Check for updates

OPEN

Stochastic master equation for early protein aggregation in the transthyretin amyloid disease

Ruo-Nan Liu & Yan-Mei Kang✉

It is significant to understand the earliest molecular events occurring in the nucleation of the amyloid aggregation cascade for the prevention of amyloid related diseases such as transthyretin amyloid disease. We develop chemical master equation for the aggregation of monomers into oligomers using reaction rate law in chemical kinetics. For this stochastic model, lognormal moment closure method is applied to track the evolution of relevant statistical moments and its high accuracy is confirmed by the results obtained from Gillespie's stochastic simulation algorithm. Our results show that the formation of oligomers is highly dependent on the number of monomers. Furthermore, the misfolding rate also has an important impact on the process of oligomers formation. The quantitative investigation should be helpful for shedding more light on the mechanism of amyloid fibril nucleation.

The aggregation of soluble proteins or protein fragments into non-soluble fibrillary polymers is a hallmark of a range of increasingly prevalent and devastating human disorders such as Alzheimer's disease, Parkinson's disease, Huntington's disease and familial amyloid polyneuropathy[1–4]. In these amyloidoses, dozens of proteins or protein components of disease-associated amyloid deposits have been identified so far, for instance, amyloid-beta peptide, alpha-synuclein, Huntingtin and transthyretin. Among these components, transthyretin, as a homotetrameric protein which is mainly synthesized in the liver, the choroid plexus and the retina, is implicated in several amyloid pathologies including familial amyloid polyneuropathy, familial amyloid cardiomyopathy, senile systemic amyloidosis and central nervous system selective amyloidosis[4–6].

Increasing studies suggest oligomers, the aggregation intermediate species, are correlated with the cellular toxicity in various forms of amyloidogenesis[7–9], which motivates the researchers to disclose how the oligomeric species are formed during the early stages of amyloid aggregation. In most of the known mechanisms for amyloid aggregation[10–19], the aggregation process is initiated by a coarse-grained "primary nucleation" reaction step, which proceeds via oligomeric intermediates[10–19]. Primary nucleation as a critical step in the amyloid formation cascade refers to the initial formation of nuclei through self-organization is characterized by the presence of a free energy barrier[20,21]. During this early stage, there are still no amyloid fibrils.

Mathematical model researches for amyloid aggregation[15,22–24] are helpful in shedding light on experimental observations and developing therapeutic strategies. Knowles et al.[15] developed an analytical solution to the kinetic of the complex self-assembly of filamentous molecular structures. Meisl et al.[22] described a framework to elucidate a molecular mechanism of protein aggregation by ways of quantitative kinetic assays and global fitting. Michaels et al.[24] presented an experimental and theoretical approach to drive the dynamics of oligomers during the aggregation of Alzheimer's Aβ42 peptide. It is worth emphasizing that most kinetic models of protein oligomers in the literature are deterministic.

Due to the intrinsic stochasticity in biochemical reaction[25–27], however, the deterministic model cannot always accurately capture the essential dynamics of amyloid aggregation[28,29]. Note that oligomers are the most toxic structures and play potential role as a target in drug discovery[30], so we take the early amyloid aggregation process of making oligomers as the main research focus, with transthyretin oligomers formed from the aggregation of at most six monomers[31,32]. To capture the stochastic effects in early amyloid aggregation, we build a mathematical model of chemical master equation, which is well accepted as probabilistic description in well-mixed and dilute

School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an 710049, Shaanxi, China. ✉email: ymkang@xjtu.edu.cn
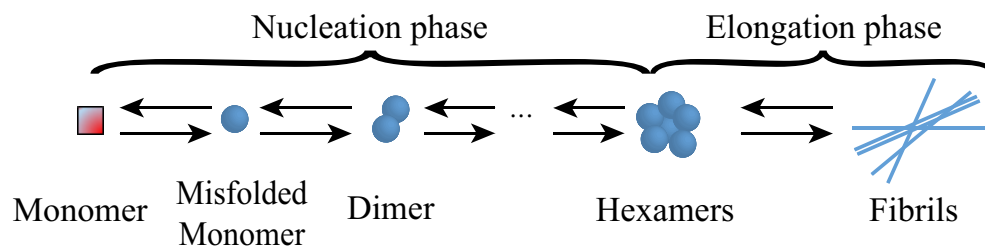
Figure 1. The two phases (nucleation phase and growth/elongation phase) of amyloid formation.

Reaction	Propensity function	Reaction	Propensity function
$M_0 \to M_1$	$K_0 M_0$	$M_1 \to M_0$	$K_6 M_1$
$M_1 + M_1 \to M_2$	$K_1 M_1 M_1$	$M_2 \to M_1 + M_1$	$K_7 M_2$
$M_1 + M_2 \to M_3$	$K_2 M_2 M_1$	$M_3 \to M_1 + M_2$	$K_8 M_3$
$M_1 + M_3 \to M_4$	$K_3 M_3 M_1$	$M_4 \to M_1 + M_3$	$K_9 M_4$
$M_1 + M_4 \to M_5$	$K_4 M_4 M_1$	$M_5 \to M_1 + M_4$	$K_{10} M_5$
$M_1 + M_5 \to M_6$	$K_5 M_5 M_1$	$M_6 \to M_1 + M_5$	$K_{11} M_6$

Table 1. Kinetics of monomers aggregation.

condition, to describe the aggregation of monomers into oligomers. And then, we use lognormal closure moment method rather than direct simulation to acquire the time dependent evolution of low-order statistical moments, and we emphasize that the semi-analytic moment method can greatly reduce the computational cost[33–36].

The paper is organized as following. In "Chemical master equation model" we describe the stochastic method of modeling the aggregation of monomers into oligomers and present moment closure method for computing the time evolution of stochastic models. In "Results", we apply the moment closure method to the stochastic model and examine the numerical results by comparing with the simulated results. Moreover, we investigate the impact of the number of monomers and the misfolding rate on the aggregation dynamics. The conclusions are drawn in last section.

Chemical master equation model

Build model of oligomers formation. Amyloid formation is considered to be a complex protein aggregation process, which consists of a range of molecular processes. In the process of protein filaments formation, primary nucleation is usually followed by growth[10–19] and self-replication through secondary pathways in some cases[11,17,18]. The nucleation phase corresponds to the period where monomers undergo conformational changes and self-associate to form the oligomeric nuclei. This phase is determined by the critical concentration of nuclei and generally considered to be thermodynamically unfavorable[21]. The growth/elongation phase represents the period in which the oligomeric nuclei acting as seeds rapidly grow and form mature fibrils. Considering the nucleation phase is an essential part of the overall aggregation process, as a large variety of oligomeric species are gradually formed[21], so we ignore elongation but focus on the process of nucleation below.

Following Refs.[32,37], we suppose that (1) the monomers undergo conformational changes into misfolded monomers which then are polymerized into small polymers involving diameter (two monomers), trimer (three monomers), etc.; (2) the small polymers polymerize (depolymerize) into bigger (smaller) structures by attaching (losing) one misfolded monomer; (3) the maximum oligomer size is limited to six since this is relevant for transthyretin oligomers[31,32]. Then, the process of making oligomers can be illustrated in Fig. 1. Let $M_0, M_1, M_2, \ldots, M_6$ denote monomers, misfolded monomers, dimers, triamers, tetramers, pentamers, hexamers, respectively. The mathematical scenario is shown in Table 1.

First, we briefly present the stochastic formulation inspired by Ref.[35]. If a model has L biochemical reactions and N molecular species, it can be written as

$$s_{l1}X_1 + \cdots + s_{lN}X_N \to r_{l1}X_1 + \cdots + r_{lN}X_N, \ 1 \le l \le L \tag{1}$$

where s_{li} and r_{li} are the coefficients of change of the molecular species X_i involved in the lth reaction. We define the stoichiometric matrix as $S = (S_{lj})$ where $S_{lj} = r_{lj} - s_{lj}$ denoting the change of the number of molecular species X_j by the lth reaction.

Under well-mixed and dilute conditions, if $P(\mathbf{x}, t)$ denotes the probability for the state vector $\mathbf{x} = (x_1, x_2, \ldots, x_N)$ at time t, then the probability evolution of the system (1) can be described by chemical master equation

$$\frac{dP(\mathbf{x},t)}{dt} = \sum_{l=1}^{L} P(\mathbf{x}-\mathbf{S}_l,t)a_l(\mathbf{x}-\mathbf{S}_l) - \sum_{l=1}^{L} P(\mathbf{x},t)a_l(\mathbf{x}) \tag{2}$$

where $\mathbf{S}_l = (S_{l1}, \ldots, S_{lN})$ is the row vector of the stoichiometric matrix and $a_l(\mathbf{x})$ is the propensity function for the lth reaction determined by mass action kinetics. For example, in chemical kinetics[38], for the reaction $s_1 X_1 + s_2 X_2 \rightarrow r_1 X_3 + r_2 X_4$, the reaction rate law expression is

$$a(x_1, x_2) = K \cdot x_1^{s_1} \cdot x_2^{s_2} \tag{3}$$

where K is the reaction constant.

Then, we establish the stochastic mathematical modeling of making oligomers. Denote the number of $M_0, M_1, M_2, M_3, M_4, M_5, M_6$ by $\mathbf{x}(t) = x_1(t), x_2(t), x_3(t), x_4(t), x_5(t), x_6(t), x_7(t)$ at time t. In modeling, the source term of monomers written as $x_1(t) \rightarrow x_1(t) + 1$ is also considered and the production rate is assumed to be k_p. For the system described in Fig. 1, the stoichiometric matrix is

$$S = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -2 & -1 & -1 & -1 & -1 & -1 & 2 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}^{T}$$

,then the chemical master equation of the model is

$$
\begin{aligned}
\frac{dP(x_1, x_2, x_3, x_4, x_5, x_6, x_7, t)}{dt} =\ & k_p P(x_1 - 1, x_2, x_3, x_4, x_5, x_6, x_7, t) \\
& + K_0(x_1 + 1)P(x_1 + 1, x_2 - 1, x_3, x_4, x_5, x_6, x_7, t) + K_1(x_2 + 2)^2 P(x_1, x_2 + 2, x_3 - 1, x_4, x_5, x_6, x_7, t) \\
& + K_2(x_3 + 1)(x_2 + 1)P(x_1, x_2 + 1, x_3 + 1, x_4 - 1, x_5, x_6, x_7, t) + K_3(x_4 + 1)(x_2 + 1) \\
& \times P(x_1, x_2 + 1, x_3, x_4 + 1, x_5 - 1, x_6, x_7, t) + K_4(x_5 + 1)(x_2 + 1)P(x_1, x_2 + 1, x_3, x_4, x_5 + 1, x_6 - 1, x_7, t) \\
& + K_5(x_6 + 1)(x_2 + 1)P(x_1, x_2 + 1, x_3, x_4, x_5, x_6 + 1, x_7 - 1, t) \\
& + K_6(x_2 + 1)P(x_1 - 1, x_2 + 1, x_3, x_4, x_5, x_6, x_7, t) + K_7(x_3 + 1)P(x_1, x_2 - 2, x_3 + 1, x_4, x_5, x_6, x_7, t) \\
& + K_8(x_4 + 1)P(x_1, x_2 - 1, x_3 - 1, x_4 + 1, x_5, x_6, x_7, t) + K_9(x_5 + 1)P(x_1, x_2 - 1, x_3, x_4 - 1, x_5 + 1, x_6, x_7, t) \\
& + K_{10}(x_6 + 1)P(x_1, x_2 - 1, x_3, x_4, x_5 - 1, x_6 + 1, x_7, t) + K_{11}(x_7 + 1)P(x_1, x_2 - 1, x_3, x_4, x_5, x_6 - 1, x_7 + 1, t) \\
& - (k_p + K_0 x_1 + K_1 x_2 x_2 + K_2 x_2 x_3 + K_3 x_3 x_4 + K_4 x_2 x_5 + K_5 x_2 x_6 + K_6 x_2 + K_7 x_3 + K_8 x_4 + K_9 x_5 + K_{10} x_6 + K_{11} x_7) \\
& \times P(x_1, x_2, x_3, x_4, x_5, x_6, x_7, t)
\end{aligned}
\tag{4}
$$

The above chemical master equation model describes the initial formation of oligomers in a stochastic sense. We remark that it is essential stochastic generalization of a time discrete deterministic model[32] and a continuous time deterministic counterpart[39] for the aggregation process. Actually, elegant analytical solutions have been developed to the time continuous model without an upper limit to oligomer size[39], and the time discrete model can be regarded as its simplification with an upper limit to the oligomer size.

Moment closure method. Although the chemical master Eq. (4) is a powerful tool for describing the stochastic dynamics of the system (1), it is difficult to find an exact analytic solution. Instead, some simulation techniques such as Gillespie's stochastic simulation algorithm (SSA) have been presented[40]. But the SSA is relatively expensive in computational cost especially when the system size is large. To overcome this shortcoming, moment closure techniques for approximating the low-order moments have become more and more popular[33,34,41–43]. This is acceptable because the first two order statistical moments (mean and variance or covariance) are sufficient for a decent description of the ensemble dynamics such as averaging behavior and evolution of the noise in the system[41].

The moment equations corresponding to Eq. (4) is not self-closed since the evolution of the Mth order moments depend on the $(M+1)$th order moments. That is, the evolution of the resultant first two order moments involves the third order moments. The so called moment closure is to approximate the involving higher order moments as nonlinear functions of lower-order moments. The frequently adopted closure schemes include Gaussian moment closure, lognormal moment closure, gamma moment closure and binominal moment closure. Here, we apply a kind of lognormal moment closure to Eq. (4) for two reasons. The first reason is that the important features of asymmetry and nonnegativity of the molecular reaction system make that population of species in bio-statistics tends to be lognormal distributed[34]. The second reason is that the lognormal closure scheme, also named the derivative matching closure as presented in Refs.[42,43], does not necessitate assumptions about the priori distribution. In this method, the nonlinear closure functions are given by matching time derivatives of the unclosed exact moment equations with those of the approximate closed moment equations at some initial point.

Given $\mathbf{m} = (m_1, m_2, \ldots, m_7) \in \mathrm{N}_{\geq 0}^7$, we define moment generating function to be

$$M(\boldsymbol{\theta}, t) = \sum_{\mathbf{x}} e^{\boldsymbol{\theta}\mathbf{x}} P(\mathbf{x}, t) = \sum_{\mathbf{m}=0}^{\infty} \mu_{\mathbf{m}} \frac{\boldsymbol{\theta}^{\mathbf{m}}}{\mathbf{m}!} \tag{5}$$

where $\mathbf{x^m} := x_1^{m_1} \cdots x_7^{m_7}, \frac{\boldsymbol{\theta^m}}{\mathbf{m!}} = \frac{\theta_1^{m_1}}{m_1!} \times \frac{\theta_2^{m_2}}{m_2!} \times \cdots \times \frac{\theta_7^{m_7}}{m_7!}, \mu_\mathbf{m} = E[\mathbf{x^m}] = \sum_\mathbf{x} \mathbf{x^m} P(\mathbf{x}, t)$ is the raw moment of $\mathbf{x}$ associated with $\mathbf{m}$ and the sum $\sum_{i=1}^{7} m_i$ is called the order of the raw moment. Then, multiplying Eq. (4) by $e^{\boldsymbol{\theta} \mathbf{x}}$ and summing over all possible values of $\mathbf{x}$, we obtain

$$\frac{d}{dt} \sum_\mathbf{x} P(\mathbf{x}, t) e^{\boldsymbol{\theta} \mathbf{x}} = \sum_{l=1}^{13} \left[(e^{\boldsymbol{\theta} \mathbf{S}_l} - 1) \sum_\mathbf{x} e^{\boldsymbol{\theta} \mathbf{x}} P(\mathbf{x}, t) a_l(\mathbf{x}) \right] \tag{6}$$

The detailed derivation can be found in supplementary method. With the moment generating function (5) in mind, the general form of moment equations is given by rewritten the Eq. (6) and extraction of the coefficients of $\theta_1, \theta_2, \ldots, \theta_7$

$$\frac{\partial}{\partial t} \mu_\mathbf{m} = \sum_{l=1}^{13} \left[\sum_\mathbf{i} a_{l,\mathbf{i}} \sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{m}} \mathbf{S}_l^\mathbf{k} \binom{\mathbf{m}}{\mathbf{k}} \mu_{\mathbf{i}+\mathbf{m}-\mathbf{k}} \right] - \sum_\mathbf{i} \sum_{l=1}^{13} a_{l,\mathbf{i}} \mu_{\mathbf{i}+\mathbf{m}} \tag{7}$$

where $a_{l,\mathbf{i}}$ is the multinomial coefficient of $a_l(\mathbf{x})$.

Since we mainly focus on the important stochastic quantities namely mean and variance, the system of $2C_7^1 + C_7^2 = 35$ equations for the first two order moments are derived shown in supplementary equations, from which we can see these equations depending on third order moments. For calculating, we truncate the infinite hierarchy (7) to a finite-dimensional system by the lognormal closure scheme. Following the lognormal closure scheme[43], the third order moments can be expressed by the nonlinear function of first two order moments. Assume $\mu_{\overline{\mathbf{m}}}(\overline{\mathbf{m}} = (\overline{m}_1, \ldots, \overline{m}_7))$ be one of the third order moments, then we can find a suitable closure function $\phi_{(\overline{\mathbf{m}})}(\cdot)$ with the separable form

$$\mu_{\overline{\mathbf{m}}} \approx \phi_{(\overline{\mathbf{m}})}(\boldsymbol{\mu}) = \prod_{p=1}^{s} (\mu_{\mathbf{m}_p})^{\gamma_p} \tag{8}$$

where $\boldsymbol{\mu} = [\mu_{\mathbf{m}_1}, \ldots, \mu_{\mathbf{m}_s}]$ is the vector of the first two order moments and $\gamma_1, \ldots, \gamma_s$ are chosen as the unique solution of the following set of linear equations

$$\binom{\overline{\mathbf{m}}}{\mathbf{m}_q} = \sum_{p=1}^{s} \gamma_p \binom{\mathbf{m}_p}{\mathbf{m}_q} \quad q = 1, \ldots, s \tag{9}$$

With the closure functions in Eq. (8) available, all the involving third-order moments can be expressed by the first two order moments, thus a self-closed moment system consisting of thirty-five ordinary differential equations can be deduced from Eq. (7). The detailed deduction based on the lognormal closure scheme (8) can be found from the supplementary. We emphasize that all the deductions are implemented by hand. We then apply the fourth order Ronge-Kutta method to the closed moment system and the first-order and the second-order moments are solved.

Results

It is generally hard to determine the unknown rate constants of oligomer growth. For simplicity, we assume all the polymerization rates are identical, namely $K_1 = K_2 = K_3 = K_4 = K_5 = K_a$ and all the polymerization rates are the same, i.e. $K_7 = K_8 = K_9 = K_{10} = K_{11} = K_b$. Although the identical assumption as an approximation to the nonlinear system is not that realistic, our numerical experience shows that the accuracy of the lognormal closure scheme does not rely on the relative variation of these rate parameters. As oligomeric intermediates in fibril formation are thermodynamically unstable[24,44,45], we specify equilibrium constant $k_e = K_a/K_b$ around 0.02, which ensures oligomers are suitably unstable compared to monomers. Since monomers are usually not produced in vitro experiments, we set the rate of monomer production $k_p = 0$. We assume the refolding rate (K_6) is larger than misfolding rate (K_0) in all the simulation. The initial conditions for the seven species are taken as $x_2(0) = \ldots = x_7(0) = 1$ and $x_1(0) = 2000$ not including Fig. 3.

Figure 2 shows the mean and the standard deviation we obtain for monomers, misfolded monomers, diamers, … and hexamers. It can be seen that the results derived from the lognormal closure are in good agreement with the results obtained from the SSA realization. This coincidence implies that the moment method is efficient and accurate in capturing the time evolution of mean and standard deviation. Thus, we only show the results obtained by the moment closure method in the other figures. As seen from the figure, all the reaction species can evolve into a steady state for suitable parameters.

Figure 3 shows that the quantitative evolution of the numbers of diamers, …, hexamers under different initial numbers of the monomers. It is shown from the picture that the initial number of monomers has same effect on different tapes of oligomers. That is, as the initial number of the monomers increases, the number of oligomers increases and the difference among initial values also undergoes a transient increase. This result implies that increasing the initial number of the monomers will speed up the process of nucleation phase. This observation supports the statement in Ref.[21] that the aggregation process is highly dependent on the number of monomers.
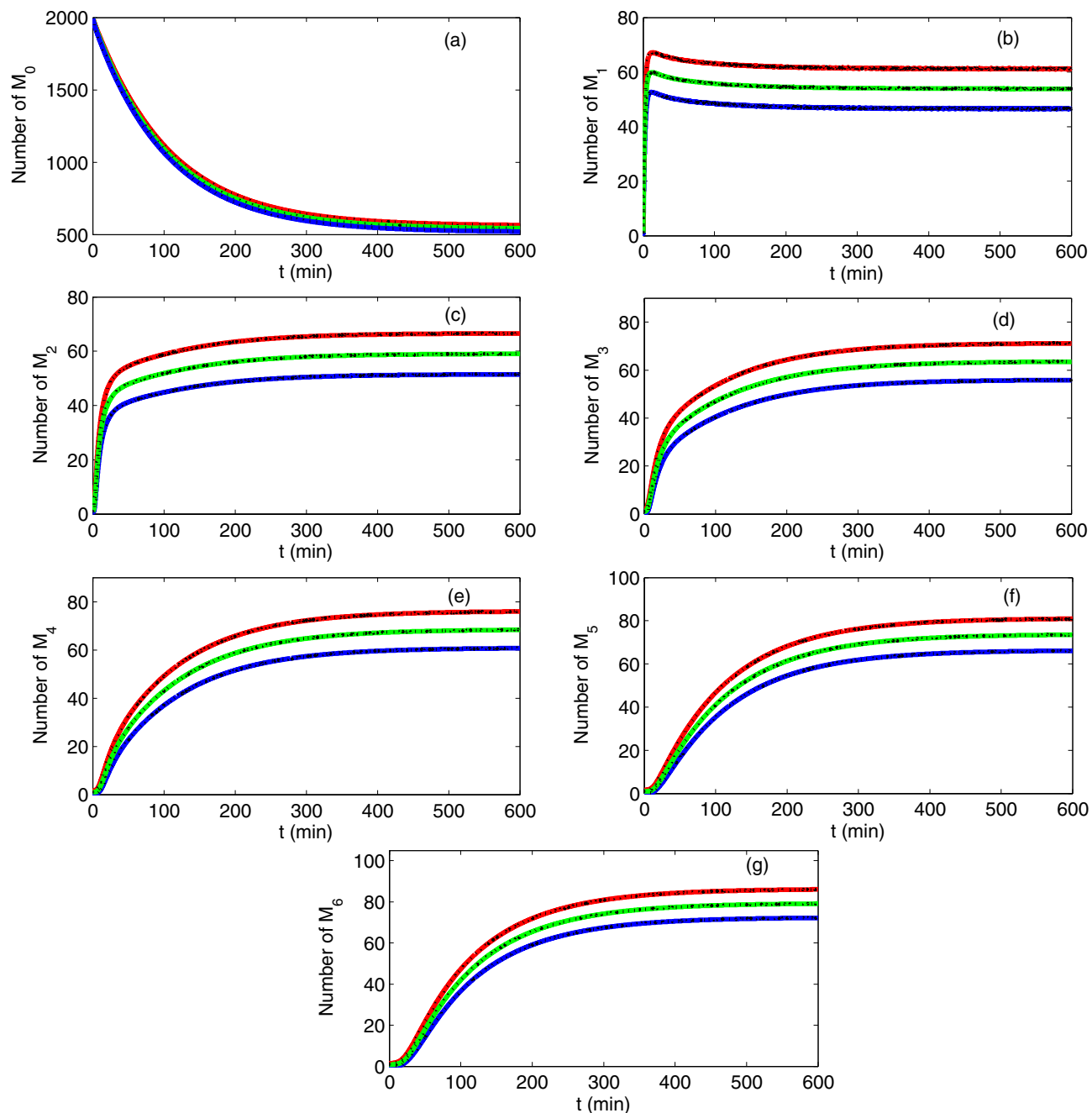
Figure 2. The time evolution of the number of $M_0, M_1, M_2, \ldots, M_6$. Comparison of the mean + standard deviation (upper red curve), mean (middle green curve) and mean − standard deviation (lower blue curve) by moment closure method (solid curves) and SSA (block dots). The initial condition is $x_1(0) = 2000$, $x_2(0) = \ldots = x_7(0) = 1$ and parameter values are $k_p = 0$, $K_0 = 0.01$ min^{-1}, $K_a = 0.002$ min^{-1}, $K_6 = 0.1$ min^{-1}, $K_b = 0.1$ min^{-1}. The results by SSA are based on 10,000 realizations.

Note that the number of the monomers of transthyretin is typically at steady state but may vary from person to person, thus the observation from Fig. 3 may be helpful in explaining why some people are more apt to suffer from transthyretin amyloid disease under same condition to some extent.

Figure 4 shows the influence of misfolding rate on the number of oligomers. This qualitative observation in Fig. 4 demonstrates that the misfolding rate has a major impact on the aggregation process, that is, as the misfolding rate increases, all the five types of oligomers grow in number, and the time for these oligomers to reach a given level also becomes shortened. Since the toxic oligomers accumulation plays a central role in amyloidoses[46],
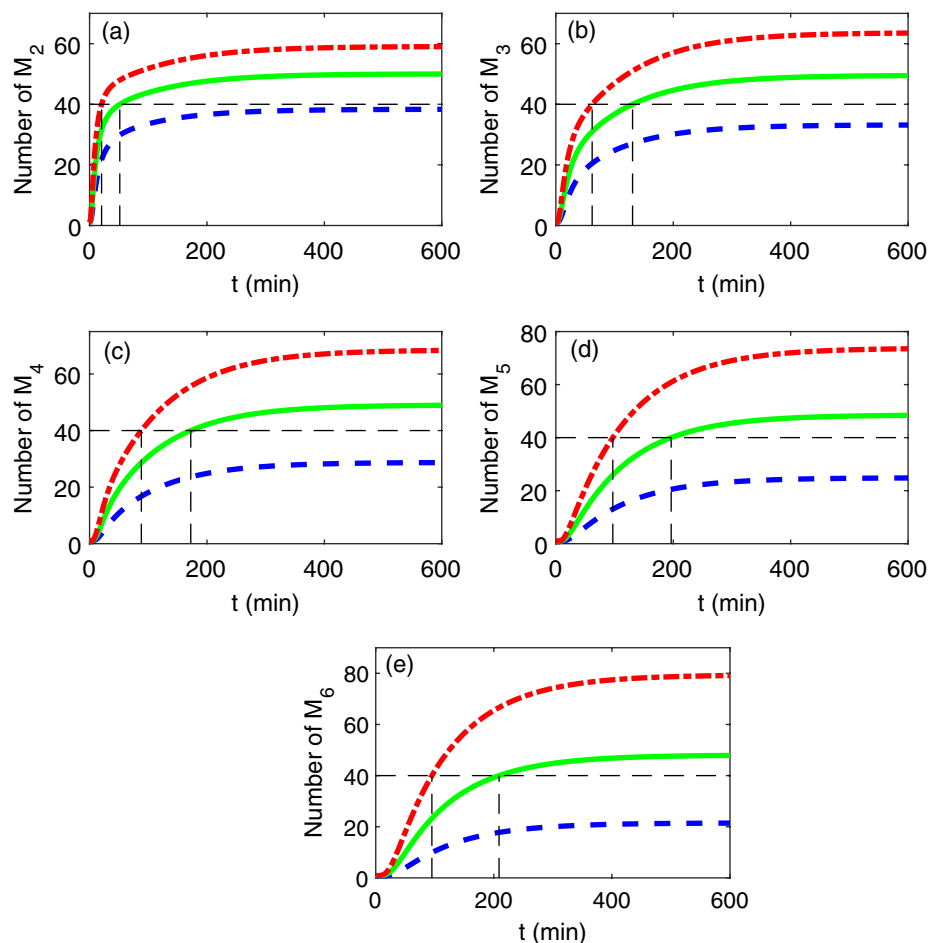
Figure 3. The time evolution of the number of different tapes of oligomers obtained by the moment closure method for different initial number of monomers $x_1(0)$: $x_1(0) = 1000$ (blue broken line), $x_1(0) = 1500$ (green solid line) and $x_1(0) = 2000$ (red dotted dash line). The horizontal dotted line represents a given level of the number of oligomers, while the vertical dotted line represents the time required to reach the given level. The other parameters are same as Fig. 2. The results by SSA are based on 10,000 realizations.

our observation suggests that decreasing the misfolding rate should be useful in slowing or inhibiting the formation of transthyretin amyloid disease. In fact, there have been some literatures towards to reduce the misfloding rate of proteins[47,48].

Conclusion

It is well known that amyloidal aggregation is the hallmark of amyloidoses. In order to quantitatively explore the process of early amyloidal aggregation, we have developed a stochastic mathematical model about oligomers aggregation from monomers using rate law in chemical kinetics. We have adopted a typical moment method based on lognormal closure to capture the statistical moments, and massive calculations show very good agreement between the semi-analytic method and the Gillespie's SSA for the stochastic model. Our results show that the aggregation of monomers into oligomers is highly dependent on the number of monomers and the misfolding rate, and decreasing the number of monomers and the misfolding rate can inhibit the formation of the toxic oligomers. Our research may be helpful in explaining the individual variation in suffering from transthyretin amyloid disease and emphasizes the importance of controlling the misfolding of protein in preventing this type of disease.

Additionally, we remark that the lognormal moment method is successful for stochastic master equation relevant to the transthyretin with an upper limit of oligomer size $N = 6$, and we expect this method would also be effective with a different upper limit which might be relevant to other amyloidogenic systems. The method in the present work can provide essential quantitative insights into the mechanism of the early steps in the aggregation reactions. The mechanism behind the formation of amyloidoses is far from clear, and in the future we will continue to investigate the kinetic of amyloid nucleation and the mechanism of amyloid fibril growth.
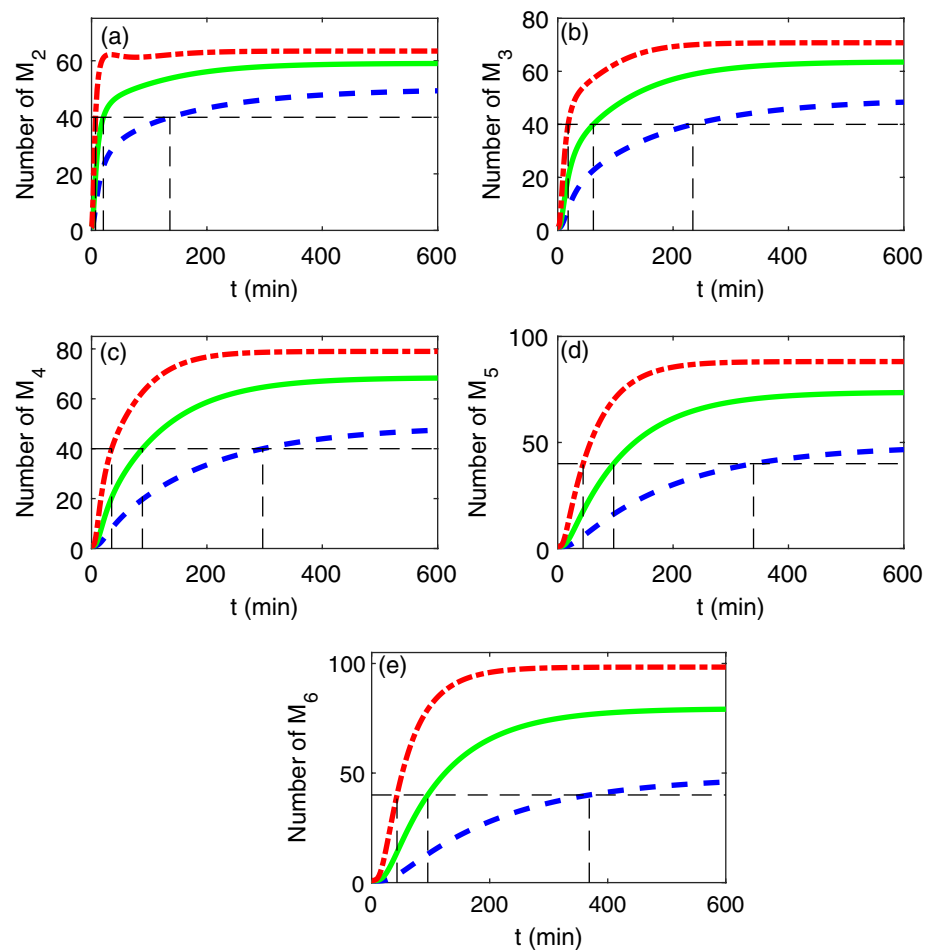
Figure 4. The time evolution of the number of different tapes of oligomers obtained by the moment closure method for different misfolding rate K_0: $K_0 = 0.005$ (blue broken line), $K_0 = 0.01$ (green solid line) and $K_0 = 0.02$ (red dotted dash line). The horizontal dotted line represents a given level of the number of oligomers, while the vertical dotted line represents the time required to reach the given level. The other parameters are same as Fig. 2. The results by SSA are based on 10,000 realizations.

References

1. Hardy, J. A. & Higgins, G. A. Alzheimer's disease: the amyloid cascade hypothesis. *Science* **256**(5054), 184–185 (1992).
2. Dauer, W. & Przedborski, S. Parkinson's disease: mechanisms and models. *Neuron* **39**, 889–909 (2003).
3. DiFiglia, M. *et al.* Aggregation of huntingtin in neuronal intranuclear inclusions and dystrophic neurites in brain. *Science* **277**, 1990–1993 (1997).
4. Brito, R. M. M., Damas, A. M. & Saraiva, M. J. S. Amyloid formation by transthyretin: from protein stability to protein aggregation. *Curr. Med. Chem. Immunol. Endocr. Metab. Agents* **3**(4), 349–360 (2003).
5. Sekijima, Y. *et al.* Energetic characteristics of the new transthyretin variant A25T may explain its atypical central nervous system pathology. *Lab. Investig.* **83**(3), 409–417 (2003).
6. Hurshman, A. R. *et al.* Transthyretin aggregation under partially denaturing conditions is a downhill polymerization?. *Biochemistry* **43**(23), 7365–7381 (2004).
7. Klein, W. L., Krafft, G. A. & Finch, C. E. Targeting small Abeta oligomers: the solution to an Alzheimer's disease conundrum?. *Trends Neurosci.* **24**, 219–224 (2001).
8. Bucciantini, M. *et al.* Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature* **416**, 507–511 (2002).
9. Kayed, R. *et al.* Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis. *Science* **300**, 486–489 (2003).
10. Oosawa, F. & Kasai, M. A theory of linear and helical aggregations of macromolecules. *J. Mol. Biol.* **4**, 10–21 (1962).
11. Ferrone, F. A. *et al.* Kinetic studies on photolysis-induced gelation of sickle cell hemoglobin suggest a new mechanism. *Biophys. J.* **32**, 361–380 (1980).

12. Jarrett, J. T. & Lansbury, P. T. Seeding one-dimensional crystallization of amyloid: a pathogenic mechanism in Alzheimer's disease and scrapie. *Cell* **73**, 1055 (1993).
13. Serio, T. R. *et al.* Nucleated conformational conversion and the replication of conformational information by a Prion determinant. *Science* **289**, 1317 (2000).
14. Collins, S. R. *et al.* Mechanism of prion propagation: amyloid growth occurs by monomer addition. *PLoS Biol* **2**, e321 (2004).
15. Knowles, T. P. J. *et al.* An analytical solution to the kinetics of breakable filament assembly. *Science* **326**, 1533 (2009).
16. Cohen, S. I. A. *et al.* Nucleated polymerisation in the presence of pre-formed seed filaments. *Int. J. Mol. Sci.* **12**, 5844–5852 (2011).
17. Cohen, S. I. A. *et al.* Proliferation of amyloid-beta 42 aggregates occurs through a secondary nucleation mechanism. *Proc. Natl. Acad. Sci. U.S.A.* **110**(24), 9758–9763 (2013).
18. Meisl, G. *et al.* Differences in nucleation behavior underlie the contrasting aggregation kinetics of the Aβ40 and Aβ42 peptides. *Proc. Natl. Acad. Sci. U.S.A.* **111**(26), 9384–9389 (2014).
19. Dear, A. J. *et al.* Statistical mechanics of globular oligomer formation by protein molecules. *J. Phys. Chem. B* **122**(49), 11721–11730 (2018).
20. Saric, A. *et al.* Kinetics of spontaneous filament nucleation via oligomers: insights from theory and simulation. *J. Chem. Phys.* **145**, 211926 (2016).
21. Morel, B. & Conejero-Lara, F. Early mechanisms of amyloid fibril nucleation in model and disease-related proteins. *BBA Proteins Proteomics* **1867**, 140264 (2019).
22. Meisl, G. *et al.* Molecular mechanisms of protein aggregation from global fitting of kinetic models. *Nat. Protoc.* **11**(2), 252–272 (2016).
23. Iljina, M. *et al.* Quantifying co-oligomer formation by α-synuclein. *ACS Nano* **12**, 10855–10866 (2018).
24. Michaels, T. C. T. *et al.* Dynamics of oligomer populations formed during the aggregation of Alzheimer's A beta 42 peptide. *Nat. Chem.* **12**(5), 445–451 (2020).
25. Grima, R. & Schnell, S. Modelling reaction kinetics inside cells. *Essays Biochem.* **45**, 41–46 (2008).
26. Wilkinson, D. J. Stochastic modelling for quantitative description of heterogeneous biological systems. *Nat. Rev. Genet.* **10**(2), 122–133 (2009).
27. Zhang, J., Nie, Q. & Zhou, T. A moment-convergence method for stochastic analysis of biochemical reaction networks. *J. Chem. Phys.* **144**(19), 194109 (2016).
28. Michaels, T. C. T., Dear, A. J. & Knowles, T. P. J. Stochastic calculus of protein filament formation under spatial confinement. *New J. Phys.* **20**(5), 055007 (2018).
29. Michaels, T. C. T. *et al.* Fluctuations in the kinetics of linear protein self-assembly. *Phys. Rev. Lett.* **116**(25), 258103 (2016).
30. Benseny-Cases, N. *et al.* In situ structural characterization of early amyloid aggregates in Alzheimer's disease transgenic mice and *Octodon degus. Sci. Rep.* **10**, 5888 (2020).
31. Reixach, N., Deechongkit, S., Jiang, X., Kelly, J. W. & Buxbaum, J. N. Tissue damage in the amyloidoses: Transthyretin monomers and nonnative oligomers are the major cytotoxic species in tissue culture. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 2817–2822 (2004).
32. Dayeh, M. A., George, L. & Saber, E. A discrete mathematical model for the aggregation of β-amyloid. *PLoS ONE* **13**(5), e0196402 (2018).
33. Ale, A., Kirk, P. & Stumpf, M. P. H. A general moment expansion method for stochastic kinetic models. *J. Chem. Phys.* **138**(17), 174101 (2013).
34. Lakatos, E., Ale, A., Kirk, P. D. & Stump, M. P. H. Multivariate moment closure techniques for stochastic kinetic models. *J. Chem. Phys.* **143**(9), 094107 (2015).
35. Lee, H., Lee, S. & Lee, C. H. Stochastic methods for epidemic models: an application to the 2009 H1N1 influenza outbreak in Korea. *Appl. Math. Comput.* **286**, 232–249 (2016).
36. Kang, Y. M. & Chen, X. Application of Gaussian moment method to a gene autoregulation model of rational vector field. *Mod. Phys. Lett. B* **30**(20), 1650264 (2016).
37. Ciuperca, I. S. *et al.* Alzheimer's disease and prion: an in vitro mathematical model. *Discrete Contin. Dyn. B* **24**(10), 5225–5260 (2019).
38. Hasegawa, K., Yamach, M. & Naiki, H. Kinetic modeling and determination of reaction constants of Alzheimer's beta amyloid fibril extension and dissociation using surface plasmon resonance. *Biochemistry* **41**, 13489–13498 (2002).
39. Michaels, T. C. T., Garcia, G. A. & Knowles, T. P. J. Asymptotic solutions of the Oosawa model for the length distribution of biofilaments. *J. Chem. Phys.* **140**(19), 194906 (2014).
40. Gillespie, D. T. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* **81**(25), 2340–2361 (1977).
41. Lee, C. H., Kim, K. H. & Kim, P. A moment closure method for stochastic reaction networks. *J. Chem. Phys.* **130**(13), 134107 (2009).
42. Hespanha, J. P. & Singh, A. Stochastic models for chemically reacting systems using polynomial stochastic hybrid systems. *Int. J. Robust Nonlin.* **15**(15), 669–689 (2005).
43. Singh, A. & Hespanha, J. P. Approximate moment dynamics for chemically reacting systems. *IEEE Trans. Autom. Control* **56**(2), 414–418 (2011).
44. Magnus, K. *et al.* Oligomer diversity during the aggregation of the repeat-region of tau. *ACS Chem. Neurosci.* **9**, 3060–3071 (2018).
45. Yang, J. *et al.* Direct observation of oligomerization by single molecule fluorescence reveals a multistep aggregation mechanism for the yeast prion protein ure2. *J. Am. Chem. Soc.* **140**(7), 2493–2503 (2018).
46. Chiti, F. & Dobson, C. M. Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.* **75**(1), 333–366 (2006).
47. Chabry, J., Caughey, B. & Chesebro, B. Specific inhibition of in vitro formation of protease-resistant prion protein by synthetic peptides. *J. Biol. Chem.* **273**(21), 13203–13207 (1998).
48. Miles, L. A. *et al.* Bapineuzumab captures the N-terminus of the Alzheimer's disease amyloid-beta peptide in a helical conformation. *Sci. Rep.* **3**, 1302 (2013).

Acknowledgements

Author contributions

K.Y.-M. guided and financed the research, L.R.-N. did the calculation, and they both contributed to language wring and organization.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-69319-x.

Correspondence and requests for materials should be addressed to Y.-M.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.