

## SUPPORTING INFORMATION

### From A $\beta$ Filament to Fibril: Molecular Mechanism of Surface-Activated Secondary Nucleation from All-Atom MD Simulations

Nadine Schwierz<sup>a</sup>, Christina V. Frost<sup>b</sup>, Phillip L. Geissler<sup>c</sup> and Martin Zacharias<sup>b</sup>

<sup>a</sup>Department of Theoretical Biophysics, Max Planck Institute for Biophysics, 60438 Frankfurt am Main, Germany; <sup>b</sup>Physik Department, Technische Universität München, 85748 Garching, Germany; <sup>c</sup>Chemistry Department, University of California, Berkeley, California 94720, United States;

#### Equilibration of the PMFs

In order to test the convergence of the potentials of mean force (PMFs), the 100 ns simulation data from each umbrella window is divided into two blocks, respectively, with the first 10 ns in each window being discarded for equilibration. Separate PMFs are calculated for both data blocks and are compared to the PMF of the entire data set. The results are shown in Figure S1. Overall, the results are in good agreement and the errors derived from this splitting analysis amount to  $\pm 1 k_B T$  for monomer,  $\pm 0.2 k_B T$  for dimer,  $\pm 1.4 k_B T$  for trimer, and  $\pm 7.4 k_B T$  for tetramer association. All free energy results presented in the main text are based on the 100 ns total simulation time per umbrella window.

#### Radius of gyration and interface hydrogen bonds

To measure the compactness of the monomer, the radius of gyration  $R_G$  is calculated with respect to the center of mass of the molecule (Figure S2)A.  $R_G$  is calculated as an average over the last 50 ns in each umbrella window and error bars correspond to standard deviations. Monomeric A $\beta$  in bulk can be regarded as a disordered, highly flexible peptide without a well-defined three-dimensional structure, but with a preference for some conformations. The calculated bulk values of  $R_G$  for the monomer around 0.9-1.2 nm are in good agreement with the range of 0.9-1.2 nm obtained from various simulation studies using different force fields [1–3]. Compared to the monomer, the cross- $\beta$  structure of the oligomers is less compact.

Hydrogen bonds are calculated at the interface between the primary filament seed and the laterally associating monomer or oligomer. For the analysis, the standard hydrogen bond definition of Gromacs is used. It corresponds to a donor-acceptor distance  $< 0.35$  nm and donor-hydrogen-acceptor angle  $< 30^\circ$  [4]. Figure S2B shows the resulting average number of hydrogen bonds as a function of the seed-oligomer separation

coordinate  $\zeta$ . The low number of hydrogen bonds even in the bound state suggests that they are not a driving force of lateral fibril growth, in contrast to filament elongation which results in 20-25 interface hydrogen bonds on average.

#### INTERFACE CONTACTS

Interface contacts formed between the seed and the laterally associating monomer or oligomer are calculated as a function of the distance coordinate  $\zeta$ . They are further classified into native and non-native contacts according to the NMR fibril structure. Both classes are additionally decomposed based on the A $\beta$  secondary structure region involved in the contact, including N-terminal strand (NT), loop region (Loop) and C-terminal strand (CT). Contacts are defined to exist when the distance between any two atoms in the corresponding residues is  $< 0.5$  nm. Figure S3 shows the resulting maximum number of contacts per umbrella window as a function of  $\zeta$ . For both monomer and oligomers, non-native contacts start forming in the far intermediate region and are maximized in the binding site, accompanied by a smaller, but likewise increasing number of native contacts.

#### COMMITTOR ANALYSIS

Figure S4 shows the splitting probability for binding and unbinding as a function of  $\zeta$ . The behavior is as expected, with the probabilities switching from zero to unity within the boundaries of the bound and unbound state. The intersection with the solid black line indicates the maximum transition path probability  $\phi_{1/2}$ . At this separation, the probability of going to the bound state or the unbound state is equal. The next step for analyzing the quality of the reaction coordinate is to run unbiased trajectories which connect the bound and unbound state, enabling the calculation of from which the transition path probability can be calculated. The theoretically expected agreement between the resulting maximum transition path probability and the splitting probability would allow to reliably justify the quality of the reaction coordinate. However, based on our results, the time required to go for instance from the transition state to the fully unbound state amounts to several seconds, which is out of reach with the current computational resources.

- 
- [1] Raffa, D. F.; Rauk, A. *J. Chem. Phys.* **2007**, *111*, 3789–3799.
  - [2] Massi, F.; Peng, J. W.; Lee, J. P.; Straub, J. E. *Biophys. J.* **2001**, *80*, 31–44.
  - [3] Baumketner, A.; Bernstein, S. L.; Wyttenbach, T.; Bitan, G.; Teplow, D. B.; Bowers, M. T.; Shea, J. E. *Prot. Sci.* **2006**, *15*, 420–428.
  - [4] Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. *J. Comput. Chem.* **2005**, *26*, 1701 – 1718.

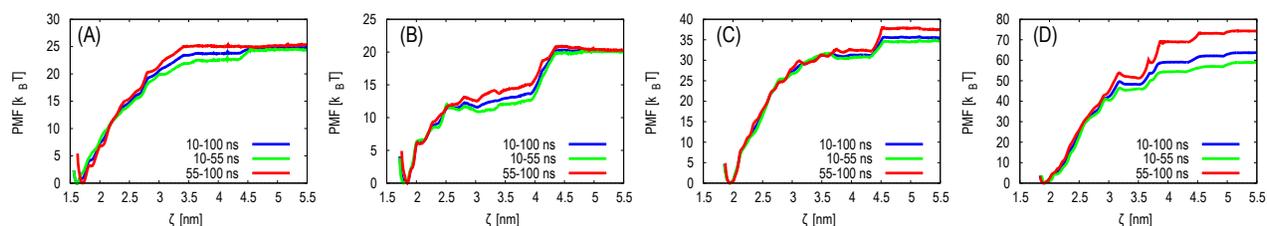


FIG. S1. Potentials of mean force in dependence of the simulation time for fibril growth by lateral (A) monomer, (B) dimer, (C) trimer, and (D) tetramer association. The simulation data in each umbrella window is divided into two blocks discarding the first 10 ns for equilibration: 10-55 ns (green) and 55-100 ns (red). The blue curve is the PMF resulting from the total simulation time of 100 ns.

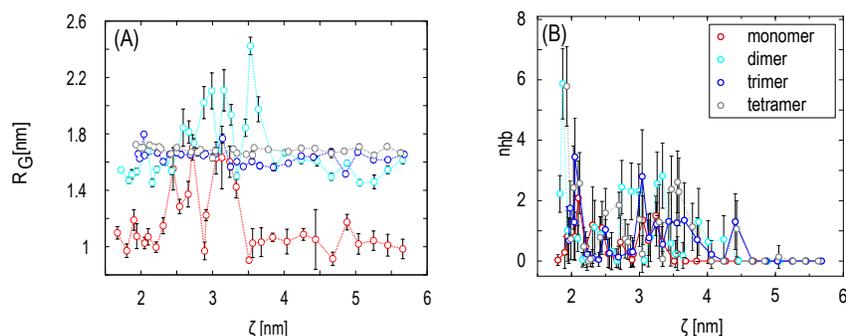


FIG. S2. (A) Radius of gyration for monomer, dimer, trimer, and tetramer as function of the seed-oligomer separation  $\zeta$ . (B) Average number of interface hydrogen bonds between the filament seed and the monomer, dimer, trimer, and tetramer as function of the seed-oligomer separation  $\zeta$ . Errorbars correspond to the standard deviation in the respective umbrella window.

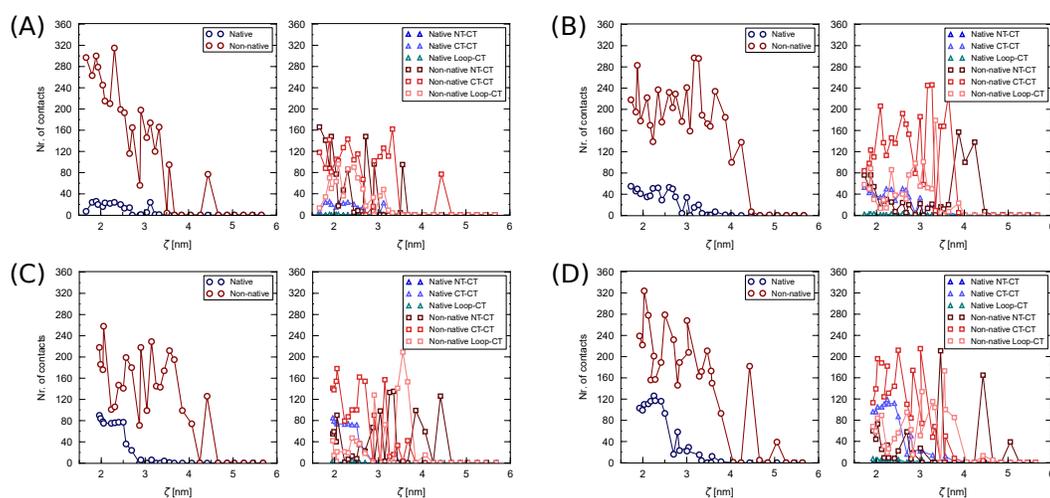


FIG. S3. Maximum number of interface contacts as a function of the distance coordinate  $\zeta$  for the laterally associating monomer (A), dimer (B), trimer (C) and tetramer (D). The pairs of plots contain a classification into native and non-native contacts (on the left) and a further decomposition according to the involved secondary structure region (on the right), with the lateral monomer/oligomer named first. 'NT-CT' hence means the contact exists between the NT-sheet of the oligomer and the CT-sheet of the seed.

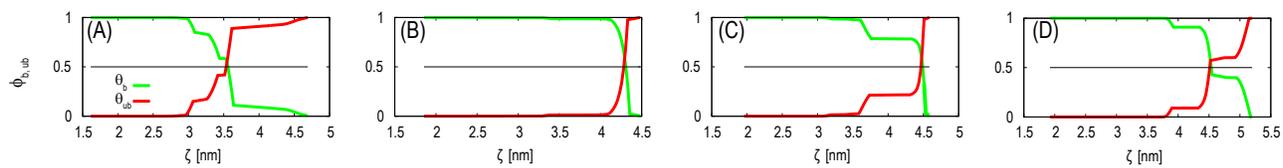


FIG. S4. Splitting probability to go from a given value of  $\zeta$  to the bound state (green) and probability to go to the unbound state (red). The intersection with solid line indicates the maximum transition path probability  $\phi_{1/2}$ .