

# Theoretical Analysis of Amyloid Polymorphism

Maëlle Ranouil\*

Department of Biotechnology and Biomedicine, Denmark Technical University

(Dated: January 13, 2026)

Amyloid polymorphism refers to the ability of a single protein sequence to adopt multiple distinct amyloid structures. Variability can occur at several structural levels, including the conformation of the monomers, the architecture of protofilaments, and their organization within mature fibrils. Amyloid polymorphism has been observed in both ex vivo and in vitro experiments, and emerges as a key readout of the structural pathways explored during amyloid fibrillation. In this report, we present a description of protein aggregation for a system of two polymorphs, in both bulk and small volumes in which the stochastic nature of the system cannot be neglected. We analyse the evolution of the system depending on the values of the kinetic constants of each polymorph. Finally, we try to link this work to the experimental results previously published.

## I. INTRODUCTION

Amyloid fibril formation is a ubiquitous process in biology and is implicated in a wide range of neurodegenerative diseases, including Alzheimer's and Parkinson's diseases. During amyloid aggregation, soluble protein monomers self-assemble into highly ordered fibrillar structures through a complex network of nucleation and growth mechanisms. Remarkably, a single protein sequence can give rise to multiple structurally distinct amyloid fibrils, a phenomenon known as amyloid polymorphism. These polymorphs may differ at several hierarchical levels, including monomer fold, protofilament architecture, and fibril packing, and are increasingly recognized as a key determinant of disease specificity and progression.

Experimental evidence for amyloid polymorphism has been obtained both ex vivo and in vitro. Cryogenic electron microscopy (cryo-EM) studies of amyloid filaments extracted from human brain tissue have revealed that distinct fibrillar folds of proteins such as tau,  $\alpha$ -synuclein, and TDP-43 are associated with different neurodegenerative [1]. In parallel, in vitro experiments have demonstrated that multiple polymorphs can coexist within the same solution or arise under nearly identical preparation conditions.

From a theoretical perspective, amyloid aggregation has been successfully described using deterministic, mean-field models that capture the average behaviour of large systems through coupled rate equations. Such approaches have provided key insights into the roles of primary nucleation, secondary nucleation, and elongation in shaping aggregation kinetics. However, mean-field descriptions inherently neglect fluctuations and correlations, which become particularly important in small volumes or at early times. In these regimes, stochastic effects can strongly influence the aggregation outcome and may lead to variability between nominally identical experiments.

In this report, we present a theoretical analysis of amyloid polymorphism based on both deterministic and stochastic descriptions of protein aggregation. We first introduce a mean-field model describing the competition between two amyloid polymorphs in bulk solution and systematically explore how their relative kinetic parameters determine coexistence or dominance. We then extend the analysis to a stochastic framework based on the chemical master equation and the Gillespie algorithm, which allows us to capture fluctuations and describe regimes inaccessible to mean-field theory. By comparing these approaches, we aim to clarify the conditions under which polymorphism can emerge and persist, and to provide a theoretical framework that connects kinetic parameters to experimentally observed variability in amyloid aggregation.

All numerical implementations are available at [https://github.com/maelleran/amyloid\\_polymorphism\\_project.git](https://github.com/maelleran/amyloid_polymorphism_project.git).

## II. MEAN-FIELD THEORY

### A. Master Equation

We consider a system of total mass  $m_{tot}$  containing a mixture of two different fibrillar aggregates and monomeric proteins. We assume competition without interaction between the polymorphs  $A$  and  $B$ . The extension of the master equations of fibril filamentation [2] to such a system gives the system of equations

$$\frac{dM_A}{dt} = 2 k_{+A} m P_A, \quad (1)$$

$$\frac{dM_B}{dt} = 2 k_{+B} m P_B, \quad (2)$$

$$\frac{dP_A}{dt} = k_{nA} m^{n_c} + k_{2A} m^{n_2} M_A, \quad (3)$$

$$\frac{dP_B}{dt} = k_{nB} m^{n_c} + k_{2B} m^{n_2} M_B, \quad (4)$$

$$\frac{dm}{dt} = -2 k_{+A} m P_A - 2 k_{+B} m P_B, \quad (5)$$

\* maelle.ranouil@polytechnique.edu

where  $i \in \{A, B\}$  is a label for polymorph type and

- $m$  = mass concentration of monomers,
- $M_i$  = mass concentration,
- $P_i$  = fibril number concentration,
- $k_{ni}$  = primary (or first) nucleation rate constant,
- $k_{2i}$  = secondary (or second) nucleation rate constant,
- $k_{+i}$  = elongation rate constant,
- $n_c$  = primary nucleation reaction order, and
- $n_2$  = secondary nucleation reaction order.

Note that we have set the reaction orders to be the same for the two polymorphs in our analysis.

Equations 1 and 2 capture the change in fibril mass concentration of type A and B. These mass concentrations evolve a function of the monomers incorporated in their respective fibrils type  $i$  with rate  $k_{+i}$ . Equations 3 and 4 capture the evolution of the number of fibrils of each polymorph, while equation 5 pictures the evolution of the mass concentration of the free monomers. The evolution of the number of fibrils depends on the spontaneous association of free monomers to make a new fibril at rate  $k_{ni}$  and on the creation of new fibrils from secondary nucleation with associated rate  $k_{2i}$ . Finally, the mass of free monomers decreases proportionally to the number of free monomers incorporated into fibrils. Such equations can easily be extended to describe a system including more polymorphs. To do so would entail setting, for each polymorph  $j$ , similar equations for  $M_j$  and  $P_j$ , and adding accordingly the terms  $-2k_{+j}mP_j$  to the equation for  $m$ .

For simplicity, it is convenient to rewrite this system using appropriate time scales in order to obtain a set of dimensionless equations, which is what we will treat numerically. The total initial mass of monomers  $m_{tot}$  can be used as a mass scale and the quantity  $k_{+A}m_{tot}$  gives an appropriate time scale. We then rewrite the variables and constants of the system as  $m = \mu m_{tot}$ ,  $M_A = M'_A m_{tot}$ ,  $M_B = M'_B m_{tot}$ ,  $t = \tau/(k_{+A}m_{tot})$ , with  $\mu$ ,  $M'_A$ ,  $M'_B$ , and  $\tau$  real numbers. The system of equations is therefore

$$\begin{aligned}\frac{dM'_A}{d\tau} &= 2\mu P_A, \\ \frac{dM'_B}{d\tau} &= 2\frac{k_{+B}}{k_{+A}}\mu P_B, \\ \frac{dP'_A}{d\tau} &= \frac{k_{nA}}{k_{+A}}m_{tot}^{n_c-2}\mu^{n_c} + \frac{k_{2A}}{k_{+A}}m_{tot}^{n_2-1}\mu^{n_2}M'_A, \\ \frac{dP'_B}{d\tau} &= \frac{k_{nB}}{k_{+A}}m_{tot}^{n_c-2}\mu^{n_c} + \frac{k_{2B}}{k_{+A}}m_{tot}^{n_2-1}\mu^{n_2}M'_B, \\ \frac{d\mu}{d\tau} &= -2\mu P'_A - 2\frac{k_{+B}}{k_{+A}}\mu P'_B.\end{aligned}$$

Furthermore, let us set  $\alpha_{nA} = \frac{k_{nA}}{k_{+A}}$ ,  $\alpha_{2A} = \frac{k_{2A}}{k_{+A}}$ ,  $\alpha_{nB} = \frac{k_{nB}}{k_{+A}}$ ,  $\alpha_{2B} = \frac{k_{2B}}{k_{+A}}$  and  $\beta = \frac{k_{+B}}{k_{+A}}$ . Additionally, we use

$n_c = n_2 = 2$ . Then the system further simplifies to

$$\frac{dM'_A}{d\tau} = 2\mu P_A, \quad (6)$$

$$\frac{dM'_B}{d\tau} = 2\beta\mu P_B, \quad (7)$$

$$\frac{dP'_A}{d\tau} = \alpha_{nA}\mu^2 + \alpha_{2A}m_{tot}\mu^2M'_A, \quad (8)$$

$$\frac{dP'_B}{d\tau} = \alpha_{nB}\mu^2 + \alpha_{2B}m_{tot}\mu^2M'_B, \quad (9)$$

$$\frac{d\mu}{d\tau} = -2\mu P'_A - 2\beta\mu P'_B. \quad (10)$$

## B. Numerical Solutions

Equations 6-10 were solved numerically with initial conditions  $P'_A(0) = P'_B(0) = 0$ ,  $M'_A(0) = M'_B(0) = 0$  and  $\mu(0) = 1$ . We then studied the evolution of the mass concentration of both types of amyloid as a function of time.

*a. Aggregation curves* We validate the numerical implementation by setting identical parameters and constants for polymorphs  $A$  and  $B$ . Under these conditions, both species form at the same rate, and overlapping growth curves are expected. Since the two species compete symmetrically for the available monomers, each consumes half of them, leading to a final mass of  $M_A(\infty) = M_B(\infty) = \frac{1}{2}m_{tot}$ . The expected aggregation time is  $\sim 1$  h. The time evolution of the numerical solution for the set of parameter described above is shown in Figure 1.

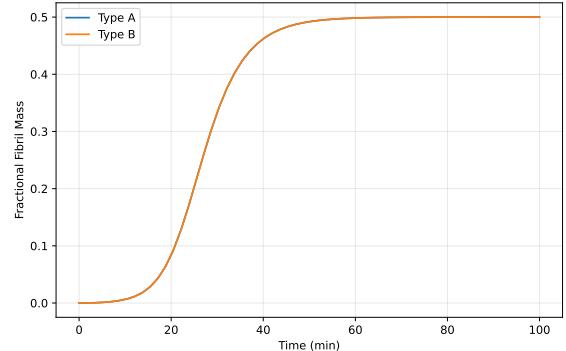


FIG. 1. Evolution of the fibril mass concentration of the system as a function of time, with  $k_{+A} = k_{+B} = 1 \times 10^6 \text{ M}^{-1} \text{s}^{-1}$ ,  $k_{2A} = k_{2B} = 1 \times 10^4 \text{ M}^{-2} \text{s}^{-1}$ ,  $k_{nB} = k_{nA} = 1 \times 10^4 \text{ M}^{-1} \text{s}^{-1}$ .

The simulations lead to overlapping aggregation curves, with a maximal mass concentration  $M_i(\infty) = \frac{1}{2}m_{tot}$  reached after approximately one hour, which matches the theoretical expectation.

In a second step, the values of the rate constants were varied. Specifically, each rate constant of B was successively doubled, and the corresponding aggregation curves were plotted, as shown in Figures 2, 3, and 4.

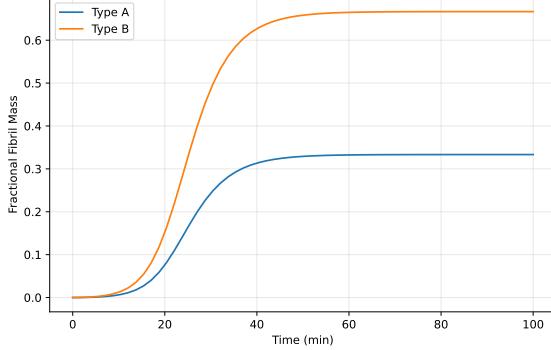


FIG. 2. Evolution of the mass concentration of type A and type B as a function of time, with  $k_{+A} = k_{+B} = 1 \times 10^6 \text{ M}^{-1} \text{s}^{-1}$ ,  $k_{2A} = k_{2B} = 1 \times 10^4 \text{ M}^{-2} \text{s}^{-1}$ ,  $k_{nB} = 2 \times 10^4 \text{ M}^{-1} \text{s}^{-1}$  and  $k_{nA} = 1 \times 10^4 \text{ M}^{-1} \text{s}^{-1}$ .

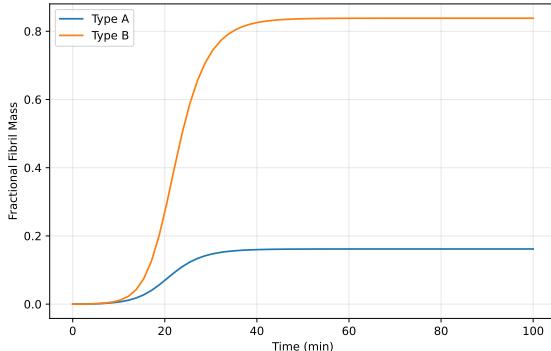


FIG. 3. Evolution of the mass concentration of type A and type B as a function of time, with  $k_{+A} = 1 \times 10^6 \text{ M}^{-1} \text{s}^{-1}$ ,  $k_{+B} = 2 \times 10^6 \text{ M}^{-1} \text{s}^{-1}$ ,  $k_{2A} = k_{2B} = 1 \times 10^4 \text{ M}^{-2} \text{s}^{-1}$ ,  $k_{nB} = k_{nA} = 1 \times 10^4 \text{ M}^{-1} \text{s}^{-1}$ .

We see that, as expected, the fractional mass for B increases faster, and in the limit, it acquires a larger fractional mass. Note that, for the given set of parameters, the doubling the rate of second nucleation is more impactful than the doubling the rate of elongation, which is itself more impactful than doubling the first nucleation rate. Note that, if this is this true for the given set of parameters, we may observed slightly different results when using other ratio between first and second nucleation or different reaction orders.

*b. Heat Maps* To identify the parameter regime in which polymorphs coexist, we systematically varied the ratios of the kinetic rate constants of the two species

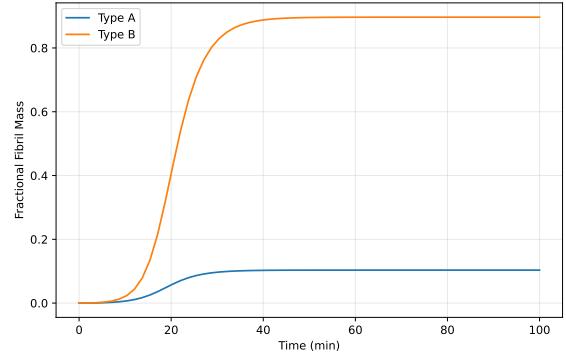


FIG. 4. Evolution of the mass concentration of type A and type B as a function of time, with  $k_{+A} = k_{+B} = 1 \times 10^6 \text{ M}^{-1} \text{s}^{-1}$ ,  $k_{2A} = 1 \times 10^4 \text{ M}^{-2} \text{s}^{-1}$ ,  $k_{2B} = 2 \times 10^4 \text{ M}^{-2} \text{s}^{-1}$ ,  $k_{nB} = 1 \times 10^4 \text{ M}^{-1} \text{s}^{-1}$  and  $k_{nA} = 1 \times 10^4 \text{ M}^{-1} \text{s}^{-1}$ .

governing secondary nucleation and elongation.

The parameters for A were constant and set to

$$\begin{aligned} k_{+A} &= 1 \times 10^6 \text{ M}^{-1} \text{s}^{-1}, \\ k_{2A} &= 1 \times 10^4 \text{ M}^{-2} \text{s}^{-1}, \\ k_{nA} &= 1 \times 10^{-4} \text{ M}^{-1} \text{s}^{-1}. \end{aligned}$$

The first nucleation rate of polymorph B was also fixed to  $k_{nB} = 10^{-4} \text{ M}^{-1} \text{s}^{-1}$ . The equations were then solved for  $k_{2B}$  ranging from  $0.1 k_{2A}$  to  $10 k_{2A}$  and  $k_{+B}$  ranging from  $0.1 k_{+A}$  to  $10 k_{+A}$ . The final mass concentration of each polymorph for each value of the couple  $\left(\frac{k_{+A}}{k_{+B}}, \frac{k_{2A}}{k_{2B}}\right)$  was recorded. We then plotted heat maps to visualize the regions of coexistence.

First, the relative mass concentration of polymorph A was plotted as a function of the ratios  $\frac{k_{+A}}{k_{+B}}$  and  $\frac{k_{2A}}{k_{2B}}$ , i.e. we plotted  $f\left(\frac{k_{+A}}{k_{+B}}, \frac{k_{2A}}{k_{2B}}\right) = \frac{M_A(\infty)}{M_A(\infty) + M_B(\infty)}$ .

Then, the minimum between the relative mass concentration of A and B as a function of the ratios  $\frac{k_{+A}}{k_{+B}}$  and  $\frac{k_{2A}}{k_{2B}}$ . That is, we plotted

$$f\left(\frac{k_{+A}}{k_{+B}}, \frac{k_{2A}}{k_{2B}}\right) = \frac{\min(M_A(\infty), M_B(\infty))}{M_A(\infty) + M_B(\infty)}.$$

This allows us to clearly see the region where both species coexist.

*c. Variation of the first nucleation rate* Finally, we plotted heat maps for height distinct values of  $k_{nB}$ . 200 values between 0.1 and 10 where taken for both  $\frac{k_{+A}}{k_{+B}}$  and  $\frac{k_{2A}}{k_{2B}}$ .

*d. conclusion* The mean-field framework makes it possible to identify a domain of kinetic rate constants in which two polymorphs can coexist in bulk solution. This coexistence region is relatively narrow: for a given

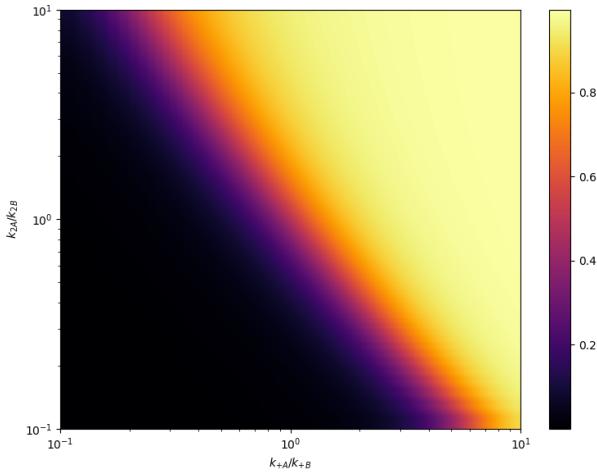


FIG. 5. Relative mass concentration of A as a function of the ratios  $\frac{k_{+A}}{k_{+B}}$  and  $\frac{k_{2A}}{k_{2B}}$ . 500 values between 0.1 and 10 where taken for both  $\frac{k_{+A}}{k_{+B}}$  and  $\frac{k_{2A}}{k_{2B}}$ .

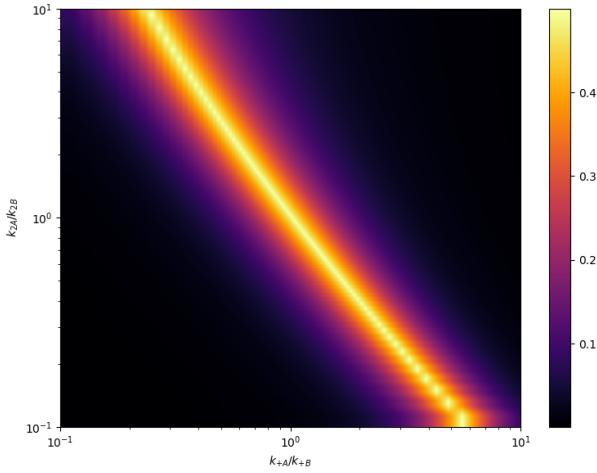


FIG. 6. Minimum of the relative mass concentration of A and B as a function of the ratios  $\frac{k_{+A}}{k_{+B}}$  and  $\frac{k_{2A}}{k_{2B}}$ . 500 values between 0.1 and 10 where taken for both  $\frac{k_{+A}}{k_{+B}}$  and  $\frac{k_{2A}}{k_{2B}}$ .

rate constant, the ratio between the rates of the two polymorphs can increase to at most about 3–3.5 before one species becomes effectively undetectable. Variations in the primary nucleation rate strongly shift this coexistence domain. In particular, a polymorph with a low primary nucleation rate must compensate through higher secondary nucleation and elongation rates in order to coexist with a polymorph that generates seeds more efficiently.

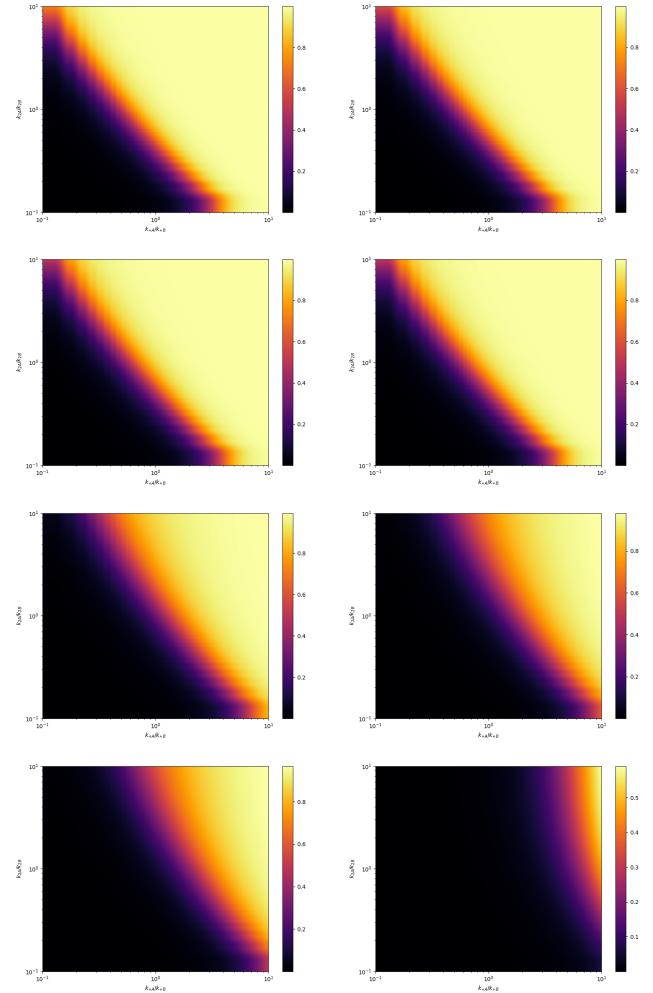


FIG. 7. Relative mass concentration of A as a function of the ratios  $\frac{k_{+A}}{k_{+B}}$  and  $\frac{k_{2A}}{k_{2B}}$ . From top left to bottom right:  $k_{nB} = 0.04 k_{nA}$ ,  $k_{nB} = 0.07 k_{nA}$ ,  $k_{nB} = 0.1 k_{nA}$ ,  $k_{nB} = 2 k_{nA}$ ,  $k_{nB} = 6 k_{nA}$ ,  $k_{nB} = 10 k_{nA}$ ,  $k_{nB} = 14 k_{nA}$ .

### III. STOCHASTIC THEORY

While mean-field theory imposes specific constraints on the kinetic rate constants required for the coexistence of polymorphs, a stochastic description of amyloid formation provides a more comprehensive framework. The stochastic theory not only accounts for polymorph coexistence but also captures additional regimes inaccessible to mean-field approaches. For example, it can describe situations in which different polymorphs emerge in separate wells of a multiwell plate, despite originating from the same initial monomer population [3], [4]. This motivates us to adopt a stochastic description of the system, which is essential for capturing both polymorph coexistence and the inherent variability of amyloid formation.

## A. Chemical Master Equation (CME)

*a. Description of the system* We will describe the two different proteins by their total number of fibrils  $n$  and the number of monomers incorporated into aggregates,  $m$ , leading to a total state vector of  $(n_A, m_A, n_B, m_B)$ . Both polymorphic types in state  $(n_i, m_i)$  are subjected to three different processes: first nucleation, second nucleation and elongation. The backwards reaction of each process also needs to be taken into account. For each type, the forward reactions and their rates can be written as

$$n_i, m_i \rightarrow n_i + 1, m_i + n_c \quad \text{at } \alpha_{1,i}(t), \quad (11)$$

$$n_i, m_i \rightarrow n_i, m_i + 1 \quad \text{at } \mu_i(t)n_i, \quad (12)$$

$$n_i, m_i \rightarrow n_i + 1, m_i \quad \text{at } \alpha_{2,i}(t)(m_i - n_{2,i}), \quad (13)$$

where  $\mu_i = 2k_{+,i}m(t)$ ,  $\alpha_{1,i} = k_{n,i}m(t)^{n_{c,i}}N_AV$ , and  $\alpha_{2,i} = k_{2,i}m(t)^{n_{2,i}}$ , with  $m(t)$  the mass of free available monomers. [5] For early times of the reaction, the variation of the concentration in monomers is negligible and we can assume that  $m(t) \simeq m_{tot}$ . Therefore, the rates of the reaction can be considered as constant and can be written as follow

$$\mu_i = 2k_{+,i}m_{tot}, \quad (14)$$

$$\alpha_{1,i} = k_{n,i}m_{tot}^{n_{c,i}}N_AV, \quad (15)$$

$$\alpha_{2,i} = k_{2,i}m_{tot}^{n_{2,i}}. \quad (16)$$

*b. Master Equations* Let us first consider a system of given volume  $V$  containing a mixture of fibrillar aggregates and monomeric proteins in solution and described by a vector  $(n, m)$ . It follows the previous system description but polymorphism is neglected. In the early stage reaction, the probability distribution function  $P(n, m, t)$  of the state evolve according to the master equation [5],

$$\begin{aligned} \frac{\partial P(n, m, t)}{\partial t} &= \alpha_1 P(n - 1, m - n_c, t) - \alpha_1 P(n, m, t) \\ &\quad + \mu n P(n, m - 1, t) - \mu n P(n, m, t) \\ &\quad + \alpha_2(m - n_2)P(n - 1, m - n_2, t) \\ &\quad - \alpha_2 m P(n, m, t). \end{aligned}$$

Here, the early stage of the reaction corresponds to times greater than  $\kappa^{-1}$  with  $\kappa = \sqrt{k_{+}k_2m_{tot}^{n_2+1}}$  being the characteristic time scale for aggregate proliferation [6] but sufficiently short for the constant monomer approximation to hold. The expression of  $\mu$ ,  $\alpha_1$ ,  $\alpha_2$  follows Equations 14 to 16.

Now, consider a system with two polymorphs under the same conditions and neglect any interaction between the two polymorphs. For each polymorph, the same processes will occur and, in the early times of the reaction, the approximation  $c(t) \simeq c_{tot}$  remains valid. Based on the previous equation and the general framework of the

chemical master equation, one should have the following master equation for the probability distribution function  $P(n, m, t)$  of the state  $(n_A, m_A, n_B, m_B)$ .

$$\frac{\partial P(n_A, m_A, n_B, m_B, t)}{\partial t} = P_A + P_B,$$

where

$$\begin{aligned} P_A &= \alpha_{1,i}P(n_A - 1, m_A - n_c, n_B, m_B, t) \\ &\quad - \alpha_{1,A}P(n_A, m_A, n_B, m_B, t) \\ &\quad + \mu_A n_A P(n_A, m_A - 1, n_B, m_B, t) \\ &\quad - \mu_A n_A P(n_A, m_A, n_B, m_B, t) \\ &\quad + \alpha_{2,A}(m_A - n_2)P(n_A - 1, m_A - n_2, n_B, m_B, t) \\ &\quad - \alpha_{2,A}m_A P(n_A, m_A, n_B, m_B, t) \end{aligned}$$

and  $P_B$  follows the same structure with the indices  $A$  and  $B$  swapped. This equation can easily be generalized to describe a system of  $r$  polymorphs following similar constraints and assumptions. We have

$$\frac{\partial P(n_1, m_1, \dots, n_r, m_r, t)}{\partial t} = \sum_{i=1}^r P_i,$$

with

$$\begin{aligned} P_i &= \alpha_{1,i}P(n_1, m_1, \dots, n_i - 1, m_i - n_{c,i}, n_{i+1}, m_{i+1}, \dots, t) \\ &\quad - \alpha_{1,i}P(n_1, m_1, \dots, n_r, m_r, t) \\ &\quad + \mu_i n_i P(n_1, m_1, \dots, n_i, m_i - 1, n_{i+1}, m_{i+1}, \dots, t) \\ &\quad - \mu_i n_i P(n_1, m_1, \dots, n_r, m_r, t) \\ &\quad + \alpha_{2,i}(m_i - n_{2,i}) \\ P(n_1, m_1, \dots, n_i - 1, m_i - n_{2,i}, n_{i+1}, m_{i+1}, \dots, t) \\ &\quad - \alpha_{2,i}m_i P(n_1, m_1, \dots, n_r, m_r, t) \end{aligned}$$

for early times.

## B. Gillespie Algorithm [7]

The equations defined in the previous section can now be solved numerically using the Gillespie algorithm.

This algorithm uses a Monte Carlo procedure (see the end of the document) to numerically simulate the time evolution of a given chemical system. It accounts for the inherent fluctuations and correlations that are ignored in the deterministic formulation. Let us define the state  $(X_1, X_2, \dots, X_N)$  where  $X_i$  is the number of entities of a chemical specie  $i$  and denote by  $R_\mu$  with  $\mu \in \{1, \dots, M\}$  the reactions that can occur in the system. We introduce the function  $P$  such that  $P(\tau, \mu) d\tau$  is the probability that, given the state  $(X_1, \dots, X_N)$  at time  $t$  the next reaction in  $V$  will occur in the infinitesimal time interval  $(t + \tau, t + \tau + d\tau)$  and will be an  $R_\mu$  reaction. The general expression for this function is

$$P(\tau, \mu) = \begin{cases} a_\mu \exp(-a_0\tau), & \text{if } 0 \leq \tau < \infty \\ & \text{and } \mu = 1, \dots, M, \\ 0 & \text{otherwise,} \end{cases}$$

where  $a_\mu = h_\mu c_\mu$  is the number of distinct  $R_\mu$  reactant combinations available in the state times the stochastic reaction constant of  $R_\mu$ , with  $\mu = 1, \dots, M$ , and

$$a_0 = \sum_{\nu=1}^6 a_\nu = \sum_{\nu=1}^M h_\nu c_\nu.$$

We now need to generate a pair  $(\tau, \mu)$  from the set of random pairs whose probability function is  $P(\tau, \mu)$ . Given two random numbers  $r_1, r_2$  from the unit interval uniform distribution, we take

$$\tau = (1/a_0) \ln(1/r_1), \quad (17)$$

and  $\mu$  to be the integer such that

$$\sum_{\nu=1}^{\mu-1} a_\nu < r_2 a_0 \leq \sum_{\nu=1}^\mu a_\nu. \quad (18)$$

Finally the Gillespie algorithm is as follows.

*Step 0:* Input the desired values for the M reaction constants  $c_1, c_2, \dots, c_M$  and the N initial molecular population numbers  $X_1, \dots, X_N$ . Set the time variable to be 0.  
*Step 1:* Calculate the M quantities  $a_1, \dots, a_M$  for the current molecular population numbers. Also calculate and store  $a_0$  the sum of the M  $a_\nu$  values.  
*Step 2:* Generate two random numbers in the uniform interval  $[0, 1]$  and calculate  $\tau$  and  $\mu$  according to equations 17 and 18.

*Step 3:* Using the  $\tau$  and the  $\mu$  obtained increase  $t$  by  $\tau$  and adjust the molecular population levels to reflect the occurrence of one  $R_\mu$ .

Repeat

#### IV. NUMERICAL RESULTS

The algorithm described above was implemented for systems with and without polymorphism.

*a. Aggregation curves for a single protein conformation* First, a simple state  $n, m$  is considered, with reactions for elongation, first and second nucleation described in equations 11-13. The associated  $a_0(t)$  is  $a_0 = n\mu(t) + m\alpha_2(t) + \alpha_1(t)$ . The Gillespie algorithm is implemented for an initial mass of free monomers  $m_0 = 1 \mu\text{M}$ . The aggregation curves obtained are shown in figures 8 and 9.

Generating complete aggregation curves using the Gillespie algorithm can be computationally demanding. Indeed, the time step  $\tau$  is inversely proportional to  $a_0$ , causing the time increments to become extremely small as the number of fibrils increases. To bypass this limitation, we consider a system in which the supply of free monomers is effectively unlimited. In this approximation, the reaction does not reach completion and the characteristic sigmoidal aggregation curve is not observed. However, this approach remains valid at

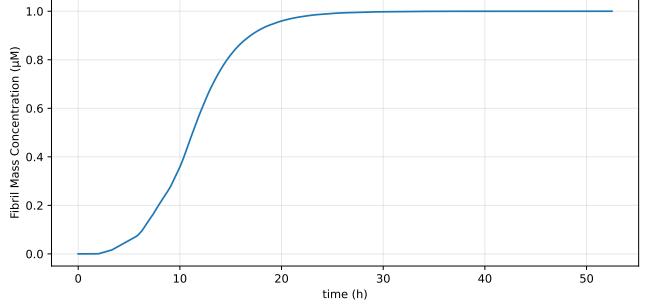


FIG. 8. Evolution of the fibril mass concentration as a function of time. One trajectory was obtained with the Gillespie algorithm for  $k_+ = 10^6 \text{ M}^{-1}\text{s}^{-1}$ ,  $k_2 = 10^4 \text{ M}^{-2}\text{s}^{-1}$ ,  $k_n = 10^{-4} \text{ M}^{-1}\text{s}^{-1}$ ,  $n_2 = n_c = 2$ ,  $m_{\text{tot}} = 10^{-6} \text{ M}$ ,  $V = 10^{-12} \text{ L}$ .

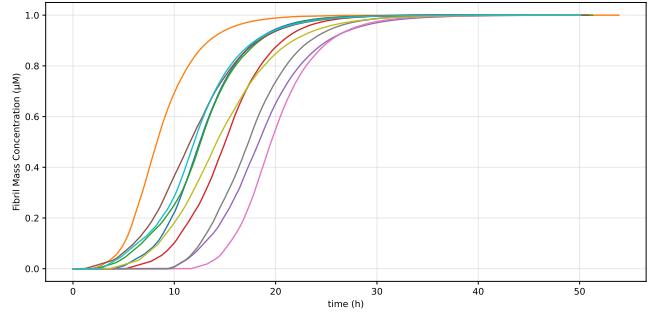


FIG. 9. Evolution of the fibril mass concentration as a function of time. 10 trajectories were obtained with the Gillespie algorithm for  $k_+ = 10^6 \text{ M}^{-1}\text{s}^{-1}$ ,  $k_2 = 10^4 \text{ M}^{-2}\text{s}^{-1}$ ,  $k_n = 10^{-4} \text{ M}^{-1}\text{s}^{-1}$ ,  $n_2 = n_c = 2$ ,  $m_{\text{tot}} = 10^{-6} \text{ M}$ ,  $V = 10^{-12} \text{ L}$ .

early times, as it corresponds to assuming a constant concentration of free monomers. Consequently, it still allows access to key information about the system dynamics.

Hence, we can consider the kinetic rates as described in Equations 14 to 16, and implement a Gillespie Algorithm at early times followed by a tau-leaping approximation at later times in order to speed up the reaction. These algorithms were run a large number of times, and the resulting trajectories, mean fibril mass concentration, as well as the distributions of fibril mass concentration and fibril number were plotted in figure 10 to 12.

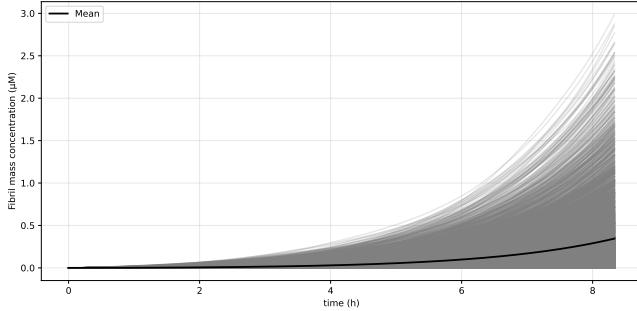


FIG. 10. Evolution of the fibril mass concentration as a function of time. 5000 trajectories were obtained with the Gillespie algorithm for  $k_+ = 10^6 \text{ M}^{-1}\text{s}^{-1}$ ,  $k_2 = 10^4 \text{ M}^{-2}\text{s}^{-1}$ ,  $k_n = 10^{-4} \text{ M}^{-1}\text{s}^{-1}$ ,  $n_2 = n_c = 2$ ,  $m_{\text{tot}} = 10^{-6} \text{ M}$ ,  $V = 10^{-12} \text{ L}$ .

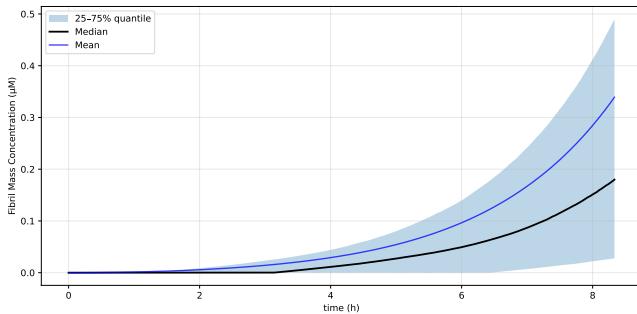


FIG. 11. Evolution of the fibril mass concentration as a function of time. 5000 trajectories obtained with the Gillespie algorithm for  $k_+ = 10^6 \text{ M}^{-1}\text{s}^{-1}$ ,  $k_2 = 10^4 \text{ M}^{-2}\text{s}^{-1}$ ,  $k_n = 10^{-4} \text{ M}^{-1}\text{s}^{-1}$ ,  $n_2 = n_c = 2$ ,  $m_{\text{tot}} = 10^{-6} \text{ M}$ ,  $V = 10^{-12} \text{ L}$ .

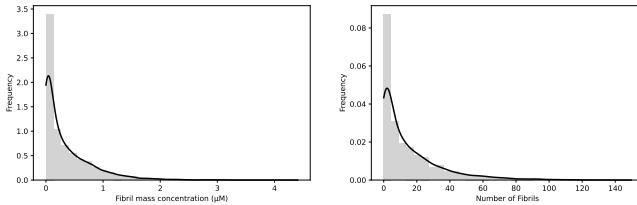


FIG. 12. left: Distribution of fibril mass concentration. Right: Distribution of fibril number. Both graphs were obtained with the Gillespie algorithm for  $k_+ = 10^6 \text{ M}^{-1}\text{s}^{-1}$ ,  $k_2 = 10^4 \text{ M}^{-2}\text{s}^{-1}$ ,  $k_n = 10^{-4} \text{ M}^{-1}\text{s}^{-1}$ ,  $n_2 = n_c = 2$ ,  $m_{\text{tot}} = 10^{-6} \text{ M}$ ,  $V = 10^{-12} \text{ L}$ .

b. *Aggregation curves for two morphs* In the case of two polymorphs, we consider the state  $(n_A, m_A, n_B, m_B)$  and the six reactions described in equations 11-13 can occur. We therefore have the following expression for  $a_0$

$$a_0 = \sum_{i=A,B} n_i \mu_i + m_i \alpha_{2,i} + \alpha_{1,i}.$$

As done above, a Gillespie algorithm is implemented for early times and switched to  $\tau$ -leaping at later times. Two distinct regimes are observed.

First, we identify a regime in which the two polymorphs coexist. In this regime, individual stochastic trajectories show increasing aggregation curves for both polymorphs, indicating that each species is able to elongate over time. Both polymorphs therefore grow simultaneously, without one being suppressed by the other. This behaviour is illustrated in figure 13.

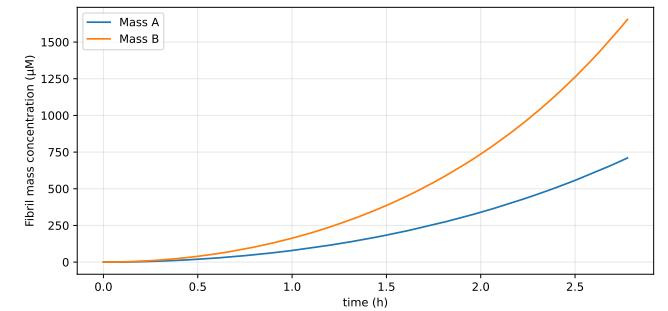


FIG. 13. Evolution of the fibril mass concentration as a function of time. The plot was obtained for the parameters  $k_{+A} = 10^6 \text{ M}^{-1}\text{s}^{-1}$ ,  $k_{+B} = 2 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$ ,  $k_{2i} = 10^4 \text{ M}^{-2}\text{s}^{-1}$ ,  $k_{ni} = 10^{-4} \text{ M}^{-1}\text{s}^{-1}$ ,  $n_2 = n_c = 2$ ,  $m_{\text{tot}} = 10^{-6} \text{ M}$ ,  $V = 10^{-12} \text{ L}$ .

When the algorithm is run over a large number of trajectories, rare events with strongly divergent behavior can be observed. However, the majority of trajectories remain within a similar range of values, resulting in relatively narrow confidence intervals. These intervals tend to broaden as the ratio between secondary and primary nucleation rates increases. In this regime, we performed simulations with  $N = 2000$  trajectories, and the resulting aggregation trajectories are shown in Fig. 14. Note that if we plotted only the fibril mass concentration over time, similar results are observed for the number of fibrils.

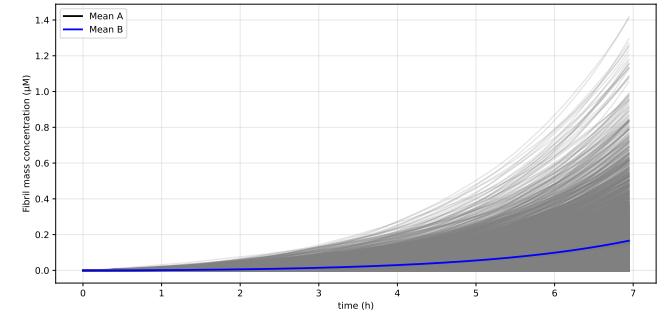


FIG. 14. Evolution of the fibril mass concentration as a function of time. 2000 trajectories were obtained with the Gillespie algorithm for  $k_{+i} = 10^6 \text{ M}^{-1}\text{s}^{-1}$ ,  $k_{2i} = 10^4 \text{ M}^{-2}\text{s}^{-1}$ ,  $k_{ni} = 10^{-4} \text{ M}^{-1}\text{s}^{-1}$ ,  $n_2 = n_c = 2$ ,  $m_{\text{tot}} = 10^{-6} \text{ M}$ ,  $V = 10^{-12} \text{ L}$ .

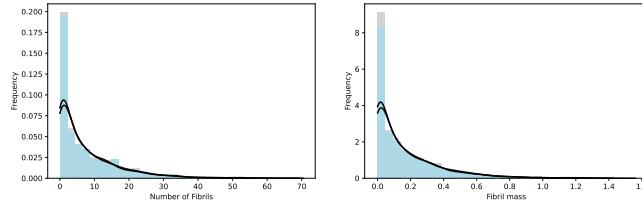


FIG. 15. left: Distribution of fibril mass concentration. Right: Distribution of fibril number. Both graphs were obtained with the Gillespie algorithm for  $k_{+i} = 10^6 \text{ M}^{-1} \text{s}^{-1}$ ,  $k_{2i} = 10^4 \text{ M}^{-2} \text{s}^{-1}$ ,  $k_{ni} = 10^{-4} \text{ M}^{-1} \text{s}^{-1}$ ,  $n_2 = n_c = 2$ ,  $m_{\text{tot}} = 10^{-6} \text{ M}$ ,  $V = 10^{-12} \text{ L}$ .

The second regime exhibits a markedly different behaviour. In this case, one polymorph rapidly dominates, effectively suppressing the growth of the other. Which polymorph prevails varies stochastically between simulations, with the probability of dominance determined by the respective kinetic rates of each species. This regime arises when there is a large disparity between the primary and secondary nucleation rates. Specifically, the secondary nucleation rate must be high enough that, once a seed of a polymorph forms through a rare primary nucleation event, secondary nucleation quickly amplifies it. Meanwhile, primary nucleation of the competing polymorph is sufficiently rare that it cannot establish a significant population and is effectively suppressed.

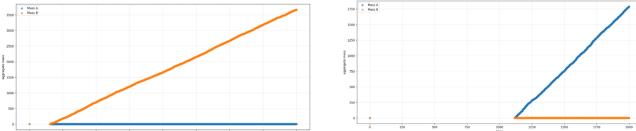


FIG. 16. Fibril mass concentration as a function of time.

When the experiment is repeated, very large confidence intervals are observed, reflecting this pronounced “all-or-nothing” behaviour. However, obtaining statistically robust results in this regime is challenging. The kinetic rates required place the system in a parameter range where aggregation proceeds extremely slowly, limiting our simulations to the early stages of the process. Moreover, increasing the number of stochastic realizations to obtain a reliable mean has proven difficult, as occasional trajectories with very quickly have very small times steps and never reach completion.

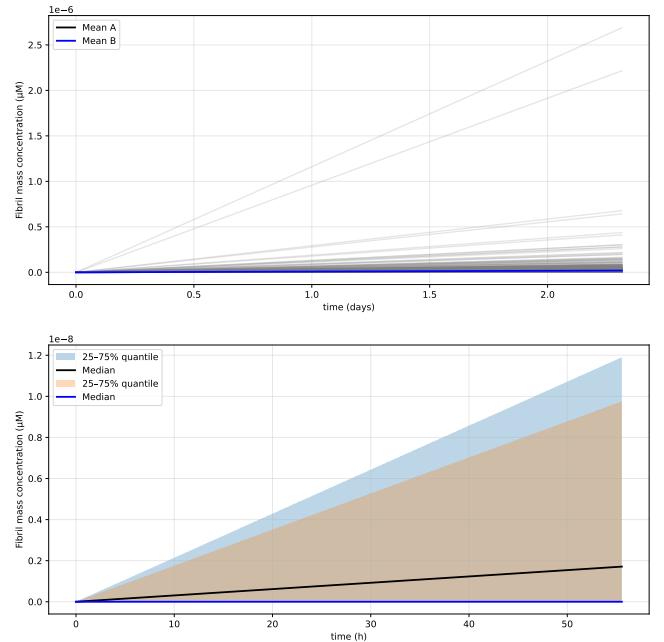


FIG. 17. Fibril mass concentration as a function of time. Both graphs were obtained with the Gillespie algorithm for 500 trajectories and the parameters  $k_{+i} = 10^7 \text{ M}^{-1} \text{s}^{-1}$ ,  $k_{2i} = 10^2 \text{ M}^{-2} \text{s}^{-1}$ ,  $k_{ni} = 10^{-8} \text{ M}^{-1} \text{s}^{-1}$ ,  $n_2 = n_c = 2$ ,  $m_{\text{tot}} = 10^{-6} \text{ M}$ ,  $V = 10^{-12} \text{ L}$ .

In this regime, the distributions of fibril mass and fibril number exhibit only two distinct outcomes, confirming the presence of a genuine switching behavior. In each simulation, either one polymorph rapidly forms one fibrils of about 3.3pM or it is completely overtaken by the competing polymorph. However, this behaviour reflects only the very early stages of the aggregation process and provides limited insight into the subsequent evolution of the system at later times.

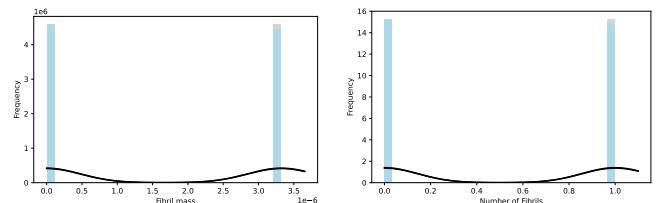


FIG. 18. left: Distribution of fibril mass concentration. Right: Distribution of fibril number. Both graphs were obtained with the Gillespie algorithm for  $k_{+i} = 10^7 \text{ M}^{-1} \text{s}^{-1}$ ,  $k_{2i} = 10^2 \text{ M}^{-2} \text{s}^{-1}$ ,  $k_{ni} = 10^{-8} \text{ M}^{-1} \text{s}^{-1}$ ,  $n_2 = n_c = 2$ ,  $m_{\text{tot}} = 10^{-6} \text{ M}$ ,  $V = 10^{-12} \text{ L}$ .

## V. CONCLUSION

In this report, we studied amyloid polymorphism using both mean-field and stochastic models of protein

aggregation. While the mean-field approach predicts coexistence under specific kinetic conditions, stochastic simulations reveal additional regimes driven by intrinsic fluctuations. In particular, we identified a coexistence regime and a switching regime characterized by an “all-

or-nothing” dominance of one polymorph, arising from rare early nucleation events. These results highlight the essential role of stochastic effects in amyloid formation and provide a framework for understanding variability observed in experiments.

- 
- [1] S. H. W. Scheres, B. Ryskeldi-Falcon, and M. Goedert, Molecular pathology of neurodegenerative diseases by cryo-em of amyloids, *Nature* **621**, 701–710 (2023).
  - [2] T. C. T. Michaels, D. Qian, A. Šarić, M. Vendruscolo, S. Linse, and T. P. J. Knowles, Amyloid formation as a protein phase transition, *Nature Reviews Physics* **5**, 379–397 (2023).
  - [3] F. De Giorgi, F. Laferrière, F. Zinghirino, E. Faggiani, A. Lends, M. Bertoni, X. Yu, A. Grélard, E. Morvan, B. Habenstein, N. Dutheil, E. Doudnikoff, J. Daniel, S. Claverol, C. Qin, A. Loquet, E. Bezard, and F. Ichas, Novel self-replicating -synuclein polymorphs that escape monitoring can spontaneously emerge and acutely spread in neurons, *Science Advances* **6**, 10.1126/sciadv.abc4364 (2020).
  - [4] T. Pálmaðóttir, C. A. Waudby, K. Bernfur, J. Christodoulou, S. Linse, and A. Malmendal, Morphology-dependent interactions between -synuclein monomers and fibrils, *International Journal of Molecular Sciences* **24**, 5191 (2023).
  - [5] J. Szavits-Nossan, K. Eden, R. J. Morris, C. E. MacPhee, M. R. Evans, and R. J. Allen, Inherent variability in the kinetics of autocatalytic protein self-assembly, *Phys. Rev. Lett.* **113**, 098101 (2014).
  - [6] 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* **135**, 10.1063/1.3608916 (2011).
  - [7] D. T. Gillespie, Exact stochastic simulation of coupled chemical reactions, *The Journal of Physical Chemistry* **81**, 2340–2361 (1977).
- 



FIG. 19. The author and her boyfriend making a pilgrimage to the Monte Carlo casino in Monaco in order to celebrate the fine numerical methods used in this project.