J|A|C|S
ARTICLES

Published on Web 10/05/2010

Amyloid Fibril Polymorphism Is under Kinetic Control

Riccardo Pellarin, Philipp Schuetz,[†] Enrico Guarnera, and Amedeo Caflisch*

Department of Biochemistry, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland

Received July 8, 2010; E-mail: caflisch@bioc.uzh.ch

Abstract: Self-assembly of proteins into amyloid aggregates displays a broad diversity of morphologies, both at the protofibrillar and final fibrillar species. These polymorphic species can coexist at fixed experimental conditions, and their relative abundance can be controlled by changing the solvent composition, or stirring the solution. However, the extent to which external conditions regulate the equilibrium of morphologically distinct species is still unknown. Here we investigate the nucleation of distinct fibril morphologies using computer simulations of a simplified model of an amyloid polypeptide. Counterintuitively, the energetically less favorable fibril morphologies nucleate more frequently than the morphologies of higher stability for models with low aggregation propensity. The free-energy profiles of the aggregation process indicate that the nucleation barrier determines the population fractions of different fibril morphologies, i.e., amyloid polymorphism is under kinetic control.

1. Introduction

Amyloid fibrils, frequent products of (poly)peptide self-assembly, are implicated in a number of neurodegenerative and systemic diseases.[1] These aggregates have the ability to self-replicate their structures: the growing edges of the fibril, consisting of the tips and the flanks, act as a templating surfaces for the depositing monomers.[2] In fact, these aggregates catalyze the conversion from the soluble to the amyloid state of the polypeptide, which is forced to adopt a fibril-specific conformation. Moreover, the amyloid aggregation is intrinsically polymorphic, both at the intermediate and the final fibrillar level.[3] Experiments based on electron and atomic force microscopy as well as solid-state NMR spectroscopy revealed that changing the sample's conditions, such as the pH[4,5] or the cosolvent concentration,[6] or introducing a mechanical perturbation[7,8] results in different fibril morphologies. Conversely, even within the same sample, a number of coexisting morphologies can be detected.[4,8–10]

The ability of amyloid fibrils to assume alternative structures might have a biological importance. Amyloid fibrils isolated from tissues of patients affected by systemic amyloidosis display a significant structural polymorphism,[11] and distinct fibril architectures have different functional activity and cytotoxicity.[4,7,12] Furthermore, this structural heterogeneity might be at the origin of species barrier[13] and strain effects in prions.[14]

Which stage of the maturation of the amyloid fibrils is competent for the structural differentiation is hitherto unclear. Using electron microscopy imaging of the dynamics of Aβ-fibril maturation, Goldsbury et al. propose an assembly mechanism consisting of multiple independent pathways that give rise to distinct fibril morphologies.[15] Tycko and co-workers, on the basis of solid-state NMR and electron microscopy of Aβ-peptide fibrils, hypothesize that the morphology that eventually dominates a solution depends on the effects of the environment on the nucleus formation and propagation.[8] Similarly, Dzwolak et al. showed the property of insulin to produce two distinct morphologies with opposite optical chiralities.[16] The authors also hypothesize a mechanism of morphogenesis, that acts at the nucleus level and whose bifurcation is regulated by the external conditions.

Amyloid aggregation has been extensively investigated by computer simulation methods. Atomistic models were employed to study the conformational space of amyloidogenic polypeptides in the monomeric state,[17,18] the very initial steps of amyloid

---

[†] Current address: Department of Electronics/Metrology/Reliability, Swiss Federal Laboratories for Materials Science and Technology, Ueberlandstrasse 129, 8600 Dübendorf, Switzerland.

(1) Merlini, G.; Bellotti, V. *New Engl. J. Med.* **2003**, *349*, 583–596.
(2) Jansen, R.; Dzwolak, W.; Winter, R. *Biophys. J.* **2005**, *88*, 1344–1353.
(3) Kodali, R.; Wetzel, R. *Curr. Opin. Struct. Biol.* **2007**, *17*, 48–57.
(4) Wasmer, C.; Soragni, A.; Sabate, R.; Lange, A.; Riek, R.; Meier, B. H. *Angew. Chem., Int. Ed.* **2008**, *47*, 5839–5841.
(5) Gosal, W. S.; Morten, I. J.; Hewitt, E. W.; Smith, D. A.; Thomson, N. H.; Radford, S. E. *J. Mol. Biol.* **2005**, *351*, 850–864.
(6) Dzwolak, W.; Grudzielanek, S.; Smirnovas, V.; Ravindra, R.; Nicolini, C.; Jansen, R.; Loksztejn, A.; Porowski, S.; Winter, R. *Biochemistry* **2005**, *44*, 8948–8958.
(7) Petkova, A. T.; Leapman, R. D.; Guo, Z.; Yau, W.-M.; Mattson, M. P.; Tycko, R. *Science* **2005**, *307*, 262–265.
(8) Paravastu, A. K.; Petkova, A. T.; Tycko, R. *Biophys. J.* **2006**, *90*, 4618–4629.
(9) Goldsbury, C. S.; Wirtz, S.; Müller, S. A.; Sunderji, S.; Wicki, P.; Aebi, U.; Frey, P. *J. Struct. Biol.* **2000**, *130*, 217–231.
(10) Meinhardt, J.; Sachse, C.; Hortschansky, P.; Grigorieff, N.; Fändrich, M. *J. Mol. Biol.* **2009**, *386*, 869–877.
(11) Jimenez, J. L.; Tennent, G.; Pepys, M.; Saibil, H. R. *J. Mol. Biol.* **2001**, *311*, 241–247.
(12) Seilheimer, B.; Bohrmann, B.; Bondolfi, L.; Müller, F.; Stüber, D.; Döbeli, H. *J. Struct. Biol.* **1997**, *119*, 59–71.
(13) Tanaka, M.; Chien, P.; Yonekura, K.; Weissman, J. S. *Cell* **2005**, *121*, 49–62.
(14) Jones, E. M.; Surewicz, W. K. *Cell* **2005**, *121*, 63–72.
(15) Goldsbury, C.; Frey, P.; Olivieri, V.; Aebi, U.; Müller, S. A. *J. Mol. Biol.* **2005**, *352*, 282–298.
(16) Dzwolak, W.; Loksztejn, A.; Galinska-Rakoczy, A.; Adachi, R.; Goto, Y.; Rupnicki, L. *J. Am. Chem. Soc.* **2007**, *129*, 7517–7522.
(17) Vitalis, A.; Wang, X.; Pappu, R. V. *Biophys. J.* **2007**, *93*, 1923–1937.

10.1021/ja106044u

formation,[19−25] and the structural stability of fibril models.[26−28] Yet, all-atom simulations are unsuitable to study the pathways and the kinetics of fibril formation. Low-resolution models, which make use of a simplified representation of protein geometry and the interactions of its components, have provided fundamental insights into the detailed mechanisms of protein oligomerization,[29−34] and fibrillation.[35−41] Nonetheless, these works do not explain the mechanisms underlying the spontaneous differentiation of fibril morphologies.

Here, using a simplified off-lattice model of an amphipathic polypeptide,[42] we investigate the production of distinct fibrillar morphologies and the underlying mechanisms. The peptide monomer has a single degree of freedom and its free-energy profile has two minima corresponding to the amyloid competent (β) and amyloid protected states (π) (see Figure 1). Hence, the free-energy difference between the two states dictates the amyloidogenic potential of the monomer. In a previous work,[43] we showed that by varying the single parameter of the model, i.e., by reducing the β-aggregation propensity, the roughness of the free-energy landscape and the heterogeneity of the fibril elongation pathways increase. At high amyloidogenic conditions (β-stable scenario), the process of fibril formation is downhill (i.e., without intermediates) and kinetically fast, whereas at low amyloidogenic conditions (β-unstable scenario), several intermediates are detected, including micellar oligomers, and protofibrils with two or three protofilaments. The fibril maturation is hierarchical, where the protofibrils transform to fibrils by templated protofilament formation. Furthermore, simulations of the same peptide model in conjunction with a vesicle made of coarse-grained lipids allowed the study of the effects of amyloid aggregation on membrane permeability[44] and the fibril degradation induced by monomer adsorption on the membrane.[45]

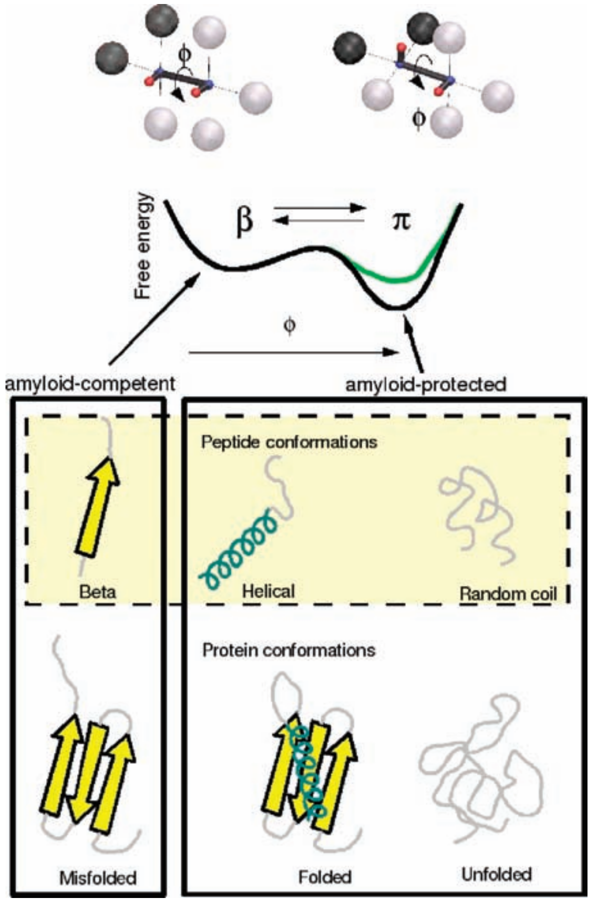

Figure 1. The model. Sticks and beads representations of the monomer in the amyloid competent state, β, and the amyloid protected state, π. The large spheres are hydrophobic (black) and hydrophilic (gray), while the two dipoles are shown with small red and blue spheres. The β- and π-states of the monomer are shown on top of the two corresponding minima of the free energy plotted as a continuous function of the dihedral angle ϕ of the two dipoles. Note that the population of monomers in the β-state decreases by lowering the free energy of the π-state, as indicated by the green and black profiles. In the β-state, the parallel orientation of the two intramolecular dipoles favors ordered aggregates, with intermolecular dipolar interactions parallel to the fibril axis. Conversely, the π-state represent the ensemble of all polypeptide conformations that are not compatible with self-assembly into a fibril. Reprinted with permission from ref 43.

In the present work, after giving evidence that the model spontaneously produces a number of distinct morphologies, we address the following questions. What is the effect of the free-energy profile of the monomer on the population of the different morphologies? Do the different morphologies have the same pathways in terms of intermediates? How are the morphologies selected by the aggregation mechanism? In the simulations with the model reflecting a peptide with low-amyloidogenicity potential, the most stable fibril morphology is observed less frequently than less stable morphologies. In other words, fibril formation is under kinetic control.

---

(18) Vitalis, A.; Lyle, N.; Pappu, R. V. *Biophys. J.* **2009**, *97*, 303–311.
(19) Gsponer, J.; Haberthr, U.; Caflisch, A. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 5154–5159.
(20) Hwang, W.; Zhang, S.; Kamm, R. D.; Karplus, M. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 12916–12921.
(21) de la Paz, M. L.; de Mori, G. M. S.; Serrano, L.; Colombo, G. *J. Mol. Biol.* **2005**, *349*, 583–596.
(22) Cecchini, M.; Curcio, R.; Pappalardo, M.; Melki, R.; Caflisch, A. *J. Mol. Biol.* **2006**, *357*, 1306–1321.
(23) Strodel, B.; Whittleston, C. S.; Wales, D. J. *J. Am. Chem. Soc.* **2007**, *129*, 16005–16014.
(24) De Simone, A.; Esposito, L.; Pedone, C.; Vitagliano, L. *Biophys. J.* **2008**, *95*, 1965–1973.
(25) Bellesia, G.; Shea, J.-E. *Biophys. J.* **2009**, *96*, 875–886.
(26) Ma, B.; Nussinov, R. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 14126–14131.
(27) Buchete, N.-V.; Tycko, R.; Hummer, G. *J. Mol. Biol.* **2005**, *353*, 804–821.
(28) Wu, C.; Bowers, M. T.; Shea, J.-E. *PLoS Comput. Biol.* **2010**, *6*, e1000693.
(29) Ding, F.; Dokholyan, N. V.; Buldyrev, S. V.; Stanley, H. E.; Shakhnovich, E. I. *J. Mol. Biol.* **2002**, *324*, 851–857.
(30) Dima, R. I.; Thirumalai, D. *Protein Sci.* **2002**, *11*, 1036–1049.
(31) Jang, H.; Hall, C. K.; Zhou, Y. *Biophys. J.* **2004**, *86*, 31–49.
(32) Urbanc, B.; Cruz, L.; Yun, S.; Buldyrev, S. V.; Bitan, G.; Teplow, D. B.; Stanley, H. E. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 17345–17350.
(33) Chebaro, Y.; Mousseau, N.; Derreumaux, P. *J. Phys. Chem. B* **2009**, *113*, 7668–7675.
(34) Williamson, T. E.; Vitalis, A.; Crick, S. L.; Pappu, R. V. *J. Mol. Biol.* **2010**, *396*, 1295–1309.
(35) Nguyen, H. D.; Hall, C. K. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 16180–16185.
(36) Nguyen, H. D.; Hall, C. K. *J. Biol. Chem.* **2005**, *280*, 9074–9082.
(37) Nguyen, H. D.; Hall, C. K. *J. Am. Chem. Soc.* **2006**, *128*, 1890–1901.
(38) Auer, S.; Meersman, F.; Dobson, C. M.; Vendruscolo, M. *PLoS Comput. Biol.* **2008**, *4*, e1000222.
(39) Bellesia, G.; Shea, J.-E. *J. Chem. Phys.* **2009**, *130*, 145103.
(40) Zhang, J.; Muthukumar, M. *J. Chem. Phys.* **2009**, *130*, 035102.
(41) Urbanc, B.; Betnel, M.; Cruz, L.; Bitan, G.; Teplow, D. B. *J. Am. Chem. Soc.* **2010**, *132*, 4266–4280.
(42) Pellarin, R.; Caflisch, A. *J. Mol. Biol.* **2006**, *360*, 882–892.
(43) Pellarin, R.; Guarnera, E.; Caflisch, A. *J. Mol. Biol.* **2007**, *374*, 917–924.
(44) Friedman, R.; Pellarin, R.; Caflisch, A. *J. Mol. Biol.* **2009**, *387*, 407–415.
(45) Friedman, R.; Pellarin, R.; Caflisch, A. *J. Phys. Chem. Lett.* **2010**, *1*, 471–474.

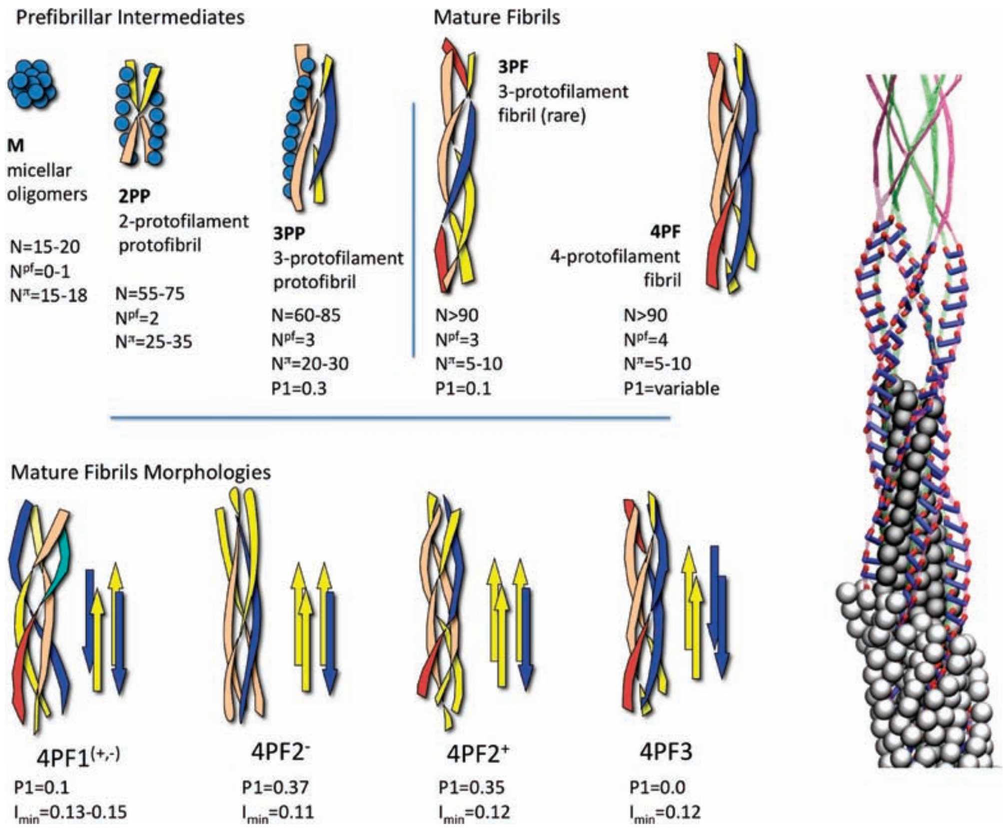

Figure 2. Morphologies of mature fibrils and prefibrillar species. (Top) The prefibrillar species are: the micellar oligomers M, consisting of π-monomers (blue beads) aggregated through hydrophobic forces; the 2-protofilament protofibril (2PP), and the 3-protofilament protofibril (3PP), which are early stages of fibril maturation, where the π-monomers are deposited onto the lateral surface of the fibril, and the β-monomers make up the protofilaments (colored ribbons). Mature fibrils display a 4-protofilament structure (4PF), or, sporadically, a 3-protofilament structure (3PF). Average values of the aggregation number N, the number of protofilaments N^{pf}, the number of π-monomers N^{π} and the polar order parameter P_1 are reported. (Bottom) The 4PF morphologies have different orientation of the protofilaments (represented by the blue and yellow arrows), organization of up and down protofilaments (see for instance 4PF3 and 4PF1), and thickness of the fibril (see for instance 4PF2− and 4PF2+). Average values of P_1 and the minimum inertia moment $I_{\min}$ are indicated. (Right) Structure of a mature 4PF2+ fibril. From top to bottom the different elements constituting the fibril are overlaid. The green and pink ribbons represent the polar interactions between the monomers. A pair of green and pink ribbons constitutes a single protofilament. The blue and red sticks represent the two-dipole system of the monomer. Note that this fibril has three protofilaments with up-oriented dipoles and one down-oriented. The dark-gray and light-gray spheres are the hydrophobic and hydrophilic beads, respectively.

2. Material and Methods

Simplified Model of Amphipathic Peptide. The amphipathic peptide model has been described extensively elsewhere.[42,43] This model does not represent any particular amino acid sequence; yet, it captures the general properties of amyloidogenic (poly)peptides and the heterogeneous phenomenology of fibril formation. The simplified peptide (termed monomer in the following) consists of 10 spheres, 4 of which carry partial charges of $\pm 0.4e$, thereby generating two dipoles that drive the cross-β association (see Figure 1). Two of the remaining six spheres have more pronounced Lennard-Jones energy minimum (−1.3 kcal/mol) than the other 4 (−0.1 kcal/mol). These energy parameters are chosen to mimic an amphipathic moment, where the former two spheres behave as a permanent hydrophobic surface of the peptide. The formation of multifilament fibrils is a result of this choice. The interaction between the protofilaments is stabilized by the hydrophobic contacts, and the hydrophilic surface prevents the association of more that 4 protofilaments in the same fibril (see Figure 2).

The conformational landscape of the isolated monomer is simplified such that only two states are considered: the amyloid-competent β and the amyloid-protected π. The interconversion between the two states involves a single degree of freedom consisting in the dihedral angle between the two dipoles (see Figure 1). In the β-state the orientation of the two dipoles is parallel, favoring ordered aggregation with intermolecular dipolar interaction along the fibril axis. Conversely, the π-state represents the ensemble of all polypeptide conformations that are not compatible with self-assembly into a fibril (e.g., random coil, partially unfolded, helical). In this state the two dipoles are orthogonal, preventing the fibrillar aggregation. Owing to the amphipathic moment, the monomers in the π-state are still free to associate into micellar oligomers, with the hydrophobic spheres partitioned in the interior of the aggregate. During the simulations, the isolated monomer undergoes a reversible isomerization between the π- and the β-states, and the energy difference between these two states, $dE = E_{\pi} - E_{\beta}$, is related to the β-aggregation propensity of the polypeptide, which is the relative probability of the β- and π-states. For instance, when $dE = 0$ kcal/mol, the π- and β-states are equally populated, whereas for $dE = -1.5$ and -2.5 kcal/mol, the π-state is about 15 and 100 times more populated than the β-state, respectively. In the following we will express β-aggregation propensity of the monomer in terms of the energy difference dE. Note that the β-aggregation propensity dE cannot be directly compared to experimental measures. Yet, it can be related to relative changes, e.g., upon single-residue mutation.[46]

Langevin Dynamics. The amyloid aggregation simulations are started by placing 125 monodispersed monomers in a cubic box with a size of 290 Å, corresponding to a concentration of 8.5 mM. The system is minimized, equilibrated and simulated by Langevin

---

(46) Christopeit, T.; Hortschansky, P.; Schroeckh, V.; Gührs, K.; Zandomeneghi, G.; Fändrich, M. *Protein Sci.* **2005**, *14*, 2125–2131.

Table 1. Thermodynamic Properties of Morphologies at $dE = -2.25$[a]

	4PF1	4PF2−	4PF2+
C_r [mM]	1.59 ± 0.15	1.21 ± 0.11	0.98 ± 0.08
Δg [kcal/mol]	-3.96	-4.13	-4.26
Δe [kcal/mol]	-14.21 ± 0.07	-15.08 ± 0.20	-15.22 ± 0.20
ΔF_{Nucl} [kcal/mol]	3.70	4.46	3.89

[a] The thermodynamic stability of the fibril is $\Delta g = kT \log(C_r)$ where C_r is the critical concentration, i.e., average concentration of dissociated monomers. Δe is the specific potential energy. The most stable morphology is 4PF2+, followed by 4PF2− and 4PF1. The nucleation barrier ΔF is evaluated from the cFEP plots (Figure 7).

dynamics using CHARMM[47] with a friction parameter of 0.01 ps^{-1} and a temperature of 310 K. Four values of the parameter dE where investigated, namely −1.5, −2.0, −2.25, and −2.5 kcal/mol, and 100 aggregation simulations with different random values for initial velocities were run for each dE value. The length of the simulations depends on the nucleation time and thus on dE. When $dE = -2.5$, a simulation length of 20 μs is needed to observe about 80 nucleation events in a pool of 100 simulations; when $dE = -1.5$, 100 simulation runs nucleate within 0.1 μs. The range of dE is sufficient to describe all the relevant aggregation phenomenology of the model, in fact all models with $dE > -2.0$ behave similarly, whereas when $dE < -2.5$ the system does not nucleate within 40 μs. A 20 μs simulation requires about 40 days on a single CPU.

The Observables Used to Identify the Fibril Morphologies. An aggregate consists of monomers whose mutual minimal distance between any sphere is less than 6 Å. Aggregated monomers are identified automatically using a clustering procedure as described previously.[42] The analysis is based on the largest aggregate (la) in the simulation box at a given simulation time, and in particular its size N_{la}, number of β-monomers contained N_{la}^{β}, number of π-monomers N_{la}^{π}, and number of protofilaments $N_{\mathrm{la}}^{\mathrm{pf}}$ (see Supporting Information).[43]

The parameter P_1 is used to measure the orientation order within an aggregate. It is calculated by using the formula:

$$P_1 = \frac{1}{N}\sum_{i=1}^{N} \hat{\mathbf{z}}_i \cdot \hat{\mathbf{d}}$$

where the dipoles of the peptides belonging to the aggregate are used as molecular vectors $\hat{\mathbf{z}}_i$, N is the total number of vectors and $\hat{\mathbf{d}}$ is the eigenvector that corresponds to the largest positive eigenvalue of the order matrix.[48] In a disordered oligomer, such as a micelle, P_1 is close to zero. In ordered structures, such as protofibrils or fibrils, the value depends on the relative orientation of the protofilaments that make up a fibril. If the protofilaments are oriented in a completely antiparallel or parallel fashion, P_1 is close to zero or one, respectively. Other orientations result in intermediate values of P_1.

The minimum moment of inertia $I_{\min}$ measures the thickness of the fibril, and it is calculated as the minimum eigenvalue of the fibril inertia tensor I_{ij}. The diagonal terms of the inertia tensor are $I_{ii} = \sum_{n=1}^{S} m_n(\sum_{k \neq i} X_{k,n}^2)$, and the off-diagonal terms are $I_{ij} = -\sum_{n=1}^{S} m_n X_{i,n} X_{j,n}$. The indexes $i,j,k = 1,2,3$ represent the coordinates (i.e., $X_1 = x$, $X_2 = y$, $X_3 = z$), n is the index of the spheres belonging to the fibril, S is the total number of spheres, and m_n is the mass of sphere n.

Critical Concentration and Specific Potential Energy of Fibrils. The quantities reported in Table 1 are calculated as follows. Snapshots of mature fibrils ($N > 90$ and $N^{\mathrm{pf}} \geq 4$) were isolated from nucleation trajectories of $dE = -2.25$ kcal/mol, and separated according to the morphology to ensure the homogeneity of the measurements. The critical concentration C_r is the average concentration of residual dissociated monomers that are in equilibrium with a mature fibril of a specific morphology. It is equivalent to the dissociation constant K_d and is connected to the free energy of monomer association $\Delta g = kT \log(K_d)$,[49,50] which is a measure of fibril stability. The specific potential energy Δe is calculated as:

$$\Delta e = \left\langle \frac{1}{N_{\mathrm{fibril}}}(E_{\mathrm{elec}} + E_{\mathrm{vdw}} + E_{\phi}) \right\rangle_{\mathrm{morph}} - e_{\mathrm{mon}}$$

where $\langle \cdot \rangle_{\mathrm{morph}}$ is the average computed over all snapshots of a given morphology, N_{fibril} is the aggregation number of the fibril, E_{elec}, E_{vdw}, and E_{ϕ} are the electrostatic, van der Waals and dihedral potential energies of the fibril, respectively, and e_{mon} is the average potential energy of the dissociated monomer. Δe expresses the average potential energy gain for a single monomer that binds to the fibril. Being an energy difference, Δe only slightly changes upon varying dE (data not shown). In fact, lowering dE corresponds to an increase of π-monomers deposited at the growing edges of the fibril that have a higher Δe with respect to the β-monomers embedded into the fibril, but they contribute only marginally to the total energy of the fibril because the number of π-monomers is about 7 in a fibril of 88 monomers when $dE = -2.5$. On the contrary, C_r changes, but the rank of values for the different morphologies is not influenced by dE (see Table S1, Supporting Information).

Aggregation State Fingerprints. Aggregation states of the system are coarse-grained using fingerprints, which are strings composed by the observables N_{la}, $N_{\mathrm{la}}^{\mathrm{pf}}$, N_{la}^{β}, P_1, and $I_{\min}$. In this way the aggregation trajectories are translated into symbolic time series that can be used to compute the free energy profile, as described below. The observables are combined in the fingerprint according to the number of protofilaments $N_{\mathrm{la}}^{\mathrm{pf}}$ to optimally distinguish between the different intermediates. When $N_{\mathrm{la}}^{\mathrm{pf}}$ is equal to 0 or 1 (early oligomerization), only N_{la} and N_{la}^{β} are employed in the fingerprint. Hence, the micellar oligomer M ($7 < N_{\mathrm{la}} < 30$) can be distinguished from the initial small oligomers ($N_{\mathrm{la}} < 7$), and from higher-order oligomers ($N_{\mathrm{la}} > 30$) with variable quantity of β-monomers. These threshold values are chosen according to the oligomer size distribution obtained at the lag phase for $dE = -2.5$ (see Figure S10, Supporting Information). When $N_{\mathrm{la}}^{\mathrm{pf}}$ is equal to 2 or 3 (protofibrillar intermediates), then the fingerprint observables are $N_{\mathrm{la}}^{\mathrm{pf}}$, N_{la}^{π}, and P_1. This set of observables is able to discern 2PP from 3PP, and 3PP from 3PF, as the main difference between 3PP and 3PF is the quantity of π-monomers which are deposited onto it, and different values of P_1. When $N_{\mathrm{la}}^{\mathrm{pf}}$ is equal to 4 or larger (mature fibrils), the quantities $N_{\mathrm{la}}^{\mathrm{pf}}$, P_1, and $I_{\min}$ are sufficient to distinguish between the four morphologies.

Cut-Based Free Energy Profile (cFEP). Krivov and Karplus have exploited an analogy between the kinetics of a complex process and equilibrium flow through a network to develop the cFEP, a projection of the free energy surface that preserves the barriers[51] and can be used for extracting protein folding pathways and molecular mechanisms from MD simulations.[52] The input for the cFEP calculation is the equilibrium transition network (ETN) whose nodes and links are the state fingerprints and the number of direct transitions (observed within 0.5 ns along the trajectories), respectively. For a node i in the network the partition function is $Z_i = \sum_j c_{ij}$, where c_{ij} is the absolute number of transitions from node i to node j observed along the time series. If the nodes of the ETN are partitioned into two sets $\mathscr{A}$ and $\mathscr{B}$, where set $\mathscr{A}$ contains the reference node A, then

---

(47) Brooks, B. R.; et al. *J. Comput. Chem.* **2009**, *30*, 1545–1614.

(48) Cecchini, M.; Rao, F.; Seeber, M.; Caflisch, A. *J. Chem. Phys.* **2004**, *121*, 10748–10756.

(49) O'Nuallain, B.; Shivaprasad, S.; Kheterpal, I.; Wetzel, R. *Biochemistry* **2005**, *44*, 12709–12718.

(50) Williams, A. D.; Portelius, E.; Kheterpal, I.; tao Guo, J.; Cook, K. D.; Xu, Y.; Wetzel, R. *J. Mol. Biol.* **2004**, *335*, 833–842.

(51) Krivov, S. V.; Karplus, M. *J. Phys. Chem. B* **2006**, *110*, 12689–12698.

(52) Krivov, S. V.; Muff, S.; Caflisch, A.; Karplus, M. *J. Phys. Chem. B* **2008**, *112*, 8701–8714.

$$\begin{aligned} Z_A &= \sum_{i\in\mathcal{A}} Z_i \\ Z_B &= \sum_{i\in\mathcal{B}} Z_i \\ Z_{AB} &= \sum_{i\in\mathcal{A},j\in\mathcal{B}} c_{ij} \end{aligned}$$

and the free energy of the barrier between the two groups is $\Delta F = -kT\log(Z_{AB}/Z)$ where Z is the partition function of the full ETN. The progress coordinate is then the normalized partition function Z_A/Z of the reactant region containing the native node A, but other progress coordinates can be used, because the cFEP is invariant with respect to arbitrary transformations of the reaction coordinate. Note that it might be possible to extract geometrically intuitive reaction coordinates from the ETN, as shown recently for the reversible folding of a β-sheet peptide,[53] but such analysis would go beyond the purpose of the present study. In practice, the cFEP is calculated (using the program WORDOM[54]) from the ETN in three steps: (1) The mean first passage time (mfpt) to a reference state is calculated analytically for each node on the ETN by solving a system of transition rate equations.[51,52] (2) Nodes are sorted by increasing values of mfpt, and for each of these values the relative partition function Z_A and the cut Z_{AB} are calculated. (3) The individual points on the profile are evaluated as $[x = Z_A/Z, y = \Delta F = -kT\log(Z_{AB}/Z)]$. The result is a one-dimensional profile that preserves the barrier height between the free energy basins.

Simplified Graphs. To isolate the free energy basins and determine the number of transitions between basins, an automatic procedure based on the cFEP has been employed. The cFEP calculated given a reference state is a curve with a number of minima and maxima. The first maximum, i.e. with the lowest Z_A/Z, is the exiting barrier from the free energy basin that contains the reference state.[52] Thus, all the states whose Z_A/Z value lies between those of the reference state and the first maximum belong to the same basin. Owing to the kinetic superposition, originating from the presence of states with comparable mfpt from the reference state, all the states whose Z_A/Z is beyond the first maximum cannot be unambiguously assigned to a free energy basin.[51,52] To overcome this difficulty, a recursive procedure has been conceived to correctly assign the states to the basins. Once the first basin is identified, a new reference state, corresponding to the first minimum of the cFEP beyond the first barrier, is chosen. All the states belonging to the detected basin are removed from the fingerprint time series and replaced by a“stop-state” that is used to separate distinct trajectories. This step is necessary to ensure that already assigned states are not reassigned to other basins. The resulting time series consists of fragments of trajectories containing the transitions among the remaining states. Next, the cFEP is recalculated on the resulting fingerprint time series with respect to the new reference state, and the procedure is repeated until all the states are associated to a basin. Finally, the original trajectories of state fingerprints are transformed into time series of basin index, and simplified graphs, depicting the interbasin transitions, are produced. Two examples, based on random walkers on a predetermined 2D energy surface, are reported in the Supporting Information (see Figures S1 and S2). In both cases the method correctly partitions the energy surface into basins, and the corresponding simplified graph reproduces the expected transitions.

3. Results

Hereafter, the unit dimension of the β-aggregation propensity dE, that is kcal/mol, will be omitted for simplicity. Most of the analysis focuses on the more frequent morphologies, i.e., 4PF1, 4PF2+, and 4PF2−.

Morphological Species Classification and Seeding Propagation. The protofibrillar morphologies are displayed in Figure 2 (note that the π-monomers are schematically represented by blue beads, and protofilaments consisting of aggregated β-monomers are depicted as ribbons). Spherical oligomers are micellar (M) assemblies of π-monomers with hydrophilic and hydrophobic spheres arranged outside and inside, respectively. They are populated during the lag phase of simulations with $dE = -2.25$ and $dE = -2.5$, and have a size of 15−20 monomers (see Figure S10, Supporting Information), whereas 2-protofilaments protofibrils (2PP) and 3-protofilaments protofibrils (3PP) appear after the nucleation as immature fibrils with one or more deposits of π-monomers and filaments consisting of ordered β-monomers. The mature fibrils (4PF) is made up of 4 protofilaments. In previous simulation studies with the same coarse-grained model, the mechanism of nucleation from the micellar oligomer and the hierarchical formation of ordered protofilaments was elucidated. The hydrophobic collapse drives the formation of the micellar oligomers, whose kinetic stability is proportional to the stability of the π-state of the monomer.[42] In fact, for high-amyloidogenic potentials the nucleus size is submicellar (about 4 and 15 monomers for $dE = -1.5$ and -2.0, respectively; see Figure S11, Supporting Information), and the fibril nucleation is faster than the micelle nucleation. For low-amyloidogenic potentials the nucleation occurs within a micelle and the nucleus is larger than the micelle (about 20 and 40 monomers for $dE = -2.25$ and -2.5, respectively; see Figure S11, Supporting Information). In this last case, the nucleation is a cooperative interconversion of monomers from the π-state to the β-state and exploits the spatial proximity of monomers associated into oligomers. These results are in accord with very recent discontinuous molecular dynamics simulations of Aβ oligomerization.[41] After the nucleation, the earliest species are the 2PP and 3PP protofibrils.[43] The transitions from 2PP to 3PP, from 2PP to 4PF, and from 3PP to 4PF occur at the lateral surface of a protofibril by collective interconversion of a file of previously deposited π-monomers.[43] This mechanism is more frequent for the model with low β-aggregation propensity, due to the frustration of the conformational landscape of the monomer. The number of protofilaments N_{la}^{pf} can exceed 4 only in reactive potentials ($dE > -1.5$), where multiple nucleations occur in the simulation box, and the nascent fibrils merge through the tips to a fibril with structural defects, i.e., with more than 4 protofilaments. On the other hand, aggregates consisting of one protofilament (1PP) are instable, and result from transient fluctuations of the micellar state. The minimum number of protofilaments for a (meta)-stable protofibril is two.[43]

As detailed in the Materials and Methods section, the polar order parameter (P_1) and the minimum moment of inertia (I_{min}) are the most appropriate variables to discriminate different fibril morphologies. The P_1 vs I_{min} 2D frequency histogram calculated on the trajectories with $dE = -2.25$ shows distinct peaks (Figure 3), which correspond to distinct morphologies. They are termed 4PF1, 4PF2, and 4PF3, according to the value of P_1. It is remarkable that the same pattern of peaks is found for the other potentials (see Figure S3, Supporting Information), suggesting that the morphologies are not a consequence of a particular choice of dE but rather originate from the intrinsic geometry of the monomer in the β-state. The 4PF1 and 4PF2 fibrils have two subpopulations of different thickness (I_{min}): 4PF1−, 4PF1+, 4PF2−, 4PF2+. A ribbon representation of the five fibril

---

(53) Qi, B.; Muff, S.; Caflisch, A.; Dinner, A. R. *J. Phys. Chem. B* **2010**, *114*, 6979–6989.

(54) Seeber, M.; Cecchini, M.; Rao, F.; Settanni, G.; Caflisch, A. *Bioinformatics* **2007**, *23*, 2625–2627.

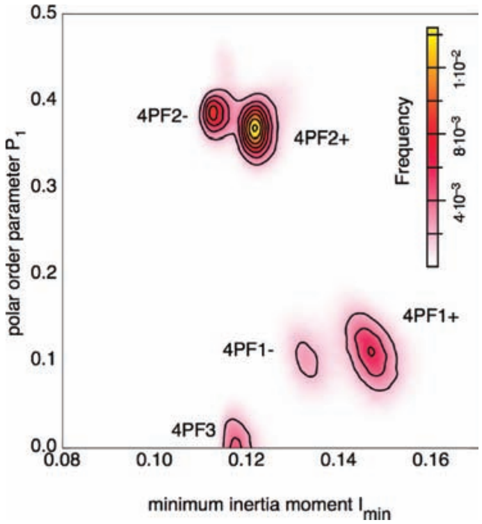

Figure 3. Two-dimensional frequency histogram of the polar order parameter P_1 and the minimum inertia moment I_{min} for $dE = -2.25$. Highest frequency is colored with yellow, and the contour lines represent the linear spacing of frequency values.

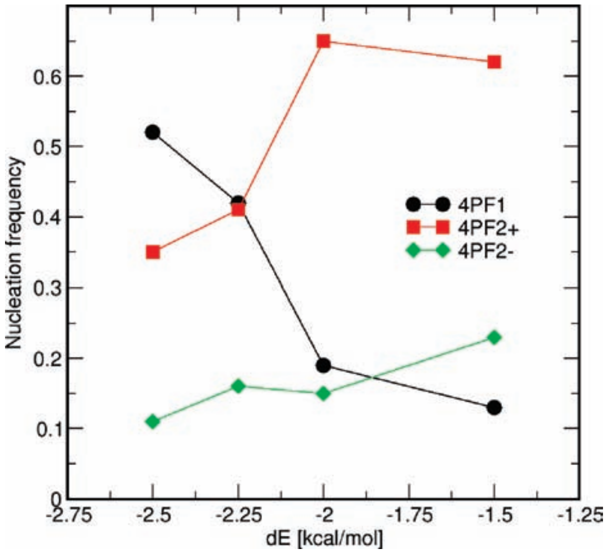

Figure 4. Relative nucleation frequency of three main morphologies. For 4PF1, 4PF2−, and 4PF2+ morphologies, the nucleation frequency, i.e. the relative number of events of formation of a particular morphology is calculated over a set of 100 independent simulations. The nucleation frequency is evaluated for different values of the amyloidogenic potential *dE*.

morphologies is shown in Figure 2. The orientation of the protofilaments, and the thickness of the fibril, are the main features that distinguish the morphologies. The fibrils of type 4PF2(+,−) have three protofilaments oriented up and one oriented down. Type 4PF1 and 4PF3 fibrils have two protofilaments oriented up and two down in alternate (4PF1) or adjacent (4PF3) arrangement. The 4PF1− and 4PF1+ architectures interconvert several times on a 10-μs time scale (see Figure S4, Supporting Information), and thus they can be considered two structural states of the same morphology, termed simply 4PF1. Conversely, although their structural properties are very similar, the 4PF2+ and 4PF2− morphologies do not interconvert within the time scale of the simulations (7.5−15 μs, see the red and green time series in Figure S4, Supporting Information).

Simulations of seeding, analogous to *in vitro* experiments,[7] were carried out to test the self-templating properties of individual morphologies. Small fragments of fibrils, consisting of about 50 monomers, were extracted from already formed fibrils of morphologies 4PF1, 4PF2+, and 4PF3. Note that the fibril fragments are larger than the average nucleus size, as the latter consists of about 40 monomers when *dE* is −2.5 (see Figure S11, Supporting Information). For each morphology, a simulation was carried out starting from four replicas of the same fragment randomly placed in a simulation box with a total number of 1000 monomers at a concentration of 8.5 mM. The initially dissociated monomers (about 800) aggregate to the fibril fragments. The morphologies are perfectly templated, and at the end of the 2-μs-long simulations each box contains four homogeneous fibrils (data not shown).

Thermodynamic Stability of Fibril Morphologies. The end of the fibrillation reaction corresponds to a dynamic equilibrium between the mature fibril and the residual dissociated monomers. The concentration of the residual monomers (the critical concentrations of fibril formation C_r), the specific potential energies Δe, and the free energy of monomer association Δg of different relevant morphologies are reported in Table 1 for potential $dE = -2.25$ and in Table S1 (Supporting Information) for all *dE* values. The rank of stability (Δg values) for the most common morphologies is 4PF2+ > 4PF2 > 4PF1. This rank correlates with Δe, indicating that the optimal packing and the electrostatic interactions are responsible for the thermodynamic stability of the morphology.

The Energy Difference *dE* Regulates the Relative Population of Morphologies. The percentage of nucleation events that produces a given morphology is reported in Figure 4 (4PF3 morphology is omitted in this and the forthcoming analyses due to its very low nucleation frequency). The nucleation frequency of a fibril morphology is calculated as the fraction of simulations that produced a specific morphology over the total number of simulations that nucleate. For the most reactive potential ($dE = -1.5$) the morphology that nucleates most frequently is the 4PF2+, followed by 4PF2− and 4PF1. This rank correlates with the corresponding fibril stability (Table S1, Supporting Information); thus, for highly amyloidogenic potentials, the selection of the final morphology is driven by the thermodynamic stability of the fibril. Decreasing the amyloidogenicity (i.e., decreasing *dE*), this correlation is lost. Remarkably, for the least amyloidogenic potential ($dE = -2.5$), the morphology that nucleates most is the least stable (4PF1). Hence, we hypothesize that when the amyloid aggregation is unfavored with respect to the folded state (β-unstable regime) the morphologies are selected kinetically, i.e., by the nucleation barriers. The following cFEP analysis was performed to substantiate this hypothesis.

The Free Energy Profiles. The cFEPs calculated for the runs resulting in fibril morphology of type 4PF1 are shown in Figure 5 (for the other morphologies see Figure S5, Supporting Information). Upon lowering *dE*, the downhill-shaped cFEP of high-amyloidogenic potential ($dE = -1.5$, -2.0) changes to a rough landscape, with presence of multiple barriers for the low-amyloidogenic potentials ($dE = -2.25$, -2.5). For $dE = -1.5$, -2.0, and -2.25 the free energy landscape is dominated by the final fibrillar state (red circles), with the presence of a 3PP minimum (blue circles) for $dE = -2.0$ and -2.25 potentials. The cFEP calculated for the potential $dE = -2.5$ clearly shows hierarchical assembly of the fibril. Here, the barriers separate the various intermediates with different number of protofilaments, as shown by the different colors reported in Figure 5, in accord with a previous network analysis (Figure 3 of ref 43).

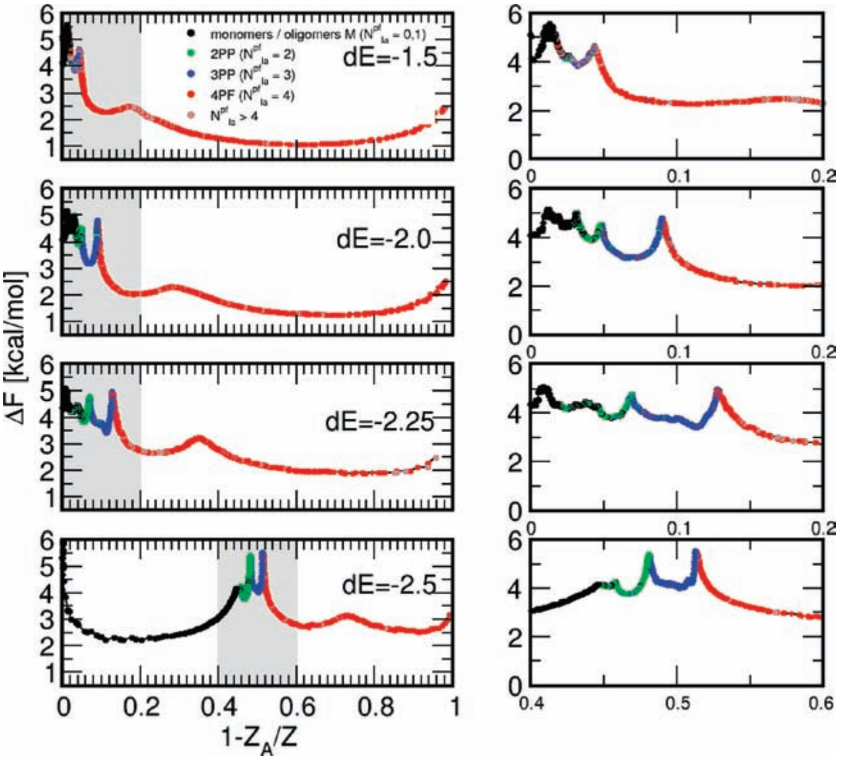

Figure 5. Cut-based free energy profiles of fibrillation. The mfpt-cFEPs calculated for the runs that produced 4PF1 morphology are shown here, while those for the 4PF2(+,−) morphologies are in Figure S5 (Supporting Information). Every state is represented as a circle, whose color reflects the number of protofilaments contained as shown in the legend in the top, left. The reference state is the most populated 4PF1 fingerprint. Note that the *x*-axis is one minus the relative partition function, i.e., it has the reference state on the right of the plot and not on the left as in the conventional representation of cFEP. Following the *x*-axis, which is proportional to the mftp to the final fibrillar state (reference state), the basin corresponding to the lag phase oligomers is kinetically most distant (black circles), followed by the 2PP (green circles), and the 3PP states (blue circles). The low barrier in the 4PF basin reflects the frequent interconversion between the morphologies 4PF1+ and 4PF1−. The plots on the right are magnifications of the area marked in gray.

Also, the different shapes of the cFEP for different *dE* values reflect the change of the nucleus size (see Figure S11, Supporting Information). When $dE = -2.5$ the nucleus size is much larger than for higher amyloidogenicity models. Consequently, the oligomers assembled at the lag-phase are more populated, which yields an enhancement of the width of the black basin on the left of the plot.

The Pathways of Fibril Formation Depend on the Morphology. Using the iterative procedure described in the Materials and Methods section, the free energy basins were isolated, and the fingerprint trajectories were transformed into time series of basin labels. This allows one to create simplified graphs that describe the process of fibrillation (see Figure 6 and Figure S6, Supporting Information). For $dE = -1.5$ and -2.0, irrespective of the final fibrillar morphology, the process follows a simple one-step reaction, where the initial state, mainly composed of a mix of monomeric and prefibrillar species, quickly transform to a final fibril state (Figure 6A and 6B). The same is observed for $dE = -2.25$, the only difference being the presence of the on-pathway intermediate 3PP (Figure 6C). In contrast, for $dE = -2.5$, a different behavior emerges for the 4PF1 and 4PF2(+,−) morphologies. While 4PF2(+,−) follow a three-stage reaction similar to that obtained with $dE = -2.25$ potential (Figure 6E), 4PF1 morphology has an additional intermediate, consisting of 2PP protofibrils (Figure 6D). Remarkably, this intermediate is off-pathway; i.e. from the micellar state, the reaction can follow two main pathways: either through the 3PP intermediate directly or through the 2PP and then 3PP intermediates. Therefore, intermediates are not the same for different fibril morphologies. As revealed by the simplified graphs, the 4PF1 fibrils have two intermediate protofibrils (2PP and 3PP) that exchange, while 4PF2(+,−) have only one (3PP). This implies that the 3PP species which are competent to form 4PF2 fibrils must be different from those competent to form 4PF1 morphologies. However, the two intermediates have the same order parameter, P_1, and similar average numbers of associated π-monomers and cannot be distinguished. Visual investigation of the two protofibrils revealed that the two 3PP species, 3PP1 and 3PP2, have two different arrangements of the protofilaments, which could not be distinguished by P_1 (see Figure S7, Supporting Information). Interestingly, the 3PP1 and 3PP2 protofibrils were observed to be competent to form 4PF1 and 4PF2(+,−) fibrils, respectively.

The Nucleation Barriers. The estimation of the free energy barrier between the oligomeric state and the protofibril intermediates is necessary to understand the origin of the population of the morphologies when the peptide has a low amyloidogenic potential. As can be seen from the simplified graph analysis (Figure 6), the model with $dE = -2.5$ fibrillates very frequently through one or more intermediates and seldom proceeds directly from the oligomeric stage to the final fibril. Thus, to accurately calculate the free energy barrier between the oligomer state and the intermediates using the cFEP approach, one has to remove from the trajectories the states not explored along the interested reaction, i.e., all the final fibril states. The cFEPs calculated from the most populated oligomeric state are reported in Figure 7 for 4PF1, 4PF2+, and 4PF2− morphologies. In these plots, the basin on the left of the first barrier contains all oligomeric states that are explored during the lag phase, whereas the protofibril intermediates are located on the right of the first

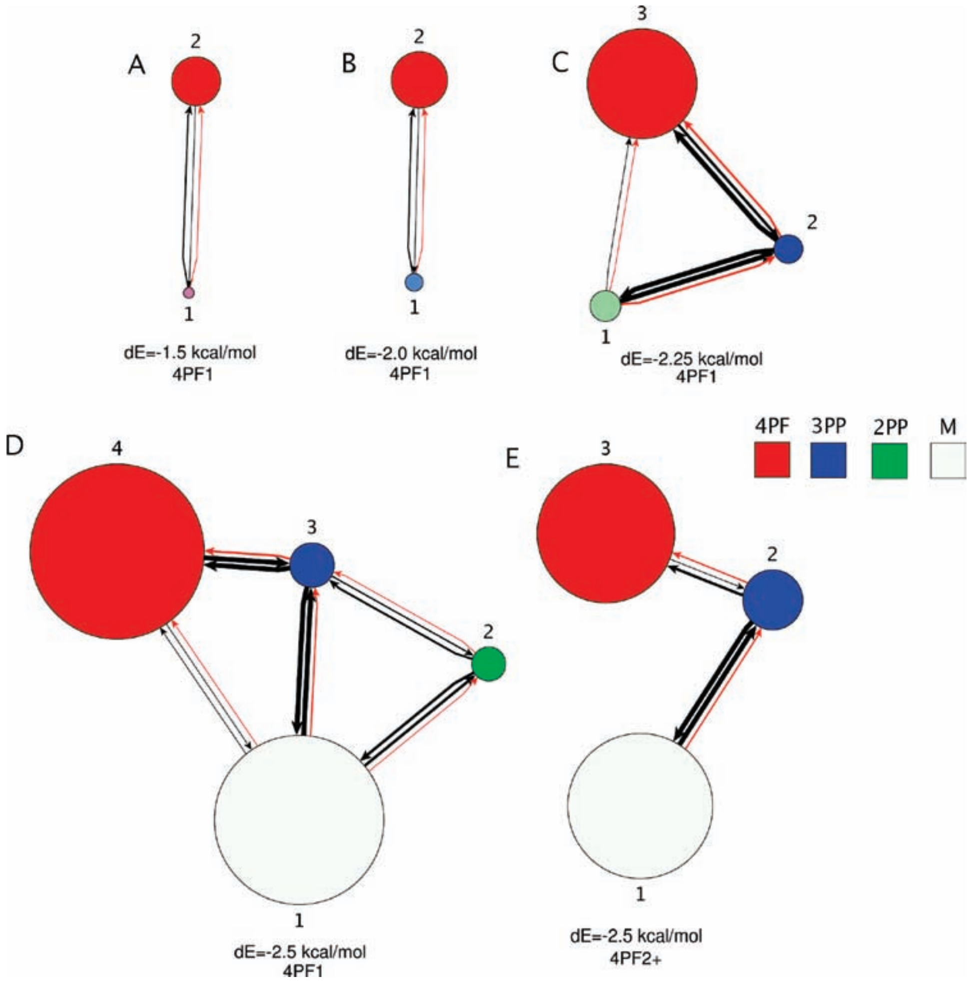

Figure 6. Simplified networks of interbasin transitions. Selected simplified graphs describing the process of fibrillation for different potentials and different fibril morphologies (see Figure S6 (Supporting Information)for complete sets of graphs). The size of the nodes is proportional to the statistical weight of the corresponding cFEP basin, while the thickness of the links is proportional to the number of transitions. The red link is the difference between the forward and backward black links. The color of the nodes is obtained by mixing the four colors indicated in the legend, which are weighted according to the population of the corresponding states.

barrier. Note that 4PF1 has two intermediates, 2PP and 3PP, as already observed in the simplified graphs. The first maximum can be interpreted as the nucleation barrier, i.e., the activation energy ΔF_{Nucl} needed by an oligomer to be promoted to a protofibril. The values of ΔF_{Nucl}, reported in Figure 7 and in Table 1, indicate that 4PF1 and 4PF2+ morphologies have roughly the same nucleation rate while formation of 4PF2− is slower, in agreement with the nucleation frequency analysis (see $dE = -2.5$ data points in Figure 4).

Long Time Behavior. Owing to the finite size of the nucleation simulations, the behavior of an infinite system with complex equilibrium between different morphologies cannot be directly derived. A polymerization master equation, based on the monomer-wise fibril elongation, has been developed to investigate the long time behavior of the system (see Supporting Information.). The master equation solution for the $dE = -2.5$ model predicts that after the nucleation, the morphologies coexist for a limited time, and, at the long time scale, the less stable morphologies disappear in favor of the most stable one (Figure S8, Supporting Information). This result is supported by a morphology coexistence simulation with the $dE = -2.5$ model (see Figure S9, Supporting Information), which starts from multiple replicas of fibrils with the three different morphologies in the same box. During this simulation the four 4PF1 fibrils shrink and disappear whereas those of 4PF2(+,−) elongate. However, the simulation shows that fibrils of 4PF2+ and 4PF2− morphologies are able to merge, suggesting that the equilibrium between the different morphologies cannot be merely approximated by the equilibrium of the fibrils with the dissociated monomers.

4. Conclusions

The fibril formation process has been investigated by multiple Langevin-dynamics simulations of an amphipathic polypeptide model whose simplicity allows for the sampling of hundreds of fibril formation events. The polypeptide model has only two states but displays great complexity in the assembly mechanism and heterogeneity of fibril morphologies. Different fibril architectures, distinguished by the arrangement of the internal protofilaments, spontaneously emerge from the initially monodispersed state and have distinct nucleation probabilities and energetic stability.

By varying a single parameter of the model, namely the relative stability of the amyloid-competent and amyloid-protected states of the polypeptide (i.e., the β-aggregation propensity dE), it is possible to shed light on the process of morphology differentiation and the morphology population

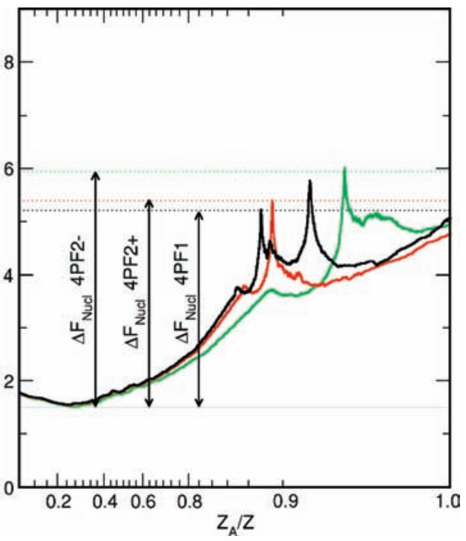

Figure 7. Nucleation barriers. The cFEPs for 4PF1, 4PF2+, and 4PF2− morphologies at $dE = -2.5$ are evaluated using the most populated oligomeric state M as reference state, and neglecting the 4PF states to better resolve the nucleation process. The cFEPs calculated using the runs that yielded 4PF2+ (red) and 4PF2− (green) have two minima, corresponding to the M state ($Z_A/Z < 0.9$) and the 3PP protofibril, while 4PF1 (black) has three minima, the M, the 3PP ($Z_A/Z \approx 0.9$), and the 2PP. A logit scale ($x' = \log(x/(1-x))$) is applied to the x-axis to magnify the region close to $Z_A/Z = 1$.

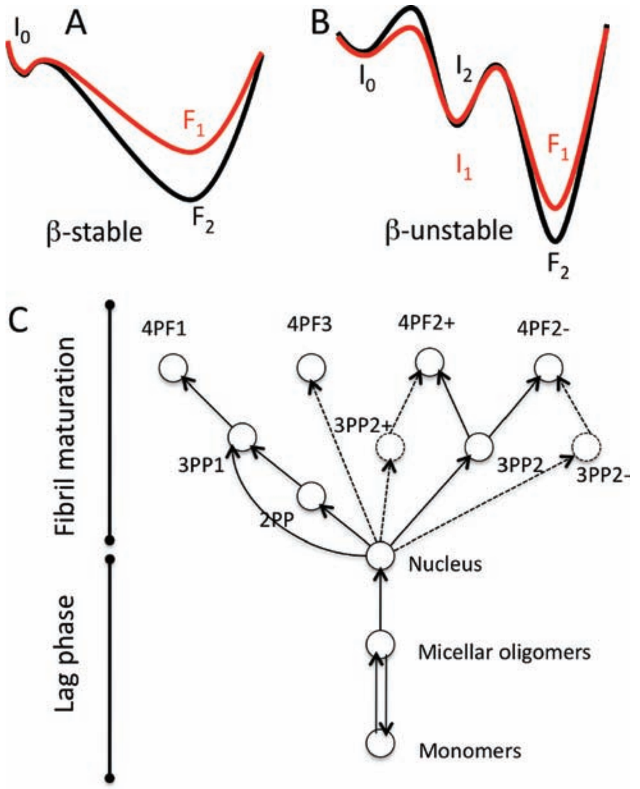

Figure 8. Pathways of morphology differentiation. (A) The process of fibril formation is downhill with a negligible nucleation barrier at high amyloidogenic conditions. The selection of the fibril morphology is thermodynamically driven, and the most stable fibril F_2 nucleates more frequently. (B) At low amyloidogenic conditions, the nucleation barrier is significantly higher. Since the process of fibril maturation is out-of-equilibrium and irreversible, the nucleation rates of the intermediates I_1 and I_2, which are competent to fibrils F_1 and F_2, respectively, determine the final morphology population. Starting from the state I_0, the intermediate I_1 is kinetically more accessible than I_2. Thus, the morphology F_1 will emerge more frequently, albeit the fibril F_2 is energetically more stable. (C) For the low amyloidogenic scenario ($dE = -2.5$), the process of morphology differentiation can be represented by a branched tree. During the lag phase, the micellar oligomers are in equilibrium with the dispersed monomers. The early morphology differentiation occurs at the nucleation step, where the formation of the protofibrillar intermediates is regulated by the structural bifurcation of the nucleus. The 2PP and 3PP1 intermediates are competent to 4PF1 fibrils, while the 3PP2 intermediate is competent to 4PF2(+,−) fibrils. Alternatively, the presence of 3PP2+ and 3PP2− intermediates that are directly competent to 4PF2+ and 4PF2− fibrils, respectively, can be hypothesized, although this was not observed. Also, the pathway of formation of 4PF3 fibrils was not investigated in detail, due to the small number of nucleation events of this morphology.

change upon varying external conditions. The simulation results show that the populations are sensibly and nontrivially influenced by the β-aggregation propensity, and two main mechanisms for fibril morphogenesis emerge. When the peptide is highly prone to aggregate ($dE = -1.5, -2.0$, Figure 8A), the morphogenesis is under thermodynamic control, meaning that the morphology with the highest stability will emerge with the highest probability. In contrast, when the peptide has a low amyloidogenic potential ($dE = -2.25, -2.5$, Figure 8B), the fibril morphogenesis is under kinetic control. The morphologies that nucleate more readily are not necessarily the most stable ones, but those whose precursors are kinetically more accessible, as revealed by the free energy profiles of the fibrillation. In the simulations with $dE = -2.5$ the fibril morphology that nucleates most frequently is the least stable one. Although this scenario is likely to depend on the details of the peptide model that was employed, the findings regarding the origin of morphologies should hold in general. An alternative scenario might arise when the kinetically accessible intermediate is a precursor of the most stable fibril, although it was not observed in the present simulations.

As remarked in previous works,[3,7,55] amyloid polymorphism can originate from (1) variation in the number of protofilaments,[10,56] (2) distinct modes of lateral association of protofilaments without significant variation in molecular structure,[10,57,58] or (3) variations of the polypeptide conformation at the protofilament level.[7,59] The present model, having a single fibril-competent structure, is limited to the first and second cases. The third case could be investigated using a more realistic coarse-grained or atomistic polypeptide model which accounts for the flexibility of the backbone. Nevertheless, the simplified model has the advantage of restricting the variety of fibril morphologies to a few very well-defined species, allowing the precise and rigorous calculation of the cFEPs. A recent systematic cryo-EM study of Aβ-fibril polymorphism[10] revealed that the competitive inter- and intrachain interaction patterns yield a continuous spectrum of coexisting fibril morphologies. In another work, microcrystal fibril structures were reported for a number of small peptides.[60] Subtle changes in the side-chain orientation resulted in different packing of the steric zipper in the crystal, and hence fibril polymorphism. These experimental findings imply that atomistic

---

(55) Fändrich, M.; Meinhardt, J.; Grigorieff, N. *Prion* **2009**, *3*, 89–93.

(56) Paravastu, A. K.; Leapman, R. D.; Yau, W.-M.; Tycko, R. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 18349–18354.

(57) Jimenez, J. L.; Nettleton, E. J.; Bouchard, M.; Robinson, C. V.; Dobson, C. M.; Saibil, H. R. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 9196–9201.

(58) White, H. E.; Hodgkinson, J. L.; Jahn, T. R.; Cohen-Krausz, S.; Gosal, W. S.; Müller, S.; Orlova, E. V.; Radford, S. E.; Saibil, H. R. *J. Mol. Biol.* **2009**, *389*, 48–57.

(59) Schmidt, M.; Sachse, C.; Richter, W.; Xu, C.; Fändrich, M.; Grigorieff, N. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 19813–19818.

simulations of polypeptide aggregation might be limited by sampling problems, as they could only partially explore the conformational landscape of the spontaneous morphology differentiation process. Moreover, the results discussed in this work can be generalized to flexible peptides since they are based on the activation energies for the formation of the fibril intermediates, regardless of the origin of the activation energy itself. In fact, the β- and π-conformations might be considered as ensembles of fibril-competent and -incompetent structures, respectively, and the dE parameter as the effective free energy difference between the two states. By direct comparison with experimental results, the change of the β-aggregation propensity of the polypeptide model can be interpreted as an experimental change of external conditions that modulates the amyloid competent conformation stability, such as the temperature, the pressure, the pH, and the cosolvents concentration. Alternatively, mutations of amyloidogenic polypeptide sequences can be emulated quantitatively by changes of dE.

This work provides a clear picture of the fibrillation pathway. Morphology differentiation is a hierarchical out-of-equilibrium process which can be illustrated by a branched tree (Figure 8C). The emergence of a specific fibril morphology is a stochastic event that is selected by the earliest protofibrils/oligomers. Indeed, the metastable intermediates compete for one or more mature fibril morphologies. Therefore, the population of a given morphology depends on the production rate of the earliest morphology-competent intermediate. Moreover, the diagram in Figure 8C is an oversimplification, as the nucleus and the micellar oligomers could be structurally heterogeneous, although this was not investigated in the present work. The multiple-pathways process observed here has a close similarity with the scenario described by Goldsbury et al.,[15] where two different morphologies of Aβ have distinct maturation pathways, either with or without the presence of metastable protofibrils. In the absence of on-pathway intermediates, e.g., for downhill fibrillation reactions, the selection process does not hold, and the model predicts that the most stable morphologies are nucleated more frequently. Once they are formed, fibrils are highly ordered structures that, owing to their cooperative character and the very high energy barriers, cannot directly interconvert between the different morphologies.[6,56] During their elongation, fibrils are persistently in equilibrium with dissociated monomers.[49,61] In a recent work, using an on-lattice simplified model, Zhang et al. found that the elongation phase of amyloid fibrils is regulated by Ostwald ripening phenomena, where bigger fibrils grow at the expense of smaller ones.[40] We do not directly see such effects, although our results are similar, as it is found that the most stable morphologies are growing at the expense of less stable ones. In fact, as predicted by the master equation and the morphology coexistence simulations presented here, the equilibrium with dissociated monomers dynamically influences the population of morphologies that have different dissociation constants. However, this work is not conclusive. The final fibril morphologies will be influenced also by other processes, such as secondary nucleation,[62] fragmentation,[63,64] fibril−fibril coalescence,[2] and conformational switching within the individual fibril.[65] In particular, fibril fragmentation accelerates the aggregation kinetics of the more brittle morphologies by increasing the number of growing surfaces. In these terms, fibril fragmentation can be considered as a form of kinetic control.

Amyloid fibril polymorphism is strongly connected to prion strain effects and aggregation cytotoxicity, and a putative structure−activity relationship has been hypothesized.[66] Specific strains emerge at selected environmental conditions[67] in a manner that was termed "survival-of-the-fittest",[68] and different amyloid conformations and morphologies of the prions result in distinct transmissible states,[69,70] infectivity changes,[4] and species-barrier phenomena.[14] Furthermore, the infectivity of the prions and the toxicity of neuropathological amyloids have presumably the same mechanism, which is the diffusion of amyloid material in the cellular tissue.[71,72] This is in accord with the result that the brittleness of fibril aggregates was shown to correlate with the prion strain infectivity[67] and amyloid cytoxicity.[73] In a recent work, Prusiner and co-workers generated a number of PrP amyloid fibrils with different stabilities in vitro by modulating the concentration of urea, the temperature, and the sequence.[74] Remarkably, when inoculating the resulting synthetic prion strains into model mice, they observed that the incubation time for the onset of the neurological symptoms, which is an inverse measure of infectivity, correlates with the structural stability of the amyloid fibrils. The present work can help to interpret these results. Change of the amyloidogenic potential, i.e., the environmental conditions, modulates the relative nucleation of stable fibrils (4PF2+) versus less stable ones (4PF1), and thus the infectivity, by selecting the opportune intermediate. What is the biological origin of the prion strains? We can speculate that there is a genetically encoded influence on the thermodynamic stability of the aggregation-prone state of the prion and/or the kinetic accessibility of the morphology-competent intermediates, which results in the emergence of the prion strain.

Amyloid fibrils display exceptional strength and stability,[75] and are highly resistant to degradation, which makes these

---

(60) Sawaya, M. R.; Sambashivan, S.; Nelson, R.; Ivanova, M. I.; Sievers, S. A.; Apostol, M. I.; Thompson, M. J.; Balbirnie, M.; Wiltzius, J. J. W.; McFarlane, H. T.; Madsen, A.; Riekel, C.; Eisenberg, D. *Nature* **2007**, *447*, 453–457.

(61) Carulla, N.; Caddy, G. L.; Hall, D. R.; Zurdo, J.; Gairi, M.; Feliz, M.; Giralt, E.; Robinson, C. V.; Dobson, C. M. *Nature* **2005**, *436*, 554–558.

(62) Ruschak, A. M.; Miranker, A. D. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 12341–12346.

(63) Xue, W.-F.; Homans, S. W.; Radford, S. E. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 8926–8931.

(64) Knowles, T. P. J.; Waudby, C. A.; Devlin, G. L.; Cohen, S. I. A.; Aguzzi, A.; Vendruscolo, M.; Terentjev, E. M.; Welland, M. E.; Dobson, C. M. *Science* **2009**, *326*, 1533–1537.

(65) Makarava, N.; Ostapchenko, V. G.; Savtchenko, R.; Baskakov, I. V. *J. Biol. Chem.* **2009**, *284*, 14386–14395.

(66) Maji, S. K.; Wang, L.; Greenwald, J.; Riek, R. *FEBS Lett.* **2009**, *583*, 2610–2617.

(67) Tanaka, M.; Collins, S. R.; Toyama, B. H.; Weissman, J. S. *Nature* **2006**, *442*, 585–589.

(68) Pedersen, J. S.; Otzen, D. E. *Protein Sci.* **2008**, *17*, 2–10.

(69) Tanaka, M.; Chien, P.; Naber, N.; Cooke, R.; Weissman, J. S. *Nature* **2004**, *428*, 323–328.

(70) Diaz-Avalos, R.; King, C.-Y.; Wall, J.; Simon, M.; Caspar, D. L. D. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 10165–10170.

(71) Li, J.-Y.; Englund, E.; Holton, J. L.; Soulet, D.; Hagell, P.; Lees, A. J.; Lashley, T.; Quinn, N. P.; Rehncrona, S.; Björklund, A.; Widner, H.; Revesz, T.; Lindvall, O.; Brundin, P. *Nat. Med.* **2008**, *14*, 501–503.

(72) Ren, P.-H.; Lauckner, J. E.; Kachirskaia, I.; Heuser, J. E.; Melki, R.; Kopito, R. R. *Nat. Cell Biol.* **2009**, *11*, 219–225.

(73) Xue, W.-F.; Hellewell, A. L.; Gosal, W. S.; Homans, S. W.; Hewitt, E. W.; Radford, S. E. *J. Biol. Chem.* **2009**, *284*, 34272–34282.

(74) Colby, D. W.; Giles, K.; Legname, G.; Wille, H.; Baskakov, I. V.; DeArmond, S. J.; Prusiner, S. B. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 20417–20422.

(75) Knowles, T. P.; Fitzpatrick, A. W.; Meehan, S.; Mott, H. R.; Vendruscolo, M.; Dobson, C. M.; Welland, M. E. *Science* **2007**, *318*, 1900–1903.

assemblies an attractive biomaterial.[76] The surface of amyloid fibrils can be engineered to have a specialized function, as for instance being employed as depots for chemical compounds,[77] or acting as electrical conducting nanowires.[78] However, to generate homogeneous material and reduce the amount of side products, the precise control of the fibril polymorphism is necessary. The present work might help explain this issue as it sheds light on the actual origin of the fibril morphology differentiation.

Acknowledgment. We thank Dr. S. Krivov for the help with the cFEP analysis, Dr. S. Muff for interesting discussions, and A. Magno for the nucleus size calculation. The simulations were performed on the Schrödinger cluster of the University of Zurich, and we gratefully acknowledge the support of C. Bolliger and A. Godknecht. This work was supported by a Swiss National Science Foundation Grant to A.C.

Supporting Information Available: Progress variables, aggregation state fingerprints, stability of morphologies, simple examples to illustrate basin-isolation procedure, structural characterization of fibril morphologies, interconversion between 4PF1+ and 4PF1−, cFEPs for 4PF2+ and 4PF2− morphologies, simplified networks of interbasin transitions, 3PP protofibrils polymorphism, chemical master equation of the fibril elongation, simulation of morphologies coexistence, oligomer size histogram, nucleus size, complete reference 47. This material is available free of charge via the Internet at http://pubs.acs.org.

JA106044U

---

(76) Cherny, I.; Gazit, E. *Angew. Chem., Int. Ed.* **2008**, *47*, 4062–4069.
(77) Maji, S. K.; Schubert, D.; Rivier, C.; Lee, S.; Rivier, J. E.; Riek, R. *PLoS Biol.* **2008**, *6*, e17.
(78) Scheibel, T.; Parthasarathy, R.; Sawicki, G.; Lin, X.-M.; Jaeger, H.; Lindquist, S. L. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 4527–4532.