (1997).
[75] M. Fändrich and C. M. Dobson, EMBO J. **21**, 5682 (2002).
[76] S. Auer, C. M. Dobson, and M. Vendruscolo, HFSP J. **1**, 137 (2007).
[77] B. Ma and R. Nussinov, Protein Sci. **11**, 2335 (2002).
[78] R. Pellarin and A. Caflisch, J. Mol. Biol. **360**, 882 (2006).
[79] B. van Rooijen, K. van Leijenhorst-Groener, M. Claessens, and V. Subramaniam, J. Mol. Biol. **394**, 826 (2009).
[80] H.-Y. Kim, M.-K. Cho, A. Kumar, E. Maier, C. Siebenhaar, S. Becker, C. O. Fernandez, H. A. Lashuel, R. Benz, A. Lange, and M. Zweckstetter, J. Am. Chem. Sci. **131**, 17482 (2009).
[81] M. M. Apetri, N. C. Maiti, M. G. Zagorski, P. R. Carey, and V. E. Anderson, J. Mol. Biol. **355**, 63 (2006).
[82] W. Hoyer, D. Cherny, V. Subramaniam, and T. M. Jovin, J. Mol. Biol. **340**, 127 (2004).
[83] S. R. Aragon and D. Flamik, Macromolecules **42**, 6290 (2009).
[84] G. H. Koenderink and A. P. Philipse, Langmuir **16**, 5631 (2000).