Supplementary Information

I. Coarse-grained Simulations

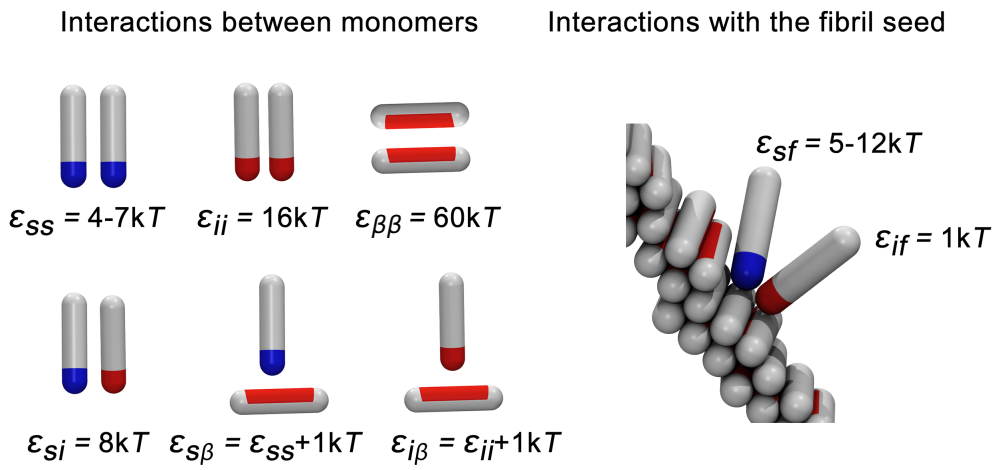

FIG. S1: Model: Possible interactions in the system, and their values.

A. Rate of spontaneous fibril formation and self-replication

The rates for spontaneous fibril formation through primary nucleation and self-replication are calculated from the respective average lag time for nucleation, $\langle t_{lag} \rangle$. The lag time is defined as the number of MC steps needed for the first nucleus consisting of at least two β-prone proteins to appear in the simulation. In our simulations the appearance of such a nucleus always leads to further fibril growth. In the case of primary nucleation, such an event takes place within an oligomer that was formed in the solution, while in the case of secondary nucleation, the event takes place in an oligomer which was formed at the surface of the preformed fibril. The average lag time is calculated from $4-6$ repetitions of the same system with different random seeds, and is expressed in the units of 10^8 Monte Carlo steps.

The lag time we measure is in fact the average over $N = 4-6$ independent realizations of the time needed for the β-nucleus to escape from the potential well, which is stochastic by its nature. Using the equivalence between this average first exit time $\langle t_{lag} \rangle$ and the inverse of the associated Kramers rate r [1], we have computed from first principles the rate of nucleation from the average time of formation of the first nucleus as:

$$r = \frac{1}{\langle t_{lag} \rangle}. \tag{S1}$$

B. Fraction of self-replication events

The fraction of the self-replication events in the system (Fig. 2) is calculated from the rates of the primary and secondary nucleation $\eta_{\mathrm{self-replication}} = \frac{r(\mathrm{secondary})}{r(\mathrm{secondary})+r(\mathrm{primary})}$.

C. Choice of the intermediate state

The rate of self-replication for Aβ was observed to be ~ 8 orders of magnitude faster than the rate of spontaneous formation [2, 3]. We have found in our simulations that, with two states only, we cannot achieve a self-replication

rate that is significantly faster than primary nucleation, under any condition (Fig. S2). This has taught us that the secondary nucleus had to be energetically different from the primary one. This observation is consistent with the idea that a modified energetic landscape is a necessary condition for the catalysis. Experiments have indeed reported that the replication reaction of Aβ peptides produces oligomers which differ from the fully developed β-sheet structures (Methods, [2]). It is likely that these oligomers contain proteins in the range of states between the soluble and the β-prone state. We have opted for the simplest possible case by introducing exactly one intermediate state. We have assigned the possibility of conversion into this intermediate state to the fibril-adsorbed protein, since the protein conformation in the adsorbed state is in general different from that in solution [4, 5], which has also been reported for Aβ peptides interacting with various surfaces [6, 7].

We have found that a significant increase in the rate of self-replication, compared to that of spontaneous formation, can be achieved if the intermediate conformation binds more strongly to its own kind than to the soluble species or the fibril. This leads to oligomer detachment, as experimentally observed. Since the self-interaction of β-prone proteins is stronger than the self-interaction of soluble species, it is reasonable that intermediate species will also have stronger self-interaction than the soluble species. Fig. S2 shows the rate of self-replication with and without the intermediate state (denoted with $\epsilon_{ii} = 0$), as well as for different values of the self-interaction, ϵ_{ii}, between species in the intermediate state. It is apparent that the more favourable the interaction between the intermediate state, the faster the self-replication becomes. This result is valid up to the point where this interaction compares to that between the β-prone proteins, however, we have not explored that limit. We have opted for $\epsilon_{ii} = 16\mathrm{k}T$ and have kept it constant in our simulations.

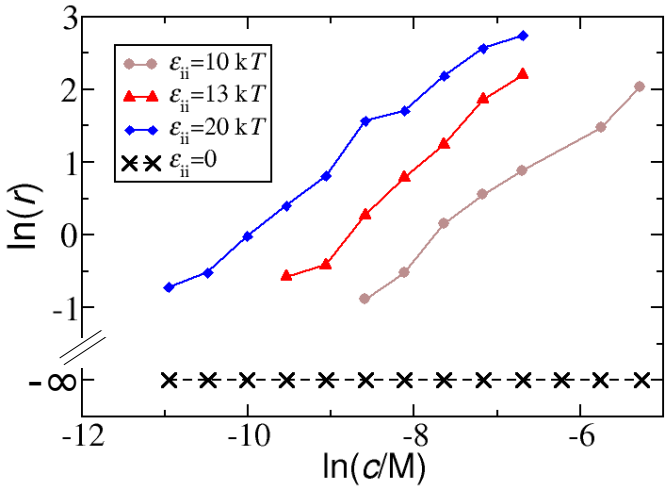

FIG. S2: The rate of self-replication for a two state system (dashed line, denoted $\epsilon_{ii} = 0$), which did not yield a single nucleation event within the simulation time, and for three different self-interaction values of the intermediate state, ϵ_{ii}. Data collected at $\epsilon_{ss} = 4\mathrm{k}T$ and $\epsilon_{sf} = 8\mathrm{k}T$.

D. Surface coverage

The surface coverage was quantified as the number of monomeric proteins whose attractive patches are in contact with the preformed fibril, normalized by the maximum number of such monomers. The maximum surface coverage was obtained from the fit of the surface coverage to the Langmuir isotherm, at $\epsilon_{sf} = 12\mathrm{k}T$ and $\epsilon_{ss} = 0\mathrm{k}T$.

E. Free energy for oligomer conversion and detachment

We employ the standard umbrella sampling technique [8] to obtain the free energy barrier $\Delta F_c(N)$ for conversion of a micelle of a size N, comprised of soluble proteins, into a micelle consisting of proteins in the intermediate conformation on the surface of the preformed fibril. The position of the center of the micelle is restrained to stay

within distance of 2σ form the fibril's surface, and a biasing harmonic potential is applied, which ensures that the number of monomers in the i-form, N_i, oscillates between exactly $N_i = 0$ and $N_i = N$. The free energy difference between the micelle of N proteins in the "i" state and N proteins in the "s" state is then calculated from the relative probabilities for each of these micelles to appear, corrected for the harmonic bias. This gives us the free energy for detachment of a micelle of the size N.

F. Free energy for oligomer formation on the fibril surface

In solution, the probability of forming an oligomer increases exponentially with the monomer concentration $P(N^*) \sim e^{\Delta f(N^*)/kT} c^{N^*}$, where $\Delta f(N^*)$ is a concentration-independent free-energy. On a finite surface we replace the monomer concentration with the monomer coverage of the surface. The probability of forming an oligomer onto the fibril surface is thus $P(N^*) \sim e^{\Delta f(N^*)/kT}(Kc/(1+Kc))^{N^*}$, where K is the monomer-surface binding constant ($K \sim \epsilon_{sf}$). The free energy for formation of such an oligomer would then scale as $\Delta F(N^*) \sim -N^* ln(Kc/(1+Kc))$. This however holds only at low concentrations, and should deviate at higher surface coverages, when oligomers start mutually interacting.

The free energy change for oligomerisation on the surface of the fibril is obtained from simulating the system of soluble monomers which are able to adsorb on the preformed fibril, but cannot change their conformation. As in all our simulations, the chemical potential of free (non-adsorbed) monomers in solution is kept constant. We collected the size distribution of oligomers which are in contact with the fibril $P(N)$. The free energy for formation of such an oligomer of the size N is then $\Delta F_o(N) = -logP(N) + F^0$, where the zero-energy level F^0 is attributed to free monomers in the solution.

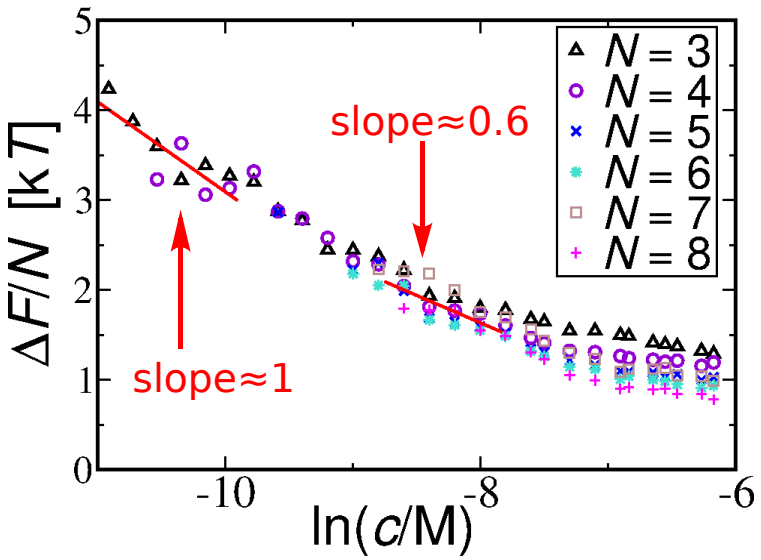

FIG. S3: The free energy per particle for formation of an oligomer of size N on the fibril surface versus the monomer concentration; $\epsilon_{ss} = 4\mathrm{k}T$ and $\epsilon_{sf} = 6\mathrm{k}T$ are kept constant. The lowest concentration range at which nucleation is observed in Fig. 4b ($-9 < \ln(c) < -8$) is marked by an arrow, where the slope is ≈ 0.6. The slope at vanishing monomer concentration approaches ≈ 1.

II. Derivation of integrated rate law

This derivation of the integrated rate equations describing the aggregation of monomeric protein into fibrils, via a surface catalysed secondary nucleation mechanism, closely follows that in Meisl et al. [3]. The difference being that the saturation of the secondary nucleation rate is now captured by $\left(\frac{c}{1+c/K_M}\right)^{n_2}$ rather than $\frac{c^{n_2}}{1+c^{n_2}/K_M}$, where c is

the total concentration of free monomer, and K_M is the inverse of the monomer-fibril equilibrium binding constant ($K_M = K^{-1}$). The choice of this slightly modified description was motivated by our finding that the nucleation process in our simulations proceeds by attachment to the surface, followed by a reactive encounter, therefore the rate is expected to be proportional to a power of the concentration of bound species, which is given in its simplest description by Langmuir as $\left(\frac{c}{1+c/K_M}\right)$. Numerically these two models lead to very similar results and the fits to experimental data are equally good, allowing no distinction. However, the detailed molecular insight from our simulations motivate us to use this new description to fit the data and enable a direct comparison of the simulation and experimental data.

The differential equations describing the aggregation are given by:

$$\frac{dP(t)}{dt} = k_2 M(t) \left(\frac{m(t)}{1+m(t)/K_M}\right)^{n_2} + k_n m(t)^{n_c} \tag{S2}$$

$$\frac{dM(t)}{dt} = 2m(t)k_+ P(t) \tag{S3}$$

where $m(t)$ is the free monomer concentration at a given time, $M(t)$ is the fibril mass concentration, $P(t)$ is the fibril number concentration, k_2, k_n and k_+ are the rate constants of secondary nucleation, primary nucleation and elongation respectively, and n_2 and n_c are the reaction orders of secondary nucleation and primary nucleation respectively.

We linearise these equations by setting $m(t) = c$ where c is the initial monomer concentration and solve them to yield $P_0(t)$ and $M_0(t)$. These correspond to the early time solution where monomer depletion is insignificant. The solution is equivalent to that for the previous form of secondary nucleation only with a change of constants $k_2 \frac{1}{1+c^{n_2}/K_M} \to k_2 \left(\frac{1}{1+c/K_M}\right)^{n_2}$ in the final solution.

Now a fixed point operator, obtained by integrating equation (S3), is applied to an initial guess for $P_i(t) = \frac{P_0(t)}{1+P_0(t)/P(\infty)}$ (see Cohen et al. [2]), yielding

$$M(t) \approx e^{-2k_+ \int_0^t P_i(\tau)d\tau} \int_0^t 2k_+ c e^{-2k_+ \int_0^\tau P_i(\bar{\tau})d\bar{\tau}} P_i(\tau)d\tau \tag{S4}$$

where $P_0(t)$ is the early time linearised solution to equation (S2) and $P(\infty)$ is the long time limit of the aggregate number.

The change in secondary nucleation mechanism affects only this initial guess, the fixed point operator remains unchanged.

$P(\infty)$ needs to be derived explicitly: First divide equation (S3) by $m(t)$ and use $\frac{dM(t)}{dt} = -\frac{dm(t)}{dt}$ to give

$$\frac{1}{m(t)}\frac{dm(t)}{dt} = -2k_+ P(t) \tag{S5}$$

Then differentiating and substituting equation (S2) yields

$$\begin{aligned}\frac{d}{dt}\left(\frac{1}{m(t)}\frac{dm(t)}{dt}\right) &= -2k_+k_n m(t)^{n_c} - 2k_+k_2 c\left(\frac{m(t)}{1+m(t)/K_M}\right)^{n_2} \\ &+ 2k_+k_2 m(t)\left(\frac{m(t)}{1+m(t)/K_M}\right)^{n_2}\end{aligned} \tag{S6}$$

We now try to find an expression for $\frac{1}{m(t)}\frac{dm(t)}{dt}$ and then we will use equation (S5) to obtain $P(t)$. We multiply both sides by $\frac{1}{m(t)}\frac{dm(t)}{dt}$ and use $\frac{df(m(t))}{dt} = \frac{df(m)}{dm}\frac{dm(t)}{dt}$:

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\left(\frac{1}{m(t)}\frac{dm(t)}{dt}\right)^2 &= \frac{d}{dt}\left(-\frac{2k_+k_n m(t)^{n_c}}{n_c} - 2k_+k_2 c\int \frac{m(t)^{n_2-1}}{(1+m(t)/K_M)^{n_2}}dm(t)\right. \\ &\left. + 2k_+k_2\int \frac{m(t)^{n_2}}{(1+m(t)/K_M)^{n_2}}dm(t)\right)\end{aligned} \tag{S7}$$

We perform the integrals with respect to m

$$\frac{d}{dt}\left(\frac{1}{m(t)}\frac{dm(t)}{dt}\right)^2 = -\frac{d}{dt}A(t) \tag{S8}$$

where $A(t)$ is given by

$$\begin{aligned} A(t) &= \frac{4k_+k_n m(t)^{n_c}}{n_c} + 4k_+k_2 c\frac{m(t)^{n_2}}{n_2}\left({}_2F_1\left[n_2, n_2, n_2+1, -\frac{m(t)}{K_M}\right]\right) \\ &+ 4k_+k_2\frac{m(t)^{n_2+1}}{n_2+1}\left({}_2F_1\left[n_2, n_2+1, n_2+2, -\frac{m(t)}{K_M}\right]\right) \end{aligned} \tag{S9}$$

where ${}_2F_1$ is the ordinary hypergeometric function.

Now substitute equation (S5) and integrate from 0 to τ with respect to t:

$$2k_+P(\tau) = \sqrt{A(0) - A(\tau)} \tag{S10}$$

P_∞ is then obtained by taking the long time limit $\tau \to \infty$ and using the fact that $\lim_{t\to\infty} m(t) = 0$

$$2k_+P(\infty) = \sqrt{A(0)} \tag{S11}$$

Following Meisl et al. [3] the full solution is then given by

$$\frac{M(t)}{M(\infty)} = 1 - e^{-k_\infty t}\left(\frac{B_- + C_+e^{\kappa t}}{B_+ + C_+e^{\kappa t}}\cdot\frac{B_+ + C_+}{B_- + C_+}\right)^{\frac{k_\infty}{\kappa\bar{k}_\infty}} \tag{S12}$$

where the definitions of the parameters are

$$\kappa = \sqrt{2ck_+k_2\left(\frac{c}{1+c/K_M}\right)^{n_2}} \tag{S13}$$

$$\lambda = \sqrt{2k_+k_n c^{n_c}} \tag{S14}$$

$$C_\pm = \pm\frac{\lambda^2}{2\kappa^2} \tag{S15}$$

$$k_\infty = 2k_+P(\infty) \tag{S16}$$

$$\bar{k}_\infty = \sqrt{k_\infty^2 - 2C_+C_-\kappa^2} \tag{S17}$$

$$B_\pm = \frac{k_\infty \pm \bar{k}_\infty}{2\kappa} \tag{S18}$$

c is the initial monomer concentration, and $P(\infty)$, $M(\infty)$ are the aggregate number and mass concentration at the start of the reaction and in the long time limit.

A. Determining secondary nucleation rate

The rate at which secondary nuclei are formed is given by

$$r_{\text{sec}} = k_2M(t)\left(\frac{m(t)}{1+m(t)/K_M}\right)^{n_2} \tag{S19}$$

In the context of our simulations the total mass of fibrils, M, is fixed, so we instead consider the rate at which nuclei are produced per mol of fibrils:

$$r_{\text{sec}} = k_2\left(\frac{m(t)}{1+m(t)/K_M}\right)^{n_2} \tag{S20}$$

We therefore need to determine 3 parameters in order to be able to compute r_{sec}: k_2, n_2 and K_M. Global fits of the integrated rate equation, Eq. S12, with 4 free global parameters, were performed using the AmyloFit interface [9], also available online [13]. The best fit is shown in Fig. S4 and yields the parameters $k_+k_n = 0.4\ \text{M}^{-2}\text{s}^{-2}$, $k_+k_2 = 9\cdot10^{10}\ \text{M}^{-n_2-1}\text{s}^{-2}$, $n_2 = 2.3$ and $K_M = 3.8\ \mu$M. The rate of elongation, k_+, was estimated separately from seeded experiments in Meisl et al.[3] as $k_+ = 3\cdot10^5\text{M}^-1\text{s}^-1$, yielding $k_2 = 3\cdot10^5\ \text{M}^{-n_2}\text{s}^{-1}$. Note that the values of the rate constants are approximate within at least an order of magnitude, the error on the reaction rates was estimated to be at least 20%.

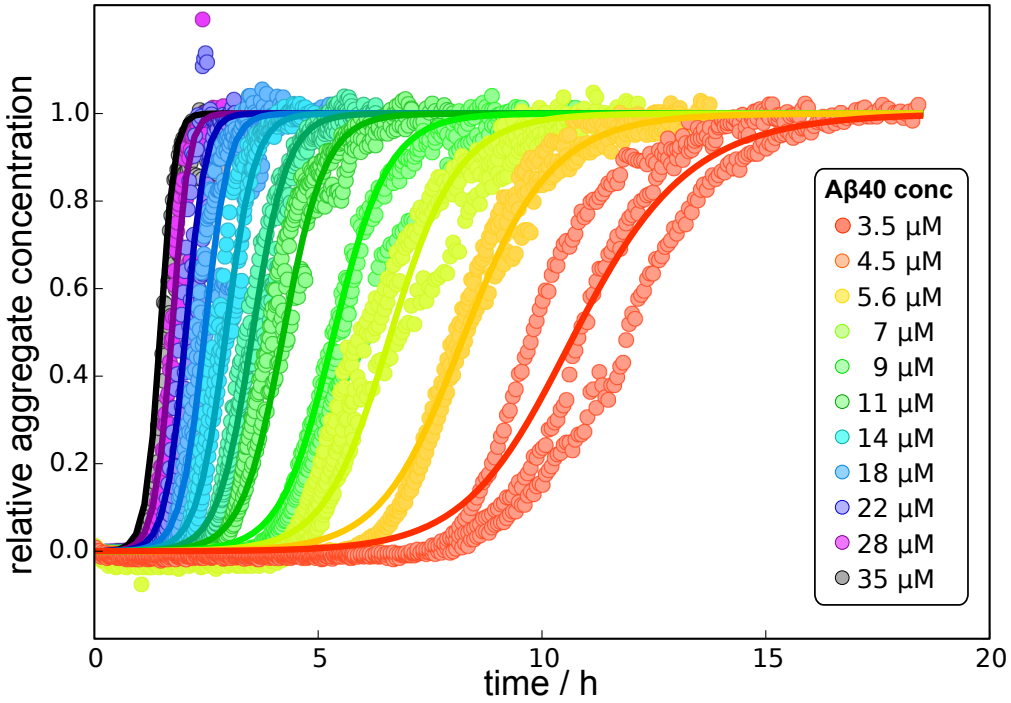

FIG. S4: Global fit to experimental data for Aβ40 aggregation from Ref. [3]. The global fit of Eq. S12 to the unseeded aggregation data with 4 free global parameters: k_+k_n, k_+k_2, n_2 and K_M.

III. Surface plasmon resonance measurement of Aβ40 peptide adsorption onto surfaces of its fibrils

When surface-bound fibrils are exposed to a solution of monomeric peptide, monomers simultaneously attach both to the fibrils' ends, which we refer to as the elongation, and to their surface, which we call simply adsorption. The two processes can be easily distinguished due to their very different kinetics and thermodynamics. The elongation of fibrils leads to a linear increase in surface-bound mass, while the rate of surface-adsorption is expected to decrease exponentially, as the available binding sites are being occupied. Reversely, upon washing of the saturated surface with buffer, the rate of fibril dissociation is expected to be linear and very slow due to the high thermodynamic stability of the β-sheet rich fibrils [10], while the surface-bound peptide molecules are expected to show an exponential detachment behaviour, at much higher rates due to their lower binding free energy. Therefore, short contact of the amyloid fibrils with monomer is likely to bias the binding behaviour towards surface attachment and against elongation. Aβ40 amyloid seed fibrils were firstly left to grow for about 30 min in monomer solution in order to obtain substantial coverage. Then we washed the surface extensively (ca. 1h) with buffer, followed by a series of short (30s) injections of monomeric Aβ 40 at concentrations between 4 and 44 μM. After each injection, the sensor surface was washed for at least 45 min with buffer. The kinetic traces of detachment showed the expected behaviour, i.e. a superposition of a linear and an exponential dissociation (Fig. S5). The amplitude of the exponential part was taken to correspond to the peptide that had been attached to the surface of the fibrils, and plotted against the monomer concentration to obtain the Langmuir absorption isotherm, with an equilibrium constant of $K^{-1} = 15\mu$M, as shown in Fig. 4C and Fig. S6. It is interesting to note here that the affinity of Aβ 40 monomers for fibril ends is 100 times higher than for the surface binding sites, corresponding to a difference in binding free energy of almost 5 kT.

Protocol The Aβ 40 peptide was expressed and purified as described previously [11]. The purified peptide was lyophilised and stored at -20°C. For the use in the SPR experiments, the peptide was dissolved in 10 mM NaOH at a concentration of 40 or 80 μM and 500 μl were injected into a Superdex 200 Increase 10/300 GL gel filtration column (GE Healthcare, Little Chalfont, UK) that had been equilibrated with 20 mM phosphate buffer at pH 7.4, with 0.2 mM EDTA and 0.02% (w/v) sodium azide added. The peptide was collected in fractions of 250 μl and only the central fractions of the monomer were used. The fractions were stored on ice for up to 24h until use. For the concentration determination, the absorption profile of the SEC chromatogram was used, with an absorption coefficient of 1200 at 280 nm. A solution of 21 μM of monomeric peptide was incubated for 24h at 37°C to form fibrils. For the attachment of the fibrils to the surface of the sensor, the fibrils were diluted 5 fold into 10 mM acetate buffer

at pH 4.0 [12] and sonicated for 30 s with a Sonopuls 2070 probe sonicator (Bandelin, Berlin, Germany) at 10% power and 30% pulses. The SPR experiments were performed with a Biacore3000 instrument (GE Healthcare), using C3 sensors. The carboxylic acid groups on the sensor surface were activated with a mixture of EDC and NHS to enable standard amide coupling chemistry. The injection of fibrils led to an increase of ca. 3000 RU. The subsequent incubation with monomer added ca. 9000 RU. For the data analysis, the linear part of the dissociation curves were fitted to a linear function which was extrapolated to the beginning of the dissociation. The difference amplitude was taken to correspond to the monomer binding to the fibril surface. The amplitudes were plotted as a function of monomer concentration, and fitted to $A(c) = \frac{A(\infty)Kc}{1+Kc}$ and then plotted as$A(c)/A(\infty) = \frac{Kc}{1+Kc}$, with K being the binding constant and c the monomer concentration.

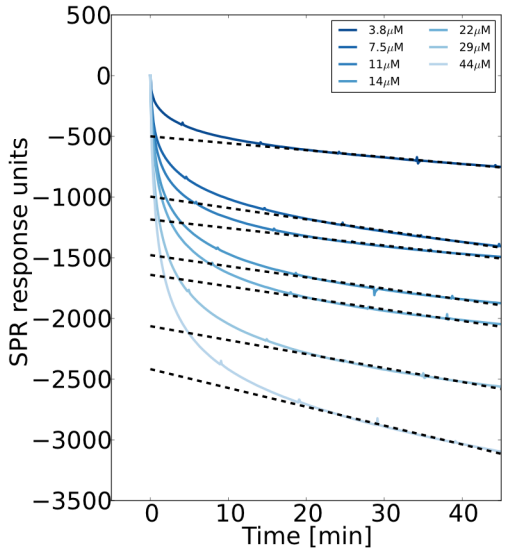

FIG. S5: The raw SPR dissociation data.

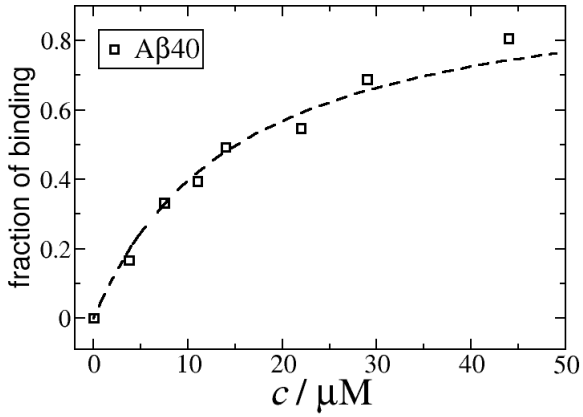

FIG. S6: Fraction of the peptides bound onto the surface of Aβ40 fibrils, at the same conditions as the kinetic experiments in Fig. 4A, versus the concentration of the soluble monomer. The dashed line is the fit to the Langmuir isotherm with $K^{-1} = 15\mu$M.

[1] P Reimann, GJ Schmid, and P Hänggi. Universal equivalence of mean first-passage time and kramers rate. *Physical Review E*, 60(1):R1, 1999.

[2] Samuel IA Cohen, Sara Linse, Leila M Luheshi, Erik Hellstrand, Duncan A White, Luke Rajah, Daniel E Otzen, Michele Vendruscolo, Christopher M Dobson, and Tuomas PJ Knowles. Proliferation of amyloid-β42 aggregates occurs through a secondary nucleation mechanism. *Proceedings of the National Academy of Sciences*, 110(24):9758–9763, 2013.

[3] Georg Meisl, Xiaoting Yang, Erik Hellstrand, Birgitta Frohm, Julius B Kirkegaard, Samuel IA Cohen, Christopher M Dobson, Sara Linse, and Tuomas PJ Knowles. Differences in nucleation behavior underlie the contrasting aggregation kinetics of the Aβ40 and Aβ42 peptides. *Proceedings of the National Academy of Sciences*, 111(26):9384–9389, 2014.

[4] Jeffrey J Gray. The interaction of proteins with solid surfaces. *Current opinion in structural biology*, 14(1):110–115, 2004.

[5] Paul Roach, David Farrar, and Carole C Perry. Interpretation of protein adsorption: surface-induced conformational changes. *Journal of the American Chemical Society*, 127(22):8168–8173, 2005.

[6] Carla E Giacomelli and Willem Norde. Conformational changes of the Amyloid β-peptide (1–40) adsorbed on solid surfaces. *Macromolecular bioscience*, 5(5):401–407, 2005.

[7] Christopher Aisenbrey, Tomasz Borowik, Roberth Byström, Marcus Bokvist, Fredrick Lindström, Hanna Misiak, Marc-Antoine Sani, and Gerhard Gröbner. How is protein aggregation in amyloidogenic diseases modulated by biological membranes? *European Biophysics Journal*, 37(3):247–255, 2008.

[8] Glenn M Torrie and John P Valleau. Nonphysical sampling distributions in monte carlo free-energy estimation: Umbrella sampling. *Journal of Computational Physics*, 23(2):187–199, 1977.

[9] Georg Meisl, Julius B Kirkegaard, Paolo Arosio, Thomas CT Michaels, Michele Vendruscolo, Christopher M Dobson, Sara Linse, and Tuomas PJ Knowles. Molecular mechanisms of protein aggregation from global fitting of kinetic models. *Nature protocols*, 11(2):252–272, 2016.

[10] Erik Hellstrand, Barry Boland, Dominic M Walsh, and Sara Linse. Amyloid β-protein aggregation produces highly reproducible kinetic data and occurs by a two-phase process. *ACS chemical neuroscience*, 1(1):13–18, 2009.

[11] Dominic M Walsh, Eva Thulin, Aedín M Minogue, Niklas Gustavsson, Eric Pang, David B Teplow, and Sara Linse. A facile method for expression and purification of the Alzheimer's disease-associated amyloid β-peptide. *Febs Journal*, 276(5):1266–1281, 2009.

[12] Kazuhiro Hasegawa, Kenjiro Ono, Masahito Yamada, and Hironobu Naiki. Kinetic modeling and determination of reaction constants of Alzheimer's β-amyloid fibril extension and dissociation using surface plasmon resonance. *Biochemistry*, 41(46):13489–13498, 2002.

[13] http://www.amylofit.ch.cam.ac.uk/