Journal Name

ARTICLE TYPE

Cite this: DOI: 10.1039/xxxxxxxxxx

Estimation of the lag time in a subsequent monomer addition model for fibril elongation

Suzanne K. Shoffner [a] and Santiago Schnell [a,b,c,*]

Received Date
Accepted Date

DOI: 10.1039/xxxxxxxxxx

www.rsc.org/journalname

Fibrillogenesis, the production or development of protein fibers, has been linked to protein folding diseases. The progress curve of fibrils or aggregates typically takes on a sigmoidal shape with a lag phase, a rapid growth phase, and a final plateau regime. The study of the lag phase and the estimation of its critical timescale provide insight into the factors regulating the fibrillation process. However, methods to estimate a quantitative expression for the lag time rely on empirical expressions, which cannot connect the lag time to kinetic parameters associated with the reaction mechanisms of protein fibrillation. Here we introduce an approach for the estimation of the lag time using the governing rate equations of the elementary reactions of a subsequent monomer addition model for protein fibrillation as a case study. We show that the lag time is given by the sum of the critical timescales for each fibril intermediate in the subsequent monomer addition mechanism and therefore reveals causal connectivity between intermediate species. Furthermore, we find that protein fibrillation can exhibit a lag phase without a nucleation process. Our approach could be valuable for investigating the effects of intrinsic and extrinsic factors to the protein fibrillation reaction mechanism and provides physicochemical insights into parameters regulating the lag phase.

1 Introduction

Protein folding is vital to normal functioning of the cell. While most proteins have one or more native conformations, some proteins misfold into a non-native conformation, causing accumulation and ultimately the formation of amorphous aggregates, or in the case of amyloidogenic proteins, mature amyloid fibrils.[1,2] Fibrils generally provide a more stable conformation than the anomalous state due to the stabilization of cross-β-sheets by the peptide backbone.[3] Though the formation of dimers, trimers, and other larger oligomeric complexes are part of normal, healthy cell functioning, aberrant protein aggregates can be toxic and have been shown to have pathological consequences.[4] Protein aggregation has been linked with over 50 protein folding diseases and disorders, including type II diabetes and Alzheimer's, Parkinson's, and Huntington's disease.[2] Though significant progress has been made toward understanding the reaction mechanisms of protein aggregation for some diseases (i.e. Alzheimer's disease via Aβ propagation[5]), the large majority of proteins aggregate with mechanisms that remain to be identified. Fibrillogenesis is a complex multistep process, generally beginning with monomers or other small molecules that collide and bond to form larger molecules including oligomers and protofibrils, until the fibril sizes formed have reached equilibrium.[6] The equilibrium fibril size distribution can vary from strongly skewed distributions to broad distributions depending on a number of factors, including elapsed time, fragmentation effects, and aggregate merging.[7,8] Estimation of the rates of aggregation reactions will be important for not only for identification and overall understanding of aggregation mechanisms, but also for developing pharmacological treatment and strategies for disease prevention.

In a typical protein aggregation kinetic experiment, the time course of protein fibrillation is measured by the absorbance of light at one or more wavelengths using dyes and extrinsic fluorescent probes.[9] Time course data often follow the characteristic shape of a sigmoidal curve and are separated into three regions: a lag phase, a fast growth phase, and a plateau phase. The lag phase is of particular interest because it provides critical information about the factors regulating the fibrillation process. A major challenge is determining which molecular events regulate the lag phase in fibril formation.[10] There has been much debate on whether nucleation, growth, or both contribute to the lag phase of the sigmoidal-shaped curve.[11,12] Knowles et al.[13] emphasize

[a] *Department of Molecular & Integrative Physiology, University of Michigan Medical School, Ann Arbor, MI 48109, USA.*
[b] *Department of Computational Medicine and Bioinformatics, University of Michigan Medical School, Ann Arbor, MI 48109, USA.*
[c] *Brehm Center for Diabetes Research, University of Michigan Medical School, Ann Arbor, MI 48105, USA.*
[*] *Corresponding author: Fax: 734-232-8162; Tel: 734-615-8733; E-mail: schnells@umich.edu*

that the lag phase does not correspond to a particular molecular event and cannot be attributed to primary nucleation events alone.

The lag phase is studied with an empirical logistic (sigmoidal) function, which is used to estimate phenomenological parameters from fibrillation time course data. The basic logistic function gives the characteristic sigmoidal shape, but is limited to describing symmetrical progress curves.[14] The Gompertz function and the Richards function exhibit the asymmetrical sigmoidal shape commonly observed with fibrillogenesis.[15,16] The Richards function adds an additional parameter, v, to account for asymmetry:

$$F(t) = \frac{F_{\max}}{\left(1 + v\,\exp^{-k(t-t_m)}\right)^{1/v}} \tag{1}$$

where $F(t)$ is the fluorescence intensity at time t, $F_{\max}$ is the steady-state fluorescence at the plateau of the progress curve, t_m is the inflection time at which the growth rate reaches its maximum ($v_{\max}$), and v is a parameter that alters the asymmetry of the curve. A geometric representation of these parameters is shown in Figure 1. Eqn (1) and variations thereof are often used to estimate the empirical parameters for progress curves of protein aggregation experiments.[17–19]

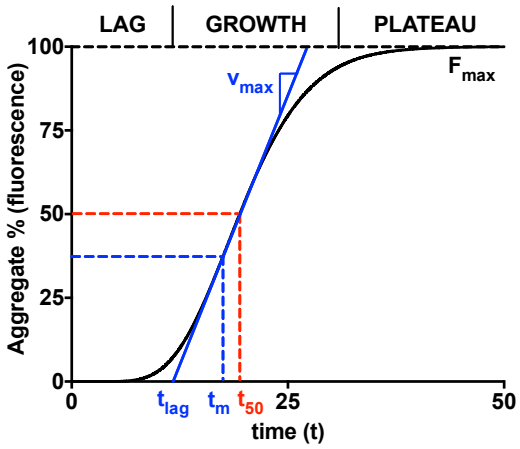

Fig. 1 The increasing concentration of aggregates as a function of time displays the characteristic sigmoidal curve for amyloid fibril formation. Aggregate concentration is represented as fluorescent intensity percentage of the final aggregates ($F_{\max}$). The curve is typically divided into the lag phase, the growth phase, and the plateau phase. The half-time t_{50} is the time at which half of the plateau aggregates are formed. The inflection time t_m is when the growth rate reaches its maximum, v_{max}. The lag time is then typically estimated by extending the tangent at t_m down to the time axis.

The lag time (t_{lag}) is typically estimated by extending the tangent at t_m to the initial baseline. The lag time is then given in terms of the empirical parameters:

$$t_{\text{lag}} = t_m - (1+v)/k\,. \tag{2}$$

Another type of critical timescale is the amount of time needed for the percentage of aggregate formed to reach a certain threshold. For example, t_{50} ($t_{1/2}$), often called the halftime, refers to the amount of time required to reach half of the maximum fluorescence intensity or percentage of the final aggregate. Occasionally, researchers will choose an arbitrary value, such as 10%, for the threshold. However, having multiple definitions for critical lag times introduces variability and uncertainty in estimates reported in the literature. At the same time, phenomenological time lag estimates are unable to provide a relationship between the model parameters and the elementary reaction steps governing the reaction mechanisms.

In amyloid studies, there can be billions of monomers combining to form oligomers, and the reaction scheme for this process may be very complex.[10] There are many possibilities for fragmentation and association when considering the reversible association of polymers.[20] Polymer fragmentation may also lead to secondary nucleation, since fragmented polymers may serve as nuclei for elongation.[13] Fragmentation and other molecular events, such as inhibition, and off-pathway aggregation, have been predicted to contribute significantly to the lag phase of fibril formation. For example, Pagano et al.[18] show that targeting the initial steps of fibrillation via kinetic inhibition reduces the concentration of early intermediate size oligomers. The onset of fibrillation relates to the concentration of unbound protein species in the presence of an inhibitor. In another study by Liu et al.,[19] the lag time for Aβ40 fibrillogenesis is slightly elongated in the presence of a protein aggregation inhibitor, chitosan. Inhibition kinetics may therefore be particularly relevant for research on aggregation prevention strategies. The lag time may also be affected by off-pathway aggregation that leads to the formation of toxic deposits. For example, a study by Crespo et al.[8] suggests that competitive off-pathway steps may be favored for monomer addition and therefore increase the lag time for fibrillation. Additionally, environmental factors such as pH and temperature can have an effect on the shape of the curve and the resulting estimated lag time.[21]

Reaction kinetic and thermodynamic models have sought to describe the process mechanistically. They range in complexity, from simpler subsequent monomer addition models to complex nucleation and elongation models.[22,23] In this paper, we propose a protocol for deriving an expression for the lag time from the governing rates of a subsequent monomer addition model post-nucleation. The goal is to find an analytical expression for the lag time in terms of the reaction parameters that is based on reaction kinetics rather than on empirical sigmoidal equations. By focusing on the elongation and growth stage of the process, we are able to determine whether growth alone is sufficient to produce the characteristic sigmoidal shape. This work introduces a novel approach to derive an expression for the lag time that may be more meaningful molecularly to the protein aggregation community, while allowing us to introduce the underlying molecular details of a subsequent monomer addition mechanism.

2 A subsequent monomer addition dock-lock reaction mechanism for fibril elongation

Since the initiation process by which native monomers misfold and nucleate is complex and largely based on random events, in

this paper, we focus on the highly organized elongation of the fibrils. Our model is a classic dock-lock mechanism for protein aggregation with one intermediate, where the monomer M first reversibly binds to a stem fibril species F_0, creating the complex C_1. The complex undergoes an irreversible conformational change to form fibril F_1. M reacts with each fibril, forming complexes and fibrils of increasing size until the final fibril (F_n) is synthesized. The reaction scheme of this fibril elongation reaction mechanism is given by

$$\begin{aligned} M + F_0 &\underset{k_-}{\overset{k_+}{\rightleftharpoons}} C_1 \quad \xrightarrow{k_c} F_1 \\ M + F_1 &\underset{k_-}{\overset{k_+}{\rightleftharpoons}} C_2 \quad \xrightarrow{k_c} F_2 \\ &\vdots \\ M + F_{n-1} &\underset{k_-}{\overset{k_+}{\rightleftharpoons}} C_n \quad \xrightarrow{k_c} F_n \end{aligned} \tag{3}$$

where k_+ and k_- are the on- and off- rate constants, respectively. The synthesis rate for the new fibrils is given by k_c. The term 'fibril' is used loosely to refer to the highly structured aggregate formed; however, the reaction scheme is also applicable to the formation of amorphous aggregates via monomer addition. The growth in size of the fibril chain is analogous to the growth in size of an amorphous aggregate, assuming that fragmentation is negligible and that addition occurs in any position. Therefore, the term 'fibril' can refer to either the growth of an amorphous aggregate or the growing ends of a highly structured fibril. Applying the law of mass action to reaction scheme (3) yields a nonlinear system of $2n+2$ ordinary differential equations (ODEs) of the form:

$$\frac{\mathrm{d}m}{\mathrm{d}t} = \textstyle\sum_{i=1}^{n} k_- c_i - \sum_{i=0}^{n-1} k_+ m f_i \tag{4}$$

$$\frac{\mathrm{d}f_0}{\mathrm{d}t} = -k_+ m f_0 + k_- c_1 \tag{5}$$

$$\frac{\mathrm{d}c_i}{\mathrm{d}t} = k_+ m f_{i-1} - (k_- + k_c) c_i \quad \text{with } i = 1 \ldots n \tag{6}$$

$$\frac{\mathrm{d}f_i}{\mathrm{d}t} = k_c c_i + k_- c_{i+1} - k_+ m f_i \quad \text{with } i = 1 \ldots n-1 \tag{7}$$

$$\frac{\mathrm{d}f_n}{\mathrm{d}t} = k_c c_n \tag{8}$$

where the lowercase indicates the concentration of that species. The initial conditions for the system are given by $(m, f_0, c_i, f_i)(t = 0) = (m_0, f_0^*, 0, 0)$ with $i = 1 \ldots n$.

The above system has two conservation laws: the total fibril and total monomer are conserved in the free and bound state. Assuming that the reaction is a closed and isolated thermodynamic system, we obtain a mass conservation law for the total fibril, given by a sum of the stem fibril species f_0 (free state) and the complex c_i and fibril f_i (bound states):

$$\frac{\mathrm{d}f_0}{\mathrm{d}t} + \sum_{i=1}^{n} \left(\frac{\mathrm{d}c_i}{\mathrm{d}t} + \frac{\mathrm{d}f_i}{\mathrm{d}t} \right) = 0 \tag{9}$$

which implies that

$$f_0 = f_t - \sum_{i=1}^{n} (c_i + f_i) \tag{10}$$

where f_t is the total concentration of fibril. Given that the monomer is also carried through the process in the free or bound form, there is a second conservation law for free monomer and the monomer attached to a complex or fibril:

$$\frac{\mathrm{d}m}{\mathrm{d}t} + \sum_{i=1}^{n} \left[i \left(\frac{\mathrm{d}c_i}{\mathrm{d}t} + \frac{\mathrm{d}f_i}{\mathrm{d}t} \right) \right] = 0 \tag{11}$$

which implies that

$$m_0 = m + \sum_{i=1}^{n} \left[i (c_i + f_i) \right] . \tag{12}$$

We can use these conserved quantities to simplify the ordinary differential equation system, thereby reducing its dimension. We combine the rate constants into physically meaningful parameters, where $K_m^* = (k_- + k_c)/k_+$ is the apparent dissociation constant of the monomer from the complex, $K_m = k_-/k_+$ is the dissociation constant of the monomer from the complex, and $K = k_c/k_+$ is the fibrillation constant.[24]

3 Estimation of the fibrillation time lag

From the physicochemical point of view, the intermediate complex species are short-lived and react quickly during an initial transient, t_c, of reaction mechanism (3). During this period, we assume (**Assumption I**) that the concentration of the monomer and the concentration of the stem fibril remain approximately constant. For $t < t_c$,

$$m \approx m_0 \tag{13}$$

$$f_0 \approx f_0^* \tag{14}$$

This assumption is known as the reactant stationary approximation, and the conditions for its validity are presented in Section 4.1. Using the reactant stationary approximation, we can estimate the timescale for a significant increase in the concentration of the intermediate fibrils (t_{f_i}) using a mathematical scaling and simplification technique similar to Segel.[25] However, we first need to derive a solution for the time course of the complex during the initial transient ($t < t_c$). We substitute eqn (13) into eqn (6) and obtain:

$$c_i(t) = c_{i_{\max}} \left[1 - \exp\left(-\frac{t}{t_c} \right) \right] \tag{15}$$

where $c_{i_{\max}}$ is the maximum concentration of the intermediate complex i during the reaction

$$c_{i_{\max}} = \frac{m_0 f_{i-1}}{K_m^*} \tag{16}$$

and t_c is the critical timescale of the intermediate complexes

$$t_c = (k_- + k_c)^{-1} . \tag{17}$$

After the initial transient ($t > t_c$), the complexes and the monomer start to be depleted. As the reaction progresses, the concentrations of each intermediate complex become significantly smaller than the initial monomer concentration, such that

$$c_i \ll m_0 \quad \text{with} \quad i = 1 \ldots n \; . \tag{18}$$

By applying eqn (18) to the conservation law for the monomer eqn (12), we find an expression for the concentration of the monomer after the initial transient

$$m \approx m_0 - \sum_{i=1}^{n} i f_i \; . \tag{19}$$

We can then approximate the depletion of the monomer at the timescale for each fibril (t_{f_i}) using the reaction stoichiometry. We assume (**Assumption II**) that the stem fibril F_0 is the limiting reactant and that the monomer is in excess. Since there is a stoichiometric ratio of 1:1 between the reactant F_0 and the first fibril product F_1, we know that the maximum number of molecules of F_1 produced will be equal to the the initial number of stem fibril molecules if the stem fibril is the limiting reactant. In the next step of the reaction, the monomer is again assumed to be in excess, and the maximum number of F_2 molecules produced from the limiting reactant F_1 will again be equal to the initial number of stem fibril molecules. This will continue for each step of the reaction, such that m is depleted by approximately the concentration of f_0^* for each step. Therefore, the approximate final steady-state concentration for the monomer will be equal to

$$m(t \to \infty) \approx m_0 - n f_0^* \tag{20}$$

and m at the timescale for each fibril t_{f_i} can be approximated by the following expression:

$$m(t_{f_i}) \approx m_0 - i f_0^* \; . \tag{21}$$

We then substituted eqn (15) into eqn (7), using eqn (21) for the depletion of m. The time course for the fibril after the complex reaches its maximum is then given by

$$f_i(t) = c_{i_{\max}} t_{f_i} k_c \left[1 - \exp\left(-\frac{t}{t_{f_i}} \right) \right] \tag{22}$$

where t_{f_i} is the timescale of the intermediate fibrils

$$t_{f_i} = \left[k_+ \left(1 - \frac{K_m}{K_m^*} \right) (m_0 - i f_0^*) \right]^{-1} \quad \text{with} \quad i = 1 \ldots n-1 \; . \tag{23}$$

The timescales for each species (t_c and t_f) are defined as the absolute value of the reciprocal of the eigenvalue for that species.[26] They are only dependent on kinetic parameters of the reaction mechanism and the initial concentration of monomer or stem fibril. The underlying physicochemical principle behind **Assumption II** is that each preceding intermediate must form before the subsequent fibril size is formed. Given that the fibrils are ultimately depleted and the monomer is in excess, we can then estimate the critical lag time for significant production of the final fibril f_n by summing the timescales for all preceding fibril intermediates:

$$t_{\text{lag}} = \sum_{i=1}^{n} t_{f_i} = \sum_{i=1}^{n} \frac{1}{k_+ (1 - \frac{K_m}{K_m^*})(m_0 - i f_0^*)} \; . \tag{24}$$

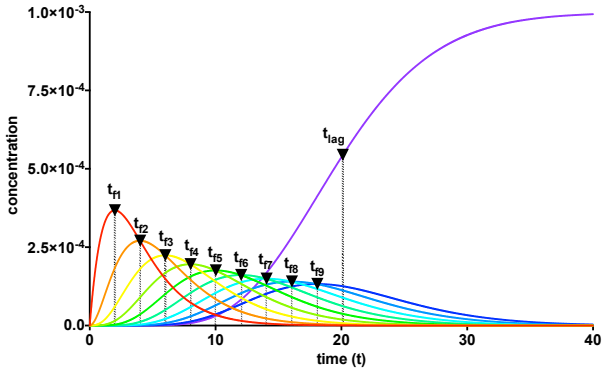

Fig. 2 The lag phase for fibril elongation is characterized by the cumulative sum of the timescales for each of the intermediate size fibrils, analogous to causal connectivity of falling dominos. Parameter values are: $K_m^* = 1000$, $K_m = 500$, $k_+ = 1$, $n = 10$, $m_0 = 1$, and $f_0 = 0.001$.

The lag time describes causal connectivity of intermediate fibrils in the reaction mechanism, showing that as the number of reaction steps between the final product and the initially perturbed species (the monomer and stem fibril species) increase, the separation time between the maxima of the intermediate fibrils increase. The subsequent monomer addition can be compared to a set of falling dominos; after an initial perturbation of the reactant, each subsequent species must wait for the previous species to form (or fall). The waiting time for each domino, given that the previous domino has just begun to fall, is analogous to the timescale of each intermediate fibril t_{f_i}. The lag time, equal to the total time before the last domino falls, is the sum of the waiting times for every domino, analogous to a significant formation of fibril f_n at t_{lag}. Upon plotting the intermediate fibrils, the 'domino effect' is observed, showing that the lag time is equal to the sum of the timescales required for each previous fibril to form. The estimated lag time appears similar to a halftime t_{50} definition, though it has a direct relationship with the physicochemistry and kinetics parameters of the reaction mechanism 3. The timescales for each fibril, calculated using eqn (23), and the lag time for the equilibrium fibril, calculated using eqn (24), are shown in Figure 2. The system of ODEs was solved using a variable order stiff differential equation solver in MATLAB®.

4 Conditions for validity of the lag time expression

In deriving the lag time, eqn (24), we made two assumptions. **Assumption I.** During the initial transient, $t < t_c$, the reactants (monomer and stem fibril) follow the reactant stationary approximation.[26] **Assumption II.** After the initial transient, $t > t_c$, the stem fibril is the limiting reactant and the monomer is in excess. Under what experimental conditions are these assumptions valid?

4.1 Reactant stationary approximation

The reactant stationary approximation states that the depletion of the reactant is negligible during the initial transient, and therefore, the reactant concentrations will remain approximately constant during this time.[26] We can mathematically derive expressions for $m \approx m_0$ and $f_0 \approx f_0^*$ during the initial transient by writing that the decrease in reactant concentration (Δm and Δf_0) is less than the product of the timescale of the initial transient (t_c) and the initial maximal rate of reactant depletion.[25] For the monomer, this expression is:

$$\left|\frac{\Delta m}{m_0}\right| = \frac{t_c}{m_0}\left|\frac{\mathrm{d}m}{\mathrm{d}t}\right|_{\max} \ll 1 \ . \tag{25}$$

Substituting in eqn (17) and the maximal depletion rate from eqn (4), we find that the reactant stationary approximation for the monomer is given by the following inequality:

$$= \frac{|-k_+ m_0 f_0^*|}{(k_- + k_c)m_0} = \frac{f_0^*}{K_m^*} \ll 1 \ . \tag{26}$$

Similarly, we assume that the stem fibril f_0 remains approximately constant during the initial transient ($t < t_c$):

$$\left|\frac{\Delta f_0}{f_0^*}\right| = \frac{t_c}{f_0^*}\left|\frac{\mathrm{d}f_0}{\mathrm{d}t}\right|_{\max} \ll 1 \ . \tag{27}$$

Substituting in eqn (17) and the maximal depletion rate from eqn (5), we obtain the reactant stationary approximation for the stem fibril:

$$= \frac{|-k_+ m_0 f_0^*|}{(k_- + k_c)f_0^*} = \frac{m_0}{K_m^*} \ll 1 \ . \tag{28}$$

The reaction must satisfy the reactant stationary approximation conditions for the monomer and the stem fibril, (eqns (26) & (28)), in order for the estimation of the lag time to be valid. Numerical confirmation for the validity of this approximation is shown in Figure 3. We guarantee that eqn (26) and eqn (28) were met for all simulations, such that $f_0^*/K_m^* < 0.1$ and $m_0/K_m^* < 0.1$.

4.2 Monomer in excess and stem fibril as the limiting reactant

Under **Assumption II**, we consider that the seed fibril is the limiting reactant and the monomer is in excess during the time course of the reaction. The excess monomer condition ensures that there is stoichiometrically a sufficient amount of monomer for the reaction to finish to completion. The maximum amount of fibril created from the first step is stoichiometrically equal to the amount of limiting reagent. If the monomer is the limiting reagent, the intermediate fibrils are formed but not depleted. In the particular case we are investigating, intermediate size fibrils are short-lived and ultimately depleted. To guarantee that this happens, we assume that the monomer is always in excess, taken from eqn (20), such that $m_0 \gg n f_0^*$. Then to test the strength of this condition, we add a factor α to describe the order of magnitude difference necessary to have stable conditions

$$\frac{m_0}{n f_0^*} \geq \alpha \ . \tag{29}$$

Eqn (29) allows us to test the validity of the lag time expression under different conditions of the ratio of reactant in excess to limiting reactant (m_0 to f_0^*), given that the conditions, eqns (26) and (28), for the reactant stationary approximation are met. With m_0 fixed to one and the length n fixed, for large values of f_0^*, such that $\alpha < 10$, the percentage of the maximum product reached at the lag time increases very rapidly and the excess monomer condition is unstable. When f_0^* is small such that $\alpha > 10$, the condition is very strong and there is very little variation in the lag time with changes to the ratio f_0^*/m_0. Additionally, when the monomer is greatly in excess (more than one order of magnitude larger), the percentage of the maximum concentration of f_n approaches a constant value of approximately 55%. Therefore, when the excess monomer condition is met such that $m_0/(nf_0) \geq \alpha$ for $\alpha > 10$, the percentage of the maximum concentration of f_n reached approaches a constant for each size n. As n increases, the percentage of the maximum reached at t_{lag} approaches 50%, as shown in Figure 3.

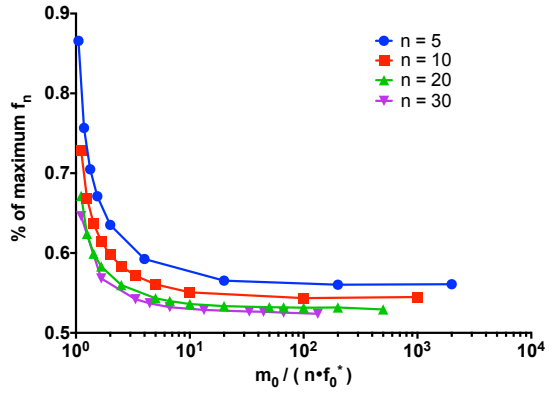

Fig. 3 Numerical confirmation of the conditions for validity of the time lag expression. The reactant stationary approximation is always met, such that $f_0^*/K_m^* < 0.1$ and $m_0/K_m^* < 0.1$. For the excess monomer condition ($m_0/(nf_0^*) \geq \alpha$), the variation in the percent of the maximum reached at t_{lag} is small. For $\alpha > 10^1$, the percentage is approximately 55%. The percentage of the maximum reached decreases at a decreasing rate as the length of the fibril n increases.

5 Factors affecting the length of the lag time

Now we focus our attention to reaction kinetic factors influencing the length of the lag phase by systematically varying each of the parameters in eqn (24).

The elongation and size of the longest fibril is determined by n. As expected, the longer the fibril, the longer it takes to form. The lag time increases linearly with fibril length, as shown in Figure 4A. For computational purposes, we selected small fibril sizes ($n = 10, 20, 30, 40, 50$), though fibrils can be composed of hundreds of thousands of monomers.

Changes to the kinetic constants also have an effect on the lag time. When examining how the ratio of the dissociation constant of the monomer from the complex K_m to the apparent dissociation constant of the monomer from the complex K_m^* alters the lag time, we considered how much of an effect the monomer-fibril association-dissocation reaction has with respect to the entire re-

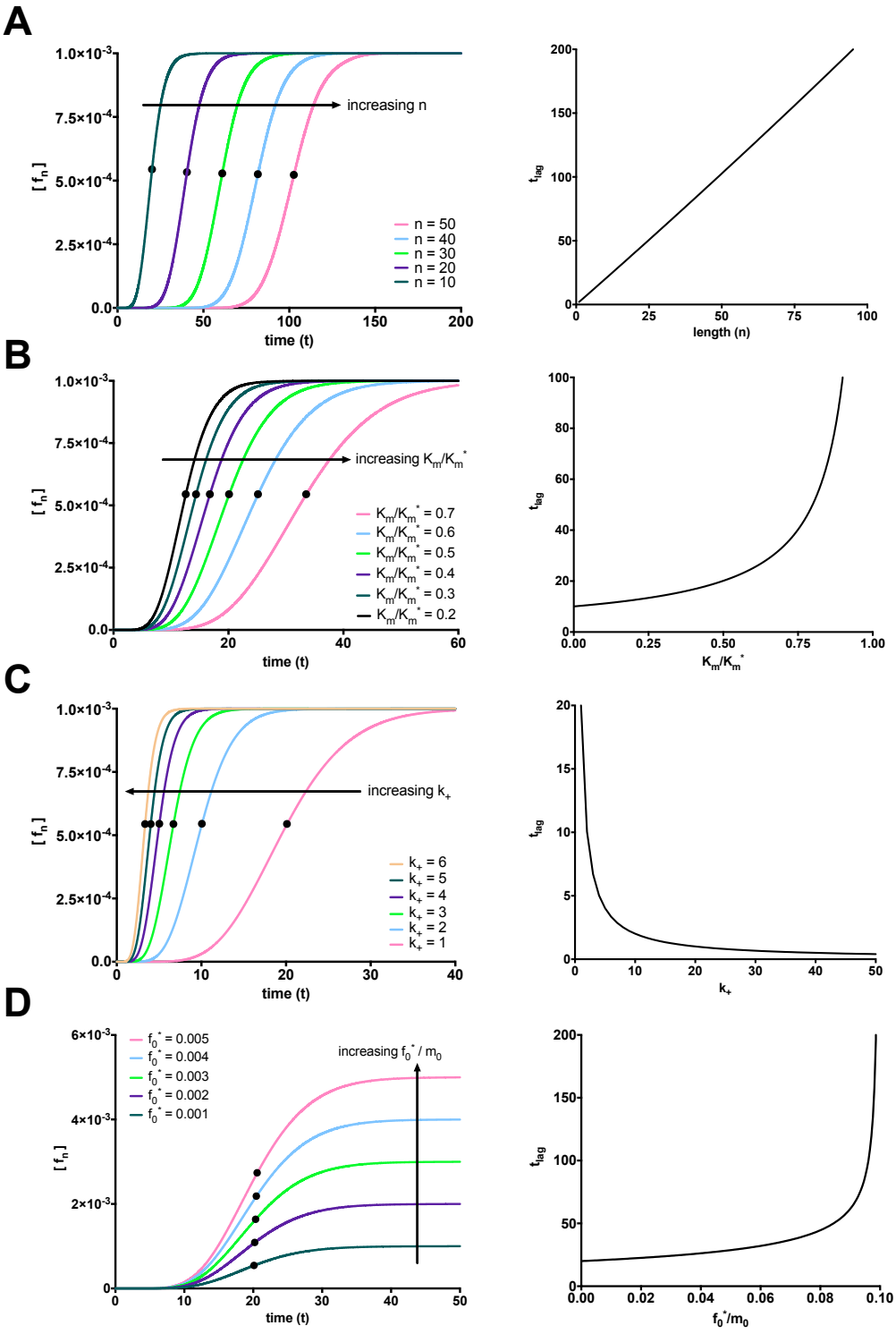

Fig. 4 Factors affecting the length of the fibrillation lag time. (A) The time lag increases linearly with increasing fibril length. (Parameter values: $K_m^* = 1000$, $K_m = 500$, $k_+ = 1$, $m_0 = 1$, and $f_0 = 0.001$). (B) The time lag increases exponentially with increasing ratio of K_m to K_m^*. (Parameter values: $K_m^* = 2000$, $k_+ = 1$, $n = 10$, $m_0 = 1$, and $f_0 = 0.001$). (C) The time lag decreases exponentially with increasing k_+. (Parameter values: $K_m^* = 1000$, $K_m = 500$, $n = 10$, $m_0 = 1$, and $f_0 = 0.001$). (D) Increasing the ratio of f_0^* to m_0 increases the plateau for the final fibril species (f_n) and has little effect on the lag time in the range of f_0^*/m_0 required for the condition of validity for eqn (29). (Parameter values: $K_m^* = 1000$, $K_m = 500$, $k_+ = 1$, $n = 10$, and $m_0 = 1$).

action mechanism. Increases to K_m/K_m^* lead to an exponential increase in the lag time. When the overall reaction mechanism is shifted towards monomer-fibril complex association-dissociation, the reaction mechanism is shifted left and will take longer to form the longest fibril. The lag time exhibits an exponential increase with increasing $\frac{K_m}{K_m^*}$, as shown in Figure 4B.

The on- rate constant (k_+) also affects the length of the lag time. An increase in the constant for association between a monomer and fibril (k_+) shifts the overall reaction to the right and ultimately promotes fibril formation. The lag time exhibits an exponential decrease with increasing k_+, as shown in Figure 4C.

We now turn our attention to the initial concentration of stem fibril f_0^* and the initial concentration of monomer m_0. Note that increasing the ratio of f_0^* to m_0 can violate the condition for monomer in excess and the stem fibril as the limiting reactant eqn (29). Therefore, we can only increase f_0^* within a range where this condition is not violated. In this range, according to the conservation law for the total fibril (eqn (10)), $f_0^* = f_t$ and $f_n(t \to \infty) = f_t$. The final fibril f_n approaches the initial concentration of the limiting reactant, the stem fibril f_0^*. Increasing f_0^* directly leads to a proportional increase in the plateau for the final fibril. The increase in $F_{\max}$ with increasing f_0^*/m_0 is illustrated in Figure 4D. When the excess monomer condition is met, there is a negligible, yet slight positive correlation between the lag time and the initial ratio of stem fibril to monomer.

6 Discussion

We have shown that it is possible to derive a lag time expression as a function of the kinetic parameters by writing the rate equations governing the reaction mechanism. In this work, we use the subsequent monomer addition reaction mechanism (3) as a case study. For this reaction mechanism, we define the time lag as the sum of all of the timescales for the intermediate fibril species. This expression is a result of the causal connectivity in the subsequent monomer addition model, in which a monomer is docked and locked into each intermediate, delaying the formation of the final fibril.

In our subsequent monomer addition model, the fibrillation process occurs without being preceded by nucleation. Many current nucleation-dependent polymerization models rely on high order kinetics and unlikely collisions between multiple monomers to create a nucleus.[12,13,27] It is widely believed that the time lag is a result of the nucleation process.[6,12] Interestingly, we found that the characteristic 'S' shape sigmoidal curve typical of a lag time fibrillation process can be observed independent of primary and secondary nucleation processes. This finding is in agreement with earlier work by Rangarajan and de Levie[28] and Baldassarre et al.[29] We observe a lag phase in the absence of nucleation, which shows that the lag time is not necessarily a 'wait time' for the nucleus to form or the result of secondary nucleation. Though elongation is sufficient to produce a lag phase, it is not necessarily the sole contributor to the lag time. The contributions of nucleation and elongation to the lag time should be further explored by investigating other reaction mechanisms.

The duration of the lag time exhibits a nonlinear relationship with respect to the kinetic parameters for the subsequent monomer reaction mechanism. In the first association-dissociation reaction, the equilibrium constant depends nonlinearly on the concentrations of the monomer, fibril, and intermediate complex. Therefore, small changes to the concentrations will trigger larger changes in the equilibrium constant, and consequently, in the lag time. For example, the lag time increases exponentially with increases to the ratio between the dissociation constant of the monomer from the complex K_m and the apparent dissociation constant of the monomer from the complex K_m^*, while the lag time decreases exponentially with increases to the on- rate constant, k_+ (see Figure 4B-C.).

We found that the lag time does not change significantly with increasing concentration of the stem fibril f_0^*. If a nucleation process is associated to our reaction mechanism, the nucleating species could be analogous to the stem fibril species. Several studies suggest that increasing the concentration of seed available (via secondary nucleation, sonication-induced fragmentation, or increased initial concentration) decreases the lag time of fibrillogenesis.[30–32] However, in our case, there is a slight increase in the lag time with increases in the ratio of f_0^* to m_0 when m_0 is in excess. The observed increase in lag time can likely be attributed to differences in the nature of reaction mechanism (3). In our model, the stem fibrils are depleted and not replenished to remain at a steady-state value. The addition of a nucleation process will change the derivation of the lag time.

Our approach to derive an analytical expression for the lag time could be used to investigate the factors affecting the length of the lag time for reaction mechanisms with off-path aggregation, inhibition kinetics, fragmentation, higher-order oligomerization, or other alternative pathways. Since the timescales of each reacting species are dependent on the kinetic parameters for a reaction mechanism, our protocol for estimating a lag time will serve as a basis for understanding the factors affecting complex protein aggregation and polymerization processes.

Acknowledgements

This work is supported by the University of Michigan Protein Folding Diseases Initiative.

References

1 F. Chiti and C. M. Dobson, *Annu Rev Biochem*, 2006, **75**, 333–366.
2 T. P. J. Knowles, M. Vendruscolo and C. M. Dobson, *Nat Rev Mol Cell Biol*, 2014, **15**, 384–396.
3 J. I. Guijarro, M. Sunde, J. A. Jones, I. D. Campbell and C. M. Dobson, *Proc Natl Acad Sci U S A*, 1998, **95**, 4224–4228.
4 R. W. Carrell and D. A. Lomas, *Lancet*, 1997, **350**, 134–138.
5 W. P. Esler, E. R. Stimson, J. M. Jennings, H. V. Vinters, J. R. Ghilardi, J. P. Lee, P. W. Mantyh, and J. E. Maggio, *Biochemistry*, 2000, **39**, 6288–6295.
6 A. M. Morris, M. A. Watzky and R. G. Finke, *Biochim Biophys Acta*, 2009, **1794**, 375–397.
7 J. S. Schreck and J.-M. Yuan, *J Phys Chem B*, 2013, **117**, 6574–6583.

8 R. Crespo, E. Villar-Alvarez, P. Taboada, F. A. Rocha, A. M. Damas and P. M. Martins, *J Biol Chem*, 2015, DOI: 10.1074/jbc.M115.699348.

9 J. T. Pedersen and N. H. H. Heegaard, *Anal Chem*, 2013, **85**, 4215–4227.

10 P. Arosio, T. P. J. Knowles and S. Linse, *Phys Chem Chem Phys*, 2015, **17**, 7606–7618.

11 M. Firestone, R. D. Levie and S. Rangarajan, *J Theor Biol*, 1983, **104**, 535–552.

12 M. A. Watzky and R. G. Finke, *J Am Chem Soc*, 1997, **119**, 10382–10400.

13 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, *Science*, 2009, **326**, 1533–1537.

14 P. Verhulst, *Correspondence Mathematiques et Physiques*, 1838, **10**, 113–121.

15 B. Gompertz, *Phil Trans R Soc Lond*, 1825, **182**, 513–585.

16 F. J. Richards, *J Exp Bot*, 1959, **10**, 290–301.

17 J. S. Pedersen, D. Dikov, J. L. Flink, H. A. Hjuler, G. Christiansen and D. E. Otzen, *J Mol Biol*, 2006, **355**, 501–523.

18 R. S. Pagano, M. Lopez Medus, G. E. Gomez, P. M. Couto, M. S. Labanda, L. Landolfo, C. D'Alessio and J. J. Caramelo, *Biophys J*, 2014, **107**, 711–720.

19 H. Liu, B. Ojha, C. Morris, M. Jiang, E. P. Wojcikiewicz, P. P. N. Rao and D. Du, *Biomacromolecules*, 2015, **16**, 2363–2373.

20 A. Wegner and P. Savko, *Biochemistry*, 1982, **21**, 1909–1913.

21 E. Sahin, W. F. t. Weiss, A. M. Kroetsch, K. R. King, R. K. Kessler, T. K. Das and C. J. Roberts, *J Pharm Sci*, 2012, **101**, 1678–1687.

22 A. Wegner and J. Engel, *Biophys Chem*, 1975, **3**, 215–225.

23 J. M. Andrews and C. J. Roberts, *J Phys Chem B*, 2007, **111**, 7897–7913.

24 S. Schnell and P. K. Maini, *Bull Math Biol*, 2000, **62**, 483–499.

25 L. A. Segel, *Bull Math Biol*, 1988, **50**, 579–593.

26 S. M. Hanson and S. Schnell, *J Phys Chem A*, 2008, **112**, 8654–8658.

27 F. A. Ferrone, J. Hofrichter, H. R. Sunshine and W. A. Eaton, *Biophys J*, 1980, **32**, 361–380.

28 S. Rangarajan and R. de Levie, *J Theor Biol*, 1983, **104**, 553–570.

29 M. Baldassarre, B. Maggiore, A. Scirè and F. Tanfani, *Biochimie*, 2015, **118**, 244–252.

30 Y. Ohhashi, M. Kihara, H. Naiki and Y. Goto, *J Biol Chem*, 2005, **280**, 32843–32848.

31 C. Dutta, M. Yang, F. Long, R. Shahbazian-Yassar and A. Tiwari, *J Phys Chem B*, 2015, **119**, 15089–15099.

32 V. V. Shvadchak, M. M. A. E. Claessens and V. Subramaniam, *J Phys Chem B*, 2015, **119**, 1912–1918.