UNIVERZA V MARIBORU
FAKULTETA ZA NARAVOSLOVJE IN MATEMATIKO

Oddelek za fiziko

MAGISTRSKO DELO

Samo Curk

Maribor, 2018

UNIVERZA V MARIBORU
FAKULTETA ZA NARAVOSLOVJE IN MATEMATIKO
Oddelek za fiziko

Magistrsko delo

INHIBICIJA SAMO-REPLIKACIJE PROTEINSKIH FIBRILOV

Master Thesis

INHIBITION OF SELF-REPLICATION OF PROTEIIN FIBRILS

Mentor:
dr. Marko Gosak

Kandidat:
Samo Curk

Somentorica:
dr. Anđela Šarić

Maribor, 2018

I would like to thank Anđela for welcoming me to Cambridge and guiding me on this project. Thank you Dr. Gosak for your help in writing the thesis. Thanks to my family for your support and especially to Tine for letting me follow in your footsteps. Thanks to my fiancée Eva for being amazing, for all the fun we have together, and for tolerating my soul-consuming aspirations in Science. Finally, big thanks to Julka and Livica for taking me on walks and filling my life with joy.

UNIVERZA V MARIBORU
FAKULTETA ZA NARAVOSLOVJE IN MATEMATIKO

IZJAVA

Podpisani Samo Curk, rojen 15.10.1993 Študent Fakultete za naravoslovje in matematiko Univerze v Mariboru, študijskega programa Fizika, izjavljam, da je magistrsko delo z naslovom Inhibicija samo-replikacije proteinskih fibrilov pri mentorju dr. Marko Gosaku in mentorici dr. Anđeli Šarić avtorsko delo. V magistrskem delu so uporabljeni viri in literatura korektno navedeni; teksti in druge oblike zapisov niso uporabljeni brez navedbe avtorjev.

Maribor, _____ Podpis: _____

UNIVERZA V MARIBORU
FAKULTETA ZA NARAVOSLOVJE IN MATEMATIKO

Curk S.: Inhibicija samo-replikacije proteinskih fibrilov
Magistrsko delo, Univerza v Mariboru, Fakulteta za naravoslovje in matematiko, Oddelek za fiziko, 2018.

IZVLEČEK

Proteinski fibrili nastanejo z agregacijo delno zvitih proteinov in so odgovorni ali pomembno vplivajo na mnogo hudih človeških bolezni, kot sta Alzeimerjeva ali diabetes tipa II. Samo-replikacija fibrilov je proces, v katerem obstoječi proteinski fibrili katalizirajo nastanek novih fibrilov na način, da ponudijo vezavno površino, kjer se proteini lažje srečajo in agregirajo. To prispeva k naglem in eksponentnem napredovanju bolezni. V magistrski nalogi predstavimo Monte Carlo simulacije računalniškega modela agregacije, v katerega vpeljemo inhibitorje. To so delci, ki se lahko vežejo na površino fibrilov in tako inhibirajo oziroma upočasnijo proces samo-replikacije. Takšen način inhibicije se izkaže za zelo učinkovit, ampak zaradi odbojne interakcije med delci naletimo tudi na pojav makromolekularnega gnečenja, ki povzroči, da se pri določeni pokritosti površine s proteini hitrost samo-replikacije poveča. Edinstven deskriptor hitrosti replikacije najdemo v povprečni velikosti skupka na površino vezanih proteinov, ki nosi informacijo o celotni porazdelitvi agregacijskih skupkov. Predstavimo teorije, ki uspešno razložijo vse značilnosti opažanega obnašanja. S pomočjo mrežnega modela napovemo, katere interakcije med delci na površini imajo največji inhibicijski potencial.

Ključne besede: agregacija amiloidov, samo-sestavljanje proteinov, inhibicija, računalniške simulacije, metoda Monte Carlo, nukleacijski mehanizem, gnečenje makromolekul, krajina proste energije, statistična mehanika, mehka snov, fizikalna kemija

UNIVERZA V MARIBORU
FAKULTETA ZA NARAVOSLOVJE IN MATEMATIKO

Curk S.: Inhibition of self-replication of protein fibrils.
Master Thesis, University of Maribor, Faculty of Natural Sciences and Mathematics, Department of Physics, 2018.

ABSTRACT

Protein fibrils are formed by a process called amyloid aggregation and are implicated in many debilitating human diseases such as Alzheimer's or Type II Diabetes. Self-replication of fibrils is a process by which existing protein fibrils catalyse the formation of new fibrils by offering a surface on which proteins can bind, and therefore facilitate aggregation. This leads to exponential growth of fibril mass and fast propagation of amyloid diseases. In this thesis, we present simulations of a minimal but fairly complex computational model of aggregation with added inhibitory particles that can bind to the fibril surface. It turns out the mechanism of inhibition where inhibitors compete with proteins for the surface is very promising. However, we also find a manifestation of a macromolecular crowding effect which actually promotes self-replication at given protein coverage of the fibril surface. We find a unique descriptor for the rate of replication in the average protein aggregate size. We present theories that successfully explain all characteristics of observed simulation behaviour. By employing a lattice model, we predict which inter-particle interactions on the fibril surface have the largest inhibitory potential.

Key words: Amyloid Aggregation, Self-assembly, Protein Fibrils, Inhibition, Course-grained Simulation, Monte Carlo Method, Nucleation Mechanism, Macromolecular Crowding, Free Energy Landscape, Statistical Mechanics, Soft Matter, Physical Chemistry

Contents

1 Introduction

Amyloid fibrils are a form of protein assembly that results from the aggregation of normally soluble proteins. Fibrils are elongated thread-like structures and are a dominant type of protein aggregate that is associated with more than 50 increasingly prevalent human diseases which are at present incurable [1, 2]. All these diseases share a common morphology even though the soluble proteins that make up the aggregates as well as the area of pathogenic influence vary substantially across diseases. Diseases most associated with amyloid aggregation are Parkinson's disease with its α-synuclein protein, Diabetes II with amylin, various types of Amyloidosis, and chiefly Alzheimer's disease with its Aβ peptide which accounts for up to 80% of all dementia cases [3].

The aggregation reaction ordinarily involves many molecular steps [1, 4–7]. A first important step is *spontaneous* or *primary nucleation* by which the first fibril nuclei are formed from soluble monomers in solution. These nuclei then grow by an attachment of monomeric proteins to fibril ends by a process called *elongation*, eventually into mature fibrils that are observable under the microscope. These two processes are enough to explain the early time signature of the aggregation reaction. But recent experiments revealed that once first amyloid fibrils are generated, their surfaces catalyse the formation of new fibril nuclei via a secondary *self-replication* process, which leads to exponential growth of fibril mass and therefore to fast propagation of amyloid diseases [5, 8–12].

Particularly in Alzheimer's disease, this secondary auto-catalytic process has been found to be the main source of new fibril nuclei [8, 9, 13, 14]. Even more importantly, it is a dominant source of oligomers, intermediate protein aggregates that do not yet posses the a β-sheet structure characteristic of mature fibrils but are able to convert to a growth-competent fibril nucleus. These intermediate species are important as it has been found that fully grown protein fibrils do not significantly interact with the mammal brain as previously believed and are quite inert. Oligomers, in fact, have been found to be toxic to neuronal cells and are now believed to be the main cause of neuronal death in Alzheimer's patients [1].

In order to slow down or eliminate the self-replication process that generates the majority of toxic oligomers, we can try to cover the binding sites on the catalytic fibril surface. Recent experiments with Alzheimer's Aβ aggregation have revealed an inhibitory particle, a type of chaperone, that non-selectively binds to the fibril surface and in effect significantly slows down the self-replication process both in vitro and in vivo [13, 15–22]. The molecular mechanism of inhibition, however, has not yet been sufficiently explored and will form the subject of this master's thesis.

The aggregation reaction is by its nature very rapid, heterogeneous, and involves several molecular steps. Just developing an experimental protocol for getting reproducible kinetic data on the time dependence of fibril mass has been a breakthrough in the field of amyloid disease. It is therefore no wonder that the exact molecular mechanisms driving primary spontaneous nucleation, elongation, or self-replication remain hardly accessible to experimental analysis of macroscopic data. This is where computer simulations come of use. Computational models along with theoretical reasoning can provide a window to the behaviour of aggregation reactions on a molecular scale. In silico, one can isolate individual molecular steps of the aggregation reaction and utilise simple models that successfully reproduce and subsequently also interpret experimental findings [23–30]. In order to investigate the inhibition mechanism for thwarting self-replication, we will therefore build upon a computational model that has already been used to characterise the mechanism of self-replication [24] and upgrade the model by incorporating inhibitory particles.

In section 2 we will discuss in detail the course-grained computational model that we use to investigate the secondary nucleation mechanism as well as its inhibition. In 2.1 we discuss the main characteristics of our course-grained model for amyloid aggregation and then specifically focus on the interactions between various particles in 2.1.1, the choice of interaction parameters in 2.1.2, and finally our setup for the Monte Carlo simulation in 2.1.3. We then discuss the results of our simulations in section 3. First, we discuss the nucleation mechanism and behaviour in the absence of inhibitors in section 3.1. To test the code, we attempt and succeed at reproducing some of the results of a previous study [24] in the limit of no inhibitor particles (section 3.1.1). We then perform many additional simulations at previously unexplored external simulated conditions that provide novel insights into the nucleation mechanism (section 3.1.2). Especially, we provide further evidence that the fibril surface coverage by monomeric proteins determines the rate of nucleation. Additionally, we identify the most important structural steps in the nucleation pathway and discuss which of those steps are rate-determining. Next we perform extensive simulations with inhibitors that are capable to bind to the fibril surface in section 3.2. As hypothesized, we find that the surface-bound inhibitors reduce the rate of nucleation. We then look at the binding isotherms of both protein particles and inhibitors and find the binding to be non-ideal (section 3.2.1). Then, by plotting the rate of nucleation against monomer coverage, we interestingly find that inhibitors perturb the nucleation mechanism by crowding the fibril surface (section 3.2.2) and that monomer coverage does not uniquely determine the nucleation rate. We successfully account for this crowding effect by acknowledging that inhibitors reduce the free energy of oligomerisation on the fibril surface which is a crucial step

in the nucleation pathway. We also separately measure the cluster distribution of monomeric proteins and find that inhibitors indeed promote the formation of larger oligomers. Next, in section 3.2.3 we find the average oligomer size on the fibril to be a unique determinant of the nucleation rate, both in inhibitor's presence and its absence on the fibril surface. In section 3.2.4 we successfully develop a theory that explains the relationship between average oligomer size and the nucleation rate by using and reinterpreting concepts from classical nucleation theory. Finally, we develop and employ a theoretical lattice model in section 3.3 that takes into account interactions between particles on the fibril surface and is used to successfully fit the measured binding isotherms (section 3.3.1), and is able to capture and better explain the crowding effect using the concept of surface pressure (section 3.3.2). Lastly, having developed a correspondence between simulation data and the lattice model, we can investigate the inhibitor design that is most effective at thwarting self-replication of protein fibrils (section 3.3.3).

2 Methods

To investigate the inhibition mechanism of surface-catalysed nucleation we employ a course-grained computational model developed by dr. Šarić et al. which has been used to characterise both the spontaneous nucleation mechanism [23, 25], and the self-replication mechanism [24] of the amyloid aggregation reaction. We then extend this model by introducing inhibitory particles that can slow down the self-replication process by attaching to the fibril surface.

Two simulation techniques are widely used in the molecular realm: molecular dynamics and the Monte Carlo method. Molecular dynamics is used to study classical many-body systems by integrating equations of motion that are governed by forces between particles. It probes the time-evolution of the system and can be used to calculate thermal averages as well as transport properties [31]. Monte Carlo, on the other hand, is a stochastic integration technique used to calculate thermodynamic averages by sampling the phase space with discrete random moves. Mostly, all other things being equal, molecular dynamics would be the simulation method of choice. But when we are dealing with discrete degrees of freedom such as Ising spins where we are not interested in the internal dynamics of the transition from one state to another the use of Monte Carlo is necessary. Our course-grained model involves conformational changes between discrete protein states which makes the Monte Carlo method more suited for the task.

2.1 Course-grained model

Our computational model is a course-grained Monte Carlo simulation written in the C programming language. This model has initially been used to capture the creation of amorphous oligomers, spontaneous nucleation of fibrils, and elongation of fibrils [23, 26, 27], and has been further modified to capture nucleation catalysed by the fibril surface [24]. The self-assembly into various morphologies proceeds by action of non-specific anisotropic interactions between various monomeric species. The justification for using such a simplified model comes from experimental evidence that the mode of attraction between unfolded proteins seems to be largely indifferent to the change of amino acid structure of aggregation prone proteins as long as the changed residue has similar properties in terms of hydrophobicity etc. [32, 33]. The main advantage of such a minimal model is that, being sufficiently cheap in terms of computational time, we can explore a wide range of protein concentrations [30].

In its present state the computational model involves a fibril at the center of the simulation box and a solution of monomeric proteins. These monomers are modelled as directional hard spherocylinders with an attractive patch that drives assembly

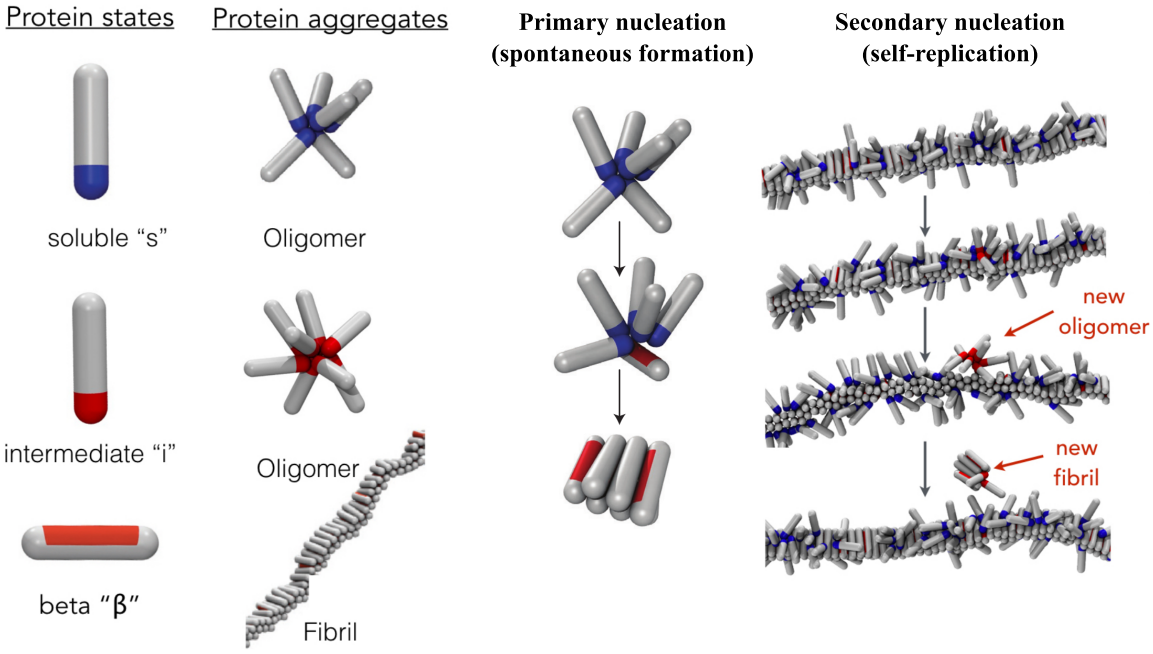

Figure 1. Coarse-grained computer model for amyloid aggregation. Monomers can primarily exist in two states: a soluble state that forms oligomers and a β-sheet state that make up fibrils. When bound to the fibril surface, soluble monomers can also convert into an intermediate state which forms more tightly bound oligomers which have a tendency to detach from the fibril surface. Primary nucleation, at physiologically relevant protein concentrations, proceeds through soluble monomer oligomerisation and subsequent conversion to a β-nucleus that continues to grow in solution by elongation. Secondary nucleation is an auto-catalytic process and proceeds by protein adsorption and oligomerisation on the surface of an existing mature fibril, conversion into an intermediate oligomer, oligomer detachment, and finally conversion into a β-nucleus in a fashion similar to primary nucleation.

(see Fig. 1). Additionally, monomers can take three different conformational states: a soluble 's'-state can form micellar-like oligomers and can bind to the fibril surface, an intermediate 'i'-state is a prefibrilar state that can form more strongly bound oligomers, binds only very weakly to the surface, and has a higher internal free energy, and a β-sheet rich state which has a very high internal free energy and is able to form fibrils. Monomers in this β-state have a very strong mutual side-patch interaction as evidenced by the immense structural integrity of fibril aggregates [34]. Monomeric particles are allowed to swap between these conformations but the conversion from the lower free energy 's'-state to the higher free energy intermediate and β-states is thermodynamically unfavourable as well as kinetically slow. Therefore, these conversions to β-state happen only in large monomer oligomers where per-particle energy becomes low enough to overcome the entropic barrier associated

with protein refolding to a more β-sheet rich state. When two β-state particles form together, we call that a β-nucleus which can only irreversibly grow further by elongation.

Nucleation can happen by way of primary or secondary nucleation. Primary nucleation involves a random density fluctuation where soluble monomers form an energetically favourable but high-entropy oligomer. If this oligomer is sufficiently long-lived it can facilitate nucleation by providing an environment in which a conversion to β-state is energetically favourable. Secondary nucleation involves the concomitant adsorption of many soluble monomers to some binding site to make a fibril-bound surface-oligomer. Again, if sufficiently long-lived, this oligomer can convert into an intermediate oligomer that has negligible surface affinity but is very tightly bound. This 'i'-oligomer can then detach from the fibril surface and further grow in solution and/or convert to a β-nucleus in a manner similar to primary nucleation except that the intermediate oligomer is more thermodynamically stable than its soluble oligomer counterpart. Both nucleation processes are therefore multi-step because they involve at least both oligomerisation and conversion where monomers attach/detach or convert one by one.

In this work, we focus on the surface-mediated secondary nucleation because this process is the source of the majority of toxic oligomers that accompany amyloid self-replication. As remarked previously, a way to slow down or extinguish the secondary nucleation process is by covering the fibril surface with inhibitory particles that can attach to catalytic binding sites and prevent soluble proteins from approaching the surface. We implement this kind of inhibition in our simulations by introducing a new particle, called inhibitor. It is modelled as the same spherocylindrical shape as a monomer and also has an attractive patch at one end with which it can bind to the fibril surface or possibly to other soluble monomers and inhibitors. This has been done in order to preserve simplicity in measuring and comparing monomer and inhibitor coverages which led to simpler analysis of the inhibition effect.

2.1.1 Interactions in the model

In our model, each particle is represented by a spherocylinder with a length to width ratio of $L/\sigma = 3$ which mimics the elementary β-sheet unit in Aβ peptides with dimensions $2\,\mathrm{nm} \times 6\,\mathrm{nm}$. The unit $\sigma = 2\,\mathrm{nm}$ defines the natural distance measure in our model against which every other distance is compared. A hard core repulsion between all particles forbids *any* distance between any two points on spherocylinder centerlines to be smaller than σ.

The attractive potential V_{ss} between two monomers in the soluble 's'-state is

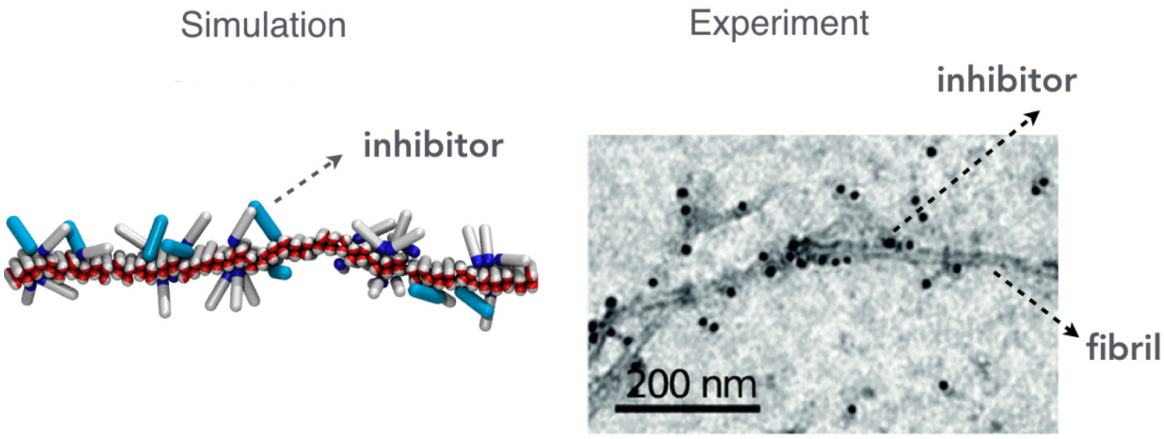

Figure 2. Inhibition of self-replication. Inhibitory particles, depicted in left figure as turquoise spherocylinders, can bind to the fibril surface. Inhibitors compete for the surface binding sites with soluble proteins and therefore slow the rate of secondary nucleation. The right figure shows a transmission electron microscopy image of Brichos chaperones functionalised by nanogold-conjugated antibodies clearly showing binding of the chaperone to Aβ42 fibrils (from ref. [13]).

implemented as:

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

where r is the distance between the centers of the attractive tips located at the spherocylinder's ends, $r_{cut} = 1.3\sigma$ is the cutoff distance, and ϵ_{ss} is the maximal interaction energy between two solubles. An attractive patch is added to only one of the two spherocylinder poles to ensure formation of finite micellar-like oligomers where the tips of soluble monomers are all in contact at the oligomer center (see Fig. 1). The tip-to-tip interaction between other two monomeric particles is implemented in the same way (Eq. (1)) but with different interaction strengths. So we have $\epsilon_{ss} \to \epsilon_{ii}$ and $\epsilon_{ss} \to \epsilon_{si}$ for the interactions between two intermediate monomers and the interaction between a 's'-state monomer and an 'i'-state monomer, respectively, and possibly $\epsilon_{ss} \to \epsilon_{ii}$ and $\epsilon_{ss} \to \epsilon_{sI}$ for the interaction between two inhibitors and between a soluble monomer and an inhibitor, respectively.

The attractive side-patch of the β-state monomer is $L_b = 0.7L$ long and spans an angle of 180°. If patches of two β-state monomers face each other their interaction $V_{\beta\beta}$ is:

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

where $2\epsilon_{\beta\beta}$ is the maximal interaction strength between two β-particles, ϕ is the angle between the long axes of the particles, d is the shortest distance between the axes of the patches, and r is the distance between the centers of the patches. The first term ensures that proteins in the β-form pack parallel to each other, mimicking the hydrogen-bond interactions between β-sheets, while the second term promotes compactness of the fibrils [34–36]. This is also the interaction that holds together the mature fibril at the centre of our simulation box. To make fibrils thermodynamically stable, $\epsilon_{\beta\beta}$ has to be by far the strongest of all the interactions in the system.

The cross-interaction $V_{s\beta}$ between the soluble and the β-sheet configuration is implemented simply as a potential well:

$$V_{s\beta}(d) = \begin{cases} -\epsilon_{s\beta} & \text{if } d < r_{cut} \\ 0 & \text{if } d > r_{cut}, \end{cases} \tag{3}$$

where d is the shortest distance between the centre of the attractive tip and the axis of the β-patch, and $\epsilon_{s\beta}$ is the interaction strength. The soluble tip has to face the 180° opening of the β-particle side patch. The i-β interaction is described in the same way, with $\epsilon_{s\beta} \to \epsilon_{i\beta}$. The interactions described above (Eqs. (1),(2),(3)) are all required to have spontaneous nucleation at physiological conditions.

For secondary nucleation and its inhibition we additionally need binding to the fibril surface. Adsorption of the soluble protein onto the preformed fibril is given by:

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

where d is the shortest distance between the centre of the attractive tip on the soluble protein and the body of the fibril particle which again extends $L_f = 0.7L$ in length. Intermediate 'i' protein and inhibitor adsorptions onto the fibril are described in the same way (Eq. (4)), with $\epsilon_{sf} \to \epsilon_{if}$, and $\epsilon_{sf} \to \epsilon_{If}$, respectively. We note here that the fibril surface is not uniform as it is made of spherocylindrical monomeric particles (see Figs. 1 and 2).

2.1.2 Choice of interaction parameters

As this model contains many different interactions and particle species it can exhibit many different phenomenologies depending on interaction parameters [26]. We choose parameter values that make surface assisted secondary nucleation possible in reasonable computer time and keeps the necessary oligomerisation on the fibril and subsequent oligomer detachment.

Throughout this work we set the internal free energy of the β-state to 20 kT, the internal free energy of the intermediate state to 10 kT, and of the soluble state to zero. The free energy penalty of conversion from a soluble to β-particle is therefore $\Delta\mu_{s\to\beta} = 20\,\mathrm{kT}$ while the conversion penalties to and from the intermediate state are $\Delta\mu_{s\to i} = \Delta\mu_{i\to\beta} = 10\,\mathrm{kT}$. These numbers follow the fact that aggregation-prone peptides such as Aβ are typically not found in the β-sheet configuration in solution [37,38].

We choose a relatively low value of 's'-state binding interaction $\epsilon_{ss} = 4\,\mathrm{kT}$ even though the nucleation rate is faster at larger ϵ_{ss} where oligomers would be larger and more long-lived. This was done in order to hinder spontaneous nucleation in solution whilst still retaining high rates of surface-catalysed nucleation so we do not need to track both phenomena at the same time. As the 'i'-state represents the conformation with more β-sheet content than the soluble state, we set its interaction strength at $\epsilon_{ii} = 16\,\mathrm{kT}$. This high value promotes the conversion of a soluble oligomer to an intermediate oligomer as well as ensures the stability of the detaching 'i'-oligomer. The strength of the interaction between monomers in 'i'-state and 's'-state is set to $\epsilon_{si} = 8\,\mathrm{kT}$ which is somewhere in between ϵ_{ss} and ϵ_{ii}.

The interaction between two β-particles is the strongest by far: $\epsilon_{\beta\beta} = 60\,\mathrm{kT}$. This high number ensures that the β-nucleus is well defined as a thermodynamically stable β-dimer, leaving no ambiguity in the measurements of nucleation rate. The cross-interaction strengths between the soluble or intermediate and the β-state are $\epsilon_{s\beta} = \epsilon_{ss} + 1\,\mathrm{kT}$ and $\epsilon_{i\beta} = \epsilon_{ii} + 1\,\mathrm{kT}$. This means nucleation can happen only in large oligomers because the free energy penalty of converting one particle to β-state is offset by many new hydrophobic contributions. For example, in a detached 'i'-oligomer that contains eight 'i'-state monomers the free energy change of converting one monomer to β-state is approximately $\Delta F_{8i\to 7i1\beta} \approx \Delta\mu_{i\to\beta} - 7(\epsilon_{i\beta} - \epsilon_{ii}) = 3\,\mathrm{kT}$. So such a conversion, if attempted, has about $e^{-3} \approx 5\%$ chance of succeeding.

The interaction strength ϵ_{sf} that promotes adsorption to the fibril surface has also been carefully chosen. If this value is too low (4 kT) there is negligible adsorption to the fibril and no surface-catalysis. But if too high (say 10 kT) there is no detachment of oligomer from the surface and therefore no geometrical rearrangement that allows for conversion to β-state. At certain conditions [24], the maximum rate of nucleation is achieved at $\epsilon_{sf}^{max} = 8\,\mathrm{kT}$. We opt for the lower value $\epsilon_{sf} = 6\,\mathrm{kT}$ in order to avoid possible non-monotonic behaviour of nucleation rate when we change parameters such as soluble chemical potential and inhibitor properties. The interaction between 'i'-state monomers and fibril particles is set to $\epsilon_{if} = 1\,\mathrm{kT}$. This low value promotes the detachment of an 'i'-oligomer from the fibril while hindering the first conversion of a monomer from soluble to intermediate state. To again

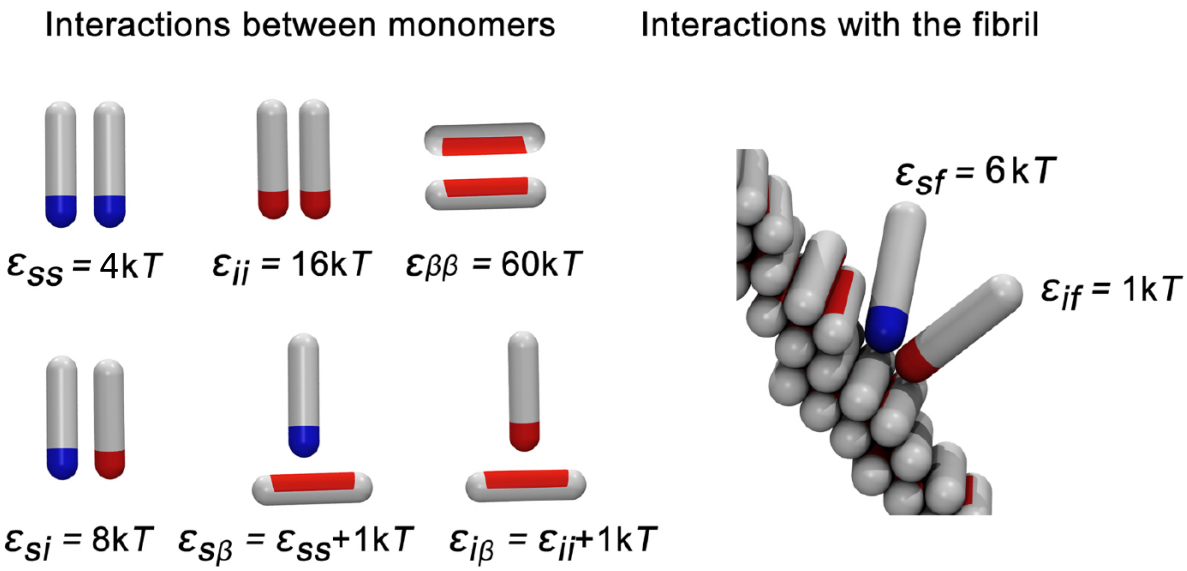

Figure 3. Interparticle interaction strengths. There are in total eight non-zero attractive interactions in the non-inhibited computer model: six between various monomer states (ϵ_{ss}, ϵ_{ii}, $\epsilon_{\beta\beta}$, ϵ_{si}, $\epsilon_{s\beta}$, and $\epsilon_{i\beta}$) where 's' stands for soluble, 'i' for intermediate, and 'β' for β-state, and two between soluble and intermediate monomer states (ϵ_{sf} and ϵ_{if}). The parameter values were chosen such that secondary nucleation on the fibril surface is favoured to spontaneous nucleation in solution and such that only large surface-oligomers have a chance to detach and nucleate.

make a quick back of the envelope calculation, the free energy change to convert one soluble state particle to intermediate state in a surface oligomer of six solubles is $\Delta F_{6s\to 5s1i} \approx \Delta\mu_{s\to i} + 5(\epsilon_{ss} - \epsilon_{si}) + 2(\epsilon_{sf} - \epsilon_{if}) = 0\,\mathrm{kT}$. So, if attempted, there is a fair chance a large oligomer will be able to nucleate.

Finally, we vary the inhibitor binding strength between $\epsilon_{If} = 6{-}8\,\mathrm{kT}$, mimicking the experimental situation where inhibitors have a larger affinity for the fibril surface than unfolded proteins [13]. We nominally set both the inter-inhibitor binding as well as the interaction strength between inhibitors and monomers to zero, that is $\epsilon_{II} = \epsilon_{sI} = 0$. Later, when investigating optimal inhibitor design, we also try non-zero values of these parameters (section 3.3.3) but the majority of this work is done with no cooperative inhibitor binding and without mixed oligomers.

2.1.3 MC simulation

Simulations were performed in a periodic cubic box in a semi-grand canonical ensemble where we kept the volume of the box V, the temperature T, the chemical potential of monomers μ_m, and total number of inhibitors N_I constant. Such a

scheme was chosen to avoid the depletion of monomers and inhibitors from solution due to adsorption onto the fibril surface and to have good control over the number of particles simulated. All simulations start with a simulation box of size $150\sigma \times 150\sigma \times 150\sigma$, a preformed fibril at the center of the box and 600 randomly distributed soluble monomers. The fibril contains 92 tightly bound side-patched particles and is capped at both ends (unable to grow by elongation) as we want to keep the amount of binding sites constant. We also randomly distribute a certain amount of inhibitors $N_I = \chi \cdot 600$, depending on the parameter χ that we input in the program. We then scale the simulation box, making it larger or smaller, to match the specified soluble chemical potential μ_m by making grand-canonical exchange moves that add or removes soluble monomers from anywhere in the simulation box, excluding the $D = 12\sigma$ wide exclusion zone around the preformed fibril. During scaling, we also reinsert inhibitors that find themselves outside the simulation box by placing them randomly inside the box. We make those replacements during scaling and not after because inhibitors in solution interact with monomers via volume exclusion, thus influencing μ_m. We settled for 20 rescaling moves and 60 exchange grand-canonical moves that removed or added soluble monomers according to μ_m and box volume.

We note here that we could take a more elegant, fully grand-canonical route in preparing the simulation by first specifying a fixed box size and then distribute monomers and inhibitors inside the box according to their chemical potentials. This way we would have direct control on the real chemical potentials of both species separately. Unfortunately, the effect of this route is that we would, depending on the values of chemical potentials, sometimes simulate 2000+ particles which would be very expensive in terms of computational time or simulate only 100− particles which could potentially introduce finite-size effects.

After the simulation is initialised we run the 'equilibration simulation' for an indefinite number of Monte Carlo steps. In such a simulation we let the soluble monomers and inhibitors explore the configuration space until the system reaches equilibrium - some solubles and inhibitors adsorb to the fibril, some form oligomers, most remain in solution. At this point we do not allow for possible conversions between monomer states because we first want equilibrium conditions on the fibril surface where the surface and solution chemical potentials of a given species are equated. We find this dynamic equilibrium state by looking at the graphs of the number of soluble monomers and inhibitors adsorbed to the fibril against simulation 'time' (see Fig. 4). When the adsorbed particle numbers are observed to fluctuate around a mean, the system is on average in its lowest free energy state.

After we have reached equilibrium we manually stop the simulation and input

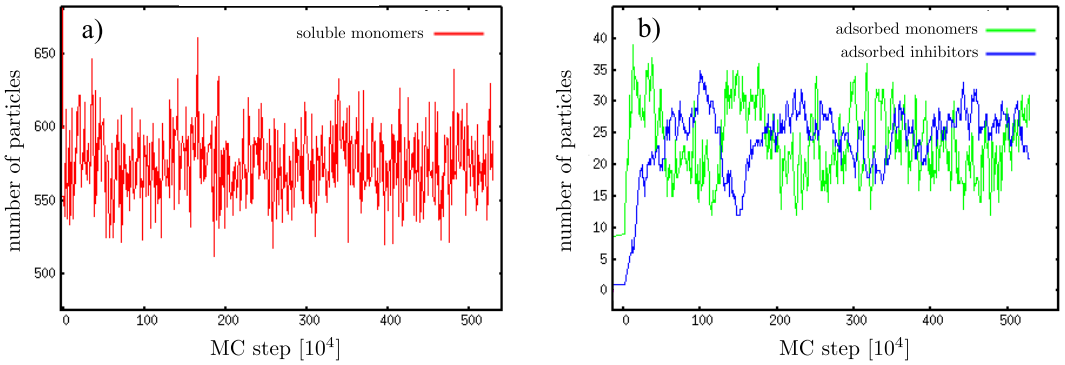

Figure 4. Reaching dynamic equilibrium state. a) the number of soluble monomers in the solution (red) is kept constant on average by grand-canonical exchange moves. b) Equilibration of chemical potentials between solution and fibril surface is achieved when both the number of fibril-bound monomers (green) and inhibitors (blue) start fluctuating around a mean.

its last configuration into a new 'rate simulation'. This simulation first runs for a further 500.000 Monte Carlo steps so we can measure and average the equilibrium properties such as monomer and inhibitor coverages of the fibril surface, monomer oligomer distribution on the surface, monomer and inhibitor number densities in solution, average per-particle binding energies of both species on the fibril, and monomer cooperative binding energies in oligomers.

After the measurement phase we switch on the Monte Carlo routines that govern conversion between different monomer states. During conversion, the position of the particle and the orientation of the spherocylinder's long axis are conserved. If the peptide conversion is between an 's' or 'i'-state to a β-state, we also randomly assign the orientation of the β-sheet side-patch. The probability of attempting a conversion move was set to $p_{swap} = 1/5000$, which mimics the slow conversion of the soluble unfolded protein into a β-sheet prone configuration.

Overall, each Monte Carlo step involves N translational and orientational moves, where N is the total number of particles in the system (inhibitors, monomers in various states, and fibril particles). Additionally, for β-state monomers and fibril particles there is an additional random rotation around the spherocylinder's long axis. The chance of these translational and orientational moves being accepted is governed by the Metropolis algorithm. Also, with a chance of $p_{exchange} = 1/30.000$ in each step, we perform N_s grand-canonical exchange moves, where N_s is the number of soluble monomers outside the exclusion zone. Finally, each step, all monomers are given a chance to change their conformation between the 's', 'i' or β-state. The

simulation terminates when we observe two mutually bound β-particles that define the β-nucleus.

Depending on the values of monomer and inhibitor chemical potentials and inhibitor binding affinity, the initialisation phase takes up to two hours, equilibration phase from one to four days, and nucleation phase from just hours to infinite time but we ran the rate simulations for at most 14 days. The highly stochastic nature of the nucleation rate measurements forces us to make many repetitions for a given set of parameters. Therefore all parameter sweeps were performed on the computational cluster 'Dexter' that contains 480 computing nodes and is located at the Department of Chemistry, University of Cambridge.

3 Results

The results section is structured as follows: first we perform simulations at conditions that have been used in a previous study [24] where no inhibitors are present and confirm that our code is working properly in the limit of no inhibitor particles. We discuss the relevant measured variables such as monomer coverage, oligomer distribution on the fibril, and nucleation rates. Then we introduce inhibitors in our simulations and observe how they influence monomer coverage on the fibril. We find that binding isotherms cannot globally be fitted with the Langmuir competitive binding model because of volume-exclusion interactions between monomers and inhibitors on the fibril. Looking at the rate of nucleation under the influence of inhibitory particles we discover that the reported nucleation mechanism where the rate is simply dependent on monomer coverage [24] fails when inhibitors are present on the fibril due to macromolecular crowding. We therefore look for a more general metric that is able to uniquely describe the rate of secondary nucleation and finding it to be the average oligomer size on the fibril. Having found this unique determinant of the rate we build a theory that uses concepts from the classical nucleation theory. Using our theory, we successfully fit our simulation data. Finally, we develop a statistical-mechanical lattice model that is able to explain the non-ideal binding isotherms, provide some insight into macromolecular crowding, and make useful predictions for effective inhibitor design.

3.1 Absence of inhibitor

Due to implementing a new particle into the existing simulation scheme for studying secondary nucleation, we made quite a few changes to the program written in C. Therefore, we needed to make sure the underlying code for the secondary nucleation remained intact. We did that by reproducing some of the results published in [24]. We performed a parameter sweep across a broad range of monomer chemical potentials $\mu_m \in [0.71, 7.04]\,\mathrm{kT}$ in the absence of inhibitor particles and at exactly the same binding strength parameter values that were used in the previous study (see Fig. 3).

3.1.1 Monomer coverage and the rate

In our simulations, the fibril surface is not uniform but is made up of 92 fibril particles with a Lennard-Jones type interaction (Eq. (4)). Binding sites are therefore not very well defined; in simulations, an adsorbed particle would mostly dock with its tip into a pocket between two fibril particles (see Fig. 3, left) as this is the place with the lowest potential energy. Alternatively, it could sit on top of one fibril particle while

still interacting with the neighbouring two fibril particles. Additionally, depending on the strength of interaction between two adsorbed soluble monomers (ϵ_{ss}), the adsorbed particles may stack upon one another. This can pose a difficulty in defining the monomer coverage but because most of our simulations are done at a constant temperature (T), and constant binding strengths ϵ_{ss} and ϵ_{sf}, we will simply define monomer coverage (θ_m) as:

$$\theta_m = \left.\frac{N_{mon}(\mu_m)}{N_{mon}(\mu_m \to \infty)}\right|_{T,\epsilon_{ss},\epsilon_{sf}}, \tag{5}$$

where N_{mon} is the number of soluble monomers adsorbed to the fibril. We can evaluate the maximum value of adsorbed monomers by fitting the data to the Langmuir isotherm:

$$\theta_m = \frac{c_{mon}/K_m}{1 + c_{mon}/K_m}, \tag{6}$$

where c_{mon} is the measured concentration and K_m is the monomer dissociation constant. We measure the concentration by simply counting the number of soluble monomers in the bulk volume (outside the fibril exclusion zone), dividing by that same volume, and averaging over multiple equilibrium configurations. By comparing the two coverage equations (Eq. (5), (6)) and fitting to measured data (Fig. 5a), we get $N_{mon}(\mu_m \to \infty) = 156 \pm 2$, which is more than twice the number of fibril particles, and $K_m = (4.42 \pm 0.07) \cdot 10^{-3}\,\sigma^{-3}$. To convert from the simulation number density units into SI units, we need to multiply by 0.208 mol/L to get $K_m = (0.92 \pm 0.02)$ mM. We note however, that the exact numbers are not important to our discussion as they crucially depend on our choice of interaction parameters. We will therefore always use simulation units throughout this work and will simply use a shorthand $\sigma^{-3} \equiv M$.

The rate of secondary nucleation is calculated from the average lag time for nucleation $< t_{lag} >$ which is defined as the average number of Monte Carlo steps needed to obtain a β-nucleus. Regardless of the value of the monomer chemical potential, we average over 9 successful realisations of the stochastic process (black circles in Fig. 5) with different random seeds but starting from the same equilibrium configuration. The lag time is expressed in the units of 10^5 MC steps. Using the equivalence between the average first exit time $< t_{lag} >$ across the free energy barrier and the inverse of the associated Kramers rate (r), we have:

$$r = \frac{1}{< t_{lag} >}. \tag{7}$$

Plotting the logarithm of the secondary rate against monomer concentration in

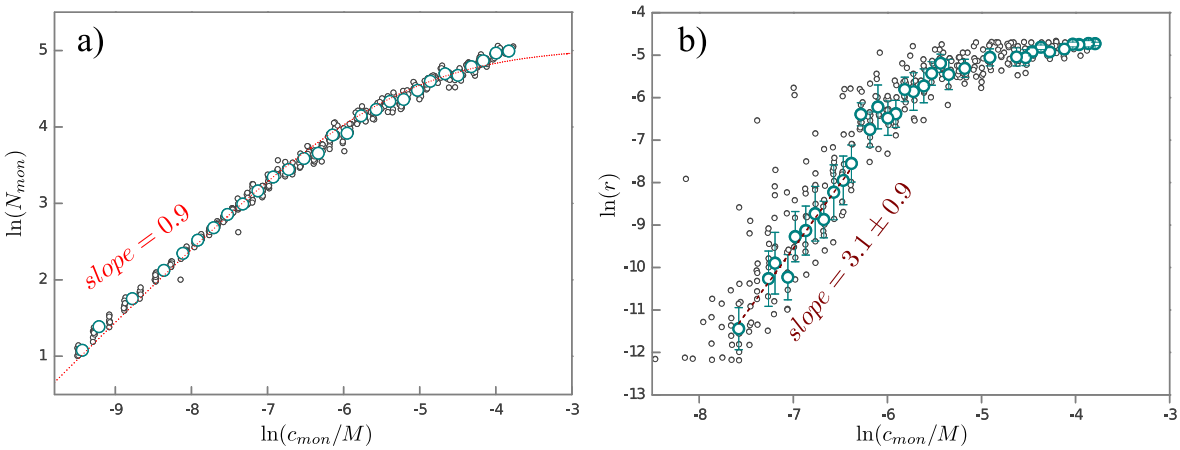

Figure 5. Monomer adsorption and nucleation rate in absence of inhibitor. The left graph (a) shows the variation of the number of monomers adsorbed (N_{mon}) with respect to monomer concentration (c_{mon}) in log scale. It shows a linear regime at low concentrations and a saturation regime at higher concentrations. The graph is fitted with a Langmuir isotherm (6) yielding the maximum number of monomers adsorbed 156 and a dissociation constant $K_m = 4.42\,\mathrm{mM}$. The fit is not perfect because of volume exclusion and cooperative binding interactions on the fibril and mostly serves as a guide to demonstrate the saturation effect. On the right (b), we plot the rate of secondary nucleation (r) with respect to monomer concentration (c_{mon}). This graph also demonstrates a linear and a saturation regime indicating that the rate is crucially dependent on monomer coverage. Each black circle stands for a particular realisation of the stochastic process while cyan circles with error bars represent the ensemble average over nine realisations with the same parameter set but different initial conditions.

Fig. (5b), we immediately see that the rate exhibits saturation at higher concentrations. The scaling exponent $\gamma_s = \frac{\mathrm{d}\ln(r)}{\mathrm{d}\ln(c_{mon})}$ is highly dependent on monomer concentration, ranging from about 3.0 at low concentrations to 0.5 at high concentrations. This indicates that the dominant aggregation mechanism changes with concentration as was experimentally found for Aβ40 amyloid aggregation [9]. Comparing the two graphs in Figure 5, we can see that the change in the scaling factor γ_s follows the trend in the change of fibril coverage. The change of γ_s, therefore, seems to be caused by the surface saturation. This is a plausible conclusion since secondary nucleation ought to be a surface-catalysed phenomena so monomer coverage, not monomer concentration, directly influences the rate of secondary nucleation. This claim has been further confirmed by varying the protein surface affinity at constant monomer concentration, both in simulations (changing ϵ_{sf}) and in experiments (changing the salinity to influence ionic screening) [24].

Yet another striking confirmation that monomer coverage determines the rate

of secondary nucleation comes from studying the variation of rate and monomer coverage with temperature [11, 39, 40]. In these simulations, we keep the monomer chemical potential constant and only vary the temperature which influences both the adsorption to the fibril and the success of conversions. These simulations were performed to study the multi-step nucleation pathways of primary and secondary nucleation in relation to Kramers rate theory. The results have recently been published in [41] but will not be a topic of this thesis due to lack of space. In Fig. 6a we overlay three different datasets on a parametric plot of the rate of secondary nucleation against monomer coverage. One dataset (blue) is the same as in Fig. 5 where we vary the monomer chemical potential at constant temperature $T = 1$, while the other two (green, red) datasets are obtained by varying the temperature at constant chemical potentials $\mu_m(green) = 4\,\mathrm{kT}$ and $\mu_m(red) = 3\,\mathrm{kT}$, corresponding to monomer concentrations $c_{mon}(green) = 2.0\,\mathrm{mM}$ and $c_{mon}(red) = 0.72\,\mathrm{mM}$, respectively (see Fig. 6b).

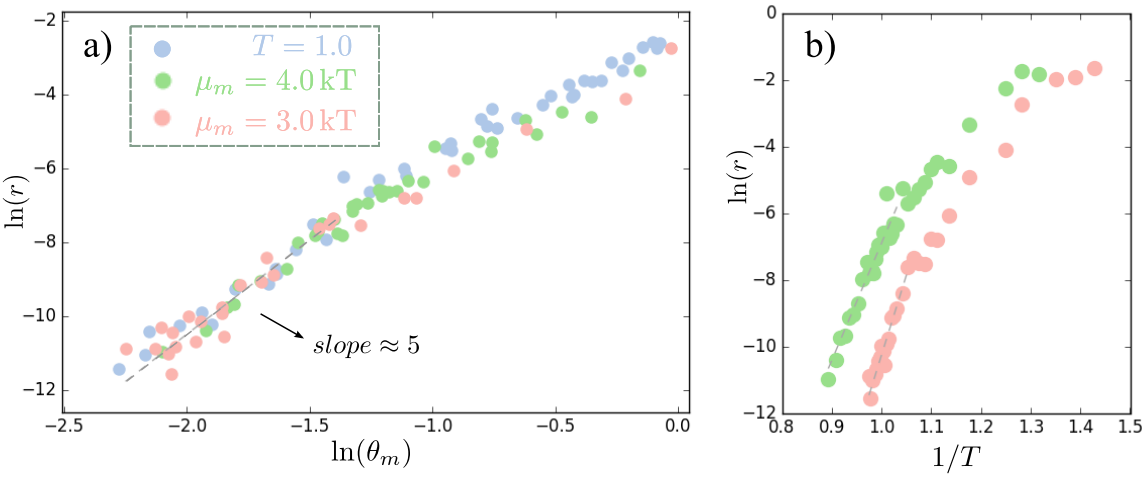

Figure 6. Coverage of the fibril surface determines the nucleation rate. a) parametric plot of secondary rate (r) against monomer coverage (θ_m) for three different parameter sweeps. Blue points are obtained by varying monomer chemical potential at a constant temperature, while green and red are obtained by varying the temperature at constant chemical potential. All three datasets collapse on the same curve indicating that monomer coverage solely determines the rate of self-replication. b) The variation of the rate with inverse temperature for two chemical potentials. The rate increases with lowering the temperature as more solubles adsorb to the surface.

We can see that all three datasets collapse on the same line at low to intermediate monomer coverages and differ only slightly in the saturation regime. We can therefore write with some confidence the relation between monomer coverage and

secondary rate as:

$$r \sim \theta_m^{N^*}, \tag{8}$$

where N^* is the surface reaction order for the self-replication reaction and is expected to be related to the size of a nucleating oligomer at the top of the highest relative free energy barrier.

3.1.2 Nucleation mechanism

Having demonstrated that the rate of nucleation is entirely dependent on fibril surface coverage by soluble monomers, we now turn to explain the molecular steps of the underlying nucleation mechanism. As mentioned earlier, nucleation in our model proceeds through four visually different and easily defined molecular steps (Fig. 1d): oligomerisation on the fibril surface where many solubles form a micelle-like surface-oligomer by adsorbing in the vicinity of each other, conversion of that surface oligomer into an 'i'-oligomer that contains only monomers in the intermediate state, detachment of the 'i'-oligomer and finally conversion to a β-nucleus. Each of these steps contain many smaller steps where monomers attach or convert one-by-one. Needless to say, the free energy landscape for such a nucleation path is very complex. Additionally, nucleation can take many different paths along the multi-dimensional free energy landscape, a problem we will tackle in section 3.2.4.

Nevertheless, we can still determine which of the major steps influences the rate of nucleation most. We do that by measuring the scaling exponents for different steps in the nucleation reaction for different values of the conversion attempt probability (p_{swap}). The result of increasing this swap probability is, in effect, that all conversion free energy barriers are lowered by a constant amount while adsorption and oligomerisation processes remain unaffected. We measured the time lags for several states along the nucleation pathway: the simulation time it takes to observe the first monomer conversion to the intermediate state, the time it takes to observe two mutually bound 'i'-state monomers, three mutually bound 'i'-state monomers, the time of first conversion to β-state, and the time to form a β-nucleus or nucleation time.

By analysing the data and observing the simulation trajectories we find that neither the first nor the second conversion of adsorbed monomer to intermediate state necessitates a nucleation event. Only when at least three monomers are mutually bound in their intermediate state do we always achieve nucleation. This indicates that a surface oligomer with three 'i'-state monomers ('i_3'-oligomer) is already a stable thermodynamic species. The subsequent conversion to a fully intermediate-state oligomer and detachment are even more thermodynamically favourable and

will necessarily happen.

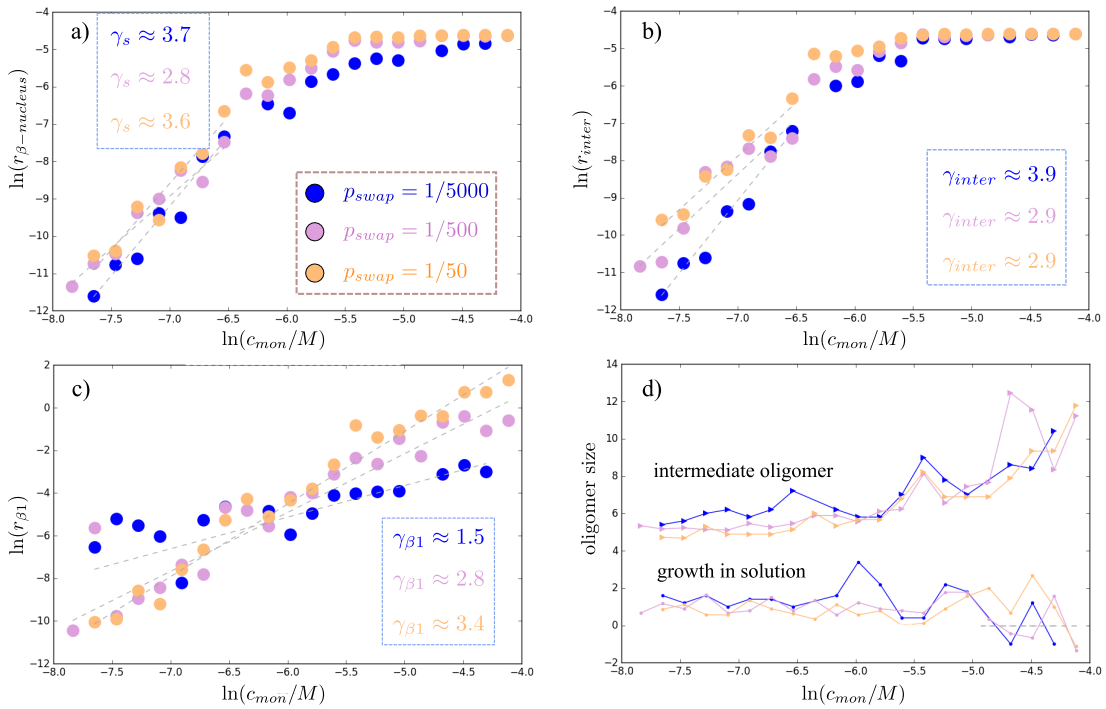

Figure 7. Identifying the rate-limiting step by varying the conversion attempt probability p_{swap}. We plot the rate of nucleation (a), rate of formation of a surface-bound 'i_3'-oligomer (b), and the rate of conversion between a detached 'i'-oligomer and an oligomer that has one monomer transformed to β-state (c). By comparing the scaling exponents (γ_s, γ_{inter}, and $\gamma_{\beta 1}$) we can infer that the rate of nucleation is mostly dominated by the formation of an 'i_3'-oligomer, making it the rate determining step. In (d) we plot the nucleating oligomer size when in 'i_3'-oligomer form and by how much it grows in solution after it detaches. All data is measured at the same parameter values as in Fig. (5) while each point on the graph is the result of averaging over 5 simulation runs.

The next slow step is the conversion of one 'i'-state monomer to a β-sheet particle inside a detached oligomer. This conversion is also unfavourable as already sketched in section 2.1.2. To facilitate easier conversion to β-state, the oligomer grows further in solution by conscripting nearby monomers. In principle, each monomer addition is favourable until we reach a geometrical constraint because only a finite number of monomers (about 12) can take full advantage of micellar interactions. In an enlarged oligomer, conversion to β-state is more available and such a state more long-lived; enough, in fact, to facilitate another conversion 'i'$\rightarrow \beta$. This final conversion to a β-nucleus is very favourable because of the very strong β-β hydrophobic interaction. All subsequent conversions and monomer additions go thermodynamically downhill

from there in a type of irreversible polymerisation reaction.

In short, we have identified three main checkpoints on the nucleation pathway: the surface oligomer that contains three 'i'-state particles, the detached oligomer that contains one β-sheet particle and the final β-nucleus state. We plot the inverse of the average simulation time lapsed between these checkpoints in Fig. 7. It shows the rate of nucleation (a), the rate of creation of an 'i_3'-oligomer (b), and the rate of transformation from that oligomer to a detached oligomer with one β-sheet particle. By comparing these graphs we immediately see how the rate of nucleation follows very closely the rate of 'i_3'-oligomer formation; especially we can appreciate that the saturation of 'i_3'-oligomer formation rate at high monomer concentrations causes the saturation effect in the nucleation rate. But depending on parameter values, the conversion of one particle to β-state also plays a major role. This is especially true for the dataset with $p_{swap} = 1/50$ where the rate of conversion to β-state is slower than the rate of 'i_3'-oligomer formation at low monomer concentration. In this regime, the scaling of nucleation rate with concentration is governed by β-conversion scaling as can be seen by comparing scaling exponents (γ_s and $\gamma_{\beta 1}$ in Fig. 7). Another exception is that at very high surface saturation the free energy barrier for making an intermediate oligomer is nigh non-existent so again the first conversion to β-state becomes the rate limiting step.

To conclude this section, we have successfully replicated the monomer coverage and nucleation rate data from previous work [24] where the authors have found, both in experiments and *in silico*, that the rate of self-replication is governed by monomer coverage of the fibril surface. We have further provided evidence for this assertion by performing simulations where we varied temperature instead of monomer chemical potential. Making a parametric plot of nucleation rate against monomer coverage, all data collapsed on the same line. Finally, we have identified and characterised the major molecular steps along the nucleation pathway. We have found that the formation of a partially converted surface-oligomer is the most important and could be considered a rate-limiting step in our simulations while the conversion from the detached 'i'-oligomer to β-nucleus also requires mounting a free energy barrier and can also be rate-limiting under certain conditions. Incidentally, it has been only very recently demonstrated [42] through a combination of novel experiments and chemical kinetics theory that this latter step is indeed very rare and that only a small number of detached oligomers ultimately form a fibril nucleus. This means that many toxic oligomers are recycled before ultimately converting to a growth competent fibril nucleus. These simulations aided in providing a justification that such a mechanism is indeed possible.

3.2 Influence of inhibitor

We now turn to investigate the influence of surface-bound inhibitory particles on the nucleation pathway. In this section, we report the results of simulations that we performed at various inhibitory conditions: we explored how binding of monomer and the rate of nucleation changes if we add inhibitors at a constant monomer chemical potential or if we keep a constant ratio of the monomer and inhibitor number densities in solution. Additionally, we explored the influence of the strength of inhibitor-fibril binding affinity on the nucleation pathway. The compiled data is the result of over a thousand well-equilibrated simulations that together took well over a 100.000 hours of computer time. We will analyse and explain the results in separate subsections.

We start by simulating a fibril in solution of monomers and inhibitors that both have the same chemical potential ($c_{mon} = c_I$). An inhibitor also binds to the fibril surface with exactly the same binding strength as soluble monomer ($\epsilon_{If} = 6\,\mathrm{kT}$) but has no cooperative binding interaction ($\epsilon_{II} = 0$). We performed a sweep across several chemical potentials of both species and obtain a green curve in Fig. (8a,c) where we plot it alongside the blue curve which is obtained by simulating without inhibitors (same data as in section 3.1). We can immediately see that for all monomer concentrations the green curve lies below the blue curve. This indicates that the presence of inhibitors always reduces monomer coverage and the rate of nucleation.

Next, we performed simulations by keeping the monomer chemical potential constant at ($\mu_m = 4.08$ or $\ln(c_{mon}/M) = -6.28$) while changing the concentration of inhibitors in solution from 0 to two times the concentration of monomers ($c_I/c_{mon} = 2.0$). We did that for three different inhibitor-fibril interaction strengths: for $\epsilon_{sf} = 6\,\mathrm{kT}$, represented by red, for $\epsilon_{sf} = 7.2\,\mathrm{kT}$, plotted in violet, and $\epsilon_{sf} = 8\,\mathrm{kT}$, plotted with the orange colour. Although this difference in binding strength appears only minor, it is amplified by the fact that one inhibitor particle binds to $2-3$ fibril particles at the same time. Looking at Fig. 8b,d, we again see that both the monomer coverage and the rate of secondary nucleation monotonically drop when increasing the concentration of inhibitors in solution. There is a small kink in nucleation rate for very low inhibitor concentrations in red data but this is well within the stochastic error. Inhibitors with higher affinity to the surface better inhibit self-replication at a given inhibitor concentration than those with lower affinity.

Comparing graphs that show monomer coverage with those that show the rate of nucleation, we can again appreciate that the rate seems to be determined by the amount of monomers adsorbed on the surface which has been succinctly expressed by equation (8). The data suggests that the inhibitory effect of inhibitory particles simply stems from their ability to occupy some binding sites on the fibril, making

them inaccessible to monomer particles, and therefore driving down monomer coverage and the associated secondary rate. In short, the mechanism of inhibition seems mainly to be that monomers have to compete with inhibitors for the same binding sites.

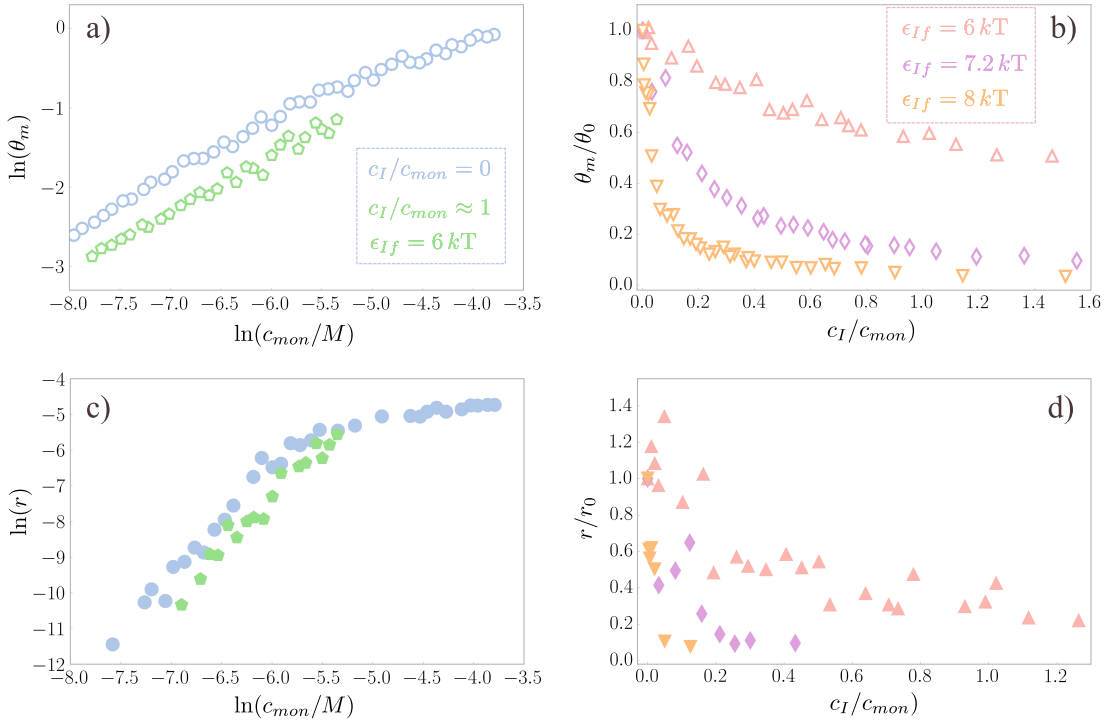

Figure 8. Inhibitors drive down the rate of nucleation. a,b) In the presence of inhibitors, monomer coverage (θ_m) drops for all combinations of inhibitor (c_I) and monomer concentrations (c_{mon}) and all probed inhibitor-fibril binding strengths (ϵ_{sf}). c,d) Following the trend of monomer coverage, the rate of secondary nucleation (r) also significantly decreases by action of inhibitors.

3.2.1 Binding isotherms

Competition between two or more species for the same binding sites is a well known and studied phenomena. We employ here an ideal competitive-Langmuir model to see if it can explain the binding behaviour of monomers and inhibitors at various simulated conditions. The Langmuir adsorption theory describes binding to a lattice with M binding sites, all of which are equivalent, distinguishable, independent, and also do not facilitate interactions between adsorbed particles [43]. If these conditions are met, then, regardless of particular arrangement of lattice sites, the competitive

binding between monomers and inhibitors can be described by the following binding isotherms:

$$\theta_m = \frac{c_{mon}/K_m}{1 + c_{mon}/K_m + c_I/K_I}, \tag{9}$$

and

$$\theta_I = \frac{c_I/K_I}{1 + c_{mon}/K_m + c_I/K_I}, \tag{10}$$

where K_m and K_I are monomer and inhibitor dissociation constants and θ_I is inhibitor coverage of the lattice sites. We have already used a one-species Langmuir curve (Eq. (6)) to successfully fit the monomer adsorption data in Fig. (5). So we have already obtained the value of K_m in that fit as well as the maximal possible number of monomers adsorbed ($N_{mon}(\mu_m \to \infty)$). We normalise all our measured adsorption data with that maximal number, including the inhibitor adsorption data to get the inhibitor coverage as:

$$\theta_I = \frac{N_I}{N_{mon}(\mu_m \to \infty)}, \tag{11}$$

where N_I is the measured number of inhibitors adsorbed on the fibril surface. We plot the normalised monomer and inhibitor coverages in Fig. 9. We make two Langmuir fits for each dataset: the violet line represents a fit to monomer coverage (Eq. (9)) with K_I the only fitted parameter, and the green line shows a fit to inhibitor coverage (Eq. (10)) with again K_I the only fitted parameter. The black curve, on the other hand, shows the Langmuir curve for inhibitor coverage (Eq. (10)) if we use the fitted value K_I, obtained when fitting the violet curve.

Going in order, we can see in Fig. 9a that the fit to monomer coverage is quite good; we get a value of $K_I^{mon}(\epsilon_{If} = 6\,\mathrm{kT}) = (1.3 \pm 0.1)\mathrm{mM}$. Instead, if we try to fit K_I to inhibitor coverage, we get an equally good fit yielding $K_I^{inh}(\epsilon_{If} = 6\,\mathrm{kT}) = (8.3 \pm 0.1)\mathrm{mM}$. There is a very large discrepancy between these two fits (K_I^{mon} and K_I^{inh}) as can be seen by plugging K_I^{mon} in the formula for inhibitor coverage (Eq. (10)); we get a black curve which completely misses all inhibitor coverage data points. So both coverages can be successfully fitted separately but not at the same time with only one shared fitted parameter.

Looking next at Fig. 9b, we can see that while monomer coverage fit works fine, yielding $K_I^{mon}(\epsilon_{If} = 7.2\,\mathrm{kT}) = (0.17 \pm 0.01)\mathrm{mM}$, we can no longer adequately fit the inhibitor coverage data (the closest fit gives $K_I^{inh} = (2.0 \pm 0.1)\mathrm{mM}$). Comparing the fitted K_I's for $\epsilon_{If} = 6\,\mathrm{kT}$ and $\epsilon_{If} = 7.2\,\mathrm{kT}$ we can appreciate that just a small increase in binding strength ($1.2\,\mathrm{kT}$) makes the dissociation constant fall by an order of magnitude. Incidentally, we can successfully perform a fit to inhibitor coverage data if both K_I and K_m are free parameters. So the shape of the curve is still

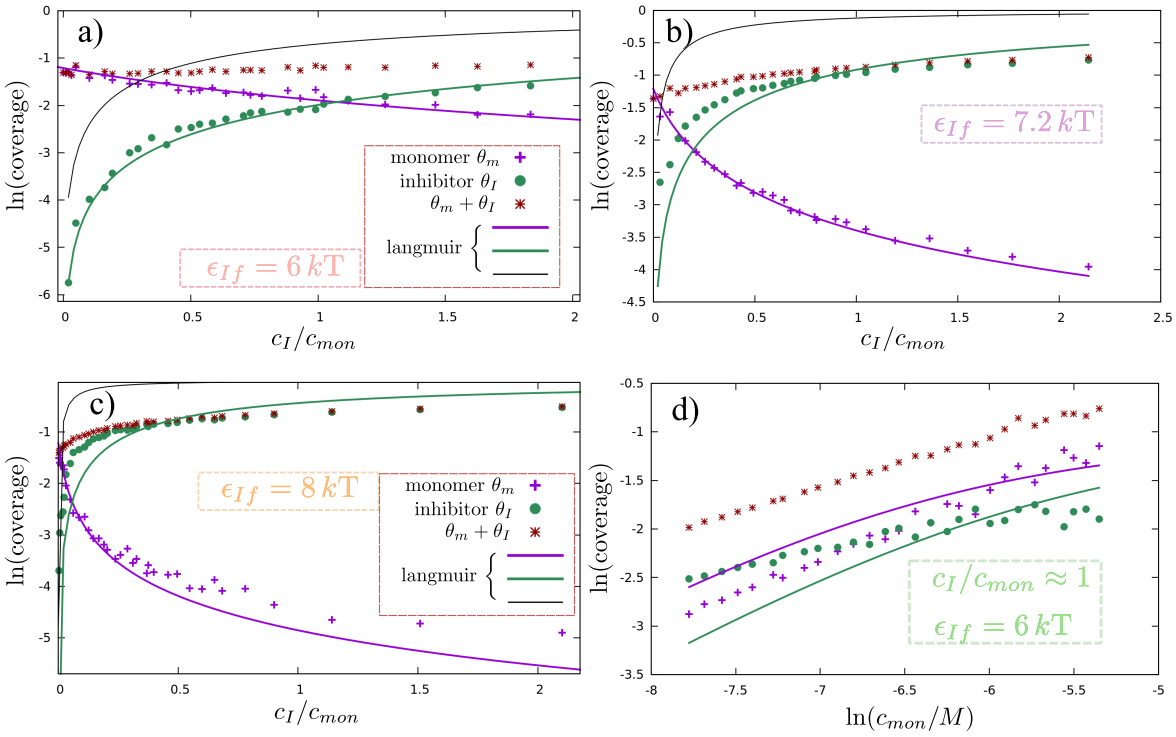

Figure 9. Binding isotherms. Monomer (violet, θ_m) and inhibitor (green, θ_I) coverages for four different datasets. a,b,c) Keeping the monomer concentration constant, at $\ln(c_{mon}/M) = -6.28$, we vary the concentration of inhibitors (c_I) with binding strengths $\epsilon_{If} = 6\,\mathrm{kT}$ (a), $\epsilon_{If} = 7.2\,\mathrm{kT}$ (b), and $\epsilon_{If} = 8\,\mathrm{kT}$ (c). In d) we vary both monomer and inhibitor concentration, keeping the ratio $c_I/c_{mon} = 1$ constant. The violet, green and black lines represent fits to the competitive-Langmuir isotherms (Eq. (9) and Eq. (10)).

approximately Langmuir but the dissociation constants effectively change in that regard.

Next, in Fig. 9c, we see that neither inhibitor nor monomer coverage can be well fitted with a Langmuir curve. The monomer coverage fit gives $K_I^{mon}(\epsilon_{If} = 8\,\mathrm{kT}) = (0.035 \pm 0.003)\mathrm{mM}$ and the inhibitor fit $K_I^{inh} = (0.72 \pm 0.06)\mathrm{mM}$. Separate fits with two loose parameters again work but are very under-determined and therefore do not hold much meaning. Finally, in Fig. 9d, we can see that for our constant concentration ratio data, the Langmuir fits fail completely. What is more, the fits fail even when performed with two free parameters.

We can therefore conclude from Figure 9 that the Langmuir binding isotherms serve as a good starting point - they describe surface saturation behaviour, competition between species, and the dependence of K_I on ϵ_{If} - but are not adequate to explain the binding behaviour. The most obvious reason would be that our simulated

particles interact via volume exclusion and in the case of monomers by tip-to-tip attraction. However, one would expect that at least at very small coverages these non-ideal effects would be suppressed. But importantly, as signified by the crimson coloured points in Fig. 9, the total coverage of the fibril surface is at least 30% for any simulation. This means that the surface is quite crowded with particles that possibly interact with each other via volume exclusion and this is more than enough to introduce non-ideal binding behaviour. We will discuss a model that takes into account interactions on the fibril in section 3.3.

3.2.2 Macromolecular crowding

In the absence of inhibitors, we find that the rate of secondary nucleation is governed by soluble monomer coverage of the fibril surface (see Fig. 6a and Eq. (8)). However, with inhibitors present on the surface, we find a dramatic discrepancy from the expected behaviour, as shown in Figure 10a where we join all datasets on the same graph of rate against monomer coverage. Inhibitors on the fibril indeed decrease the amount of the fibril surface available for monomer binding, but the rate of self-replication does not decrease to the extent predicted by Equation (8). At a given monomer coverage, we find self-replication to be faster in the presence of inhibitors.

Looking at data points (red, violet, orange) where we increase inhibitor concentration at constant monomer concentration $\ln(c_0/M) = -6.28\,\mathrm{mM}$ but for different surface binding affinities, we see that all of these points fall on the same line. This might suggest that inhibitors on the fibril somehow alter the reaction order N^* which is related to the number molecules interacting in the slowest, rate-determining step. One could for example argue that the surface-bound inhibitors block the formation of larger oligomers. However, as shown in the inset of Fig. 10b, the nucleating oligomer properties remain unchanged under the influence of inhibitors. Also, blocking larger oligomers would predict better inhibition compared to simple competitive binding, not worse as seen in Fig. 10a. Additionally, the data where we vary monomer and inhibitor concentrations simultaneously (green) does not show this apparent change in reaction order.

Another possibility would be that by occupying some binding sites, inhibitors effectively increase monomer coverage by reducing the number of available sites as $\theta_m \to \theta_m\, M/(M - N_I) = \theta_m/(1 - \theta_I)$ where M is the number of all fibril binding sites and N_I is the number of adsorbed inhibitors. If we renormalise monomer coverage in that way, we get the graph in Fig. 10c. We can see that the shift in monomer coverage is not sufficient to explain the discrepancy between inhibited and non-inhibited data. Nevertheless, the shift is in the right direction. A possible justification for performing such a transformation monomer coverage would be that

at a given number of monomers adsorbed to the fibril, the presence of inhibitors redistributes adsorbed monomers in a way that we get higher local coverages at some parts of the fibril and lower local coverages on other parts of the fibril surface. This explanation is well supported by the fact that nucleation is a very localised event and the result of a fluctuation in oligomer size.

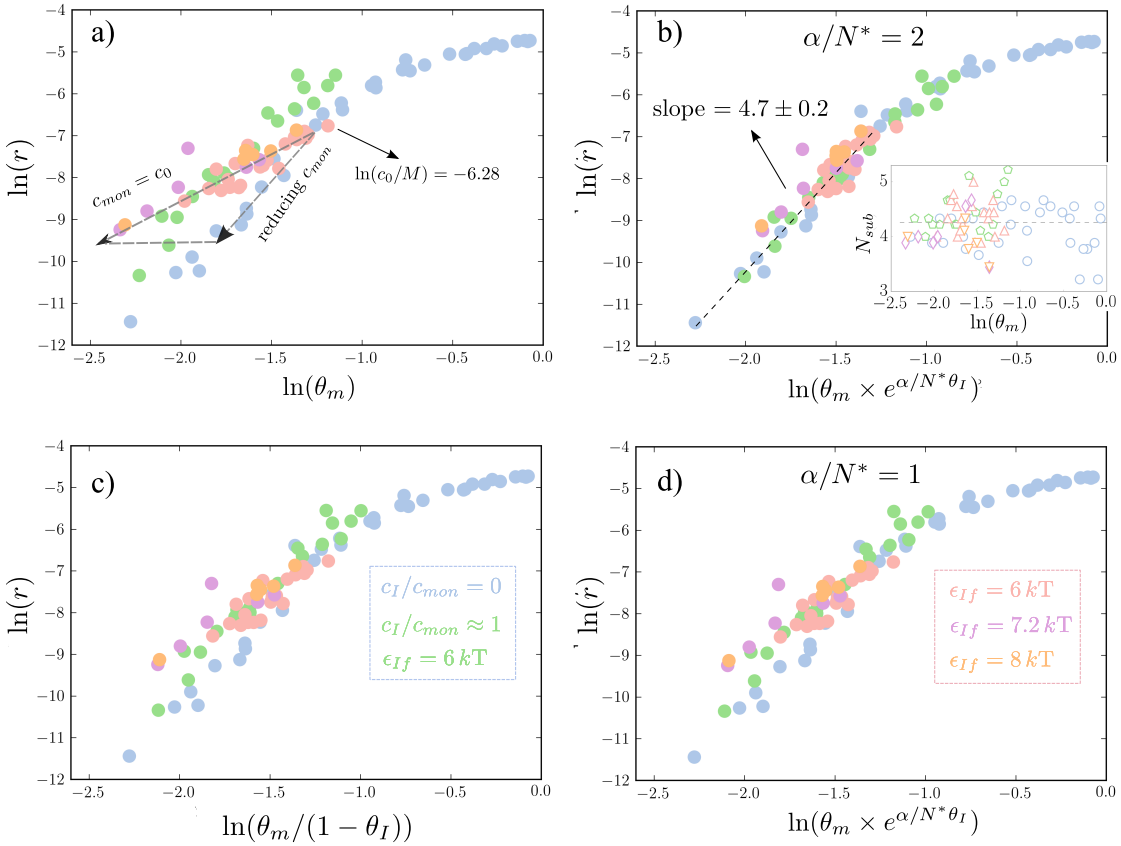

Figure 10. Inhibitors influence the rate of nucleation in a non-trivial way. a) The relation between monomer coverage (θ_m) and the rate of nucleation (r) is perturbed in the presence of inhibitor. b) This perturbation is caused by crowding between species on the surface. If we take volume-exclusion into account, all data collapses on the same curve. Inset: the presence of inhibitors does not influence the subcluster size of nucleating oligomers (N_{sub}). c) One possible but unsatisfactory explanation for branching in a) was that inhibitors redistribute monomers on the surface, resulting in higher local coverages of monomers. d) An example of a misfit, using a smaller value of parameter α.

To investigate how inhibitors influence the distribution of monomers on the fibril, we measured the cluster size distribution for various combinations of monomer and inhibitor chemical potentials. To extract valuable information from these distributions we compare the normalised distributions $n(N)$ at a given monomer coverage

but for different inhibitor coverages and binding affinities. We plot the distribution in Fig. 11 where we can clearly see that the distributions are shifted to larger oligomer sizes in the presence of inhibitors: the blue bar that shows the distribution in inhibitor absence is higher for sole monomers and lower for larger oligomers compared to distributions with inhibitor present. Remembering that a formation of a sufficiently large oligomer which can facilitate conversion to intermediate state is a crucial step in nucleation, we can recognise the shift in oligomer size distribution as the reason for deviating behaviour in Fig. 10a.

This influence on oligomer formation that is exerted by inhibitors on the fibril is a clear manifestation of a macromolecular crowding effect. Inhibitors have a non-negligible volume and due to volume-exclusion repel other proteins on the fibril, pushing them towards each other. As a result, even though the overall rate of self-replication drops due to competitive binding, the decrease in nucleation rate is to a lesser extent than predicted by Eq. (8) which effectively treats monomers and inhibitors as point-particles.

We phenomenologically capture this crowding effect by stating that inhibitors effectively decrease the free energy of oligomerisation for oligomers of all sizes by amount $\Delta\Delta F_{olig}$. By virtue of being surrounded by voluminous inhibitors, an assembled oligomer has to push against neighbouring inhibitors in order to disintegrate back into free monomers. By doing so, it has to perform positive work W_{olig} which has to be supplied by the free energy difference between the assembled oligomer state and a collection of free monomers (or any lower-size oligomer state in between).

We model this work term by making it simply proportional to inhibitor coverage: $W_{olig} \sim \theta_I$; more inhibitors on the fibril at given monomer coverage means more clustering of monomers. We can therefore write the reduction in the free energy of oligomerisation as:

$$\Delta\Delta F_{olig} = -kT\alpha\theta_I, \tag{12}$$

where α is an undetermined proportionality factor.

Oligomerisation precedes all nucleation events so all nucleation pathways can be modelled to be affected in the same way by inhibitors. We can therefore, using Equation (8), write the rate of nucleation (r_I) in the presence of inhibitors and at specified monomer coverage as:

$$r_I \sim \theta_m^{N^*} \cdot e^{-\Delta\Delta F_{olig}(\theta_I)/kT} \tag{13}$$

or, by additionally using the model (12), write the rate law as:

$$r_I \sim \left(\theta_m \cdot e^{\alpha/N^*\theta_I}\right)^{N^*}. \tag{14}$$

We recognise the expression $\theta_m \cdot e^{\alpha/N^*\theta_I}$ as effective monomer coverage of the fibril surface (θ_{eff}) that governs the rate of nucleation both in absence or presence of inhibitors.

Plotting the rate of nucleation against this effective coverage, we can see that for $\alpha = 2N^*$ all data collapses precisely on the same curve (Fig. 10b). Smaller (see Fig. 10d) or larger value of the fitted parameter α results in a misfit so the value of α has to be precise. This result where all data collapses on the same curve is remarkable because it seems to hold for all possible combinations of monomer and inhibitor coverages and surface binding affinities. So we can be quite confident that the rate of secondary nucleation is uniquely determined by both θ_m and θ_I at a given set of interaction parameters.

However, without knowing the meaning of the parameter α, this theory of inhibitor influence on the free energy barrier to form an oligomer is just a phenomenological explanation and its use cannot easily be extrapolated beyond our computer model. One reasonable explanation would be that in the work term (Eq. (12)), $kT\,\theta_I$ could stand for pressure while α could stand for the number of binding sites released when an oligomer disintegrates against the pressure of inhibitors. The best fitted value however, implies $\alpha \approx 10$, a number that is much greater than most oligomer sizes in our simulations.

3.2.3 Average oligomer size

There is another metric that is able to uniquely determine the nucleation rate - the average oligomer size on the fibril surface. We have learned that inhibitors shift the oligomer distribution towards larger oligomer sizes (Fig. 11b) and we know that larger oligomers better facilitate conversions of monomers to other states and eventually nucleation. Remarkably, surface-oligomer sizes follow a simple negative exponential distribution which means that the information about the whole distribution can be captured by a single number - the average oligomer size ($\overline{N}$).

We plot the rate against $\overline{N}$ in Fig. 11 and find that all points follow the same monotonic curve. This is a much stronger result than the one regarding effective monomer coverage because it is general for many different interactions on the fibril between monomers and inhibitors. It should hold as long as inter-protein interactions on the fibril influence only the oligomerisation part but not the conversion part of the nucleation process. We confirmed that by looking at the size of nucleating oligomers and found it unchanged in inhibitor's presence. There was an issue with that claim at first because we were only measuring the size of a β-nucleus which, as discussed in section 3.1.2, is not the most important step in the nucleation process. But when taking into account that a nucleating oligomer grows before finally nucleating, we

arrived at the correct number for the nucleating oligomer size ($\overline{N}_{nucl}$). Using this value $\overline{N}_{nucl}$ and the average oligomer size, we were able to develop a theory for determining the rate of self-replication without any undetermined parameters. We discuss this theory at length in the next section.

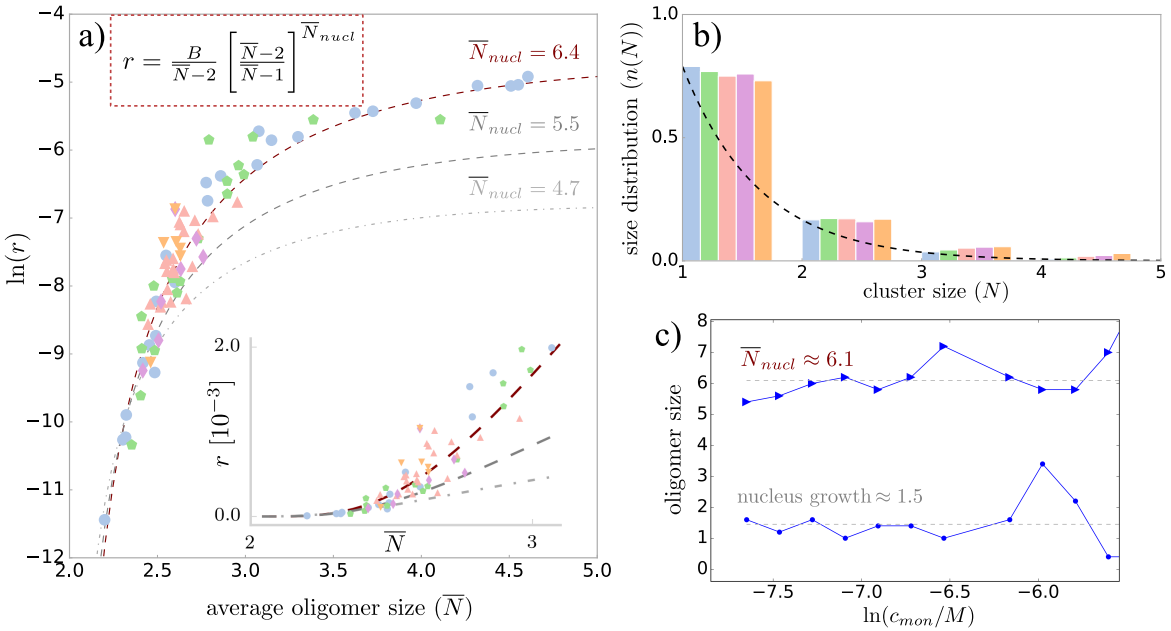

Figure 11. Average oligomer size uniquely determines the rate of self-replication. a) Plotting the rate of nucleation (r) against average oligomer size ($\overline{N}$), all data collapses on the same curve. We can fit a theoretically derived formula where the average oligomer size and the average nucleating oligomer size ($\overline{N}_{nucl}$) determine the rate of nucleation. b) The normalised monomer cluster size distribution ($n(N)$) on the surface at a given monomer coverage but different inhibitor concentrations and binding affinities. The distribution shifts towards larger oligomers with inhibitors present (red, violet, orange, green) compared to the distribution in absence of inhibitor (blue). c) The average nucleating oligomer size ($\overline{N}_{nucl}$) is constant over a range of monomer concentrations. A detached oligomer grows in solution by 1.5 monomers on average before converting to a β-nucleus.

3.2.4 Determining the rate of self-replication

We can explain the dependence of rate on the average cluster size on the fibril surface by introducing a type of two-step nucleation theory that takes into account both the oligomer growth and structural change [25, 44–46]. Nucleation is in general an activated process that accompanies most first-order phase transitions as well as many

self-assembly phenomena. It involves a growing nucleus of a new phase or structure that must overcome a free energy barrier in order to grow into a macroscopic phase at thermodynamic conditions where the emerging phase is stable and the old phase is only metastable [47, 48].

The nucleation process is usually treated by a form of classical nucleation theory (CNT) where the rate (r) is expressed as a product of an exponential factor and a frequency prefactor:

$$r = \rho Z j \exp\big(-\Delta F^*/kT\big). \tag{15}$$

Here ΔF^* is the free energy cost of creating the critical nucleus, ρ is the number density of possible nucleation sites, j the rate or flux with which molecules attach to the growing nucleus, and Z the Zeldovich factor. This factor signifies that the probability of a critical nucleus growing into a macroscopic phase is less than one. Although CNT rarely manages to provide a full quantitative explanation of nucleation phenomena [48], the simple theory in Eq(15) provides a framework for more accurate theoretical treatments.

In our simulations, nucleation proceeds in multiple steps: monomers adsorb and oligomerise on the fibril, then the oligomer converts into an intermediate oligomer, detaches, and then finally transforms into a β-nucleus. This multi-step process evolves along two independent reaction coordinates: oligomer size and oligomer structure. In contrast to the classical picture where a monomer oligomer would need to reach a certain critical size in order to nucleate, oligomers of all sizes can in principle convert to a β-nucleus. There exist several paths along the size/structure landscape and we need to sum over all of them to arrive at the true rate: $r = \sum_{paths} r_i$.

Let us first treat a single nucleation path. An oligomer of size $N = N_i$ forms on the fibril surface by a concomitant adsorption of N_i monomers on some binding site. This surface oligomer then undergoes structural or conformational change until it transforms into a N_i-size β-nucleus. We can therefore treat nucleation as a product of two consecutive activated processes: oligomerisation and conversion. Viewed in the framework of classical nucleation theory (Eq (15)) the oligomerisation part can be treated as oligomer growth towards a critical nucleus whereas conversion can be absorbed in the Zeldovich factor $Z(N_i)$ which again gives the probability that a critical nucleus actually grows into a new phase. This probability is in our simulations given as a product of a conversion attempt probability p_s and a conversion barrier: $Z(N_i) = p_s(N_i)\exp[-\Delta F_{conv}(N_i)/kT]$. Rate of secondary nucleation at a given oligomer size $r(N_i)$ is then given as:

$$r(N_i) = \rho_s\, p_s(N_i)\, j_D\, e^{-\Delta F_{olig}(N_i)/kT}\, e^{-\Delta F_{conv}(N_i)/kT}, \tag{16}$$

where ρ_s is the number density of surface binding sites, j_D is a diffusion governed flux of monomers, $\Delta F_{olig}(N_i)$ is the free energy cost to make an N_i-size surface oligomer out of free monomers and $F_{conv}(N_i)$ is the height of a conversion free energy barrier to make a N_i-size β-nucleus out of N_i-size surface oligomer. Importantly, we note that $F_{conv}(N_i)$ is not dependent on monomer or inhibitor chemical potentials but only on the interconversion dynamics and thermodynamics of a given oligomer.

Nucleation paths can take other forms from the one outlined in equation (16) where nucleation proceeds by first growth and then structural change. For example a possible (and in fact frequent) situation is that a surface oligomer of size $N = 4$ first converts into an intermediate oligomer, detaches, and then grows by one or two monomers in solution before converting to a β-nucleus. This growth in solution, however, seems to be independent of solution chemical potential (Fig. 11c). We can then treat this unorthodox path as one of many possibilities on the conversion free energy landscape of a size 4 oligomer and absorb it into $\Delta F_{conv}(N = 4)$. Therefore, to sum over all nucleation paths we need simply to sum over all oligomer sizes:

$$r \;=\; \rho_s \, j_D \sum_{N=2}^{\infty} p_s(N) \, e^{-[\Delta F_{olig}(N) + \Delta F_{conv}(N)]/kT}. \tag{17}$$

Without loss of generality we can rewrite the above sum in terms of a probability distribution over N by using:

$$n(N) \;=\; A_{olig} \, e^{-\Delta F_{olig}(N)/kT]} \tag{18}$$

$$P_c(N) \;=\; A_{conv} \, p_s(N) \, e^{-\Delta F_{conv}(N)/kT}, \tag{19}$$

where A_{olig} and A_{conv} are normalisation constants for the surface oligomer distribution $n(N)$ and conversion probability distribution $P_c(N)$, respectively. Finally, we write:

$$r \;=\; B \sum_{N=2}^{\infty} n(N) \cdot P_c(N), \tag{20}$$

with $B = \rho_s \, j_D / (A_{olig} \, A_{conv})$.

In simulations, the monomer cluster size distribution on the fibril was found to follow a simple exponential distribution across a range of coverages:

$$n(N) = (e^{1/\lambda} - 1) e^{-N/\lambda}, \tag{21}$$

where λ is a distribution parameter that increases with monomer and also inhibitor coverage. As λ is hard to measure, we instead use the average oligomer size on the

fibril ($\overline{N}$), defined as:

$$\overline{N} = \frac{\sum_{N=2}^{\infty} n(N) \cdot N}{\sum_{N=2}^{\infty} n(N)}, \tag{22}$$

and evaluated as:

$$\overline{N} = 2 + \frac{1}{e^{1/\lambda} - 1}. \tag{23}$$

We can then rewrite eq(21) as

$$n(N) = \frac{1}{\overline{N} - 2} \left(\frac{\overline{N} - 2}{\overline{N} - 1} \right)^{N}. \tag{24}$$

The charm of the following analysis is that we do not need to know the functional form of $P_{conv}(N)$ in order to make progress. We take the total derivative of nucleation rate with respect to $\overline{N}$:

$$\begin{aligned} \frac{\mathrm{d} \ln r}{\mathrm{d}\overline{N}} &= \frac{\mathrm{d} \ln B}{\mathrm{d}\overline{N}} + \frac{\mathrm{d} \ln \sum_{N=2}^{\infty} P_c(N) \cdot n(N)}{\mathrm{d}\overline{N}} \\ &= \frac{\mathrm{d} \ln B}{\mathrm{d}\overline{N}} + \frac{\sum_{N=2}^{\infty} \frac{\mathrm{d}}{\mathrm{d}\overline{N}} (P_c(N) \cdot n(N))}{\sum_{N=2}^{\infty} P_c(N) \cdot n(N)}. \end{aligned} \tag{25}$$

We now use the fact that the conversion dynamics are not dependent on the monomer and inhibitor coverages. Formally, we would write $P_c = P_c(N, \vec{\epsilon}, \vec{\mu}_{swap}, \vec{p}_{swap})$, and $\overline{N} = \overline{N}(\mu_m, \mu_I, \epsilon_{sf}, \epsilon_{ss})$, where $\vec{\epsilon}$, $\vec{\mu}_{swap}$, and $\vec{p}_{swap}$ stand for a collection of interaction energies, conformation entropy penalties, and swap probabilities between all possible monomer states, respectively. The only overlapping variables between P_c and $\overline{N}$ are ϵ_{ss} and ϵ_{sf} so we have:

$$\frac{\mathrm{d}P_c}{\mathrm{d}\overline{N}} = \frac{\partial P_c}{\partial \epsilon_{ss}} \frac{\partial \epsilon_{ss}}{\partial \overline{N}} + \frac{\partial P_c}{\partial \epsilon_{sf}} \frac{\partial \epsilon_{sf}}{\partial \overline{N}} = 0, \tag{26}$$

because we keep both ϵ_{ss} and ϵ_{sf} constant throughout our simulation runs. By using Eq. (24) and Eq. (26) we can now evaluate the expression (25) as:

$$\begin{aligned} \frac{\mathrm{d} \ln r}{\mathrm{d}\overline{N}} &= \frac{\mathrm{d} \ln B}{\mathrm{d}\overline{N}} - \frac{1}{\overline{N} - 2} + \frac{1}{(\overline{N} - 2)(\overline{N} - 1)} \frac{\sum_{N=2}^{\infty} N\, P_c(N) \cdot n(N)}{\sum_{N=2}^{\infty} P_c(N) \cdot n(N)} \\ &= \frac{\mathrm{d} \ln B}{\mathrm{d}\overline{N}} - \frac{1}{\overline{N} - 2} + \frac{\overline{N}_{nucl}}{(\overline{N} - 2)(\overline{N} - 1)}, \end{aligned} \tag{27}$$

where $\overline{N}_{nucl}$ is the average nucleating oligomer size. In general, this average nucleus size is weakly dependent on $\overline{N}$ but in a presaturation regime we find it constant in our simulations (Fig 11c). We can therefore integrate equation(27) over $\overline{N}$ and get

the rate of nucleation as:

$$r = \frac{B}{\overline{N} - 2} \left[\frac{\overline{N} - 2}{\overline{N} - 1} \right]^{\overline{N}_{nucl}}. \tag{28}$$

In Fig. 11a, we see that this formula makes for a nearly perfect fit to simulation data with B the only fitted parameter. The rate (r), average oligomer size on the fibril ($\overline{N}$) and the average nucleating oligomer size ($\overline{N}_{nucl}$) are all measured separately. Additionally, the fit is quite sensitive to the value of $\overline{N}_{nucl}$ as outlined by grey dashed lines in Fig. 11a, making the correspondence between fitted and measured $\overline{N}_{nucl}$ more convincing and valuable.

To recap, we have found that even though our simulated nucleation process involves several molecular steps and so forms a very complex and multi-dimensional free energy landscape, we could still predict the self-replication rate with a slight reinterpretation of the classical nucleation theory. We now turn to the question of whether crowding on the fibril can ever overcome the inhibitory effect of competing inhibitors, and ultimately whether we can design inhibitors that are even better at suppressing self-replication than what we have simulated with our course-grained model.

3.3 Theoretical lattice model

To investigate whether it is possible that putting inhibitors in solution can speed up the rate of nucleation we utilise a simple lattice model that captures the interactions between monomers and inhibitors on the fibril. Because of hard-core repulsions and tip-to-tip attraction between monomers, a simple Langmuir model proved insufficient to capture the binding of monomers and inhibitors to the fibril. Additionally, the Langmuir picture cannot account for the crowding effect. Therefore we utilise a more general albeit still analytical lattice model.

The fibril provides M equivalent and independent binding sites. Each site is able to accommodate at most two particles. Particles that are bound to the same site can interact while interactions between separate binding sites are not treated. In general any combination of monomers and inhibitors can adsorb to a binding site, that is: $s_m = 0, 1, 2$, $s_I = 0, 1, 2$ subject to a constraint $s_m + s_I <= 2$, where s_m is the number of monomers bound on a specific site, and s_I is the number of adsorbed inhibitors bound to that same site. If $s_m = 2$ we call that a monomer dimer.
The grand partition function for the lattice is $\Xi = \xi^M$, where by ξ we denote a

per-site grand partition function:

$$\xi = \sum_{s_m=0}^{2} \sum_{s_I=0}^{2} q(s_m, s_I) \cdot \left[e^{\mu_m/kT}\right]^{s_m} \left[e^{\mu_I/kT}\right]^{s_I}, \tag{29}$$

where μ_m, and μ_I are monomer and inhibitor chemical potentials and $q(s_m, s_I)$ is a site partition function that is dependent on the specific occupancy of both species. In our simple model, the co-occupancy of a binding site by two particles only shifts the energy levels inside a particle partition function by a constant energy term that is given by w_{mm} for the interaction between two monomers, by w_{II} for two inhibitors, and w_{mI} for the interaction between a monomer and an inhibitor on a binding site. The possible lattice site partition functions are then given as: $q(0,0) = 1$, $q(1,0) = 2\,q_m$, $q(0,1) = 2\,q_I$, $q(2,0) = q_m^2 \exp[-w_{mm}/kT]$, $q(0,2) = q_I^2 \exp[-w_{II}/kT]$, and $q(1,1) = 2\,q_m\,q_I \exp[-w_{mI}/kT]$, where q_m, and q_I are monomer and inhibitor binding partition functions. Writing out the sum in Eq. (29), we get:

$$\xi = 1 + 2x_m + 2x_I + x_m^2 e^{-w_{mm}/kT} + x_I^2 e^{-w_{II}/kT} + 2x_m x_I e^{-w_{mI}/kT}, \tag{30}$$

where we have introduced new shorthand variables $x_m = q_m\, e^{\mu_m/kT}$, and $x_I = q_I\, e^{\mu_I/kT}$.

We can now extract all relevant variables from the knowledge of our grand partition function. Average monomer occupancy of a binding site is:

$$\overline{s_m} = \frac{\overline{N_m}}{M} = \frac{1}{M} kT \frac{\partial \ln \Xi}{\partial \mu_m} = x_m \frac{\partial \ln \xi}{\partial x_m}. \tag{31}$$

This number ($\overline{s_m} \in [0, 2)$) divided by two gets us monomer coverage $\theta_m \in [0, 1)$, so we have:

$$\theta_m = \frac{1}{2}\, \frac{2x_m + 2x_m^2 e^{-w_{mm}/kT} + 2x_m x_I e^{-w_{mI}/kT}}{\xi}. \tag{32}$$

Similarly, we get for the inhibitor coverage:

$$\theta_I = \frac{1}{2}\, \frac{2x_I + 2x_I^2 e^{-w_{II}/kT} + 2x_m x_I e^{-w_{mI}/kT}}{\xi}. \tag{33}$$

In this theoretical model, the rate of secondary nucleation is captured by the probability of dimerisation P_2. This is the probability of two monomers occupying the same binding site and is given by:

$$P_2 = \frac{x_m^2 e^{-w_{mm}/kT}}{\xi}. \tag{34}$$

We can use this probability of dimerisation as a measure for the average cluster size on the lattice which, as shown in Fig. 11, uniquely determines the rate of nucleation in our simulations. So we only need P_2 in order to find how nucleation behaves under different values of inter-protein interactions.

We now by $\xi_0 = 1 + 2x_m + x_m^2 e^{-w_{mm}/kT}$ and $P_2^0 = x_m^2 e^{-w_{mm}/kT}/\xi_0$ designate the per-site grand-partition function and dimer probability in the case of no inhibitors: $x_I = 0$. The most important result of this lattice model is that the probability of dimerisation will always be smaller if we introduce inhibitors in the solution. For $x_I > 0$ it always holds that $\xi > \xi_0$ and therefore $P_2 < P_2^0$. This indicates that having particles that compete for the same binding sites on the fibril surface always drives down the rate of nucleation. No matter the inhibitor design, the effect of competition will always exceed the possible effect of crowding or other oligomer growth-inducing interactions on the fibril.

3.3.1 Correspondence to simulation data

We learn the meaning of variables x_m and x_I by taking the limit $w_{mm}, w_{II}, w_{mI} \to 0$ of monomer coverage (Eq. (32)) and compare it to the Langmuir competitive binding isotherm θ_m^L:

$$\theta_m(\vec{w} \to 0) = \frac{x_m}{1 + x_m + x_I}, \qquad \theta_m^L = \frac{c_{mon}/K_m}{1 + c_{mon}/K_m + c_I/K_I}. \tag{35}$$

We find that x_m and x_I are simply monomer and inhibitor concentrations, reduced with respect to their dissociation constants K_m and K_I. By comparing the binding curves (Eqs. (32) and (33)) with simulation data, we can estimate the values of effective binding energies w_{mm}, w_{Im} and w_{II}.

In simulations particles cannot overlap (hard-core repulsion) and monomers have a favourable tip-to-tip interaction of $\epsilon_{ss} = 4kT$ strength. We can therefore expect $w_{mI} = w_{II}$, and $w_{mm} < w_{mI}$. We first perform a fit of the inhibitor adsorption data at three different surface binding affinities (governed by ϵ_{If}). This way we vary only K_I by keeping w_{II} necessarily constant across the three datasets. We get a positive effective interaction $w_{II} = (2.3 \pm 0.2)kT$ between inhibitors on the fibril because of repulsive interactions. We then fit the monomer adsorption data for the case of no attractive interactions between monomers ($\epsilon_{ss} = 0$) by setting $w_{mm0} = w_{II} = 2.3\,kT$ to get $K_m = (0.22 \pm 0.01)\,\mathrm{mM}$. Finally, we fit the monomer adsorption data with the usual cooperative binding interaction $\epsilon_{ss} = 4kT$ using previously fitted $K_m = 0.22\,\mathrm{mM}$ to find $w_{mm} = (0.5 \pm 0.3)kT$.

We see that our lattice model where only two particles can bind on the same binding site maps quite well to simulation data in terms of coverage. In principle,

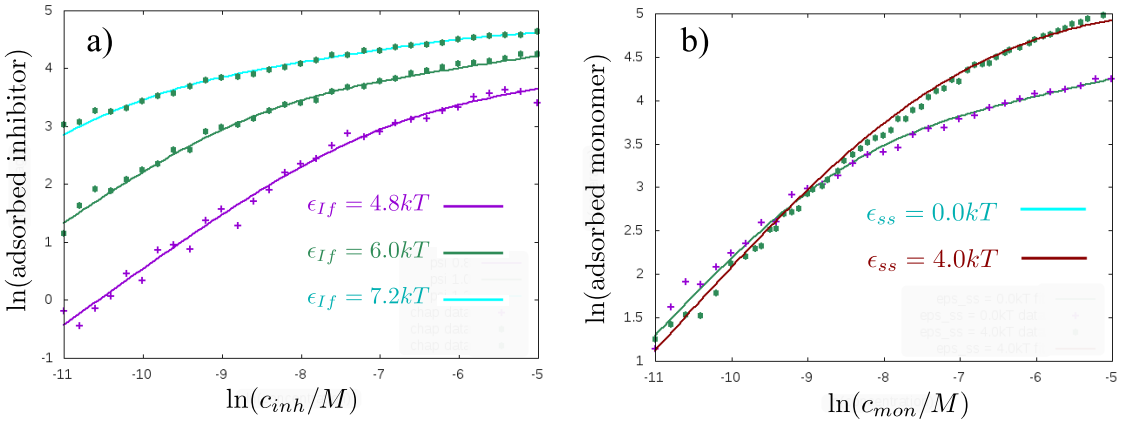

Figure 12. The binding isotherms are well described by the lattice model. a) We perform a sweep against a range of inhibitor concentrations (c_{inh}) in the absence of monomer for three different surface interaction strengths (ϵ_{If}). We globally fit these curves with equation (33) at $x_m = 0$ to find the effective interaction between hard particles as $w_{II} = (2.3 \pm 0.2)$kT. b) We also fit the monomer adsorption data for $\epsilon_{ss} = 4\,$kT with equation (32) at $x_I = 0$ and find the effective interaction between cooperatively binding monomers as $w_{mm} = (0.5 \pm 0.3)$kT.

we could extend the lattice model to include trimers, tetramers and even larger size oligomers but we find those extensions unwieldy and unnecessary because already dimers are influenced by the crowding effect.

3.3.2 Crowding and surface pressure

The main motivation for developing this lattice model was the observation of a crowding effect in simulations, meaning that in addition to depleting the fibril surface by virtue of competitive binding, inhibitors at the same time accelerate nucleation by stabilising monomer oligomers on the fibril. This is also what we observe with this lattice model. When there is repulsion between inhibitors and monomers, we get an increase in the probability of dimerisation at a given monomer coverage when inhibitors are present on the fibril (Fig 13a).

In simulations, we explained the rise of nucleation rate at a given coverage by positing that inhibitors reduce the free energy barrier of oligomer formation. This reduction is equal to the excess work that has to be provided for an oligomer to disintegrate. We modelled this excess work as proportional to inhibitor coverage as

$\Delta\Delta F(N) = -kT\alpha\theta_I$, where α was an undetermined factor. Using the lattice model, we can model this crowding effect more precisely by evaluating the excess pressure exerted by bound inhibitors on monomers that are bound into a dimer.

The combined surface pressure of monomers and inhibitors on a lattice is given by:

$$\phi = \frac{\partial(kT \ln \Xi)}{\partial M} \; = \; kT \ln \xi. \tag{36}$$

We partition this whole surface pressure into a pure monomer $\phi_0 = kT \ln \xi_0$ and an excess part ϕ_e:

$$\begin{aligned} \phi &= kT \ln(\xi_0 * \xi/\xi_0) \\ &= kT \ln \xi_0 + kT \ln(\xi/\xi_0) \\ &\equiv \phi_0 + \phi_e. \end{aligned} \tag{37}$$

For low inhibitor concentration we have $\ln \xi/\xi_0 = \overline{s_I} + \mathcal{O}(x_I^3)$, or

$$\phi_e \approx 2\, kT\, \theta_I. \tag{38}$$

By doing the same effective coverage transformation in the lattice model as in simulations, we write:

$$\theta_{eff}^L = \theta_m e^{a\, \phi_e / N^-}, \tag{39}$$

where $N^- = (\partial \ln P_2^0 / \partial \ln \theta_0)$ is the slope of the blue curve in Fig. 13, and a is a dimensionless factor. By setting $a = 1$, and using $\phi_c = kT \ln \xi/\xi_0$, all points on Fig. 13b collapse on the same line. This collapse is almost exact if we use the $\phi_c = kT \ln \xi/\xi_0$ in the effective coverage expression but holds only at low inhibitor coverages if we approximate $\phi_c \approx 2kT\theta_I$ as seen in Fig. 13c and Fig. 13d.

So remarkably, we have shown that the expression we used for the effect of crowding on the free energy of oligomerisation ($\Delta\Delta F_{olig}$)) has a sound footing in the lattice model where we have been able to analytically arrive at the expression for the excess pressure that inhibitors impose on the surface. We should note, however, that this pressure is not a force between particles in the literal sense but is an effective force that comes from the influence of monomer-inhibitor interactions on the distribution of particles on the lattice. We are unsure at this time whether crowding in our simulations works in a similar implicit way. Intuitively, we might think that inhibitors physically exert lateral pressure along the fibril surface and thus keep oligomers from breaking apart. But looking at simulation trajectories lateral movement along the surface seems rare as most redistribution happens by desorption-adsorption diffusion.

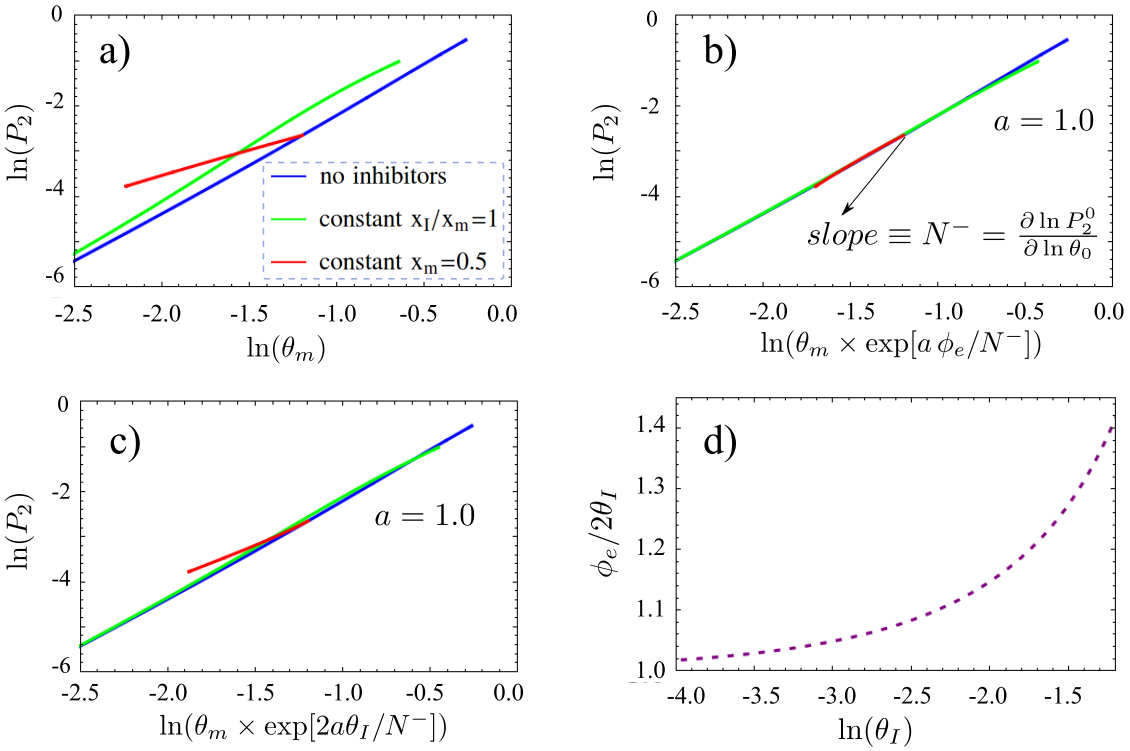

Figure 13. Crowding in the lattice model. a) Emulating the conditions of our course-grained simulations, we arrive at the same effect of crowding where bound inhibitors increase the rate of dimerisation (P_2) at a given monomer coverage (θ_m). b) Taking inhibitor surface pressure (ϕ_e) into account, all data collapses on the same line for $a = 1$. c) This collapse is still good if we use $\phi_e \approx 2kT\theta_I$ but this approximation (d) gets increasingly worse for higher inhibitor coverages (θ_I).

But we should note that this model of crowding on a lattice is still largely phenomenological because we still do not really know the meaning of the factor a. Following the picture that dimers perform work on surface-bound inhibitors, we would expect a to present the number of binding sites (or lattice volume) released when a dimer expands against the pressure of surrounding inhibitors so we expect $a \leq 2$ which is well supported by our fit ($a = 1$). But in the case of no inter-particle interactions ($w_{mI}, w_{II} = 0$) for example, we find that a has to be zero in order for the effective coverage expression (Eq. (39)) to work. Also, the collapse to the same curve in Fig. 13b is very good, but not exact, suggesting that our model of crowding might need additional enhancements in the future.

3.3.3 Effective inhibitor design

As the lattice model maps quite well to simulation data, we can use the analytical model to explore parts of parameter space that was not covered in simulations due to prohibitive computational costs and use it to make some speculations about the most

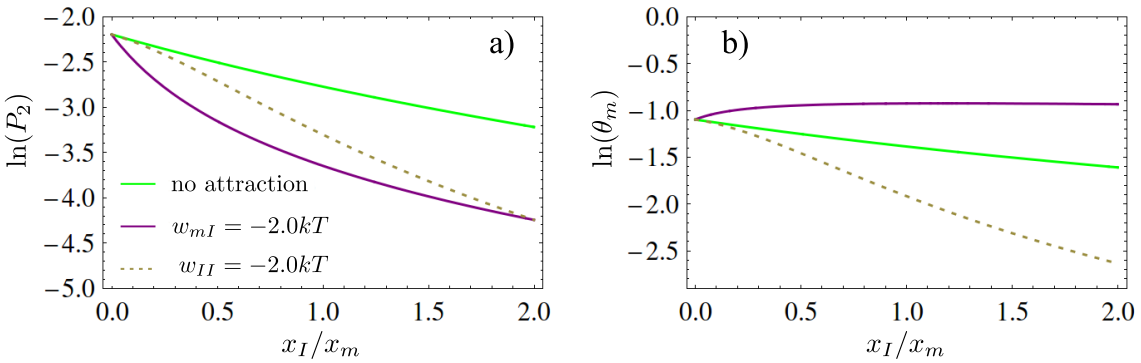

Figure 14. Effective inhibitor design. a) Adding attractive interactions with inhibitors (violet stands for attraction to monomer: $w_{mI} = -2.0kT$ and dotted for attraction between inhibitors $w_{II} = -2.0kT$) lowers the probability of dimerisation (P_2) in comparison to our basic simulation conditions (green, $w_{mI} = w_{II} = 2.3kT$). b) Monomer coverage (θ_m) actually increases when adding inhibitors to the solution if we have monomer-inhibitor attraction (violet) while it is reduced by increasing inhibitor-inhibitor attraction (dotted). Even by increasing monomer coverage, the attraction to monomer shows a greater inhibitory power than inter-inhibitor attraction due to entropy. Graphs are drawn for $x_m = 0.5$ and $w_{mm} = 0.5kT$.

effective inhibitor design. In simulations, we only have a hard-core repulsion between inhibitors and monomers $w_{Im} = w_{II} = 2.3kT$. If we also add some attractive interactions the lattice model suggests that we further decrease the probability of dimerisation (P_2) at a given set of monomer and inhibitor chemical potentials. This can be seen with expression (Eq. (34)) where lowering both w_{Im}, and w_{II} increases the denominator while keeping the numerator constant.

This result, that an attraction between inhibitors enhances inhibition, is perhaps trivial to understand because this attraction promotes more adsorption of inhibitors to the surface at a given x_I as well as promotes inhibitor dimers that then compete with monomer dimers. But analytically, we find that increasing the monomer-inhibitor attraction has an even stronger inhibitory effect. The purple line in Figure 14 ($w_{mI} = -2.0kT, w_{II} = 2.3kT$) shows that even though monomer coverage increases with inhibitor concentration (Fig. 14b) due to cooperation between adsorbed monomers and inhibitors, the amount of monomer dimers (Fig. 14a) drops even more than for the case of cooperative inhibitor binding (dotted line, $w_{mI} = 2.3kT, w_{II} = -2.0kT$) and definitely more than for the case of no attractive interactions with inhibitors (green, $w_{mI} = w_{II} = 2.3kT$).

The reason why inter-species attraction provides stronger inhibition than the attraction between inhibitors is that a state where one monomer and one inhibitor are bound to the same binding site is entropically more favourable compared to an inhibitor dimer state as or a monomer dimer state (two times as many micro-

states). The hierarchy between different modes of inhibition: crowding because of repulsions between inhibitors and monomers, just competition without interactions, inhibitor cooperative binding, and an attraction between monomers and inhibitors that limits dimer formation, is preserved across a broad parameter space of different values of x_m, x_I and w_{mm}. But for binding sites that can accommodate more than two particles and in different geometrical arrangements the effect of inter-species binding might become less trivial.

To explore this caveat, we performed new simulations with inhibitors that posses a capacity to bind to each other with the interaction strength $\epsilon_{II} = 4\,\mathrm{kT}$ (yellow in Fig. 15) and with inhibitors that can bind tip-to-tip with monomers with the same interaction strength $\epsilon_{sI} = 4\,\mathrm{kT}$ (dark grey in Fig. 15). Comparing with the red data points in Fig. 15a new simulation data confirm that increasing attraction between species makes inhibition of secondary nucleation better, but only very slightly (Fig. 15a).

We do, however, clearly see that increasing ϵ_{II} lowers the monomer coverage at given value of inhibitor concentration compared to the case $\epsilon_{II} = 0$ and that increasing ϵ_{sI} makes monomer coverage higher (inset of Fig. 15a). So on a graph of rate against monomer coverage (Fig. 15b), the grey data that stands for monomer-inhibitor attraction actually falls below the blue line, indicating that at a given monomer coverage, inhibitors effectively break oligomers apart. But the inhibitory power should be gauged by how much the nucleation rate drops at given inhibitor concentration, not how it lowers the nucleation rate at a given monomer coverage because experimentally or therapeutically the inhibitor concentration is usually the controlled variable. So overall, in simulations, it is best to design inhibitors with strong fibril surface affinity but it is even better if inhibitors are capable of cooperative binding to other surface-bound inhibitors or monomers.

Unfortunately, due to loss of access to computing clusters we were not able to simulate cooperative inhibitor binders for a larger interaction strength which would tell us if the predictions of the lattice model regarding the stronger inhibition in the case of binding to monomers are correct. What we can do in the future is to analyse different inhibitor designs on a simulated lattice model of various geometries and dimensions. If those higher-order lattice models replicate our model with only two binding sites they could provide design principles that are rigorous and invariant for different distributions of fibril binding sites.

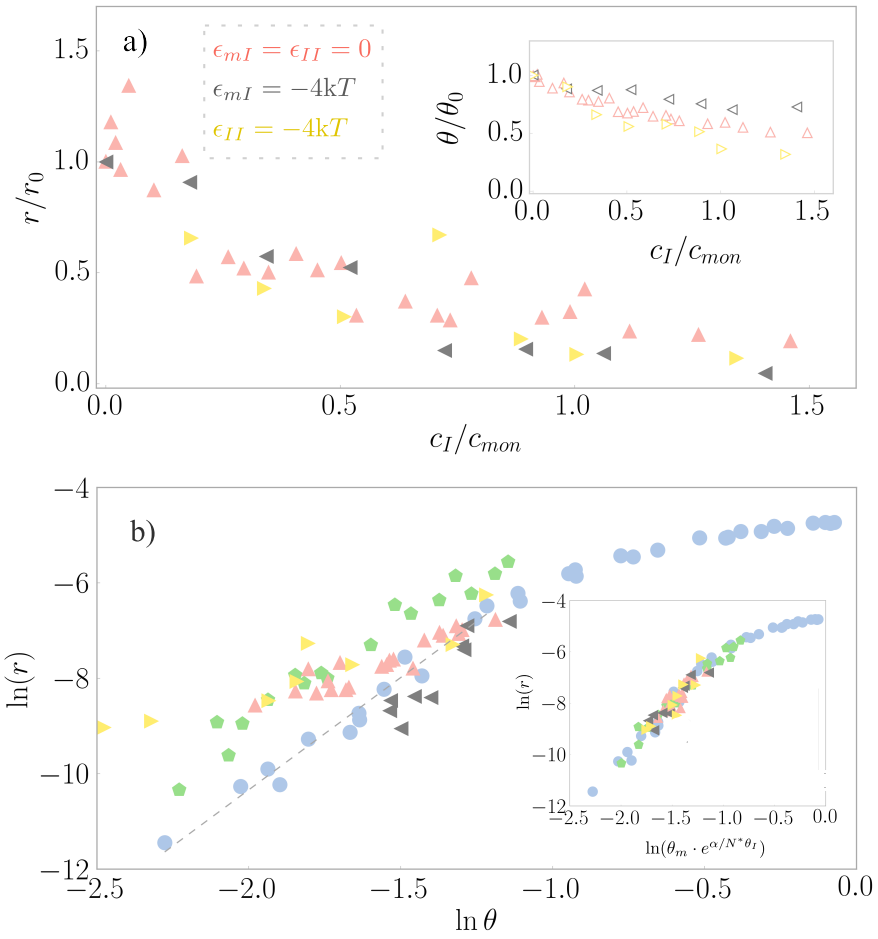

Figure 15. The effect of inhibitor attraction. a) Both for increasing the attraction between inhibitors (yellow) and for increasing the monomer-inhibitor attraction on the fibril surface (grey) the rate of self-replication (r) is slightly decreased at a given inhibitor concentration (c_{mon}) even as the monomer coverage (θ_m) actually increases in the latter case (inset). b) On the rate against coverage graph, the data for inter-inhibitor attraction (yellow) follows the line of no inhibitor attraction (red) but the data for monomer-inhibitor attraction lies below the blue line where inhibitor is absent indicating that at given monomer coverage surface-oligomers made entirely of monomers tend to be smaller. Inset shows collapse of all datasets on a single curve, using $\alpha = 2N^*$ for green, red, and yellow and using $\alpha = -0.77N^*$ for grey data.

4 Conclusions and outlook

In conclusion, we studied the effect of inhibitory particles on the self-replication pathway that is governed by protein adsorption to the amyloid fibril surface. We performed extensive computer simulations using a course-grained model where both proteins and inhibitors are modelled as hard-core spherocylinders with various anisotropic interactions that promote oligomerisation of proteins to micellar-like structures, binding of proteins and inhibitors to the fibril surface and allow refolding of proteins to β-rich conformations that promote fibril-like morphologies.

We found that surface-bound inhibitors effectively drive down the rate of secondary nucleation by competing for the fibril surface with aggregation-prone proteins. But we also get a very significant effect where inhibitors actually promote self-replication at given protein surface coverage. We identified this effect as macromolecular crowding and successfully accounted for it by using a phenomenological theory where inhibitors exert a pressure on and marginally stabilise assembled surface-oligomers. So by perturbing the nucleation pathway with surface-bound inhibitors, we actually found that monomer coverage does not uniquely determine the rate of nucleation as previously believed. Instead, the average size of a surface-bound oligomer governs the rate of nucleation both in inhibitor's presence and absence. The formation of a fibril nucleus that characterises fibril self-replication evolves along two main reaction coordinates, oligomer size and oligomer structure. Inhibitors only influence the size of oligomers on the surface by shifting the distribution towards larger oligomers while the structural conversion part of nucleation remains unaffected. Using classical nucleation theory, we found a simple formula that successfully predicts the rate of self-replication as a function of average surface-oligomer size and the average nucleating oligomer size even if the underlying free energy landscape of nucleation appears very complex.

We also studied the nucleation mechanism in the absence of inhibitor. By performing simulations with varying temperature, we reaffirmed that the fibril surface coverage by monomeric protein governs the rate of self-replication. Also, by breaking the nucleation process into several major and easily identifiable molecular steps, we identified the most significant or rate-limiting step of nucleation to be the formation and then a partial conversion of a surface-oligomer of appreciable size to a state where at least three monomers fold into an intermediate state with some β-sheet content. But if the protein folding kinetics are fast, we found nucleation to be governed mostly by conversion of detached oligomers to fibril nuclei.

To ascertain whether the crowding effect that promotes oligomerisation can ever dominate over the competitive binding capacity of inhibitors, we employed an an-

alytical statistical mechanical model that captures the interactions between bound species on the fibril surface. We found that the inhibitory effect of competitive binding always overcomes the effect of crowding so that any fibril-bound particle can serve as a valuable therapeutic agent that limits the production of toxic oligomers. Exploring the theoretical model further and comparing with simulation data, we found that the best design for an inhibitory particle would involve a very strong affinity to fibril surface as well as attractive interactions to both to surface-bound monomers and inhibitors.

Especially this inhibitor-monomer attraction that promotes binding into an alternative chain should be explored further in the future as it might slow down both the self-replication catalysed by the fibril surface as well as spontaneous nucleation in solution because they are both governed by oligomerisation in the same way. We can envision that a strong attraction between monomers and inhibitors might stabilise larger oligomers. But at the same time, only a small part of those larger oligomers would participate in the conversion to fibril nucleus. It is likely that the possible (non-)inhibitory effect of alternative chain binding would largely depend on the specific geometry of co-oligomers, both on the surface and in solution.

5 Razširjeni povzetek v slovenskem jeziku

Proteinski ali amiloidni fibrili so podolgovate nitaste strukture, ki nastanejo z agregacijo normalno topnih proteinov. So dominantna oblika proteinskih agregatov in so vpleteni v razvoj več kot 50 človeških bolezni, ki so trenutno neozdravljive in postajajo vse bolj problematične zaradi staranja prebivalstva in sodobnega načina življenja. Med njimi so Parkinsonova bolezen, diabetes tipa II in zlasti Alzheimerjeva bolezen, ki je odgovorna za do 80% vseh primerov demence.

Agregacijska reakcija je zelo heterogena in vsebuje mnogo molekularnih korakov, ki jih lahko reduciramo na le nekaj poglavitnih. Prvi pomemben korak je *spontana* oziroma *primarna nukleacija* v razstopini s katero iz topnih delno zvitih proteinov nastanejo prva jedra fibrilov. Takšna jedra nato rastejo (*elongacija*) z vezavo proteinov na oba konca eno-dimenzionalne fibrilarne strukture do makroskopskih dimenzij (reda µm). Ta dva procesa sta dovolj za opis zgodnje faze agregacijske reakcije, vendar se po oblikovanju makroskopskih fibrilov začne proces *samo-replikacije*. V tem procesu površina že izraslih fibrilov katalizira nastanek novih fibrilnih jeder, kar vodi do eksponentne rasti števila in mase fibrilov in do hitrega širjenja vpliva amiloidnih bolezni. Ta avto-katalitični proces je še posebej pomemben pri Alzheimerjevi bolezni, saj je glavni vir toksičnih oligomerjev, ki povzročajo nevronsko smrt. To so amorfni proteinski agregati, ki še nimajo lastnosti fibrilov, ampak se lahko prestrukturirajo v fibrilna jedra.

V tej magistrski nalogi se ukvarjamo z vprašanjem, kako upočasniti ali celo zaustaviti proces samo-replikacije z delci, ki jih v tej nalogi imenujemo inhibitorji. Ti delujejo tako, da tekmujejo z amiloidnimi proteini za vezavna mesta na površini in tako preprečujejo, da bi se proteini na površini srečali in posledično agregirali. Tega vprašanja se lotimo s pomočjo minimalnega računalniškega modela, ki je že bil uspešno uporabljen za opis in razlago spontane nukleacije v razstopini in samoreplikacije na površini. Naše Monte Carlo simulacije zajamejo dejstvo, da lahko protein obstaja v najmanj dveh stanjih: v topnem, delno zvitem stanju, v katerem se protein lahko združuje v micelarne skupke oziroma oligomere in v stanju z nižjo konformacijsko entropijo, z več β strukture, ki se lahko lateralno veže v podolgovate fibrile. Vse interakcije so usmerjene in anizotropne, delce modeliramo kot sferocilindre s trdo sredico. Tako proteini v topnem stanju kot inhibitorji se lahko vežejo na površino že obstoječega fibrila, ki nam v simulaciji služi kot katalitična površina. Simulacijo pripravimo tako, da glede na izbran kemijski potencial proteinov in inhibitorjev simulacijsko škatlo naselimo z razstopljenimi delci in da na sredino škatle vstavimo že formiran fibril, ki ne more rasti, ampak deluje kot katalitična površina. Nato pustimo, da se delci vežejo in zapuščajo površino, dokler ta ne

doseže dinamičnega ravnovesja z razstopino. Nato vključimo možnost, da proteini na diskreten način spremenijo vsebnost β sekundarne strukture. Nukleacijo dobimo, ko se vsaj dva proteinska delca vežeta v oligomer s polno β strukturo.

Najprej preučimo mehanizem samo-replikacije v odsotnosti inhibitorja. Reproduciramo rezultate prejšnje študije in analiziramo nove simulacije, kjer spreminjamo temperaturo pri konstantni koncentraciji proteinov. Z združitvijo osnovnih simulacij, kjer spreminjamo kemijski potencial proteinov, in novih temperaturnih simulacij potrdimo, da ima pokritost površine s proteini ključno vlogo pri avto-katalizi fibrilov. Zaženemo tudi več simulacij pri različnih hitrostih dinamike zvijanja proteinov. Ugotovimo, da je najpočasnejši molekularni korak v naši avto-katalitični reakciji tvorba dovolj velikega oligomera na površini fibrila, ki je že delno zvit in ima že nekaj β-list strukture. Po drugi strani, ob pogoju da je kinetika zvijanja proteinov hitra, najdemo najpočasnejši in tako najpomembnejši korak v nukleaciji v delni pretvorbi odcepljenega, v razstopini plavajočega, oligomera v strukturo, kjer ima en del oligomera že prečno β strukturo, ki je značilna za odrasle fibrile.

Nato se lotimo preučevanja vpliva inhibitornih delcev na mehanizem nukleacije. Ugotovimo, da na površino vezani inhibitorji učinkovito zmanjšajo hitrost samo-replikacije na način, da s proteini tekmujejo za ista vezavna mesta, ki jih zasedejo in tako zmanjšajo pokritost površine s proteini. Večja kot je vezavna energija inhibitorjev na površino, boljša je inhibicija pri določeni koncentraciji inhibitorjev. Hkrati pa dobimo tudi zelo pomemben pojav gnečenja, kjer inhibitorji dejansko pospešujejo nukleacijo pri določeni pokritosti površine z amiloidnimi proteini. Ta pojav gnečenja pojasnimo s fenomenološko teorijo, kjer inhibitorji izvajajo površinski tlak na vezane proteine v njihovi okolici in marginalno stabilizirajo že sestavljene oligomere. Torej s perturbacijo nukleacijske poti z inhibitorji ugotovimo, da pokritost fibrilne površine z agregacijskimi proteini navsezadnje ne določa hitrosti samo-replikacije na edinstven način, kot je bilo do sedaj sprejeto. S pomočjo teorije vpeljemo efektivno pokritost površine, ki je renormalizirana glede na vpliv inhibitorjev in pravilno napove hitrost samo-replikacije pri vseh kombinacijah kemijskih potencialov proteinov in inhibitorjev.

Nato poiščemo še en deskriptor hitrosti samo-replikacije, ki velja v odsotnosti kot tudi prisotnosti inhibitorjev na površini in ga najdemo v povprečni velikosti površinskih oligomerov. Oblikovanje oligomera oziroma fibrilnega jedra, ki je termodinamsko stabilno, poteka preko vsaj dveh reakcijskih spremenljivk: velikosti oligomera in njegove strukture v smislu vsebnosti β sekundarne strukture. Inhibitorji na površini vplivajo le na oligomerizacijo proteinov na način, da potisnejo porazdelitev oligomerov na fibrilni površini k večjim oligomerov, ampak ne vplivajo na konformacijski del preobrazbe skupkov proteinov v fibrilno jedro. Z uporabo

in reinterpretacijo klasične teorije nukleacije smo kljub zelo kompleksni nukleacijski poti izluščili preprosto formulo, ki uspešno napove povezavo med hitrostjo samo-replikacije na eni strani in povprečno velikostjo na površino vezanega proteinskega oligomera ter povprečno velikostjo fibrilnega jedra na drugi strani.

Nato smo raziskali, ali lahko vpliv gnečenja molekul na površini kadarkoli prevlada nad vplivom tekmovanja proteinov in inhibitorjev za ista vezavna mesta. Razvijemo statistično-mehanski mrežni model, ki zraven običajne vezave obeh vrst delcev na površino upošteva tudi interakcije med vezanimi delci in posledično zelo dobro zajame obnašanje vezavnih izoterm obeh delcev v simulacijah. Izkaže se, da tekmovalni efekt za ista vezavna mesta vedno prevlada nad gnečenjem, tako da lahko katerikoli delec, ki ima možnost vsidranja na fibrilno površino, služi v medicinske namene. Vseeno pa so nekatere interakcije med vezanimi delci bolj učinkovite v inhibiciji samo-replikacije kot druge. Teoretični mrežni model lahko uporabimo, da raziščemo smernice za oblikovanje interakcij s čim bolj učinkovitim inhibicijskim učinkom. Ugotovimo, da privlačne interakcije med inhibitorji na površini kot tudi med vezanimi proteini in inhibitorji ojačajo inhibicijski učinek. Prve zato, ker promovirajo kooperativno vezavo inhibitorjev in torej pri določeni koncentraciji inhibitorjev v razstopini dobimo več vezav inhibitorja na površino. Zanimivo pa privlačna interakcija med inhibitorji in proteini celo bolj učinkovito inhibira katalitični proces na površini, čeprav se število adsorbiranih proteinov pod vplivom inhibitorjev celo poveča in ne zmanjša. Takšna privlačna interakcija namreč spodbuja vezavo v mešane oligomere, ki so delno sestavljeni iz proteinov in delno iz inhibitorjev. Ti oligomeri so sicer v povprečju večji in entropično bolj zaželjeni, ampak ima le manjši del oligomera možnost zvitja v β strukture, tako da je nukleacija mešanega oligomera veliko počasnejša oziroma manj verjetna.

V prihodnje bi bilo vredno raziskati, kako geometrija različnih mrežnih modelov vpliva na zaključke naše študije glede oblikovanja interakcij med inhibitorji in ostalimi na površino vezanimi proteini, da bi se inhibicijska moč povečala. Lahko bi se namreč zgodilo, da se vpliv mešanih oligomerov različno obnaša pri oligomerizaciji v eni, dveh ali treh dimenzijah in pri različnem številu interakcijskih sosedov.

References

[1] T. P. J. Knowles, M. Vendruscolo, and C. M. Dobson, "The amyloid state and its association with protein misfolding diseases," *Nature Reviews Molecular Cell Biology*, vol. 15, pp. 384–396, June 2014.

[2] T. P. J. Knowles, M. Vendruscolo, and C. M. Dobson, "The physical basis of protein misfolding disorders," *Physics Today*, vol. 68, pp. 36–41, Mar. 2015.

[3] "2018 Alzheimer's disease facts and figures," *Alzheimer's & Dementia*, vol. 14, pp. 367–429, Mar. 2018.

[4] T. P. J. Knowles, C. A. Waudby, G. L. Devlin, S. I. A. Cohen, A. Aguzzi, M. Vendruscolo, E. M. Terentjev, M. E. Welland, and C. M. Dobson, "An Analytical Solution to the Kinetics of Breakable Filament Assembly," *Science*, vol. 326, pp. 1533–1537, Dec. 2009.

[5] S. I. A. Cohen, M. Vendruscolo, M. E. Welland, C. M. Dobson, E. M. Terentjev, and T. P. J. Knowles, "Nucleated polymerization with secondary pathways. I. Time evolution of the principal moments," *The Journal of Chemical Physics*, vol. 135, p. 065105, Aug. 2011.

[6] S. I. Cohen, M. Vendruscolo, C. M. Dobson, and T. P. Knowles, "From Macroscopic Measurements to Microscopic Mechanisms of Protein Aggregation," *Journal of Molecular Biology*, vol. 421, pp. 160–171, Aug. 2012.

[7] T. C. T. Michaels and T. P. J. Knowles, "Kinetic theory of protein filament growth: Self-consistent methods and perturbative techniques," *International Journal of Modern Physics B*, vol. 29, p. 1530002, Jan. 2015.

[8] S. I. A. Cohen, S. Linse, L. M. Luheshi, E. Hellstrand, D. A. White, L. Rajah, D. E. Otzen, M. Vendruscolo, C. M. Dobson, and T. P. J. Knowles, "Proliferation of amyloid- 42 aggregates occurs through a secondary nucleation mechanism," *Proceedings of the National Academy of Sciences*, vol. 110, pp. 9758–9763, June 2013.

[9] G. Meisl, X. Yang, E. Hellstrand, B. Frohm, J. B. Kirkegaard, S. I. A. Cohen, C. M. Dobson, S. Linse, and T. P. J. Knowles, "Differences in nucleation behavior underlie the contrasting aggregation kinetics of the Aβ40 and Aβ42 peptides," *Proceedings of the National Academy of Sciences*, vol. 111, pp. 9384–9389, July 2014.

[10] P. Arosio, T. P. J. Knowles, and S. Linse, "On the lag phase in amyloid fibril formation," *Physical Chemistry Chemical Physics*, vol. 17, no. 12, pp. 7606–7618, 2015.

[11] S. I. A. Cohen, R. Cukalevski, T. C. T. Michaels, A. Šarić, M. Törnquist, M. Vendruscolo, C. M. Dobson, A. K. Buell, T. P. J. Knowles, and S. Linse, "Distinct thermodynamic signatures of oligomer generation in the aggregation of the amyloid-β peptide," *Nature Chemistry*, vol. 10, pp. 523–531, May 2018.

[12] M. Törnquist, T. C. T. Michaels, K. Sanagavarapu, X. Yang, G. Meisl, S. I. A. Cohen, T. P. J. Knowles, and S. Linse, "Secondary nucleation in amyloid formation," *Chemical Communications*, July 2018.

[13] S. I. A. Cohen, P. Arosio, J. Presto, F. R. Kurudenkandy, H. Biverstål, L. Dolfe, C. Dunning, X. Yang, B. Frohm, M. Vendruscolo, J. Johansson, C. M. Dobson, A. Fisahn, T. P. J. Knowles, and S. Linse, "A molecular chaperone breaks the catalytic cycle that generates toxic Aβ oligomers," *Nature Structural & Molecular Biology*, vol. 22, pp. 207–213, Mar. 2015.

[14] R. Gaspar, G. Meisl, A. K. Buell, L. Young, C. F. Kaminski, T. P. J. Knowles, E. Sparr, and S. Linse, "Secondary nucleation of monomers on fibril surface dominates α-synuclein aggregation and provides autocatalytic amyloid amplification," *Quarterly Reviews of Biophysics*, vol. 50, 2017.

[15] M. Ankarcrona, B. Winblad, C. Monteiro, C. Fearns, E. T. Powers, J. Johansson, G. T. Westermark, J. Presto, B.-G. Ericzon, and J. W. Kelly, "Current and future treatment of amyloid diseases," *Journal of Internal Medicine*, vol. 280, pp. 177–202, Aug. 2016.

[16] F. A. Aprile, P. Arosio, G. Fusco, S. W. Chen, J. R. Kumita, A. Dhulesia, P. Tortora, T. P. J. Knowles, M. Vendruscolo, C. M. Dobson, and N. Cremades, "Inhibition of α-Synuclein Fibril Elongation by Hsp70 Is Governed by a Kinetic Binding Competition between α-Synuclein Species," *Biochemistry*, vol. 56, pp. 1177–1180, Mar. 2017.

[17] M. Del Campo, J. J. Hoozemans, L.-L. Dekkers, A. J. Rozemuller, C. Korth, A. Müller-Schiffmann, P. Scheltens, M. A. Blankenstein, C. R. Jimenez, R. Veerhuis, and C. E. Teunissen, "BRI2-BRICHOS is increased in human amyloid plaques in early stages of Alzheimer's disease," *Neurobiology of Aging*, vol. 35, pp. 1596–1604, July 2014.

[18] A. J. Doig and P. Derreumaux, "Inhibition of protein aggregation and amyloid formation by small molecules," *Current Opinion in Structural Biology*, vol. 30, pp. 50–56, Feb. 2015.

[19] L. Dolfe, B. Winblad, J. Johansson, and J. Presto, "BRICHOS binds to a designed amyloid-forming -protein and reduces proteasomal inhibition and aggresome formation," *Biochemical Journal*, vol. 473, pp. 167–178, Jan. 2016.

[20] E. Hermansson, S. Schultz, D. Crowther, S. Linse, B. Winblad, G. Westermark, J. Johansson, and J. Presto, "The chaperone domain BRICHOS prevents CNS toxicity of amyloid- peptide in Drosophila melanogaster," *Disease Models & Mechanisms*, vol. 7, pp. 659–665, June 2014.

[21] S. D. Knight, J. Presto, S. Linse, and J. Johansson, "The BRICHOS Domain, Amyloid Fibril Formation, and Their Relationship," *Biochemistry*, vol. 52, pp. 7523–7531, Oct. 2013.

[22] H. Willander, J. Presto, G. Askarieh, H. Biverstål, B. Frohm, S. D. Knight, J. Johansson, and S. Linse, "BRICHOS Domains Efficiently Delay Fibrillation of Amyloid β-Peptide," *Journal of Biological Chemistry*, vol. 287, pp. 31608–31617, Sept. 2012.

[23] A. Šarić, Y. C. Chebaro, T. P. J. Knowles, and D. Frenkel, "Crucial role of nonspecific interactions in amyloid nucleation," *Proceedings of the National Academy of Sciences*, vol. 111, pp. 17869–17874, Dec. 2014.

[24] A. Šarić, A. K. Buell, G. Meisl, T. C. T. Michaels, C. M. Dobson, S. Linse, T. P. J. Knowles, and D. Frenkel, "Physical determinants of the self-replication of protein fibrils," *Nature Physics*, vol. 12, pp. 874–880, Sept. 2016.

[25] A. Šarić, T. C. T. Michaels, A. Zaccone, T. P. J. Knowles, and D. Frenkel, "Kinetics of spontaneous filament nucleation via oligomers: Insights from theory and simulation," *The Journal of Chemical Physics*, vol. 145, p. 211926, Dec. 2016.

[26] R. Vácha and D. Frenkel, "Relation between Molecular Shape and the Morphology of Self-Assembling Aggregates: A Simulation Study," *Biophysical Journal*, vol. 101, pp. 1432–1439, Sept. 2011.

[27] N. S. Bieler, T. P. J. Knowles, D. Frenkel, and R. Vácha, "Connecting Macroscopic Observables and Microscopic Assembly Events in Amyloid Formation Using Coarse Grained Simulations," *PLoS Computational Biology*, vol. 8, p. e1002692, Oct. 2012.

[28] J. Zhang and M. Muthukumar, "Simulations of nucleation and elongation of amyloid fibrils," *The Journal of Chemical Physics*, vol. 130, Jan. 2009.

[29] I. M. Ilie, W. K. den Otter, and W. J. Briels, "A coarse grained protein model with internal degrees of freedom. Application to α-synuclein aggregation," *The Journal of Chemical Physics*, vol. 144, p. 085103, Feb. 2016.

[30] H. I. Ingólfsson, C. A. Lopez, J. J. Uusitalo, D. H. d. Jong, S. M. Gopal, X. Periole, and S. J. Marrink, "The power of coarse graining in biomolecular simulations," *Wiley Interdisciplinary Reviews: Computational Molecular Science*, vol. 4, pp. 225–248, May 2014.

[31] D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications*. Elsevier, Oct. 2001. Google-Books-ID: 5qTzldS9ROIC.

[32] W. Kim and M. H. Hecht, "Generic hydrophobic residues are sufficient to promote aggregation of the Alzheimer's Aβ42 peptide," *Proceedings of the National Academy of Sciences*, vol. 103, pp. 15824–15829, Oct. 2006.

[33] W. Kim and M. H. Hecht, "Mutations Enhance the Aggregation Propensity of the Alzheimer's Aβ Peptide," *Journal of Molecular Biology*, vol. 377, pp. 565–574, Mar. 2008.

[34] A. W. P. Fitzpatrick, G. T. Debelouchina, M. J. Bayro, D. K. Clare, M. A. Caporini, V. S. Bajaj, C. P. Jaroniec, L. Wang, V. Ladizhansky, S. A. Muller, C. E. MacPhee, C. A. Waudby, H. R. Mott, A. De Simone, T. P. J. Knowles, H. R. Saibil, M. Vendruscolo, E. V. Orlova, R. G. Griffin, and C. M. Dobson, "Atomic structure and hierarchical assembly of a cross-β amyloid fibril," *Proceedings of the National Academy of Sciences*, vol. 110, pp. 5468–5473, Apr. 2013.

[35] R. Nelson, M. R. Sawaya, M. Balbirnie, A. O. Madsen, C. Riekel, R. Grothe, and D. Eisenberg, "Structure of the cross-β spine of amyloid-like fibrils," *Nature*, vol. 435, pp. 773–778, June 2005.

[36] L. C. Serpell, "Alzheimer's amyloid fibrils: structure and assembly," *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, vol. 1502, pp. 16–30, July 2000.

[37] M. Fändrich, M. A. Fletcher, and C. M. Dobson, "Amyloid fibrils from muscle myoglobin," *Nature*, vol. 410, pp. 165–166, Mar. 2001.

[38] J. R. Allison, P. Varnai, C. M. Dobson, and M. Vendruscolo, "Determination of the Free Energy Landscape of α-Synuclein Using Spin Label Nuclear Magnetic Resonance Measurements," *Journal of the American Chemical Society*, vol. 131, pp. 18314–18326, Dec. 2009.

[39] A. K. Buell, A. Dhulesia, D. A. White, T. P. J. Knowles, C. M. Dobson, and M. E. Welland, "Detailed Analysis of the Energy Barriers for Amyloid Fibril Growth," *Angewandte Chemie International Edition*, vol. 51, pp. 5247–5251, May 2012.

[40] A. K. Buell, C. M. Dobson, and T. P. Knowles, "The physical chemistry of the amyloid phenomenon: thermodynamics and kinetics of filamentous protein aggregation," *Essays In Biochemistry*, vol. 56, pp. 11–39, Aug. 2014.

[41] T. C. T. Michaels, L. X. Liu, S. Curk, P. G. Bolhuis, A. Šarić, and T. P. J. Knowles, "Reaction rate theory for supramolecular kinetics: application to protein aggregation," *Molecular Physics*, vol. 116, pp. 3055–3065, Nov. 2018.

[42] T. C. T. Michaels, A. Šarić, S. Curk, P. Arosio, G. Meisl, A. J. Dear, S. I. A. Cohen, V. Michele, C. M. Dobson, S. Linse, and T. P. J. Knowles, "Dynamics of oligomer populations formed during the aggregation of alzheimer's aβ42 peptide," *Science*, 'in preparation'.

[43] T. L. Hill, *An Introduction to Statistical Thermodynamics*. Courier Corporation, 1986. Google-Books-ID: QttFDwAAQBAJ.

[44] P. G. Vekilov, "The two-step mechanism of nucleation of crystals in solution," *Nanoscale*, vol. 2, pp. 2346–2357, Nov. 2010.

[45] D. Gebauer and H. Cölfen, "Prenucleation clusters and non-classical nucleation," *Nano Today*, vol. 6, pp. 564–584, Dec. 2011.

[46] S. Karthika, T. K. Radhakrishnan, and P. Kalaichelvi, "A Review of Classical and Nonclassical Nucleation Theories," *Crystal Growth & Design*, vol. 16, pp. 6663–6681, Nov. 2016.

[47] P. G. Debenedetti, *Metastable Liquids: Concepts and Principles*. Princeton University Press, 1996. Google-Books-ID: tzvvsltE6Y8C.

[48] R. P. Sear, "Nucleation: theory and applications to protein solutions and colloidal suspensions," *Journal of Physics: Condensed Matter*, vol. 19, no. 3, p. 033101, 2007.