Modeling Amyloid Aggregation Kinetics: A Case Study with Sup35NM

Aditi Sharma,[1‡] Matthew A. McDonald,[1‡] Harrison B. Rose,[1] Yury O. Chernoff,[2,5] Sven H. Behrens,[1,4] and Andreas S. Bommarius[1,3]*

[1]School of Chemical & Biomolecular Engineering, [2]School of Biological Sciences, and [3]School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, Georgia 30332, USA;

[4]Polymer Science & Materials Chemistry, Exponent, Inc., Atlanta, Georgia 30326, USA;

[5]Laboratory of Amyloid Biology, St. Petersburg State University, St. Petersburg 199034, Russia.

ELECTRONIC SUPPLEMLENTARY INFORMATION

Table of Contents

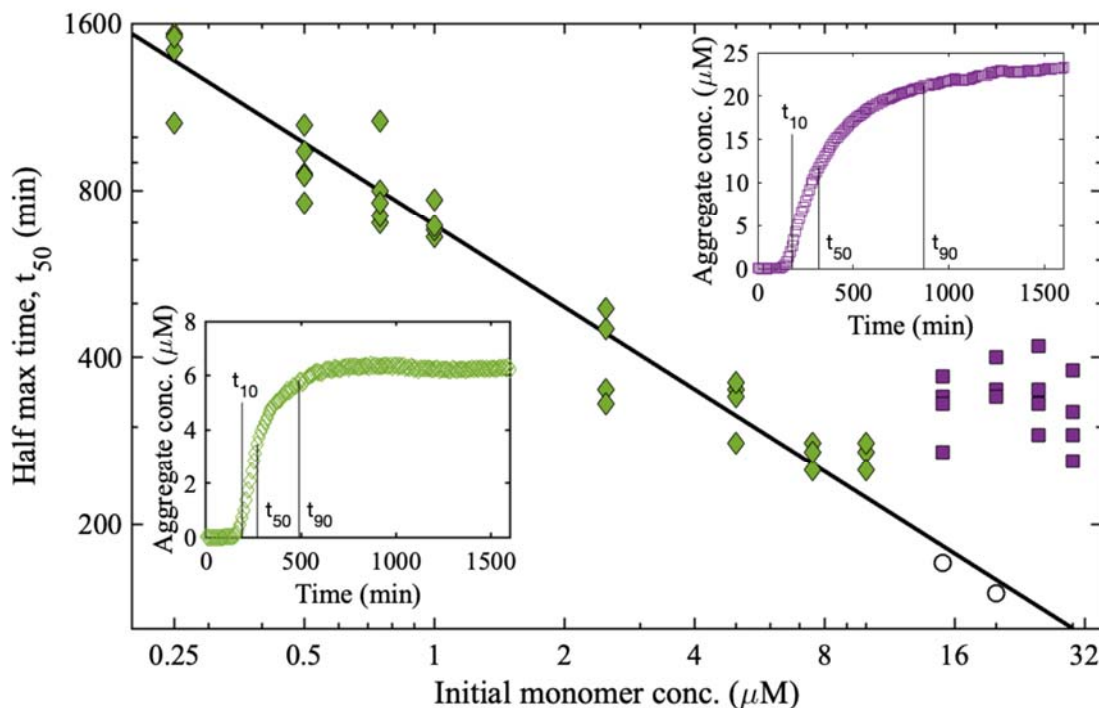## S.1. Discussion of exclusion of high monomer concentration data

Figure S1. The aggregation experiments that fit to a linear relationship for the γ value are exemplified by the green aggregation curve, while those that deviate from the linear relationship follow aggregation courses like the purple curve. The time to reach complete aggregation is much longer for the purple curve.

Figure S1 shows examples of the differences between those experiments fitting to the linear γ vs. m_0 relation and those that deviate from the linear fit. There are a number of following reasons for why this deviation could occur.

(1) Turbidity: at the highest concentrations the aggregate mass concentration exceeds 1 g/L, assuming that the aggregates are well dispersed within the solution it should create a translucent suspension that interferes with fluorescence measurement. Turbidity was noted during the high concentration experiments.

(2) Polymorphism: at high concentrations a different form of amyloid could be precipitating. Ostwald's rule of stages says that at high supersaturation the least stable form will nucleate first. Rather than well characterized fibrils, these high concentration experiments could

correspond to amorphous aggregation. That two points (one experiment each at 15 and 20 µM) fit the linear curve supports this hypothesis as nucleation is stochastic and, in these instances, the well-defined form may have nucleated first and therefore follow the γ trend.

(3) A change in mechanism that should be captured by the model but is not, however no published mechanisms can explain this observation and the authors are in no position to postulate new mechanisms when two equally plausible explanations exist.

S.2. Residuals of fit over time, for all monomer concentrations, fast shaking
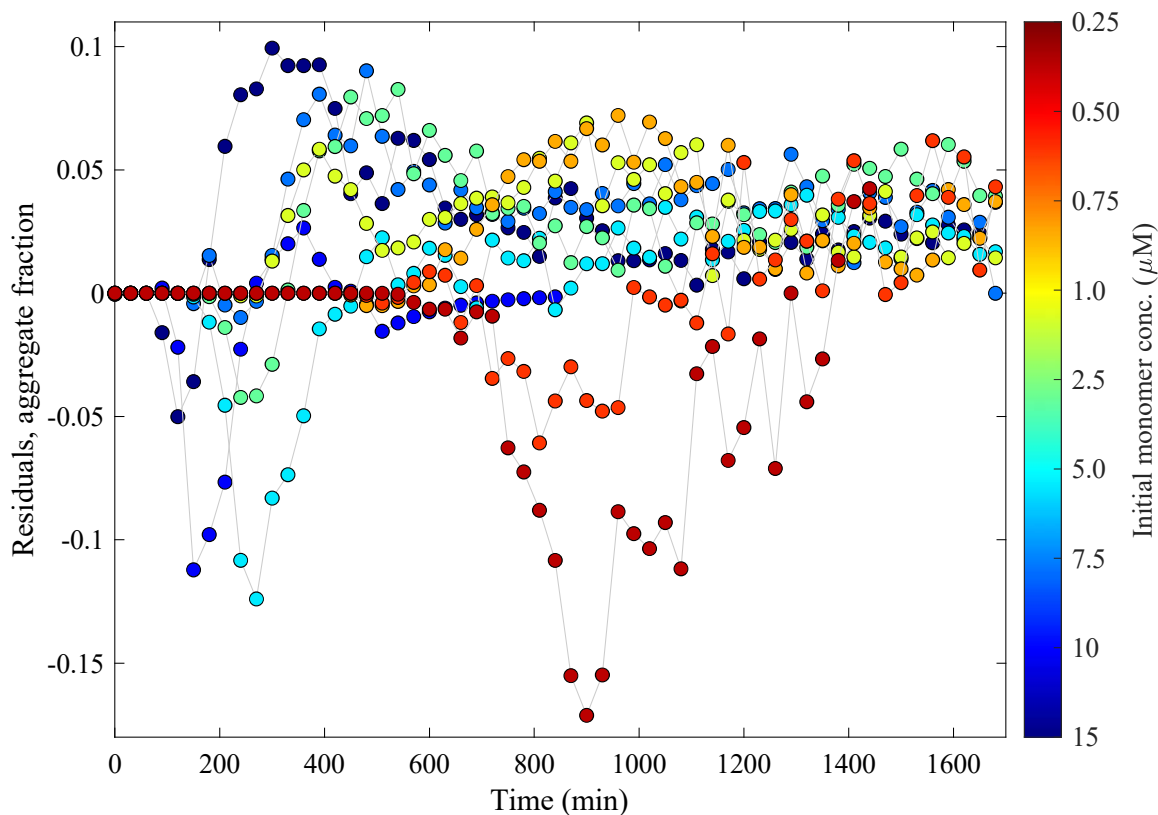

Figure S2. Residuals from the fast-shaking experiments.

The residuals from the fast-shaking experiments shows how the error is bounded by +10% and -15%, with the largest error experiments occurring at extremes of concentration (dark red and blue points). The maximum errors are less than those obtained by other authors and the

error trend, with high monomer concentrations having substantial error early and low monomer concentrations having substantial error late, agrees with the state of the art.[1]

S.3. Impact and implementation of fragmentation modeling

Fragmentation involves the fracture of a single fibril into multiple smaller fibrils. The mechanism is inherently complex requiring either detailed system knowledge or many assumptions to model accurately. The knowledge and assumptions must cover both the probability/rate of fibril fragmentation and the distribution of fragment sizes. Previous studies have looked at the impact of particle size, particle shape (aspect ratio), and agitation power input on the rate of fragmentation and the size of fragments produced.[2-5] One approach to model fragmentation is to use a selection function (not included in the PBM in the main text) to account for the length dependence of fragmentation probability, paired with a simple model of daughter fibril length distributions, $b(L,\lambda \mid L \leq \lambda)$, in the main text. A common selection function is a logistic curve such:

$$s(\lambda) = \frac{1}{1 + \exp\left[-\beta(\lambda - \alpha)\right]}$$ (S.1)

where α and β determine the length where fragmentation becomes favorable and the breadth of lengths where fragmentation can occur, respectively. Figure S3 (left) shows the $S(\lambda)$ for $\alpha = 100$ monomers and $\beta = 10\%$. In the PBM, S appears as an additional factor inside the integral term. The daughter fragment length distribution can be simply modeled with a one parameter term

$$b(L,\lambda) = \frac{q}{2\lambda} + \left(1 - \frac{q}{2}\right)\left(\frac{12L^2}{\lambda^3} - \frac{12L}{\lambda^2} + \frac{3}{\lambda}\right)$$ (S.2)

where $0 \leq q \leq 3$ is used to tune between uniform fragmentation ($q = 2$), fragmentation from the ends ($q = 0$), and fragmentation in the middle ($q = 3$). The fragmentation probability density as a

function of fibril length and q parameter is show in Figure S3 (right). The b function used in the main text would be a Dirac delta pulse at a normalized fibril length of 0.5 in Figure S3 (right). The selection function and daughter distribution are physically more intuitive than the fragmentation model implemented in the main text, however they add more parameters to a model that is already capable of over-fitting the limited data set and increase computation costs making parameter fitting more tedious.
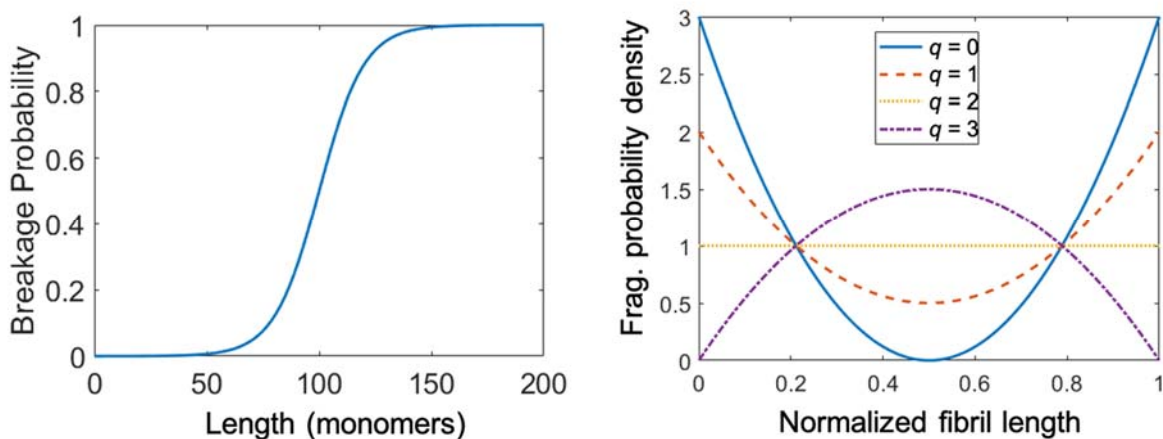

Figure S3. Left, an example selection function (eqn. S1) and, Right, an example daughter length distribution probability density (eqn. S2), normalized by total fibril length, λ, used in fragmentation models.

Fibril length distributions for a simulated fragmentation experimental are shown in Figure S4 for $q = 0$, 1, 2, and 3 and $\alpha = 100$ monomers and $\beta = 10\%$. Total length weighted counts at the end of the aggregation experiment are reported. Numerical instability causes the missing bars at the large length tail of the distributions.
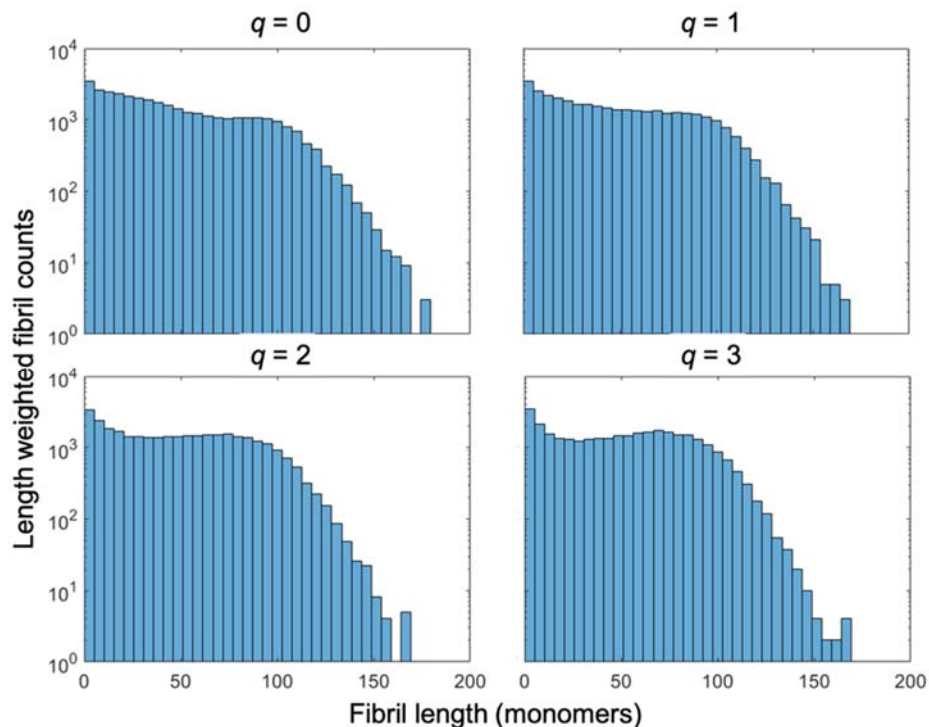
Figure 4. Differences in fibril length distribution for different values of q in eqn. S2.

S.4. Simulated aggregation experiments: comparisons between the models

Simulated experiments can be used to test the hypothesis that it is experimentally impossible to distinguish between a secondary nucleation and fragmentation dominated aggregate formation mechanism. Aggregation experiments ranging in concentration from 0.3 µM to 10 µM (nearly two orders of magnitude) were simulated using a model that only considered secondary nucleation (no fragmentation); the parameters values are in the first row of Table S1. The time course of the aggregate mass fraction, as would be measured by ThT fluorescence, generated in this simulation was used to fit a fragmentation model (no secondary nucleation); the results of the fitting are in the second row of Table S1. The goodness of fit is visually obvious looking at the trace and residuals in Figure S5, left and right, respectively.

Table S1. Parameters for the simulated true model with secondary nucleation as the only source of secondary fibrils) and parameters fit to the true model for a model with fragmentation as the only source of secondary fibrils. For explanation of variables see main text. Fragmentation is not a function of supersaturation or (therefore $n_1 = 0$) and has a linear dependence on aggregate mass (therefore $n_m = 1$).

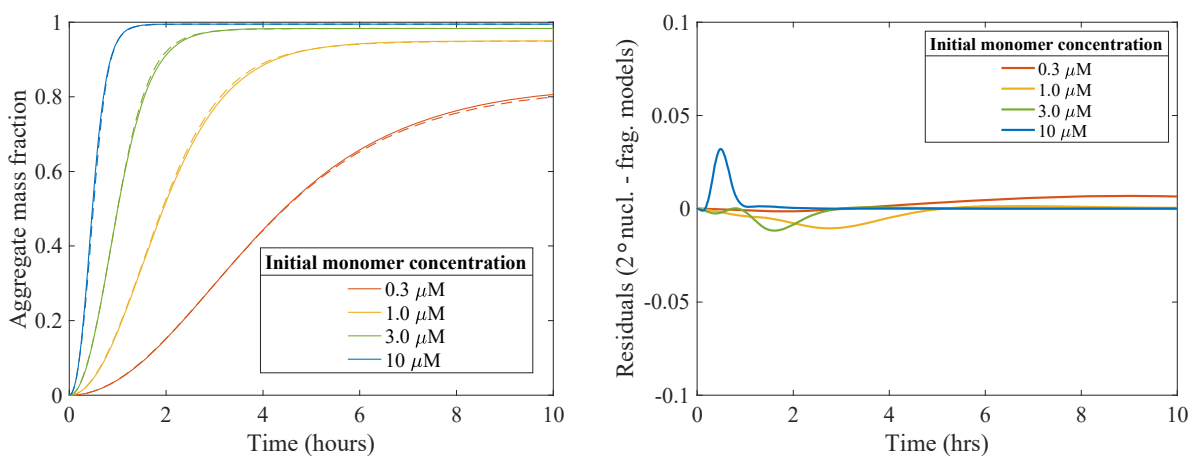|	k_E	n_E	k_{N1}	n_1	k_{N2} / k_F	n_2	n_m / n_F	c^*
Secondary nucleation	1.0	1.0	1.0×10^{-3}	1.0	1.0×10^{-3}	1.0	0.67	0.05
Fragmentation	2.83	1.41	3.43×10^{-4}	0.61	5.70×10^{-3}	NA	1.0	0.05

Figure S5. Result of fitting a fragmentation model to a simulated process with only primary and secondary nucleation. Left, simulation (solid curve) of ThT fluorescence over time for a process with only secondary nucleation and the fit of a fragmentation model (dashed curve) over nearly two orders of magnitude concentration (color scheme same as Figure S2). Right, residuals between simulated experiment and fit, note the scale compared to Figure S2.

In the reverse exercise, where a fragmentation model is used as the true process and a secondary nucleation mechanism is fit, the residuals are even smaller as there are more parameters in the secondary nucleation model than the fragmentation model. A model with both mechanisms adds even more parameters. Adding complexity to aggregation models should be done with prudence and efforts should be made to ensure that the wet experiments accompanying the models will actually inform which mechanisms best describe the observations.

The differences in the secondary nucleation and fragmentation mechanisms only become clear when the fibril length distributions are measured. Figures S6 and S7 show the final length distributions as they might be measured by AUC ($p = 2$). The shapes of the final distributions are different from those in the main text because parameter values were chosen such that primary and secondary fibril generating processes would occur at similar rates. In the main text fits to experimental data with Sup35NM showed that primary nucleation occurred at a faster rate than either secondary nucleation or fragmentation. Qualitatively, however, the results are the same. For secondary nucleating systems there are few small fibrils at the end of the experiment/simulation because after most of the monomer has been incorporated into fibrils there is no driving force for new fibril generation. For fragmentation however, the length distribution includes significant small (relatively) fibrils as new fibrils are generated constantly by fragmentation, regardless of the remaining monomer concentration. It is noteworthy that one may be able to rule out the fragmentation mechanism based only on modeling the complete length distribution with the fit parameters (second row, Table S1) as the fibril lengths become exceedingly long at high initial monomer concentration (c_0). Still, some knowledge of fibril size, unknowable from ThT fluorescence, is required.
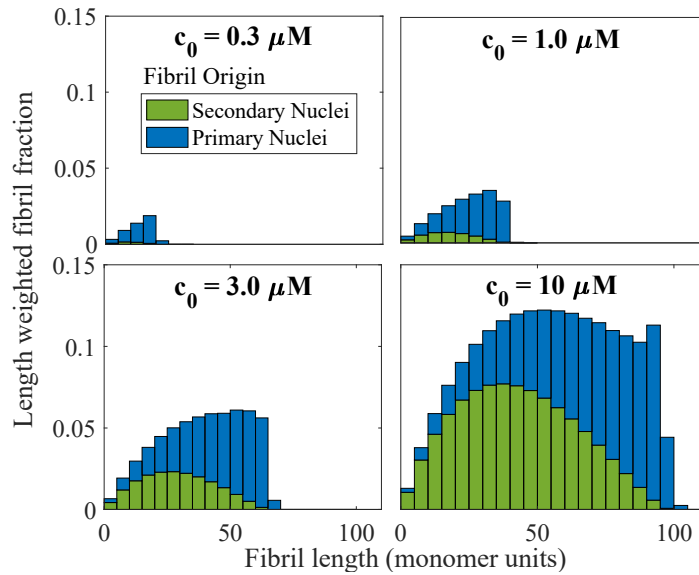
Figure S6. The simulated fibril length distributions based on the parameters in the first row of Table S1. Each panel has the same axes. Blue bars represent those fibrils originating by primary nucleation and green shows those originating from secondary nucleation. Experimentally, only the total bar height would be measurable.
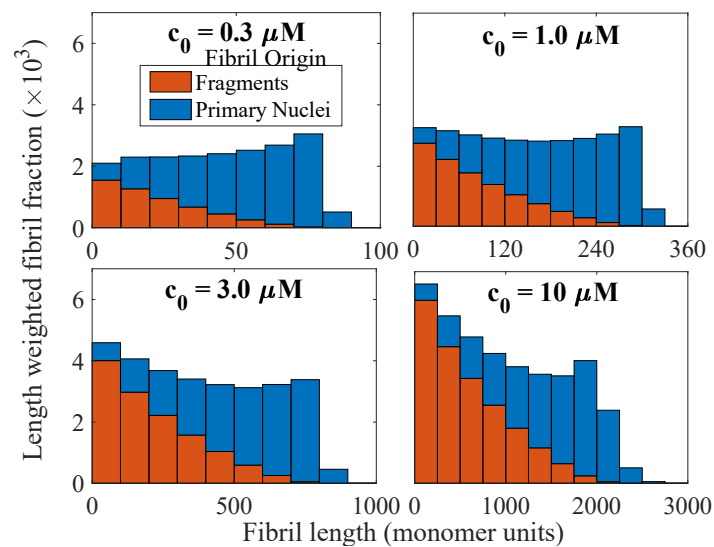

Figure S7. The simulated fibril length distributions based on the parameters in the second row of Table S1. Each panel has the same y-axes but the x-axes are scaled to show the details across the entire range of sizes. Blue bars represent those fibrils originating by primary nucleation and red shows those originating from fragmentation. Experimentally, only the total bar height would be measurable.

The fragmented length distributions in Figures S4 and S7 are similar, sharing a shoulder to the main peak indicative of fragmentation. This confirms that the overly-simple fragmentation PBM implemented in the main text (and illustrated in S7) does an adequate job representing fragmentation (more completely simulated in S4) to make the point in this study that fluorescence time courses alone are not sufficient to assign a mechanism to secondary processes for fibril generation.

S.5. References

1. Meisl, G.; Yang, X.; Hellstrand, E.; Frohm, B.; Kirkegaard, J. B.; Cohen, S. I. A.; Dobson, C. M.; Linse, S.; Knowles, T. P. J., Differences in nucleation behavior underlie the contrasting aggregation kinetics of the Aβ40 and Aβ42 peptides. *Proceedings of the National Academy of Sciences* **2014,** *111* (26), 9384-9389.
2. Salvatori, F.; Mazzotti, M., Experimental Characterization and Mathematical Modeling of Breakage of Needle-like Crystals in a Continuous Rotor-Stator Wet Mill. *Crystal Growth & Design* **2018,** *18* (10), 5957-5972.
3. Szilágyi, B.; Agachi, P. Ş.; Lakatos, B. G., Numerical analysis of crystallization of high aspect ratio crystals with breakage. *Powder Technology* **2015,** *283*, 152-162.
4. Borsos, Á.; Lakatos, B. G., Investigation and simulation of crystallization of high aspect ratio crystals with fragmentation. *Chemical Engineering Research and Design* **2014,** *92* (6), 1133-1141.
5. Xue, W.-F.; Hellewell, A. L.; Hewitt, E. W.; Radford, S. E., Fibril fragmentation in amyloid assembly and cytotoxicity: when size matters. *Prion* **2010,** *4* (1), 20-25.