Stochastic Kinetic Study of Protein Aggregation and Molecular Crowding Effects of $A\beta_{40}$ and $A\beta_{42}$

John Bridstrup and Jian-Min Yuan

Department of Physics, Drexel University, Philadelphia, PA 19104, USA

John S. Schreck

National Center for Atmospheric Research (NCAR), Computational and Information Systems Lab, Boulder, CO, USA

(Dated: September 23, 2022)

Two isoforms of $\beta$-amyloid peptides, $A\beta_{40}$ and $A\beta_{42}$, differ from each other only in the last two amino acids, I and A, at the end of $A\beta_{42}$. They, however, differ significantly in their ability in inducing Alzheimer’s disease (AD). The rate curves of fibril growth of $A\beta_{40}$ and $A\beta_{42}$ and the effects of molecular crowding have been measured in in vitro experiments. These experimental curves, on the other hand, have been fitted in terms of rate constants for elementary reaction steps using rate equation approaches. Several sets of such rate parameters have been reported in the literature. Employing a recently developed stochastic kinetic method, implemented in a browser-based simulator, $popsim$, we study to reveal the differences in the kinetic behaviors implied by these sets of rate parameters. In particular, the stochastic method is used to distinguish the kinetic behaviors between $A\beta_{40}$ and $A\beta_{42}$ isoforms. As a result, we make general comments on the usefulness of these sets of rate parameters.

I. INTRODUCTION

Protein aggregation into amyloids is often the culprit for neurodegenerative diseases such as Alzheimer’s disease (AD) and Parkinson’s disease. There are about 20 such cases [1][2]. In the earlier version of the amyloid hypothesis of Alzheimer’s disease [3], it was assumed that the self-assembly of $\beta$-amyloid peptides, such as $A\beta_{40}, A\beta_{42}$, into amyloids and plaques was the main cause of the disease, but later studies show that the most neurotoxic species are intermediate soluble oligomers and the amyloid plaque itself may even have some protective functions of the neurons [4]. But the specific sizes of the neurotoxic oligomers are still not unambiguously identified. $A\beta_{42}$ is considered to be a bio-marker for AD, for it was discovered that AD patients tend to have a lower concentration of $A\beta_{42}$ in their cerebrospinal fluids (SPF), or even better, lower $A\beta_{42}/A\beta_{40}$ ratio [5]. In fact, recently, US FDA approved (on May 4, 2022) the marketing of first diagnostic test for early detection of amyloid plaques associated with AD, Lumipulse G $\beta$-amyloid ratio (1-42/1-40) test, using SPF for 55 and older patients. The two isoforms of $\beta$-amyloid peptides, $A\beta_{40}$ and $A\beta_{42}$, differ in the extra amino acids, I and A, in the amino acid sequence of $A\beta_{42}$. The amino acid sequence for $A\beta_{42}$ is: DAEFRHDSGYEVHHQKLVF-FAEDVGSNKAIILMVGGVVIA. Since I and A are hydrophobic amino acids, it is argued that this higher overall hydrophobicity of $A\beta_{42}$ tends to yield higher aggregation rates, compared to $A\beta_{40}$. Experiments, using copper and hydrogen peroxide induced cross-linking of unmodified proteins, the PICUP method, have indeed shown that polymer formation advances at a faster rate for $A\beta_{42}$ than $A\beta_{40}$ [6]. The results are also supported by Molecular dynamics studies using all-atom molecular dynamics (MD) [6] discrete molecular dynamics (DMD) [4]. The cellular environments of proteins are crowded with other biomolecules, such as DNA, RNA, lipids, and other proteins inside cellular compartments. This crowding due to other macromolecules, called macromolecular crowding [7][11], and also confinement [12], affect the reaction rates, properties, and functions of specific proteins of interest. More specifically, for the problem of interest here, macromolecular crowding [13] or confinement in the extracellular space in the brain [14][15] may affect the aggregation rates of $\beta$-amyloid peptides and affect the rates differently for $A\beta_{40}$ and $A\beta_{42}$.

Experimental data on the rates of fibril formation have often been fitted using rate equations [13][16][17], where mass-action laws are employed for the elementary chemical reactions taking place in very small volumes, stochastic kinetic approaches are more appropriate [20][23]. This is because, the concept of concentrations may not be well-defined when the fluctuations for the numbers of chemical species are too large in a tiny volume. In fact, fluctuations for many dynamical variables all can become large in these cases. The focus of the present work is on a stochastic kinetic approach that we have recently developed [22] and apply it to study the kinetic behaviors implied by the several sets of rate parameters for $A\beta_{40}$ and $A\beta_{42}$ have been reported through such fitting [13][18][19].

More recently it has been found that in the case of biochemical reactions taking place in very small volumes, stochastic kinetic approaches are more appropriate [20][23]. This is because, the concept of concentrations may not be well-defined when the fluctuations for the numbers of chemical species are too large in a tiny volume. In fact, fluctuations for many dynamical variables all can become large in these cases. The focus of the present work is on a stochastic kinetic approach that we have recently developed [22] and apply it to study the kinetic behaviors implied by the several sets of rate parameters for $A\beta_{40}$ and $A\beta_{42}$ mentioned above. The merit of such a stochastic kinetic study partly lies in the fact that during the growth period of the aggregates, the populations of dimers, trimers, tetramers, pentamers,... all of their populations start from zero, go thru small numbers and
eventually reach their steady-state numbers. such a process is more properly studied using a stochastic approach instead of a rate equation one, as discussed above.

The organization of the article is as follows: We introduce the theoretical background of the elementary reaction mechanisms and the effects of molecular crowding considered in both rate-equation and stochastic kinetic approaches. This is followed by a brief presentation of a stochastic kinetic approach (based on the rate-equation formulation) and a browser-based stochastic numerical simulator that we have developed in Section III. In Section IV, we present and discuss the results of our stochastic kinetic calculations on five sets of rate parameters for Aβ40 and Aβ42, along with a study of the effects of molecular crowding on these two isoforms. We conclude by a comment section.

II. THEORETICAL BACKGROUND AND SIMULATION SCHEMES

In this section we introduce the reader to popsim, a stochastic simulation application developed by the authors. We also give a brief overview of the models and methods implemented. popsim is a browser-based simulation package, written in Typescript, which generates stochastic pathways based on the reaction mechanisms of several well-known models for protein aggregation using the Gillespie method \[24\] [26]. It gives the user several visualization tools for the data, as well as the option to download data for more in-depth exploration. The application is user-friendly, and has use in both in-depth research studies as well as teaching and demonstration settings.

popsim has several built-in models of protein aggregation, each composed of a combination of several individual reaction mechanisms. Each mechanism describes a single reaction or a single "type" of reaction which may occur within the system volume. In this section, we will briefly describe the various reaction mechanisms in terms of their chemical reactions and show which combinations are included in the various models. Some of the reaction details can be found also in \[13\] [22] [27] and their connections to the stochastic kinetic model pointed out below as well as in \[22\].

A. Basic reaction mechanisms

1. Monomer addition and subtraction

Monomer addition, and its reverse, monomer subtraction, is the reaction mechanism through which protein aggregates grow and shrink by gaining or losing one monomer at a time. It is described by the chemical reaction

\[ M_1 + M_r \xrightarrow{a \ b} M_{r+1}, \]

where \( M_r \) is an aggregate composed of \( r \) monomers (an \( r \)-mer) and \( a \) and \( b \) are the rate constants for the forward reaction (addition) and reverse reaction (subtraction), respectively. These two mechanisms are, by themselves, a simple model of aggregation also known as the Becker-Doring model of nucleation \[28\] [29]. The rate equations describing these reactions are

\[
\frac{dc_r}{dt} = ac_r (c_{r-1} - c_r) + b(c_{r+1} - c_r),
\]

\[
\frac{dc_1}{dt} = - \sum_{r=n_c} c_r \frac{dc_r}{dt},
\]

where \( c_r \) is the concentration of \( r \)-mers and the associated propensity functions, which are weighting factors for specific reactions to be used in the Gillespie stochastic simulation algorithm \[22\] [24], are given by

\[
W_{r+} = a' N_1 c_r,
\]

for addition and

\[
W_{r-} = b' N_r,
\]

for subtraction, where \( N_r \) is the population of \( r \)-mers and the primes signify that we are using the stochastic rate constants, defined below.

2. Coagulation and fragmentation

In real systems, aggregates may also merge with (coagulate) or break from (fragment) each other. This is described by the chemical reaction

\[ M_r + M_s \xrightarrow{k_+} M_{r+s}, \]

where \( k_+ \) and \( k_- \) are the rate constants for coagulation and fragmentation, respectively. The set of rate equations describing the change in concentration of \( r \)-mers due to these reactions is given by

\[
\frac{dc_r}{dt} = k_+ \left[ \sum_{s=n_c} c_s c_{r-s} - \sum_{s=n_c} (1 + \delta_{rs}) c_r c_s \right] + k_- \sum_{s=2}^{\infty} (1 + \delta_{rs}) c_{r+s} - (r - 3)c_r,
\]

where \( \delta_{rs} \) is the Kronecker delta function, and the propensity functions are given by

\[
W_{rs+} = k'_+ N_r (N_s - \delta_{rs})
\]

for an \( r \)- and \( s \)-mer becoming an \( (r + s) \)-mer and

\[
W_{rs-} = k'_- N_{r+s}.
\]

The \( \delta_{rs} \) is to account for the fact that two of the same sized polymers may combine. Again primes denote the stochastic rate constants defined later.
3. Primary Nucleation

Primary nucleation process, by which a critical nucleus is formed from \( n_c \) monomers, is described by the chemical reaction

\[
n_c M_1 \xrightleftharpoons[k_n^*]{k_n} M_{n_c},
\]

where \( k_n \) is the primary nucleation rate constant and \( k_n^* \) is a model-dependent composite rate constant for the dissociation of sub-critical polymers. The rate equation describing how the concentration of critical nuclei changes is

\[
dc{n_c}{dt} = k_n c_{n_c}^n - \bar{k}_2 M_{n_c}.
\]

Thus here we neglect the reverse dissociation reaction, which is often dominated by the forward reaction. The propensity function for this reaction is

\[
W_n = k'_n \prod_{s=0}^{n_c-1} (N_1 - s),
\]

When primary nucleation is included in a model, it is assumed that no polymers smaller than \( n_c \) exist, as such the rate constant \( k_n^* \) depends on all other mechanisms which would produce a sub-critical polymer.

4. Secondary nucleation

Secondary nucleation, or heterogeneous nucleation, is an auto-catalytic process where new aggregates are formed on the surface of existing aggregates. Secondary nucleation can be modeled by a two-step process \[13\] given by the chemical reactions

\[
n_2 M_1 + M_S \xrightleftharpoons[k_f]{k_b} (M_{n_2} + M_S)_{\text{bound}} \xrightarrow[k_2]{\text{bound}} M_{n_2} + M_S,
\]

where \( M_S \) is a measure of the available surface active sites to which monomers can bind and \( k_f, k_b \) and \( k_2 \) are the rate constants for surface attachment and detachment of monomers, and the production of new secondary nuclei, respectively. An in-depth discussion of this mechanism can be found in \[13\], but it is worth noting that a limiting case of this mechanism is given by the catalytic chemical reaction

\[
n_2 M_1 + M_S \xrightarrow[k_2]{\text{bound}} M_{n_2} + M_S,
\]

otherwise known as one-step secondary nucleation. In one-step nucleation, the three rate constants \( k_f, k_b \) and \( k_2 \) reduce to a single rate constant for the production of nuclei, \( k_2 \). As of this writing, \textit{popsim} includes only one-step secondary nucleation, whose propensity function is given by

\[
W_2 = k'_2 N_S \prod_{s=0}^{n_2-1} (N_1 - s),
\]

where \( N_S \) is the number of surface active sites available to monomers.

B. Available models

\textit{popsim} includes three basic models of protein aggregation. They are combinations of mechanisms described in the previous section. None of the default models include the effects of crowders or secondary nucleation, but the user may select versions of the model which do.

1. Oosawa model

The Oosawa model \[27, 30\] is a combination of primary nucleation, monomer addition and monomer subtraction. It is the simplest model available, but describes very well many aggregating polymers such as actin \[13, 27\].

2. Smoluchowski model

The generalized Smoluchowski model combines coagulation and fragmentation with the Oosawa model. Due to the additional mechanisms, there are a wider variety of systems for which this model is useful \[11, 13, 31\]. Additionally, limiting cases of this model (ie. without fragmentation, or without coagulation) can be investigated by simply setting the appropriate rate constants to 0.

3. Smoluchowski with Becker-Doring nucleation

This is the most complex of the three models. It is a combination of simple monomer addition and subtraction up to a certain size (\( r < n_c \)), above which aggregates behave according to the generalized Smoluchowski model (\( r \geq n_c \)). Effectively, this model replaces Oosawa primary nucleation with a more accurate model of nucleation \[29, 32\].

C. Effects of macromolecular crowders

The effects of crowding molecules are included through use of formulas found by applying the scaled particle theory (SPT) of fluid mixtures \[9, 13, 27, 33–36\]. We give only the resulting equations in this section. Interested readers may refer to the references for a detailed derivation. Effectively, crowding plays prominent role only for
growth processes (nucleation, coagulation, etc.). For coagulation, the rate constant is found to be

\[ k_+ = \frac{\gamma}{\alpha} k^0_+ , \]

where \( k^0_+ \) is the rate constant without crowders and \( \gamma \) and \( \alpha \) are calculated from SPT. The rate constant for monomer addition is a specific case of the previous equation and is identical in form. The primary nucleation rate constant is

\[ k_n = \left( \frac{\gamma}{\alpha} \right)^{n_c - 1} k^0_n \]

and one-step secondary nucleation

\[ k_2 = \gamma^{n_2} \Gamma k^0_2 , \]

where superscript 0 again denotes the crowderless rate constant value and \( \Gamma \) is another factor from SPT [36]. For details on how to calculate the SPT parameters, the reader may refer to the references [13, 34, 35].

popsim assumes spherical crowders, spherical monomers and spherocylindrical aggregates. The variables required for simulation are:

1. monomer radius, \( r_1 \)
2. crowder radius, \( r_c \)
3. spherocylindrical radius, \( r_{sc} \), and
4. crowder volume fraction, \( \phi \).

Each model in popsim has the option to include the effects of crowders, making it a very powerful tool for exploring the dynamics of proteins within inherently crowded living cells.

III. GILLESPIE STOCHASTIC SIMULATION ALGORITHM

popsim implements the Gillespie stochastic simulation algorithm (SSA), which is a method for generating statistically correct trajectories of a system of molecules undergoing discrete reaction steps. It is particularly useful for investigating the behavior of small numbers of reacting particles (i.e. small system volumes), for which number fluctuations are comparable to the total number of reactants and continuous rate equations are inaccurate in describing the reaction processes. The algorithm is given in [22] and included here for completeness:

1. Set the initial species populations and set \( t = 0 \).
2. Calculate the transition rate or the propensity function, \( r_i \), for each of the \( M \) possible reactions.
3. Set the total transition rate \( Q = \sum_{i=1}^{M} r_i \).
4. Generate two uniform random, (0,1), numbers, \( u_1 \) and \( u_2 \).
5. Set \( \Delta t = \frac{1}{Q} \ln(\frac{1}{u_1}) \).
6. Find \( \mu \in [1, .., M] \) such that \( \sum_{i=1}^{\mu-1} r_i < u_2 Q < \sum_{i=1}^{\mu} r_i \).
7. Set \( t = t + \Delta t \) and update molecular species based on reaction \( \mu \).
8. Return to step 2 and repeat until an end condition is met.

A. Calculating stochastic rate constants

In the previous sections, we have discussed chemical species in terms of their bulk concentration in solution. Since popsim deals with discrete populations of species, it is necessary to relate the bulk rate constants, which are volume independent, to the volume-dependent stochastic rate constants. As an example, the bulk rate of change in the concentration of nuclei due to primary nucleation is given by

\[ \frac{dc_n}{dt} = k_n c_n , \]

where all \( c \) are in units of molar concentration. Using the relationship

\[ c_i = \frac{N_i}{N_A V} , \]

where \( N_i \) is the number of species of size \( i \), \( V \) is the system volume and \( N_A \) is Avogadro’s number, the rate of change of the number of nuclei due to primary nucleation is found to be

\[ \frac{dN_n}{dt} = k_n \left( \frac{c_0}{N_0} \right)^{n_c - 1} N_1^{n_c} . \]

Again, making use of the relationship between concentration and volume, the stochastic rate constant is found to be

\[ k'_n = \left( \frac{c_0}{N_0} \right)^{n_c - 1} k_n , \]

where \( N_0 \) and \( c_0 \) are the initial number and bulk concentration of monomers, respectively. The other stochastic rate constants (’ ) are given, relative to their bulk coun-
terparts, as follows:

\[
\begin{align*}
\alpha' &= \left( \frac{c_0}{N_0} \right) \alpha \\
\kappa_+ &= \left( \frac{c_0}{N_0} \right) \kappa_+ \\
\k'_f &= \left( \frac{c_0}{N_0} \right) k_f \\
\k'_2 &= \left( \frac{c_0}{N_0} \right) k_2 \\
b' &= b \\
\k'_- &= k_- \\
\k_b &= k_b \\
\k_2 &= \bar{k}_2.
\end{align*}
\]

popsim makes these calculations automatically, thus the user need only have bulk system parameters, which are much more readily available and interpreted.

B. Data visualization in popsim

Multiple ways to visualize data are provided in popsim. Time evolutions of the average aggregate mass, \(M\), average length of the filamentous aggregates, \(<L>\), and average number of fibrils, \(P\) can be visualized. The averages are taken over many individual runs. On the other hand, \(M\) and \(L\) values of individual runs as functions of time are also available to visualize. Furthermore, the standard deviations of \(M\), \(L\), and \(P\) as functions of time can also be plotted. In popsim you can also plot the histogram of aggregate size distribution as it evolves as a function of time. Finally, the reaction frequency or propensity plot to rank each reaction mechanism can be plotted at various reaction times.

IV. RESULTS AND DISCUSSION

In this section, using Gillespie’s stochastic method we study the kinetic behaviors of five sets of rate parameters reported in the literature. These sets of rate parameters, for convenience, have been labeled as Ab40, Ab42-JSS, Ab42-Cohen, Ab40-Meisl, and Ab42-Meisl, are listed in Table I. The first three sets of rate parameters, Ab40, Ab42-JSS, and Ab42-Cohen, are given in Ref. [13] and the last two sets are listed in Ref. [18]. We present and discuss the results that we obtain by carrying out simulations for these sets of rate parameters using the stochastic kinetic method, i.e., the popsim simulator, presented in the earlier sections.

A. Ab40 versus Ab42-JSS

In this subsection, we consider the differences in kinetic behaviors represented by the two sets of reaction parameters: Ab40 and Ab42-JSS. In general we examine the average behaviors of many individual stochastic runs in popsim to determine, for instance, whether a system has reached steady-state or equilibrium. As mentioned in previous section, individual runs can behave quite different from the average. A measure of such fluctuations is standard deviation of a certain quantity. One important indicator of the kinetic behavior is the halftime, \(t_{1/2}\), which is here defined by the time when half of the protein is found to be in aggregated form. To begin with, we study the fluctuations of the aggregate mass, \(M\). In popsim you can also plot the histogram of aggregate size distribution as it evolves as a function of time. Finally, the reaction frequency or propensity plot to rank each reaction mechanism can be plotted at various reaction times.
volume increases, $t_{1/2}$ decreases, that is, average aggregation rate speeds up. A statistical interpretation of this is that $t_{1/2}$ is bounded by $(0, \infty)$, thus increased importance of fluctuations will tend to push $t_{1/2}$ to the right at small enough $N$. With $\phi$, as the amount of crowders increases $t_{1/2}$ decreases due to entropy effect [11,13]. However, comparison between the kinetic behaviors of $\beta_40$ and $\beta_42$, based on these two sets of rate parameters leads to an unexpected result that it seems that $\beta_40$ has a higher aggregation rates than $\beta_42$. It seems to lead something that is contrary to what we would have expected, that is, $\beta_42$ had a higher aggregation rates than $\beta_40$. This could mean we would not be in favor of choosing $\beta_42$ as our set of parameters, until we examine other aspects of the reaction kinetics. Using popsim we can easily obtain the time evolution of the average length of the fibrils and that of the oligomer size distribution. For $\beta_40$, the steady-state value of the average length, $< L >$, is about 41.6 for $N = 1000$ to 40,000 and the peak size of the polymer size distribution is 46 around $N = 1000$ and 51 around $N = 5000$. On the other hand, for $\beta_42$, $< L >$ approaches only about 4.12 for $N = 1000$ to $< L > = 4.43$ for $N = 10,000$ and peak size of the size distribution is 4 for $N = 1000$, 8 for $n = 3000$, and 9 for $N = 5000$. Therefore, the average size of the fibrils or oligomers is much smaller in the case of $\beta_42$. Taking a closer look at the five sets of parameters, we can attribute the size distribution to two factors: $n_c = n_2 = 3$ for $\beta_42$ and $n_c = n_2 = 2$ for the other 4 sets of parameters and the rate for monomer subtraction is non-zero for $\beta_42$, but zero for the rest of parameter sets. Furthermore, it seems that half-life, or the incubation period, is mainly determined by the rate constant for the primary nucleation. The primary nucleation rate of $\beta_40$ is the largest among all five primary nucleation rates. As mentioned above, the values of $N$ and the amount of crowders present may cause the change of $t_{1/2}$ as well. For higher $\phi$ value, for example, $\phi = 0.2$, $t_{1/2}$ of $\beta_42$ becomes smaller than that of $\beta_40$.

Next, we return to the question whether we can justify the validity of the $\beta_42$ set of rate parameters in relation to the $\beta_40$ set. According to the current version of the amyloid hypothesis the small, soluble oligomers of $\beta$-amyloid peptides may be more neurotoxic than larger fibrils, such as amyloid plaques. The presence of the high population of the smaller oligomers, such as tetramers to octamers, in the steady-state in the $\beta_42$ set that we mentioned earlier would suggest the higher toxicity of $\beta_42$ compared to $\beta_40$. This certainly lends support to the validity of $\beta_42$ as a useful set of parameters for $\beta_42$.

As a side note, at a higher $\phi$ value the smaller $t_{1/2}$ values in the $\beta_42$ set than those of the $\beta_40$ set are observed. This fact may again help justifying the $\beta_42$ set. In the brain $\beta$-amyloid peptides are present in the space outside the brain cells. This space is confined and distorted, if not crowded. The confinement effects are to a certain extent similar to the effects of molecular crowding. In this way, the lower values of half-life at higher $\phi$ may to a certain extent lend support to the $\beta_42$ set of rate parameters.

B. $\beta_40$ versus $\beta_42$-Cohen

As discussed above, the primary nucleation rate constant of $\beta_40$ is greater than those of any other set. For $\beta_42$-Cohen, that rate is about one order of magnitude smaller, thus the halftimes are larger than those of $\beta_40$ for $N$ ranges from 500 to 10,000. For the range of $\phi$ investigated, from 0 to 0.2, halftimes of $\beta_42$-Cohen stay larger than those of $\beta_40$. For details, we refer to Fig. 3, where the halftime is plotted as a function of $N$ and $\phi$. Incidentally, for these two sets of rate parameters, $n_c$ and $n_2$ are both equal to 2. As in the case of $\beta_42$, we should check other features of reaction kinetics, such as average length of the fibrils. This does not help either, for
the steady-state $< L >$ value is about 56 for Ab42-Cohen and is about 41.6 for Ab40. The peak size of the fibril size distribution for Ab42-Cohen is about 61, compared to 46 for Ab40. Based on the above observations we may suspect the usefulness of the Ab42-Cohen parameter set for the peptide Aβ42. However, a chance of saving this set of parameters is at high N values. Using the popsim simulator, for $N = 40,000$ we obtain $t_{1/2} = 3.6$ hr using the Ab42-Cohen parameter set, compared to 9.3 hr for the Ab40 set. This arises from the fact that as $N$ increases the halftime drops down faster for the Ab42-Cohen set than for the Ab40 set.

C. Ab40-Meisl versus Ab42-Meisl

In this subsection, we examine results obtained using popsim to simulate the kinetic behaviors of the Ab40-Meisl and Ab42-Meisl sets of parameters and compare their behaviors against those of the other sets of parameters. To start with, both the sizes of the primary nucleus and secondary nucleus are 2 for the Ab40-Meisl and Ab42-Meisl sets. First, from Table I we see that the reaction rate of monomer addition of Ab40-Meisl is one order of magnitude greater than that of Ab42-Cohen and more than two orders of magnitude greater than those of Ab42-JSS and Ab40. The monomer addition rate of Ab42-Meisl is another order of magnitude greater than that of Ab40-Meisl. On the other hand, the rate of primary nucleation of Ab40-Meisl is the lowest among all sets of parameters, that of Ab42-Meisl is two orders of magnitude higher, which is in the middle of the range. Since the halftime is mainly determined by the rate of primary nucleation, its value for Ab40-Meisl shown in Table II is 9364 hr, or 390 days, longest among all sets of parameters, for $N = 3000$ and that for Ab42-Meisl as shown in Table III is 61 hr, or 2.5 days. The very large rate of monomer addition means that once the primary nucleus is formed, it will grow long very quick. This yields quite different result from the previous three sets of rate parameters. The stochastic kinetic simulations using popsim show in Tables II and III that there is essentially one fibril growing to about the full length. In these tables, $< L >$ denotes the experimental data value of the average fibril length and $P$ the average number of fibrils. In the presence of crowders, as $\phi$ increases the chance of creating additional fibrils increases. A similar effect is seen as $N$ increases. However, the number of fibrils remain small. As shown in Table III, the highest average fibril number obtained is 3.46, when $\phi$ increases to 0.1. In general, as the volume fraction of crowers increases, the secondary nucleation becomes more important. The number of aggregates, $P$, increases and the average length of aggregates, $< L >$, and the size of the peak of the aggregate size distribution, decrease. As $\phi$ increases above 0.15, secondary nucleation may even overtake monomer addition as the most important reaction mechanism. Due to the overwhelmingly large magnitude, monomer addition is usually the dominating reaction mechanism for the Aβ systems in, at least, the crowderless condition.

The halftimes shown in Table II for Ab40-Meisl and Table III for Ab42-Meisl seem to support the idea that these sets of rate parameters are useful for the simulations for Aβ40 and Aβ42, because the rate of growth rate of Aβ42 would be much higher than that of Aβ40. However, a word of caution is that these sets of rate parameters were obtained through fitting experimental data by solving rate equations [13]. The rate equations, which described how concentrations of chemical species evolve, are effective only when concentrations of chemical species are well-defined. A concentration, however, is a thermodynamic quantity, which is useful, only when the number of the chemical species involved is large, for the fluctuation is inversely proportional to the number of the chemical species. The rate equation approach should not be considered justified, if throughout a reaction process the numbers of chemical species remain close to unity.

V. COMMENTS

A. System volume

Stochastic simulations inherently deal with small numbers of particles and, thus, small volumes. It is relatively simple to calculate the volume of the system in question from initial conditions by rearranging the concentration relationship to find

$$V = \frac{N_0}{N_A c_0}.$$ 

Typically the smaller the simulation volume, the larger fluctuations will be relative to total number of particles. This is where popsim and stochastic simulation become increasingly useful.
TABLE I: All 5 sets of reaction rate, length and other parameters. All quantities with superscript 0 denote rates for the crowderless cases. The unit for $k_2$ is $\mu M^{-n_2} h^{-1}$.

|       | Ab40  | Ab42-JSS | Ab42-Cohen | Ab42-Meisl | Ab42-Meisl |
|-------|-------|----------|------------|------------|------------|
| $n_c$ | 2     | 3        | 2          | 2          | 2          |
| $n_2$ | 2     | 3        | 2          | 2          | 2          |
| $C_0^r$ | 5.0 $\mu M$ | 5.0 $\mu M$ | 5.0 $\mu M$ | 3.5 $\mu M$ | 3.5 $\mu M$ |
| $a^0$ | 3.6 $\mu M^{-1} h^{-1}$ | 8.34 $\mu M^{-1} h^{-1}$ | 85.0 $\mu M^{-1} h^{-1}$ | 10 800 $\mu M^{-1} h^{-1}$ | 10 800 $\mu M^{-1} h^{-1}$ |
| $b^0$ | 0     | 0        | 0          | 0          | 0          |
| $k_c^0$ | 0     | 0        | 0          | 0          | 0          |
| $k_2^0$ | 1.43 $\times 10^{-2} h^{-1}$ | 1.43 $\times 10^{-2} h^{-1}$ | 1.43 $\times 10^{-2} h^{-1}$ | 1.43 $\times 10^{-2} h^{-1}$ | 1.43 $\times 10^{-2} h^{-1}$ |
| $k_3^0$ | 1.19 $\times 10^{-5} $ $\mu M^{-1} h^{-1}$ | 9.91 $\times 10^{-8} /\mu M^{-2} h^{-1}$ | 1.72 $\times 10^{-6} $ $\mu M^{-1} h^{-1}$ | 1.72 $\times 10^{-6} $ $\mu M^{-1} h^{-1}$ | 1.72 $\times 10^{-6} $ $\mu M^{-1} h^{-1}$ |
| $k_f^0$ | 6.30 $\times 10^{-2} h^{-1}$ | 1.74 $\times 10^{3} h^{-1}$ | 0.276 $h^{-1}$ | 0          | 0          |
| $k_b^0$ | 30.38 $h^{-1}$ | 1.78 $\times 10^{3} h^{-1}$ | 0.124 $h^{-1}$ | 0          | 0          |
| $r_c^0$ | 1.304 $\times 10^{-3}$ | 1.68 $\times 10^{-4}$ | 0.0164 | 0          | 0          |
| $r_1$ | 2.26 nm | 2.26 nm | 2.26 nm | 2.26 nm | 2.26 nm |
| $r_{sc}$ | 2.4 $r_1$ | 2.5 $r_1$ | 2.0 $r_1$ | 2.0 $r_1$ | 2.0 $r_1$ |
| $A_1$ | 2.35 $\times 10^{-8} \text{nm}^{-xi}$ | 2.35 $\times 10^{-8} \text{nm}^{-xi}$ | 2.35 $\times 10^{-8} \text{nm}^{-xi}$ | 2.35 $\times 10^{-8} \text{nm}^{-xi}$ | 2.35 $\times 10^{-8} \text{nm}^{-xi}$ |
| $\xi$ | 0.67 | 0.67 | 0.67 | 0.67 | 0.67 |

TABLE II: Results for Ab40-Meisl. $t_{1/2}$ halftime; $\langle L \rangle$, peak size, and $P$ are steady-state values of aggregate length, peak size of the size distribution, and number of aggregates.

|       | $N = 3000$ | $N = 5000$ |
|-------|------------|------------|
| $\phi$ | 0 0.05 0.1 | 0 0.05 0.1 |
| $t_{1/2}$ | 9364 4733 2263 | 5235 3261 1354 |
| $\langle L \rangle$ | 2920 2690 1890 | 4640 3730 1900 |
| Peak Size | 2236 2087 1938 | 3736 3487 3238 |
| $P$ | 1 1.21 1.89 | 1.14 1.56 2.89 |

TABLE III: Results for Ab42-Meisl.

|       | $N = 3000$ | $N = 10000$ |
|-------|------------|------------|
| $\phi$ | 0 0.05 0.1 | 0 0.05 0.1 |
| $t_{1/2}$ | 61.29 35.25 16.44 | 19.95 7.69 4.53 |
| $\langle L \rangle$ | 2970 2890 2510 | 9170 7280 3180 |
| Peak Size | 2087 2236 2236 | 7486 7486 5490 |
| $P$ | 1.02 1.08 1.35 | 1.17 1.62 3.46 |

B. Size distribution

In this article we have examined five sets of rate parameters for the kinetic behaviors of Ab40 and Ab342. It turns out quite diverse results are predicted for the two peptides. For instance, the steady-state average aggregate size predicted (for $N = 3000$) for Ab42 are, respectively, 4.43, 45.3, 2970 based on the Ab42-JSS, Ab42-Cohen, and Ab42-Meisl sets of parameters. Because of such very different kinetic behaviors predicted, it is, in principle, easy to determine which set is more correct by carrying out one additional experimental measurement, for example, the aggregate size distributions. The tools for such measurements are already available, for example, atomic force microscopy [37] or cryo-EM methods [38]. Measurements of the size distributions should be the best way in the present case to determine which sets of rate parameters are the proper ones to use.

C. Availability

popsim can be accessed by anyone at www.popsim.xyz and is being continuously developed by the authors.

VI. ACKNOWLEDGMENTS

[1] C. M. Dobson, Protein folding and misfolding, Nature 426, 884 (2003).

[2] T. Knowles, C. Waudby, G. Devlin, S. Cohen, A. Aguzzi, M. Vendruscolo, E. Terentjev, M. Welland, C. Dobson, An analytical solution to the kinetics of break-
able filament assembly, Science 326, 1533 (2009).

[3] S. D. Hardy J, The amyloid hypothesis of Alzheimer’s disease: progress and problems on the road to therapeutics, Science 297, 353 (2002).

[4] B. Urbanc, Perplexity of amyloid β-protein oligomer formation: Relevance to Alzheimer’s disease, in Biophysics and biochemistry of protein aggregation, edited by J. M. Yuan and H. X. Zhou (World Scientific, 2017) pp. 1–50.

[5] P. Lewczuk, A. Matzen, K. Blennow, L. Parnetti, J. Molinuenco, P. Eusebi, J. Kornhuber, J. Morris, and A. Fagan, Cerebrospinal fluid Aβ42/40 corresponds better than Aβ42 to amyloid pet in alzheimer’s disease, J Alzheimers Dis. 55, 813 (2017).

[6] S. Mittal, R. K. Chowhan, and L. R. Singh, Macromolecular crowding: Macromolecules friend or foe, Biochim. Biophys. Acta 1850, 1822 (2015).

[7] R. J. Ellis, Macromolecular crowding: Obvious but underappreciated, Trends Biochem. Sci. 26, 597 (2001).

[8] A. P. Minton, Molecular crowding: Analysis of effects of high concentrations of inert cosolutes on biochemical equilibria and rates in terms of volume exclusion, in Methods in enzymology, edited by M. L. Johnson (Academic Press, 1998) pp. 127–149.

[9] A. Minton, The effect of time-dependent macromolecular crowding on the kinetics of protein aggregation: A simple model for the onset of age-related neurodegenerative disease, Front. Phys. 12 (2014).

[10] G. Peng, G. Yang, J. M. Yuan, Depletion force from macromolecular crowding enhances mechanical stability of protein molecules, Polymer 47, 2564 (2006).

[11] J. S. Schreck, J. Bridstrup, and J. M. Yuan, Kinetic studies of protein aggregation with and without the presence of crowders, in Biophysics and biochemistry of protein aggregation, edited by J. M. Yuan and H. X. Zhou (World Scientific, 2017) pp. 187–220.

[12] G. Peng, J. Yuan, M. Vallieres, H. Dong, Z. Sun, Y. Wei, F. Li, and S. Lin, Effects of confinement on protein folding and protein stability, The Journal of chemical physics 118, 8042 (2003).

[13] J. S. Schreck, J. Bridstrup, and J. M. Yuan, Investigating the effects of molecular crowding on the kinetics of protein aggregation, J. Phys. Chem. B. 124, 9829 (2020).

[14] C. Nicholson, The secret world in the gaps between brain cells, Physics Today 75, 26 (2022).

[15] C. Nicholson and P. Kamali-Zare, Reduction of dimensionality in monte carlo simulation of diffusion in extra-cellular space surrounding cubic cells, Neurochem. Res. 45, 42 (2020).

[16] G. Meisl, X. Yang, B. Frohm, T. P. Knowles, and S. Linse, Quantitative analysis of intrinsic and extrinsic factors in the aggregation mechanism of alzheimer-associated aβ-peptide, Scientific reports 6, 18728 (2016).

[17] G. Meisl, J. B. Kirkegaard, P. Arosio, T. C. Michaels, M. Vendruscolo, C. M. Dobson, S. Linse, and T. P. Knowles, Molecular mechanisms of protein aggregation from global fitting of kinetic models, Nat. Protoc. 11, 252 (2016).

[18] G. Meisl, X. Yang, E. Hellstrand, B. Frohm, J. B. Kirkegaard, S. I. A. Cohen, C. M. Dobson, S. Linse, T. P. J. Knowles, Differences in nucleation behavior underlie the contrasting aggregation kinetics of the aβ40 and aβ42 peptides, Proc. Natl. Acad. Sci. U.S.A. 111, 9384 (2014).

[19] S. I. Cohen, S. Linse, L. M. Lukeshi, E. Hellstrand, D. A. White, L. Rajah, D. E. Otzen, M. Vendruscolo, C. M. Dobson, and T. P. Knowles, Proliferation of amyloid-β42 aggregates occurs through a secondary nucleation mechanism, Proc. Natl. Acad. Sci. U.S.A. 110, 9758 (2013).

[20] J. Szavits-Nossan, K. Eden, R. J. Morris, C. E. MacPhee, M. R. Evans, and R. J. Allen, Inherent variability in the kinetics of autocalytic protein self-assembly, Phys. Rev. Lett. 113, 098101 (2014).

[21] N. S. Tiwari and P. van der Schoot, Stochastic lag time in nucleated linear self-assembly, J. Chem. Phys. 144, 235101 (2016).

[22] J. Bridstrup, J. S. Schreck, and J. M. Yuan, Stochastic kinetic treatment of protein aggregation and the effects of macromolecular crowding, J. Phys. Chem. B. 125, 6068 (2021).

[23] J. L. Shen, M. Y. Tsai, N. P. Schafer, and P. G. Wolynes, Modeling protein aggregation kinetics: The method of second stochasticization, J. Phys. Chem. B 125, 1118 (2021).

[24] D. T. Gillespie, Stochastic simulation of chemical kinetics, Annu. Rev. Phys. Chem. 58, 35–55 (2007).

[25] D. T. Gillespie, A rigorous derivation of the chemical master equation, Physica A 188, 404 (1992).

[26] D. T. Gillespie, Exact stochastic simulation of coupled chemical reactions, J. Phys. Chem. 81, 2340–61 (1977).

[27] J. Bridstrup and J. M. Yuan, Effects of crowders on the equilibrium and kinetic properties of protein aggregation, Chem. Phys. Lett. 659, 252 (2016).

[28] R. Becker and W. Doring, Kinetische behandlung der koimbildung in ubersattigten dampfern, Ann. Phys 24, 719 (1935).

[29] J. Wattis, An introduction to mathematical models of coagulation-fragmentation processes: A discrete deterministic mean-field approach, Physica D 222, 1 (2006).

[30] F. Oosawa and S. Asakura, Thermodynamics of the Polymerization of Protein (Academic Press, 1975).

[31] J. S. Schreck and J. M. Yuan, A kinetic study of amyloid formation: Fibril growth and length distributions, J. Phys. Chem. B. 117, 6574 (2013).

[32] J. A. Wattis, A becker-döring model of competitive nucleation, J. Phys. A 32, 8755 (1999).

[33] H. Reiss, H. L. Frisch, J. L. Lebowitz, Statistical mechanics of rigid spheres, J. Chem. Phys. 31, 369 (1959).

[34] D. Hall and A.P. Minton, Effects of inert volume-excluding macromolecules on protein fiber formation. I. Equilibrium models, Biophys. Chem. 98, 93 (2002).

[35] D. Hall and A.P. Minton, Effects of inert volume-excluding macromolecules on protein fiber formation. II. Kinetic models for nucleated fiber growth, Biophys. Chem. 107, 299 (2004).

[36] F. A. Ferrone, M. Ivanova, R. Jasuja, Heterogeneous nucleation and crowding in sickle hemoglobin: An analytic approach, Biophys. J. 82, 399 (2002).

[37] F. Ruggeri, T. Sneidieris, M. Vendruscolo, and T. Knowles, Atomic force microscopy for single molecule characterisation of protein aggregation, Arch Biochem Biophys. 664, 134 (2019).

[38] Y. Yang, D. Arseni, W. Zhang, M. Huang, S. Lövstam, M. Schweighauser, A. Kotecha, A. Murzin, S. Peak-Chew, J. Macdonald, I. Lavenir, H. Garringer, E. Gelpi, K. Newell, G. Kovacs, R. Vidal, B. Ghetti, B. Rykseldi-Falcon, S. Scheres, and M. Goedert, Cryo-em structures of amyloid-β 42 filaments from human brains, Science
375, 167 (2022).