

Physical determinants of the self-replication of protein fibrils

Andela Šarić^{1,2,*}, Alexander K. Buell^{3†}, Georg Meisl^{1†}, Thomas C. T. Michaels¹, Christopher M. Dobson¹, Sara Linse⁴, Tuomas P. J. Knowles^{1,*} and Daan Frenkel¹

¹*Department of Chemistry, University of Cambridge
Cambridge, UK*

²*Department of Physics and Astronomy,
Institute for the Physics of Living Systems
University College London, London, UK*

³*Institute of Physical Biology, University of Duesseldorf
Duesseldorf, Germany*

⁴*Department of Biochemistry and Structural Biology, Lund University
Lund, Sweden*

† contributed equally*

The ability of biological molecules to replicate themselves, achieved with the aid of a complex cellular machinery, is the foundation of life. However, a range of aberrant processes involve the self-replication of pathological protein structures without any additional factors. One example is the autocatalytic generation of pathological protein aggregates, including amyloid fibrils, involved in neurodegenerative disorders. Here, we use computer simulations to identify the necessary requirements for the self-replication of fibrillar assemblies of proteins. We establish that a key physical determinant for this process is the affinity of proteins for the surfaces of fibrils. We find that self-replication can only take place in a very narrow regime of inter-protein interactions, implying a high level of sensitivity to system parameters and experimental conditions. We then compare our theoretical predictions with kinetic and biosensor measurements of fibrils formed from the A β peptide associated with Alzheimer's disease. Our results show a quantitative connection between the kinetics of self-replication and the surface coverage of fibrils by monomeric proteins. These findings reveal the fundamental physical requirements for the formation of supra-molecular structures able to replicate themselves, and shed light on mechanisms in play in the proliferation of protein aggregates in nature.

The molecular machinery of life is largely generated through the assembly of proteins into functional complexes. A particularly common form of protein self-assembly is that leading to linear filaments. These structures are widely used in nature, for instance as the basis of the cytoskeleton. Once formed, the vast majority of functional protein assemblies typically fulfil their biological function but do not directly catalyse the formation of further "daughter" complexes. However, certain protein structures possess the intriguing ability to promote their own replication. This phenomenon first came to prominence in the context of prions, where specific supra-molecular protein assemblies were observed to be able to multiply effectively once taken up into a variety of organisms, ranging from humans to yeast [1–3]. Such propensity to self-replicate has emerged as a more general feature of pathological protein self-assembly, observed in the context of sickle cell anemia [4, 5] as well as for amyloid fibrils implicated in medical disorders [6–8] such as Alzheimer's disease (A β peptide) [9, 10], type II diabetes (islet amyloid peptide, IAPP) [11–13], and Parkinson's disease (α -synuclein) [14, 15]. Strikingly, all of these structures are able to catalyse their replication under certain conditions. The initial fibrils are produced spontaneously from solution through primary

nucleation, followed by proliferation via heterogeneous, fibril-dependent, secondary nucleation [12]. In this type of self-replication the information about the protein conformation is transferred to the replicas, although they are not necessarily exactly identical to the parent aggregates. Spontaneous fibril formation is inherently slow, while fibril self-replication is usually many orders of magnitude faster [10]; yet a detailed microscopic understanding of either processes is currently lacking. Autocatalytic replication intrinsically introduces positive feedback into the self-assembly process that renders it challenging to control once assembly has started. As such, most functional protein complexes and fibrils do not have self-replicating properties. This finding therefore motivates a question about the fundamental ingredients necessary for fibril self-replication to occur, or indeed to be avoided.

Here, we develop a minimal computer model that is able to capture both spontaneous fibril formation in solution, and fibril-self replication. We study the necessary conditions required for self-replication to dominate over spontaneous formation, and find that strong bounds on inter-protein interactions exist for efficient self-replication that result in the high sensitivity of self-replication to environmental conditions. Indeed, it has been reported experimentally that the existence of secondary nucleation in α -synuclein, insulin, and the A β peptide strongly depends on pH [14, 16, 17], while secondary nucleation in A β also varies dramatically with salt concentration [18]. The emergence of a narrow

*a.saric@ucl.ac.uk, tpjk2@cam.ac.uk

regime that supports self-replication sheds light on why it is a relatively rare property of protein self-assembly *in vivo*, and possibly provides a physical criterion to distinguish functional from pathological assembly. Moreover, these results suggest that even pathological self-assembly, in principle, can be suppressed by moderate changes to the system to move it away from the narrow parameter space supporting efficient self-replication. Our results further infer that the secondary nucleus has to be energetically different from the primary one, pointing to two distinctive pathways.

Taking the aggregation of the Alzheimer’s $A\beta$ peptide into amyloid fibrils as a model for experimental comparison, in combination with kinetic and biosensing experiments, we show that the major characteristics of secondary nucleation can be explained in terms of the adsorption of monomeric peptides onto the surface of fibrils, and the level of surface coverage. We then demonstrate, in simulations and in experiments, that self-replication can be modulated by controlling the fibril surface coverage. Through the powerful combination of coarse-grained simulations and physical measurements, our results offer microscopic insights into the mechanism of the autocatalytic replication of protein fibrils.

Computer model

As the basis for our model we take the aggregation of peptides and proteins into amyloid fibrils, which have a common structure rich in β -sheet content. A minimal model that reproduces homogeneous fibril nucleation allows an amyloidogenic protein to exist in two states: a soluble state (denoted “ s ”) that can form finite oligomers, and a higher free-energy state that can form the β -sheet enriched fibrils (denoted “ β ”) [19, 20]. Simply considering the interaction of soluble proteins with the surface of existing fibrils captures the binding of monomers to the fibrils, but does not lower the free energy barrier for nucleation, thus does not result in efficient catalysis. To achieve a self-replication rate that is significantly faster than spontaneous formation, the structure and energy of the species involved necessarily have to differ from those observed in the absence of fibrils (Supplementary Section SI.C). The self-replication cycle in the $A\beta$ system has been shown to generate predominately small prefibrillar oligomers, whose structures differ from that of the mature fibrils (Methods, [10, 21]). Although an ensemble of such intermediate structures could exist in reality, here we consider the simplest possible case: we include one additional, intermediate (“ i ”), conformation, which can form on the fibril surface. This conformation is in-between the soluble and the β -state, and its self-interaction is stronger than its interaction with the fibril, which leads to detachment of oligomers from the parent fibril, as observed in experiments.

Amyloidogenic protein molecules in our model are represented as hard spherocylinders with attractive patches

(Fig. 1). The attractive interactions account for generic features of inter-protein interactions, such as hydrophobic interactions, hydrogen bonding, and screened electrostatic interactions. The soluble state of the protein is modelled as a spherocylinder with an attractive tip (Fig. 1a), whose self-attraction is given by the parameter ϵ_{ss} . Such particles are able to generate finite oligomers (Fig. 1b) [20]. The attractive tip can also adsorb onto the outer surface of the fibril, with interaction strength ϵ_{sf} (Supplementary Fig. S1). The intermediate conformation i is modelled with the same potential as the soluble state, but possesses a stronger self-association parameter ϵ_{ii} and a vanishing adsorption onto the fibril (Supplementary Fig. S1). The fibril forming, β -sheet prone, configuration is a hard spherocylinder with an attractive side-patch (Fig. 1a). The β -prone proteins pack parallel to one another with the maximal interaction strength $\epsilon_{\beta\beta}$, leading to fibril-like aggregates (Fig. 1b). We performed dynamic Monte Carlo (MC) simulations, allowing for the interconversion between the three protein conformations with a small probability at every MC step. The $s \rightarrow i \rightarrow \beta$ conversion is thermodynamically unfavourable, reflecting the loss of the internal conformational entropy of the protein molecule [22]. Throughout the text k denotes the Boltzmann’s constant and T is the temperature; further details are given in the Methods Section.

Spontaneous formation versus self-replication

The first question we address involves the identification of those conditions that lead to secondary nucleation being dramatically dominant over spontaneous, primary, nucleation. We have performed a series of computer experiments, in which a capped preformed fibril (incapable of further growth) was inserted into a solution of monomeric proteins, and nucleation processes were monitored. Primary nucleation takes place in two steps, whereby protein oligomers first form in solution, and then convert into β -sheet nuclei, which continue growing by monomer addition (Fig. 1c) [20, 23]. In the secondary nucleation process, proteins first adsorb onto the surface of the fibril, forming local clusters that keep growing and shrinking while still being attached to the fibril surface, as depicted in Fig. 1d. Once the oligomer of a critical size is formed, the proteins within change their conformation into the intermediate form. The oligomer then detaches into the solution, converts into the β -sheet protofibril, and grows further by monomer addition (Fig. 1d).

To investigate possible scenarios for different aggregating proteins, under various solution conditions, we measured the rates of primary and secondary nucleation at different protein concentrations and inter-protein interactions. From these measurements we calculated the fraction of self-replication events in the system for a given set of external conditions (Supplementary Sections SI.A and SI.B), Fig. 2a. Clearly, self-replication dominates over

spontaneous fibril formation at low protein concentrations and low inter-protein interactions. Indeed, proteins are typically below their critical micelle concentration at physiological conditions, which corresponds to the regime of low inter-protein interactions and low protein concentrations, where self-replication can dominate.

The reason for the dramatic dominance of self-replication in this regime is two-fold. The first contribution arises from the aided collocation of proteins on the one-dimensional surface of the fibril. This contribution is particularly important at low protein concentrations, where the probability of proteins meeting in solution and forming oligomers is very low. The second contribution lies in the decreased barrier for the secondary nuclei formation on the fibril surface, via the intermediate state (Supplementary Section SI.C). Essentially, for self-replication to dominate, the secondary nucleus has to be different from the primary one.

Strong environmental bounds for self-replication

Modulating environmental conditions and introducing protein mutations not only changes the properties of proteins interacting in solution, but also the strength of the adsorption of proteins onto the surface of fibrils, given by ϵ_{sf} in our simulations. We find that changing the protein-fibril affinity only by a few kT , the fraction of self-replication events changes non-monotonically, exhibiting a distinct region of optimal self-replication, Fig. 2b. This result is in agreement with the high sensitivity of fibril self-replication to solution composition, and can explain why it is to date observed only in a limited number of systems. Comparably, in a recent simulation, secondary nucleation of Lennard-Jones particles at a crystalline surface, when exposed to mechanical agitation, was reported to take place only in the regime of intermediate supersaturation [24].

Fig. 3a. analyses this effect in depth, at constant protein concentration. At low protein-fibril interaction strengths, proteins cover only a small fraction of the fibril surface, and the protein adsorption and oligomer formation on the fibril surface determine the reaction rate. Fig. 3b depicts the Langmuir-type isotherm for the fibril surface coverage, θ , as a function of ϵ_{sf} (Supplementary Section SI.D), indicating that the increase in the surface coverage follows the increase in the rate of self-replication in Fig. 3a. At high ϵ_{sf} , the fibril is substantially covered by proteins, however, the oligomer detachment becomes unfavourable. Nucleation will happen only after the oligomer has reached a certain size, N^* , when the energy gain due to the stronger inter-protein interactions after the conformational change overcomes the loss in the protein-fibril adsorption energy. Stronger binding to the surface hence requires larger oligomers in order to overcome the loss in the favourable adsorption energy. For very large oligomers, due to the geometric constraints, this requirement cannot be satisfied. Therefore, the con-

formational change will become unfavourable as the binding to the surface increases further (inset in Fig. 3b, Supplementary Section SI.E). In reality, in the regime of high adsorption, proteins are likely to distribute themselves evenly on fibrils in order to increase their contact area with the surface, and could form multiple layers, additionally hampering secondary nucleation. The narrow region of inter-protein interactions supporting self-replication is therefore the outcome of the balance between sufficient fibril coverage, and unhindered conformational change.

Simulated and experimental kinetics of self-replication

Our model can make a range of predictions that can be tested experimentally. Here, we seek to relate our simulations to kinetic measurements of self-replication of A β 40 amyloid fibrils, one of the two major isoforms of the A β peptide associated with Alzheimer's disease. Kinetic experiments provide the dependence of the reaction rate on monomer concentration, $r \sim c^{\gamma_s}$, where the scaling exponent γ_s is in our case the reaction order of self-replication. It reflects the monomer dependence of the dominant aggregation processes, therefore carrying information about the reaction mechanism.

Fig. 4a depicts double logarithmic plot of the rate of secondary nucleation for the A β 40 system, versus the initial monomer concentration, where the slope corresponds to the scaling exponent. The scaling exponent is highly dependent on the concentration of the monomeric peptide in solution, suggesting a possible change in the nucleation mechanism over the concentration range [25]. Fig. 4b shows the same quantities, collected in simulations, at a moderate protein-fibril affinity. The reaction order varies with the protein concentration, with a high value at low monomer concentrations ($\gamma_s \approx 3.3$), and a low value at high monomer concentrations ($\gamma_s \approx 0.5$), as observed in the A β 40 experimental data.

Due to our microscopic modelling we are able to pinpoint the processes underlying the switch in kinetic behaviour. Fig. 4d shows that the change in the reaction order follows the trend in the change of fibril coverage. Hence, the non-linear increase in surface coverage, due to surface saturation, is connected to the continuous decrease in reaction order. We find from our data that the rate of self-replication follows the surface saturation as $\ln(r) \sim N^* \ln(Kc/(1 + Kc))$, where K is the monomer-surface binding constant ($K \sim \epsilon_{sf}$) and N^* is a constant (see Methods and Supplementary Section IF for details). The reaction order, γ_s , then continuously changes between N^* , at infinite dilution, and 0 at full saturation. Interestingly, in models of nucleation where the aggregation number is the slow degree of freedom, N^* would be simply equal to the size of the nucleating oligomer, which is found to be constant over the concentration range in our simulations (inset in Fig. 4b)

Experimental verification of surface saturation

To test experimentally the prediction that the change in the apparent reaction order is governed by the change in the surface coverage, and not by a change in the nucleation mechanism, we designed a series of surface plasmon resonance (SPR) biosensing experiments that allow direct measurement of the binding of monomeric peptide molecules to the surface of amyloid fibrils, under the same conditions as the kinetic experiments. This enabled us to obtain the Langmuir absorption isotherm of A β 40 peptides onto its fibrils (Fig. 4c and Supplementary Fig. S6). Indeed, surface saturation takes place in the micromolar regime (with an equilibrium binding constant of $K^{-1} = 15\mu\text{M}$), which is exactly the regime where the change in the apparent reaction order takes place in aggregation experiments (Fig. 4a). Furthermore, this value of K is of the same order of magnitude as the value obtained from the kinetic fit to the experimental aggregation data (Methods and Supplementary Section SIII), and therefore strongly supports the hypothesis that the change in the scaling exponent is due to surface saturation.

Surface saturation controls the apparent reaction order

Finally, we show that by controlling the surface coverage by varying the strength of the inter-protein interactions, at constant monomer concentration, one can further modulate the kinetics of fibril self-replication. At constant protein concentration, the surface coverage is determined by the magnitude of protein-fibril affinity and inter-protein interactions. It is likely that both of these interaction strengths will be affected when altering experimental conditions, due to their similar physical origins. We observe that the surface coverage increases when both of these interactions are strengthened in simulations, resulting in a weaker dependence of self-replication on monomer concentration. The average scaling exponent γ_s from the simulations, as a function of ϵ_{ss} and ϵ_{sf} , is shown in Fig. 5a and Fig. 5b. We compare this behaviour to the aggregation of A β 42 at a range of salt concentrations [18], Fig. 5c. It is important to note that, in the context of our physical model, the two isoforms of A β , A β 40 and A β 42, share mechanistic similarities. An increase in ionic strength shields the electrostatic interactions and leads to an increased attraction between the negatively charged A β 42 monomers and fibrils, as well as between the monomers. Hence variation of the ionic strength offers an experimental way to vary in a controlled way the value of ϵ_{ss} and ϵ_{sf} . Indeed, the trend in the behaviour of the scaling exponents for the aggregation of A β 42 with increasing salt concentration agrees well with that found in our simulations. Therefore the large effect of ionic strength on the aggregation behaviour is in agreement with a variation of the adsorp-

tion of proteins onto their fibrils, offering a direct way to influence the self-replication process in a controlled manner.

Discussion

By developing a minimal model of protein self-replication, we have identified its dominant physical determinant to be the adsorption of monomeric proteins onto the surface of protein fibrils. Strong limits on inter-protein interactions are found for efficient self-replication, originating from the fact that changes in the interaction strength have opposing effects on the two key steps in the nucleation mechanism: oligomer formation and oligomer detachment. A narrow region of "ideal" interaction values supporting self-replication (Fig. 2b) results in its high specificity and sensitivity to environmental conditions.

An additional conformational change taking place on the fibril surface is a minimal requirement for the catalysis and detachment of oligomers from the parent fibril, which, in the context of many amyloid diseases, is a crucial step in the proliferation of pathological species [26–28]. The conformational change is at the origin of the formation of amyloid fibrils; the aggregating protein necessarily undergoes a change from its soluble form into the characteristic β -hairpin conformation. Models which attempt to achieve self-replication in (nearly) minimal colloidal systems, require an external dynamical change to permit detachment of the replicas from the parents [29, 30]. Amyloidogenic proteins naturally possess this dynamic characteristic.

A direct practical conclusion from our analysis is the ability to relate the reaction order measured in experiments to the underlying microscopic mechanism. We have found that the changes in the reaction order can be related to the change in the fibril surface coverage by the protein molecules, which we have confirmed by directly measuring the binding isotherm of monomers to the fibril surface. The characteristic concentration-dependence of the reaction order, observed in experiments, is consistent with a scheme where the rate-limiting step takes place on the fibril surface, further confirming that primary and secondary nucleation are indeed distinct processes. Whether the change in the apparent reaction order is experimentally measurable will depend on the concentration range that can be explored, as experiments might be limited to a concentration range where it appears locally constant. By measuring the fibril coverage and the apparent kinetic reaction order separately, the information about the critical size of oligomers produced via secondary nucleation may become accessible, for any protein system which exhibits this behaviour.

As a proof of principle, we have shown that by varying in a controlled manner the fibril surface coverage, by modulating the inter-protein interactions with ionic strength, one can control the kinetics of fibril self-replication. Hence, modulating the adsorption of

monomeric proteins onto the surface of protein fibrils may represent a fruitful strategy for limiting the proliferation of protein aggregates in a disease context.

Acknowledgments

We acknowledge support from the Human Frontier Science Program and Emmanuel College (A.Š.), the Leverhulme Trust and Magdalene College (A.K.B.), St. John's College (T.C.T.M.), the Biotechnology and Biological Sciences Research Council (T.P.J.K. and C. M. D.), the Frances and Augustus Newman Foundation (T.P.J.K.), the European Research Council (T.P.J.K., S.L. and D.F.), and the Engineering and Physical Sciences Research Council (D.F.).

Contributions

A.Š., T.P.J.K. and D.F. conceived the project; A.Š. designed and performed the computer simulations; A.K.B. performed the SPR measurements; G.M. performed the kinetic analysis; T.C.T.M. and S.L. contributed materials and/or analysis tools, and all authors participated in discussing results and writing the paper.

Methods

The coarse-grained model and the choice of parameters

We use the model developed in Ref. [20], extended to capture secondary nucleation. In spirit, this model is similar to the multistate Potts model of Zhang and Muthukumar [31], and the recent model of Ilie, Otter and Briels [32]. Recently, more rigorous schemes have been developed to map coarse-grained inter-protein interactions onto patchy-colloids for the purpose of studying protein aggregation by Ruff et al. [33, 34].

In our model each protein is represented by a spherocylinder that is $\sigma = 2nm$ wide and $L = 4\sigma = 8nm$ long. The hard core repulsion forbids for any distance between any two spherocylinders to be smaller than σ . The interaction between two proteins in the soluble “s” form is implemented as:

$$V_{ss}(r) = \begin{cases} -\epsilon_{ss} \left(\frac{\sigma}{r}\right)^6 & \text{if } r \leq 1.5\sigma \\ 0 & \text{if } r > 1.5\sigma \end{cases} \quad (1)$$

where r is the distance between the centers of the attractive tips located at the spherocylinders' ends. An attractive patch is added only at one spherocylinder pole to ensure formation of finite aggregates like those observed in experiments. This potential drives the formation of

micellar-like oligomers, where tips of participating proteins are in contact in the oligomer center (Fig. 1B). The parameter ϵ_{ss} controls the strength of the non-specific interactions between the soluble proteins. Using atomistic simulations we estimated ϵ_{ss} to be relatively small, on the order of $5kT$ [20]. To explore the influence of different solution conditions, we varied ϵ_{ss} between $3kT$ and $8kT$, as indicated in the text.

The interaction between two proteins in the intermediate conformation “i”, and between the soluble and the intermediate conformation is implemented using the same potential as in Eq. (1), with $\epsilon_{ss} \rightarrow \epsilon_{ii}$ and $\epsilon_{ss} \rightarrow \epsilon_{si}$, respectively. The intermediate state is designed to be between the soluble and the β -sheet forming state, corresponding to a conformation with more β -content than the soluble state, but not yet a fully folded β -hairpin. Hence, the relative strength of interactions was always preserved, with $\epsilon_{ss} < \epsilon_{si} < \epsilon_{ii}$. Their values were chosen such that nucleation is achieved within a reasonable computer time (see Supplementary Fig. S2), while preserving their relative strength; ϵ_{ii} is kept constant at $\epsilon_{ii} = 16kT$, and ϵ_{si} is kept constant at $\epsilon_{si} = 8kT$. Throughout the article k denotes the Boltzmann's constant and T is the temperature.

The attractive side-patch of the β -sheet forming configuration is $L_p = 0.6L$ long and spans an angle of 180° . If two patches face each other their interaction is:

$$V_{\beta\beta}(r) = \begin{cases} -\epsilon_{\beta\beta} \cos^2(\phi) - \epsilon_{\beta\beta} \left(\frac{\sigma}{r}\right) & \text{if } d \leq 1.5\sigma \\ 0 & \text{if } d > 1.5\sigma \end{cases} \quad (2)$$

where ϕ is the angle between the axes of the particles, d is the shortest distance between the axes of the patches, and r is distance between the centers of the patches. The first term controls that proteins in the β -sheet prone forms pack parallel to each other, mimicking the hydrogen-bond interactions between β -sheets, while the second term ensures compactness of the fibrils [22, 35, 36]. To drive the formation of thermodynamically stable fibrils, $\epsilon_{\beta\beta}$ has to be the strongest of all the interactions in the system. In this study we choose $\epsilon_{\beta\beta} = 60kT$ [37, 38]. General aggregation of patchy-spherocylinders has been studied in details in our previous work [39].

The cross-interaction between the soluble and the β -sheet-forming configuration is designed as:

$$V_{s\beta}(d) = \begin{cases} -\epsilon_{s\beta} & \text{if } d < 1.5\sigma \\ 0 & \text{if } d > 1.5\sigma \end{cases} \quad (3)$$

where d is the shortest distance between the centre of the attractive tip and the axis of the β -patch, and $\epsilon_{s\beta} = \epsilon_{ss} + 1kT$. The i - β interaction is described in the same way, with $\epsilon_{s\beta} \rightarrow \epsilon_{i\beta}$, and $\epsilon_{i\beta} = \epsilon_{ii} + 1kT$.

Protein adsorption onto the preformed fibril is given by:

$$V_{sf}(d) = \begin{cases} -\epsilon_{sf} \left(\frac{\sigma}{d}\right)^6 & \text{if } r \leq 1.5\sigma \\ 0 & \text{if } r > 1.5\sigma \end{cases} \quad (4)$$

where d is the shortest distance between the centre of the attractive tip of the soluble protein and the body of the β -sheet prone protein (there is no other angular dependence). Adsorption of the intermediate “ i ” conformation onto the fibril is described in the same way (Eq. (4)), with $\epsilon_{sf} \rightarrow \epsilon_{if}$, and $\epsilon_{if} = 1kT$. The β -sheet prone protein interacts with the preformed fibril only via volume exclusion. The model parameters are summarized in Supplementary Figure S1.

MC Scheme

MC simulations were performed with small translational and rotational moves, to approach the realistic dynamics of the system. The interconversion between the three states discussed above was carried out with a small probability $P = 0.0002$, which mimics the slow conversion of the soluble protein into a fibril-forming β -sheet prone configuration. Every conversion from the soluble to the β -state is penalized with a change in the excess chemical potential of magnitude $\Delta\mu = 20kT$, and the $s \rightarrow i$ and the $i \rightarrow \beta$ with $0.5\Delta\mu$ (Fig. 1a). These values are chosen to reflect the fact that amyloidogenic proteins with small- to mid- β -propensity, such as $A\beta$, are typically not found in the β -sheet prone conformation in solution [40, 41].

Simulations were performed in a periodic cubic box in a grand-canonical ensemble, where the chemical potential of non-adsorbed soluble proteins was kept constant. This scheme was chosen to avoid the depletion of monomers from the solution due to the adsorption onto the surface of the preformed fibril. For this purpose, we do not distinguish between the monomeric soluble species, and the soluble species that are part of an oligomer in solution. The number of soluble proteins in the beginning of each simulation was set to ~ 600 , and the box size was adjusted to match the targeted peptide concentration. Soluble proteins are added or removed from anywhere in the simulation box, according to the grand-canonical scheme [42], excluding the $r = 5\sigma$ region around the capped preformed fibril. All simulations were performed with the same size of the preformed fibril, which consists of $N = 92$ β -sheet prone proteins and is unable to grow further. We were monitoring only the first generation of replicas, and have allowed the soluble proteins to adsorb only onto the preformed fibril, and not onto its replicas.

Kinetics of self-replication

In bulk experimental systems, the overall kinetics are determined by the processes of spontaneous nucleation in solution, elongation and self-replication, all of which alter the fibril population. To compare bulk kinetic measurements to the modelling of nucleation on a single, growth-incompetent fibril used in simulations, it is necessary to dissect the macroscopic behaviour into its con-

stituent processes. This can be achieved by developing a theoretical kinetic model and global fitting to the experimental kinetic data. We have adapted a theoretical kinetic model for the aggregation of $A\beta_{40}$ [25] to include the Langmuir-like adsorption of peptide molecules onto the growing fibril, and fit it to bulk experimental kinetic data to obtain the rate of self-replication at various peptide concentrations. The details of the kinetic model as well as the global fits used to obtain this rate of secondary nucleation are shown in the Supplementary Section SII and Fig. S4.

Experimental exploration of intermediate oligomers in self-replication of $A\beta_{42}$

If the oligomers generated through secondary nucleation were of the same structure as the fibrils, their concentration, $[O]$, could be estimated from the known rate parameters for the fibrillar growth as $[O] = \frac{k_2 m_{\text{tot}}^{n_2}}{2^{n_2+1} k_+}$, where k_2 is the rate constant for secondary nucleation, k_+ is the fibril elongation rate constant, m_{tot} is the total protein concentration, and n_2 is the reaction order for secondary nucleation. [43]. Using the values for the rate constants extracted from kinetic measurements of $A\beta_{42}$ aggregation ($k_2 \approx 10^4 \text{ M}^{-2}\text{s}^{-1}$, $k_+ \approx 3 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$ and $m_{\text{tot}} = 5\mu\text{M}$) [10], we find this concentration to be $[O] \approx 0.01 \text{ pM}$. This value is at least 5 orders of magnitude smaller than the experimentally measured concentration of oligomers in the same system (nanomolar range [10]), indicating that the structure of oligomers generated via such secondary pathway is necessarily different from that of the fibrils.

Model for the change of the rate of self-replication with surface coverage

We recall that the conformational change, and the subsequent fibril nucleation, is favourable only for oligomers above a certain critical size. The free energy of formation of such an oligomer on a finite surface scales as $\Delta F(N) \sim -N \ln(Kc/(1+Kc))$ where K is the monomer-surface binding constant ($K \sim \epsilon_{sf}$) and c is the free monomer concentration (Supplementary Section SII.F). Supplementary Fig. 3 shows the free energy change for oligomer formation on the fibril surface, $\Delta F(N)$, measured from the size distribution of oligomers adsorbed onto the fibril in our simulations. As predicted, this quantity decreases with increasing protein concentration, reaching a plateau at high concentrations. If the aggregation number is the slow degree of freedom, and hence an appropriate reaction coordinate, as is the case for example in classical nucleation theory, the rate of nucleation then depends exponentially on the negative magnitude of the free energy change for the critical oligomer size

formation ($N = N^*$):

$$\ln(r) \sim -\Delta F(N^*) \sim N^* \ln(Kc/(1 + Kc)). \quad (5)$$

An arrow in the Supplementary Fig. S3 marks the lowest concentration range at which we observe nucleation ($\ln(c) \approx 8.5$). The slope at that point (≈ 0.6), multiplied by the average critical oligomer size ($N \approx 6$, inset in Fig. 4b), should give us the expected apparent reaction order in the kinetic plot $\gamma_s \approx 3.6$. The measured reaction order at the same concentration range in Fig. 4b is $\gamma_s \approx 3.3$, which agrees well with the predicted value within the error of our scaling theory and measurements.

SPR Experiments

A β 40 amyloid fibrils were attached to the surface of an SPR biosensor and exposed to a solution containing monomeric A β 40. In this case, monomers simultaneously attach both to the fibril ends and to their surfaces. However, due to their very different kinetics and thermodynamics, the two processes can readily

be distinguished (Supplementary Section SIII). The elongation of fibrils will lead to a linear increase in mass, while the rate of attachment of peptide to the surface of fibrils is expected to decrease exponentially with time as the available binding sites become occupied. Conversely, upon washing the fibrils, the surface-bound peptide molecules are expected to show an exponential detachment behaviour, at high rates due to their relatively low binding free energy, while the rate of loss from the fibril ends by monomer dissociation is expected to be linear and very slow, due to the high thermodynamic stability of the β -sheet rich fibrils [44]. By following the kinetic data of monomer detachment, we can distinguish the fast exponential from the slow linear dissociation steps (Supplementary Fig. S5), and obtain the amplitude of the exponential signal resulting from attachment to the surfaces of the fibrils, at various concentrations of the free monomers.

The data that support the plots within this paper and other findings of this study are available online at <https://www.repository.cam.ac.uk/handle/1810/255082>, as well as from the corresponding authors upon request.

-
- [1] Prusiner, S. B. Prions. *Proceedings of the National Academy of Sciences* **95**, 13363–13383 (1998).
- [2] Hall, D. & Edsles, H. Silent prions lying in wait: a two-hit model of prion/amyloid formation and infection. *Journal of molecular biology* **336**, 775–786 (2004).
- [3] Collins, S. R., Dougllass, A., Vale, R. D. & Weissman, J. S. Mechanism of prion propagation: amyloid growth occurs by monomer addition. *PLoS biology* **2**, e321 (2004).
- [4] Ferrone, F. A., Hofrichter, J. & Eaton, W. A. Kinetics of sickle hemoglobin polymerization: II. a double nucleation mechanism. *Journal of molecular biology* **183**, 611–631 (1985).
- [5] Eaton, W. A. & Hofrichter, J. Sick cell hemoglobin polymerization. *Advances in protein chemistry* **40**, 263–279 (1990).
- [6] Dobson, C. M. Protein misfolding, evolution and disease. *Trends in biochemical sciences* **24**, 329–332 (1999).
- [7] Chiti, F. & Dobson, C. M. Protein misfolding, functional amyloid, and human disease. *Annual Review of Biochemistry* **75**, 333–366 (2006).
- [8] Knowles, T. P., Vendruscolo, M. & Dobson, C. M. The amyloid state and its association with protein misfolding diseases. *Nature Reviews Molecular Cell Biology* **15**, 384–396 (2014).
- [9] Hortschansky, P., Schroeckh, V., Christopeit, T., Zandomeneghi, G. & Fändrich, M. The aggregation kinetics of Alzheimer’s β -amyloid peptide is controlled by stochastic nucleation. *Protein science* **14**, 1753–1759 (2005).
- [10] Cohen, S. I. *et al.* Proliferation of amyloid- β 42 aggregates occurs through a secondary nucleation mechanism. *Proceedings of the National Academy of Sciences* **110**, 9758–9763 (2013).
- [11] Padrick, S. B. & Miranker, A. D. Islet amyloid: phase partitioning and secondary nucleation are central to the mechanism of fibrillogenesis. *Biochemistry* **41**, 4694–4703 (2002).
- [12] Ruschak, A. M. & Miranker, A. D. Fiber-dependent amyloid formation as catalysis of an existing reaction pathway. *Proceedings of the National Academy of Sciences* **104**, 12341–12346 (2007).
- [13] Schlamadinger, D. E. & Miranker, A. D. Fiber-dependent and -independent toxicity of islet amyloid polypeptide. *Biophysical journal* **107**, 2559–2566 (2014).
- [14] Buell, A. K. *et al.* Solution conditions determine the relative importance of nucleation and growth processes in α -synuclein aggregation. *Proceedings of the National Academy of Sciences* **111**, 7671–7676 (2014).
- [15] Galvagnion, C. *et al.* Lipid vesicles trigger α -synuclein aggregation by stimulating primary nucleation. *Nature chemical biology* **11**, 229–234 (2015).
- [16] Foderà, V., Librizzi, F., Groenning, M., Van De Weert, M. & Leone, M. Secondary nucleation and accessible surface in insulin amyloid fibril formation. *The Journal of Physical Chemistry B* **112**, 3853–3858 (2008).
- [17] Meisl, G., Yang, X., Frohm, B., Knowles, T. P. & Linse, S. Quantitative analysis of intrinsic and extrinsic factors in the aggregation mechanism of Alzheimer-associated A β -peptide. *Scientific reports* **6**, 18728 (2016).
- [18] Meisl, G., Yang, X., Dobson, C. M., Linse, S. & Knowles, T. P. J. A general reaction network unifies the aggregation behaviour of the A β 42 peptide and its variants (2016). arXiv:1604.00828.
- [19] Bieler, N. S., Knowles, T. P., Frenkel, D. & Vácha, R. Connecting macroscopic observables and microscopic assembly events in amyloid formation using coarse grained simulations. *PLoS Comput Biol* **8**, e1002692 (2012).

- [20] Šarić, A., Chebaro, Y. C., Knowles, T. P. & Frenkel, D. Crucial role of nonspecific interactions in amyloid nucleation. *Proceedings of the National Academy of Sciences* **111**, 17869–17874 (2014).
- [21] Cohen, S. I. *et al.* A molecular chaperone breaks the catalytic cycle that generates toxic A β oligomers. *Nature structural & molecular biology* **22**, 207–213 (2015).
- [22] Nelson, R. *et al.* Structure of the cross- β spine of amyloid-like fibrils. *Nature* **435**, 773–778 (2005).
- [23] Serio, T. R. *et al.* Nucleated conformational conversion and the replication of conformational information by a prion determinant. *Science* **289**, 1317–1321 (2000).
- [24] Anwar, J., Khan, S. & Lindfors, L. Secondary crystal nucleation: nuclei breeding factory uncovered. *Angewandte Chemie* **127**, 14894–14897 (2015).
- [25] Meisl, G. *et al.* Differences in nucleation behavior underlie the contrasting aggregation kinetics of the A β 40 and A β 42 peptides. *Proceedings of the National Academy of Sciences* **111**, 9384–9389 (2014).
- [26] Bucciantini, M. *et al.* Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature* **416**, 507–511 (2002).
- [27] Haass, C. & Selkoe, D. J. Soluble protein oligomers in neurodegeneration: lessons from the alzheimer’s amyloid beta-peptide. *Nature Reviews Molecular Cell Biology* **8**, 101–112 (2007).
- [28] Walsh, D. M. *et al.* Naturally secreted oligomers of amyloid beta protein potently inhibit hippocampal long-term potentiation in vivo. *Nature* **416**, 535–539 (2002).
- [29] Zeravcic, Z. & Brenner, M. P. Self-replicating colloidal clusters. *Proceedings of the National Academy of Sciences* **111**, 1748–1753 (2014).
- [30] Wang, T. *et al.* Self-replication of information-bearing nanoscale patterns. *Nature* **478**, 225–228 (2011).
- [31] Zhang, J. & Muthukumar, M. Simulations of nucleation and elongation of amyloid fibrils. *The Journal of chemical physics* **130**, 035102 (2009). [OpenDataFig5c.txt](#)
- [32] Ilie, I. M., den Otter, W. K. & Briels, W. J. A coarse grained protein model with internal degrees of freedom. application to α -synuclein aggregation. *The Journal of chemical physics* **144**, 085103 (2016).
- [33] Ruff, K. M., Khan, S. J. & Pappu, R. V. A coarse-grained model for polyglutamine aggregation modulated by amphipathic flanking sequences. *Biophysical journal* **107**, 1226–1235 (2014).
- [34] Ruff, K. M., Harmon, T. S. & Pappu, R. V. Camelot: A machine learning approach for coarse-grained simulations of aggregation of block-copolymeric protein sequences. *The Journal of chemical physics* **143**, 243123 (2015).
- [35] Fitzpatrick, A. W. *et al.* Atomic structure and hierarchical assembly of a cross- β amyloid fibril. *Proceedings of the National Academy of Sciences* **110**, 5468–5473 (2013).
- [36] Serpell, L. C. Alzheimer’s amyloid fibrils: structure and assembly. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease* **1502**, 16–30 (2000).
- [37] Davis, C. H. & Berkowitz, M. L. A molecular dynamics study of the early stages of amyloid- β (1–42) oligomerization: The role of lipid membranes. *Proteins: Structure, Function, and Bioinformatics* **78**, 2533–2545 (2010).
- [38] Buell, A. K. *et al.* Detailed analysis of the energy barriers for amyloid fibril growth. *Angewandte Chemie International Edition* **51**, 5247–5251 (2012).
- [39] Vácha, R. & Frenkel, D. Relation between molecular shape and the morphology of self-assembling aggregates: a simulation study. *Biophysical journal* **101**, 1432–1439 (2011).
- [40] Fandrich, M., Fletcher, M. A. & Dobson, C. M. Amyloid fibrils from muscle myoglobin - even an ordinary globular protein can assume a rogue guise if conditions are right. *Nature* **410**, 165–166 (2001).
- [41] Allison, J. R., Varnai, P., Dobson, C. M. & Vendruscolo, M. Determination of the free energy landscape of alpha-synuclein using spin label nuclear magnetic resonance measurements. *Journal of the American Chemical Society* **131**, 18314–18326 (2009).
- [42] Frenkel, D. & Smit, B. *Understanding molecular simulation: from algorithms to applications*, vol. 1 (Academic press, 2001).
- [43] Michaels, T. C., Lazell, H. W., Arosio, P. & Knowles, T. P. Dynamics of protein aggregation and oligomer formation governed by secondary nucleation. *The Journal of chemical physics* **143**, 054901 (2015).
- [44] Hellstrand, E., Boland, B., Walsh, D. M. & Linse, S. Amyloid β -protein aggregation produces highly reproducible kinetic data and occurs by a two-phase process. *ACS chemical neuroscience* **1**, 13–18 (2009).

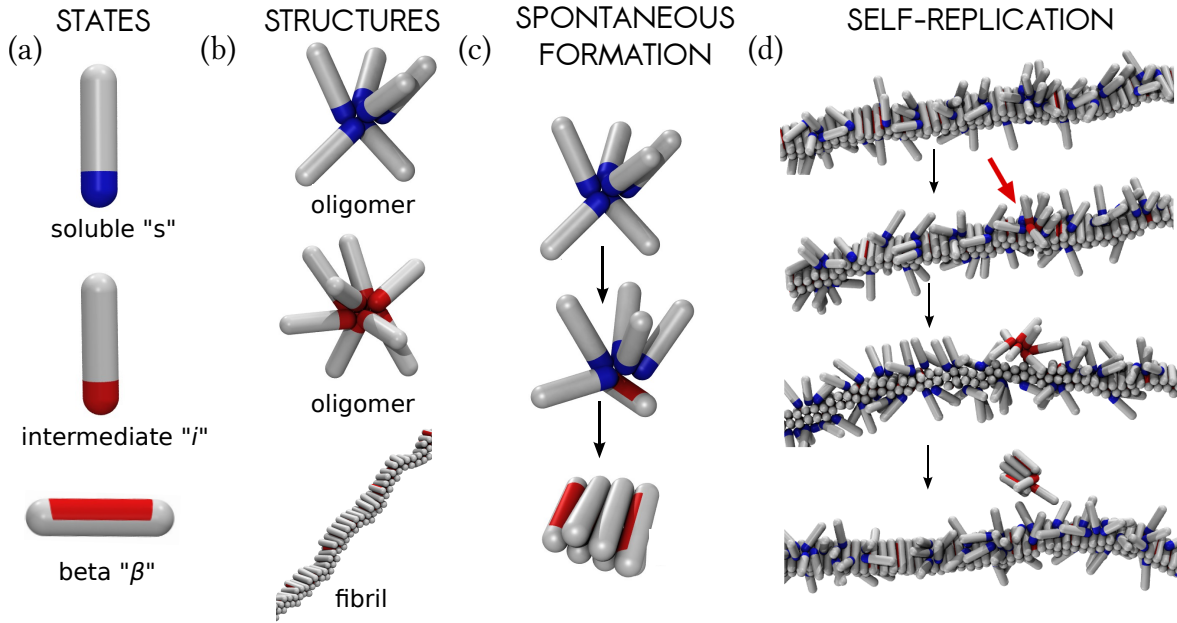


FIG. 1: **The coarse-grained model and the nucleation processes in the system.** (a) A protein is allowed to exist in three conformations. From top to bottom: soluble state ("s"), intermediate conformation ("i"), and the β -sheet prone state (" β "). (b) Aggregated proteins. From top to bottom: oligomer made of soluble proteins, oligomer made of proteins in the intermediate state, and the fibril made of proteins in the β -sheet prone state. (c) Primary nucleation takes place in two steps. Soluble proteins form finite oligomers (*top*), which can convert into a nucleus rich in β -sheet (*bottom*) that continues growing. (d) Fibril self-replication (secondary nucleation). From top to bottom: Soluble protein monomers adsorb onto the surface of a preformed fibril, locally forming oligomers. Once proteins within an oligomer convert into the intermediate conformation (depicted with red attractive tips, accentuated with the red arrow), they become more prone to self-aggregation, which in turn leads to oligomer detachment. Finally, the detached oligomer converts into a nucleus of β -sheets, and continues growing. Snapshots were taken at $\epsilon_{ss} = 4kT$, $\epsilon_{sf} = 8kT$, and $c = 50\mu M$.

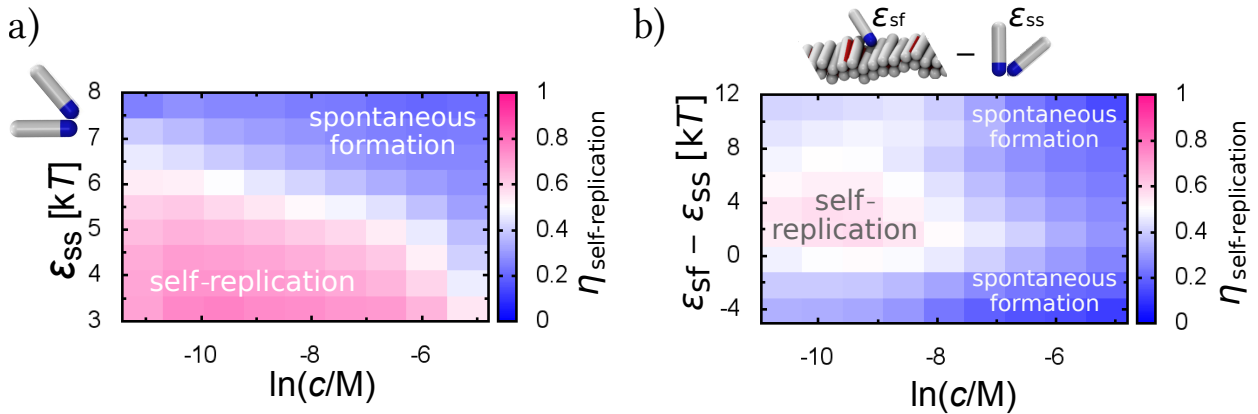


FIG. 2: **Conditions supporting fibril self-replication.** (a) Fraction of self-replication events, $\eta_{\text{self-replication}}$, in the total number of nucleation events, as a function of the protein concentration c and the interaction strength between soluble protein molecules ϵ_{ss} . The protein-fibril interaction is kept constant at $\epsilon_{sf} = 8kT$. (b) Fraction of self-replication events as a function of the protein concentration c and the difference between the protein-fibril interaction and the protein self-interaction ($\epsilon_{sf} - \epsilon_{ss}$), exhibiting a narrow regime where self-replication can be a dominant mechanism of formation. Data collected at $\epsilon_{ss} = 5kT$.

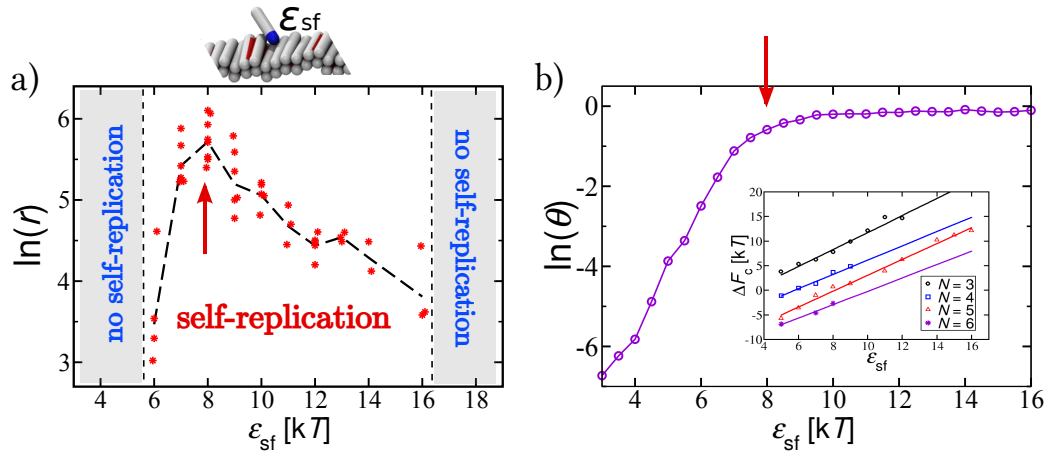


FIG. 3: **Strong bounds for self-replication.** (a) Dependence of the rate of self-replication, r , on the protein-fibril affinity, ϵ_{sf} . (b) Fractional coverage of the surface of the fibril (θ) as a function of ϵ_{sf} . Red arrows in (a) and (b) point to the area of the fastest self-replication, when the fibril is well covered with monomers. Inset: the free energy cost (ΔF_c) for the conversion of an oligomer of size N from the "s" conformation, that is attached onto the fibril, into the "i" conformation that detaches from the fibril surface. ΔF_c increases with increasing protein-fibril affinity. Data in (a) and (b) were collected at $\epsilon_{ss} = 4kT$ and $c = 0.15mM$.

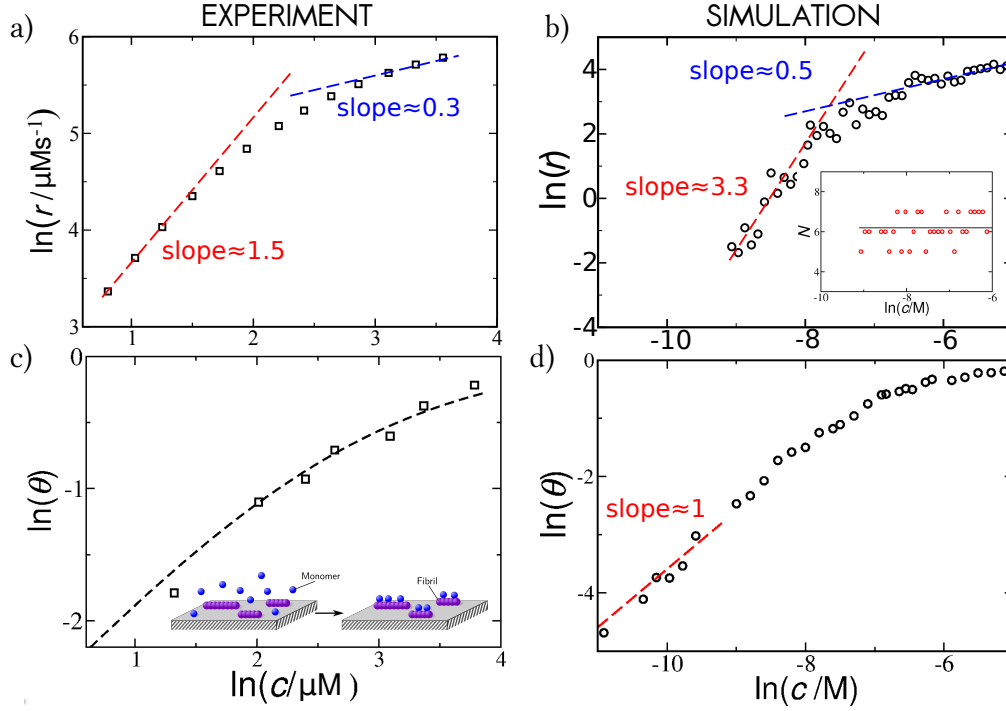


FIG. 4: **Kinetics of fibril self-replication.** (a) **Experimental results:** The rate of secondary nucleation for the A β 40 system versus the initial concentration of soluble monomers, from Ref. [25]. (b) **Simulation results:** The rate of secondary nucleation of fibrils with a moderate affinity for soluble monomers ($\epsilon_{sf} = 6kT$) as a function of the concentration of the monomeric proteins in solution. Inset: the average critical oligomer size stays constant over the entire concentration range; the solid line plots the linear fit over the concentration range. (c) **Experimental results:** Fractional coverage of the surface of A β 40 fibrils, θ , under the same conditions as the kinetic experiments in (a), versus the concentration of the monomers. The dashed line is the fit to the Langmuir isotherm with $K^{-1} = 15\mu\text{M}$. Inset: schematic representation of the adsorption of monomeric species (coloured in blue) onto the surface of fibrils (coloured in violet), measured via SPR. (d) **Simulation results:** Surface coverage θ versus the concentration of free monomers at $\epsilon_{sf} = 6kT$. Inter-protein interaction is kept constant at $\epsilon_{ss} = 4kT$ for all simulation data. The dashed lines in (a), (b), and (d) are given as the guide to the eye.

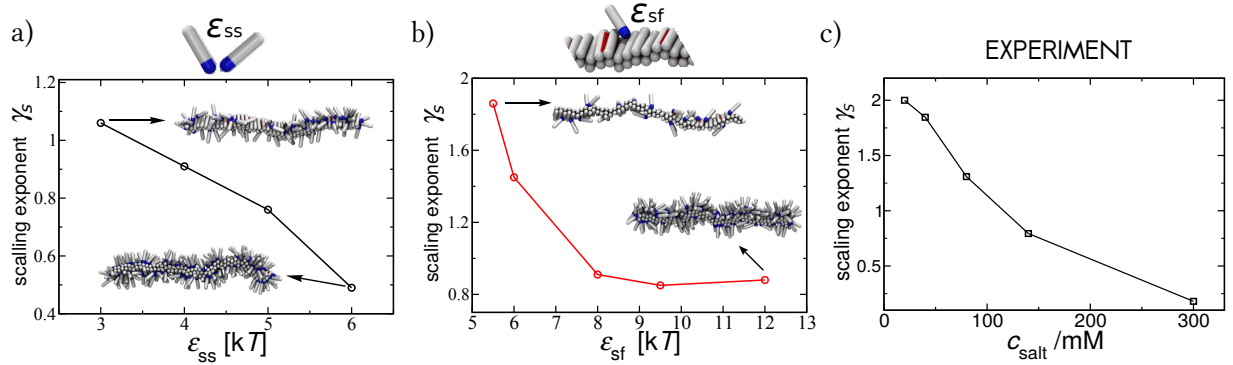


FIG. 5: **The apparent reaction order is controlled by the surface saturation.** **Simulation results:** (a) Scaling exponent for the kinetics of fibril self-replication, averaged over the range of concentrations ($20\mu\text{M} \leq c \leq 1\text{mM}$), as a function of the inter-protein interaction between soluble monomers at constant protein-fibril affinity $\epsilon_{sf} = 8kT$, and (b) as a function of the protein-fibril affinity at constant inter-protein affinity $\epsilon_{ss} = 4kT$. An increase in ϵ_{ss} and ϵ_{sf} increases the surface coverage, as shown by the representative snapshots in insets, taken at a monomer concentration $c = 0.15\text{mM}$. **Experimental results:** (c) The average scaling exponent for self-replication of A β 42 fibrils at a range of NaCl salt concentrations, whose increase is expected to increase both ϵ_{ss} and ϵ_{sf} , from Ref. [18].

Supplementary Information

I. Coarse-grained Simulations

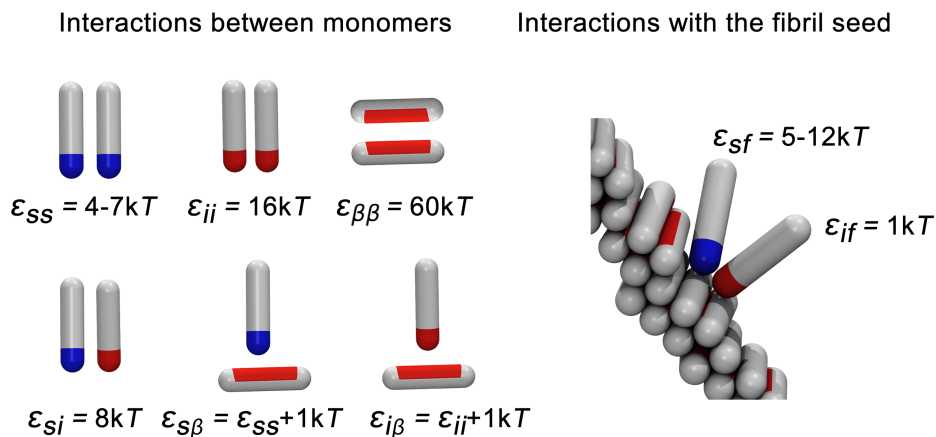


FIG. S1: Coarse-grained model: Possible interactions in the system and their values.

A. Rate of spontaneous fibril formation and self-replication

The rates for spontaneous fibril formation through primary nucleation and self-replication through secondary nucleation 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:

$$r = \frac{1}{\langle t_{lag} \rangle}. \quad (S1)$$

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 $\eta_{self-replication} = \frac{r^{(secondary)}}{r^{(secondary)} + r^{(primary)}}$.

C. Choice of the intermediate state

The rate of self-replication for $A\beta$ 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 $A\beta$ 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 $A\beta$ 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 $\epsilon_{ii} = 16kT$ and have kept it constant in our simulations.

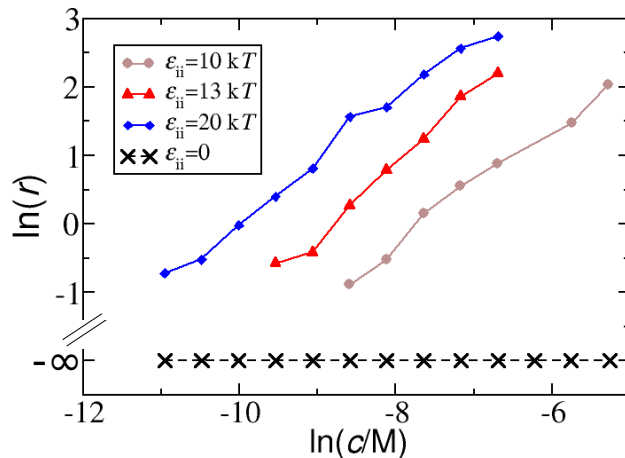


FIG. S2: The rate of self-replication for a two state system (dashed line, denoted $\epsilon_{ii} = 0$), which did not yield a single nucleation event within the simulation time, and for three different self-interaction values of the intermediate state, ϵ_{ii} . Data collected at $\epsilon_{ss} = 4kT$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

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 $\epsilon_{sf} = 12kT$ and $\epsilon_{ss} = 0kT$.

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 from 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 for forming an oligomer of size N , $P(N)$, increases exponentially with the monomer concentration $P(N) \sim e^{\Delta f(N)/kT} c^N$, where $\Delta f(N)$ is the concentration-independent free-energy. On a surface we can replace the monomer concentration with the surface coverage; hence the probability for oligomer formation on the surface is $P(N) \sim e^{\Delta f(N)/kT} (Kc/(1+Kc))^N$, where K is the monomer-surface binding constant ($K \sim \epsilon_{sf}$). The free energy change for the formation of an adsorbed oligomer then scales as $\Delta F(N) \sim -N \ln(Kc/(1+Kc))$. This holds at low concentrations, and can deviate at higher surface coverages, when oligomers start mutually interacting.

The free energy change for oligomerisation on the fibril surface is obtained by simulating the system of monomers which are able to adsorb onto the preformed fibril, but cannot change their conformation. We collected the size distribution of oligomers that are in contact with the fibril $P(N)$. The free energy for formation of an adsorbed oligomer of size N is then $\Delta F_o(N) = -\log P(N) + F^0$, where the zero-energy level F^0 is attributed to the free monomers.

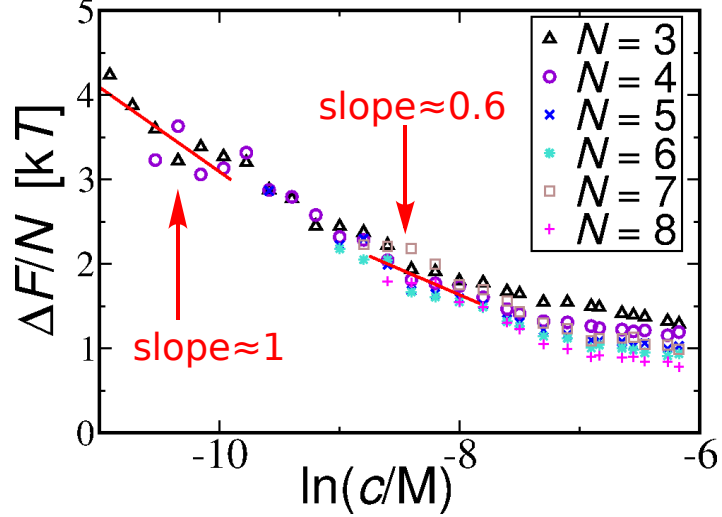


FIG. S3: The free energy per particle for formation of an oligomer of size N on the fibril surface versus the monomer concentration; $\epsilon_{ss} = 4kT$ and $\epsilon_{sf} = 6kT$ are kept constant. The lowest concentration range at which nucleation is observed in Fig. 4b ($-9 < \ln(c) < -8$) is marked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

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

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

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 $k_2 \frac{1}{1+c^{n_2}/K_M} \rightarrow k_2 \left(\frac{1}{1+c/K_M}\right)^{n_2}$ 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

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

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

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

Then differentiating and substituting equation (S2) yields

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

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 $\frac{df(m(t))}{dt} = \frac{df(m)}{dm} \frac{dm(t)}{dt}$:

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

We perform the integrals with respect to m

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

where $A(t)$ is given by

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

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

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

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

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

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

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

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

where the definitions of the parameters are

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

$$\lambda = \sqrt{2k_+k_n c^{n_c}} \quad (\text{S14})$$

$$C_\pm = \pm \frac{\lambda^2}{2\kappa^2} \quad (\text{S15})$$

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

$$\bar{k}_\infty = \sqrt{k_\infty^2 - 2C_+C_- \kappa^2} \quad (\text{S17})$$

$$B_\pm = \frac{k_\infty \pm \bar{k}_\infty}{2\kappa} \quad (\text{S18})$$

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

$$r = k_2 M(t) \left(\frac{m(t)}{1 + m(t)/K_M} \right)^{n_2} \quad (\text{S19})$$

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:

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

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

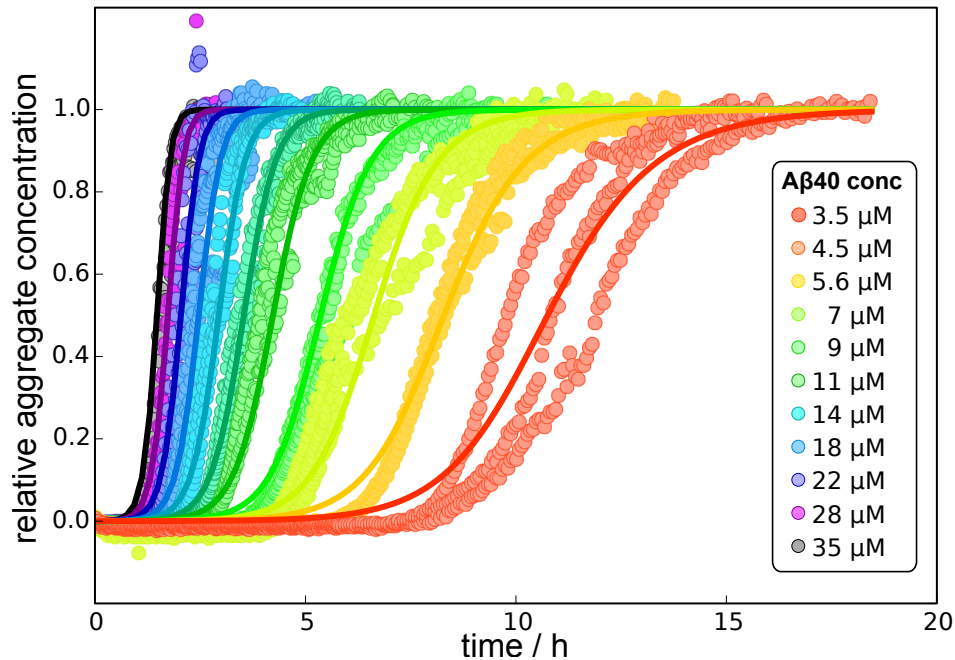


FIG. S4: Global fit to experimental data for A β 40 aggregation from Ref. [3]. The global fit of Eq. S12 to the unseeded aggregation data with 4 free global parameters: k_+k_n , k_+k_2 , n_2 and K_M .

interface [9], also available online [13]. The best fit is shown in Fig. S4 and yields the parameters $k_+k_n = 0.4 \text{ M}^{-2}\text{s}^{-2}$, $k_+k_2 = 9 \cdot 10^{10} \text{ M}^{-n_2-1}\text{s}^{-2}$, $n_2 = 2.3$ and $K_M = 3.8 \mu\text{M}$. The rate of elongation, k_+ , was estimated separately from seeded experiments in Meisl et al.[3] as $k_+ = 3 \cdot 10^5 \text{ M}^{-1}\text{s}^{-1}$, yielding $k_2 = 3 \cdot 10^5 \text{ M}^{-n_2}\text{s}^{-1}$. Note that the values of the rate constants are approximate within at least an order of magnitude, the error on the reaction order was estimated to be at least 20%. The same method was employed to obtain r for the data of varying salt concentrations. A straight line was then fitted to the double logarithmic plots of these rates versus the monomer concentration, to yield the scaling displayed in Figure 5c.

III. Surface plasmon resonance measurement of A β 40 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. A β 40 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 A β 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 $K^{-1} = 15\mu\text{M}$, as shown in Fig. 4C and Fig. S6. It is interesting to note here that the affinity of A β 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 $A\beta$ 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.

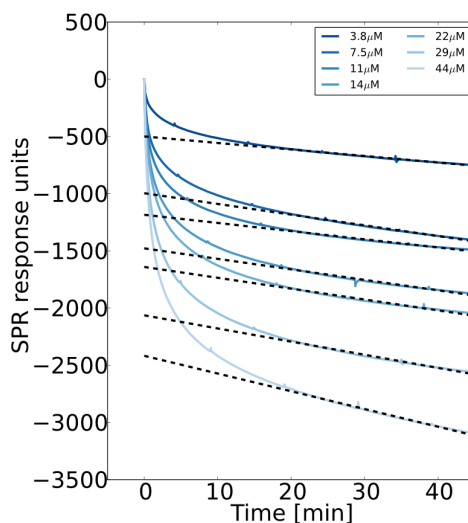


FIG. S5: The raw SPR dissociation data.

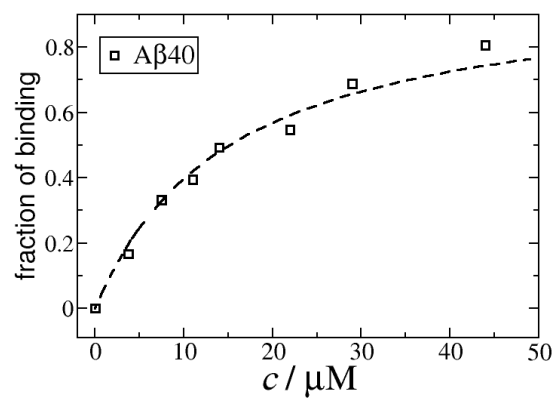


FIG. S6: Fraction of the peptides bound onto the surface of $\text{A}\beta 40$ fibrils, at the same conditions as the kinetic experiments in Fig. 4A, versus the concentration of the soluble monomer. The dashed line is the fit to the Langmuir isotherm with $K^{-1} = 15\mu\text{M}$.

-
- [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- β 42 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 A β 40 and A β 42 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/>