Supporting Information for “Investigating the Effects of Molecular Crowding on the Kinetics of Protein Aggregation”

John S. Schreck,1,2 a) John Bridstrup,3 and Jian-Min Yuan3, b)
1) National Center for Atmospheric Research, Boulder, CO 80305
2) Department of Chemistry, Drexel University, Philadelphia, PA 19104
3) Department of Physics, Drexel University, Philadelphia, PA 19104

S1. LIST OF SYMBOLS

A list of the symbols used in the main text to define the different quantities is presented in Table S1. The parameter \( \alpha \) used in Eqs. (17) and (21) may be expressed as

\[
\ln \alpha = \frac{2}{3} \left( \frac{r_{\infty}}{r_c} \right)^3 \left[ 1.5(R^2 + R + 1)z \right. + 4.5(R^2 + R)z^2 + 4.5R^2z^3 \right]
\]

where

\[
z = \frac{\phi}{\sigma - \sigma} \quad \text{(S2)}
\]

\[
R = \frac{r_{\infty}}{r_c}. \quad \text{(S3)}
\]

S2. EXPERIMENTAL DATASETS

All experimental data used in this work are listed in the supporting file data-table.xlsx. The file includes a table for each protein and includes the initial mass concentration, the crowder concentrations (if any), and the raw data extracted manually from the references listed in section S4.

S3. OBTAINING \( t_{1/2} \) FROM THIOFLAVIN T FLUORESCENCE CURVES

All half-times reported in Figs. 2 and 3 were found by fitting

\[
F(t) = F_0 + \frac{A}{1 + \exp(-k_{app}(t-t_{1/2}))} \quad \text{(S4)}
\]

to the experimental Thioflavin T (ThT) curves as a function of time. The parameters \( F_0, A, k_{app}, \) and \( t_{1/2} \) were adjusted to minimize the mean-average error between Eq. (S4) and the experimentally reported values. All fits of \( F(t) \) to the data sets were re-scaled according to \( (F(t) - \min(F(t))) / (\max(F(t)) - \min(F(t))) \) from which the half-time is obtained when this quantity equaled one half. The experimental data was similarly re-scaled using the same maximum and minimum values from the fitted \( F(t) \). During the fitting routine, we employed the basin hopping method from Scientific Python (scipy) to explore the high-dimensional parameter space to find, if possible, physically realistic rate constants.

a) Electronic mail: jsschreck@gmail.com
b) Electronic mail: yuanjm@drexel.edu

S4. RATE CONSTANTS

All non-zero fit parameters used in the figures are listed for each protein investigated. The rate constants for monomer addition and subtraction are \( k_+^0 \) and \( k_-^0 \), respectively, while we assume that merging and fragmentation mechanisms are independent of size so that \( k_+(r, s) \equiv k_+ \) and \( k_-(r, s) \equiv k_- \).

A. Actin from Rosin, et al.2

\[
n_c = 3, \ c_0 = 5 \mu M, \ k_+^0 = 455.5 \mu M^{-1}h^{-1} \quad \text{and} \quad k_-^0 = 1.02 \times 10^{-5} \mu M^2 h^{-1}
\]

\[
r^0_c = 1.8 \text{ nm}, \ r_1 = r_{sc} = 2.15 \text{ nm}, \ A_1 = 2.70 \times 10^{-9} \text{ nm}^{-2}, \quad \chi = 1.87. \quad \text{See Fig. 2.}
\]

B. Human PrP from Zhou, et al.3

\[
n_c = 1, \ c_0 = 10.0 \mu M, \ k_+ = 36.8 \mu M^{-1}h^{-1}, \ k_- = 0.044 \mu M^{-1}h^{-1},
\]

\[
\bar{k}_+ = 7.6 \times 10^{-4} \mu M^{-1}h^{-1}, \ r^0_c = 1.8 \text{ nm}, \ r_1 = r_{sc} = 3.5 \text{ nm}, \ A_1 = 2.70 \times 10^{-9} \text{ nm}^{-2}, \quad \chi = 1.87. \quad \text{See Fig. S1.}
\]

C. ApoC2 from Binger, et al.4

\[
n_c = 5, \ c_0 = 45 \mu M, \ k_+^0 = 24.49 \mu M^{-1}h^{-1}, \ k_-^0 = 1.481 \mu M^{-1}h^{-1},
\]

\[
\bar{k}_+ = 12.96 \times 10^{3} \mu M^{-1}h^{-1}, \ r^0_c = 5.70 \times 10^{-3} \mu M^{-1}h^{-1}, \ k_n = 1.54 \times 10^{-11} \mu M^4 h^{-1},
\]

\[
r_c = 1.24 \text{ nm}, \ r_1 = r_{sc} = 2.10 \text{ nm}, \ A_1 = 6.10 \times 10^{-10} \text{ nm}^{-2}, \quad \chi = 1.95. \quad \text{See Fig. S2.}
\]

D. \( \beta(1-40) \) from Lee, et al.5

\[
n_c = n_2 = 2, \ c_0 = 30.0 \mu M, \ k_+^0 = 72.05 \mu M^{-1}h^{-1},
\]

\[
k_- = 0.0033 \mu M^{-1}h^{-1}, \ k_{n} = 3.07 \times 10^{-4} \mu M^{-1}h^{-1}, \  \bar{k}_n = 3.54 \times 10^{-3} \mu M^{-1}h^{-1},
\]

\[
k^0 = 1.13 \times 10^{-19} \mu M h^{-1}, \  \bar{k}_+ = 7.47 \times 10^{-10} \mu M h^{-1}, \ k_f = 1.31 \mu M^{-1}h^{-1},
\]

\[
k_b = 0.26 h^{-1}, \ r^0_c = 1.8 \text{ nm}, \ r_1 = 2.26 \text{ nm}, \ r_{sc} = 12.3 \text{ nm}, \ A_1 = 2.35 \times 10^{-8} \text{ nm}^{-2}, \quad \chi = 0.67. \quad \text{See Fig. S3.}
\]

E. \( \beta\text{-LAC} \) from Ma, et al.6

\[
n_c = 4, \ c_0 = 82 \mu M, \ k_+^0 = 1.04 \times 10^{3} \mu M^{-1}h^{-1}, \ k_-^0 = 7.68 \times 10^{-7} \mu M^{-1}h^{-1},
\]

\[
k^0 = 2.05 \times 10^{-3} \mu M^{-1}h^{-1}, \ \bar{k}_n = 9.76 \times 10^{-9} \mu M h^{-1}, \ k_b = 4.10 \times 10^{-22} \mu M^3 h^{-1},
\]

\[
r_c = 2.01 \text{ nm}, \ A_1 = 2.77 \times 10^{-9} \text{ nm}^{-2}, \quad \chi = 1.94. \quad \text{See Fig. S4.}
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Unit</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( c_0 )</td>
<td>( \mu M )</td>
<td>Initial mass concentration of monomer protein</td>
</tr>
<tr>
<td>( \phi )</td>
<td>-</td>
<td>Volume fraction of crowders in solution</td>
</tr>
<tr>
<td>( n_c )</td>
<td>-</td>
<td>Critical nucleus size (integer) formed through a primary nucleation process</td>
</tr>
<tr>
<td>( n_2 )</td>
<td>-</td>
<td>Critical nucleus size (integer) formed through a secondary nucleation process</td>
</tr>
<tr>
<td>( k_n )</td>
<td>( \mu M^{- (n_c - 1)} h^{-1} )</td>
<td>Nucleation rate constant for a critical nuclei of size ( n_c )</td>
</tr>
<tr>
<td>( k^* )</td>
<td>( h^{-1} )</td>
<td>Dissociation rate constant for a critical nuclei of size ( n_c )</td>
</tr>
<tr>
<td>( k_+ )</td>
<td>( \mu M^{-1} h^{-1} )</td>
<td>Monomer addition rate constant</td>
</tr>
<tr>
<td>( k_- )</td>
<td>( h^{-1} )</td>
<td>Monomer subtraction rate constant</td>
</tr>
<tr>
<td>( \tilde{k}_+ (r, s) )</td>
<td>( \mu M^{-1} h^{-1} )</td>
<td>Cluster merging rate constant</td>
</tr>
<tr>
<td>( \tilde{k}_- (r + s, s) )</td>
<td>( h^{-1} )</td>
<td>Cluster breakage rate constant</td>
</tr>
<tr>
<td>( k_2 )</td>
<td>( \mu M^{- n_2} h^{-1} )</td>
<td>Nucleation rate constant for a 1-step or 2-step secondary nucleus of size ( n_2 )</td>
</tr>
<tr>
<td>( \tilde{k}_2 )</td>
<td>( \mu M^{- n_2} h^{-1} )</td>
<td>The formation rate constant of a secondary nucleus from surface-bound monomers</td>
</tr>
<tr>
<td>( k_f )</td>
<td>( \mu M^{-1} h^{-1} )</td>
<td>The rate constant for monomers binding to a fibril surface</td>
</tr>
<tr>
<td>( k_b )</td>
<td>( h^{-1} )</td>
<td>The rate constant for monomers detachment from a fibril surface</td>
</tr>
<tr>
<td>( r_1 )</td>
<td>nm</td>
<td>Radius of a protein monomer</td>
</tr>
<tr>
<td>( r_c )</td>
<td>nm</td>
<td>Radius of a crowder</td>
</tr>
<tr>
<td>( r_{sc} )</td>
<td>nm</td>
<td>Radius of a spherocylinder</td>
</tr>
<tr>
<td>( A_1 )</td>
<td>( \text{nm}^{-\chi} )</td>
<td>A constant for parameterizing the leading-order dependence of ( r_c ) on ( \phi ) in Eq. (33)</td>
</tr>
<tr>
<td>( \chi )</td>
<td>-</td>
<td>An exponent for parameterizing the leading-order dependence of ( r_c ) on ( \phi ) in Eq. (33)</td>
</tr>
<tr>
<td>( t_{1/2} )</td>
<td>h</td>
<td>The half-time is the point when ( M(t) ) reaches 1/2.</td>
</tr>
</tbody>
</table>

TABLE S1. The symbol for a given quantity is given in the first column, the quantities unit (if any) in the second column, and a description in the third column. Rate constants having an extra super script \( * \) in sections below and in the main text refer to experiments or simulations performed at crowderless conditions (\( \phi = 0 \)).

F. \( \beta_2 \)m from Luo, et al.\(^7\)

\[
n_c = 1, \quad c_0 = 1.4 \mu M, \quad k_0^0 = 1.08 \times 10^2 \mu M^{-1} h^{-1}, \quad k_0^* = 1.2 \times 10^{-6} h^{-1}, \quad \tilde{k}_+ = 4.8 \times 10^{-3} h^{-1}, \quad r_c = 1.8 \text{ nm}, \quad r_1 = r_{sc} = 2.74 \text{ nm}, \quad A_1 = 2.92 \times 10^{-9} \text{ nm}^\chi, \quad \text{and} \quad \chi = 1.82. \quad \text{See Fig. S5.}
\]

G. \( \beta_2 (1-40) \) from Meisl, et al.\(^8\)

\[
n_c = n_2 = 2, \quad c_0 = 5.0 \mu M, \quad k_0^0 = 1.19 \times 10^{-5} \mu M^{-1} h^{-1}, \quad k_0^* = 6.38 \times 10^{-2} h^{-1}, \quad k_f = 0.63 h^{-1}, \quad k_b = 30.38 h^{-1}, \quad r_c = 1.8 \text{ nm}, \quad r_1 = 2.26 \text{ nm}, \quad r_{sc} = 2.4 r_1, \quad A_1 = 2.35 \times 10^{-8} \text{ nm}^\chi, \quad \text{and} \quad \chi = 0.67. \quad \text{See Fig. 7(a).}
\]

H. \( \beta_2 (1-42) \) from Cohen, et al.\(^9\)

\[
n_c = n_2 = 2, \quad c_0 = 5.0 \mu M, \quad k_0^0 = 85.0 \mu M^{-1} h^{-1}, \quad k_0^* = 1.72 \times 10^{-6} \mu M^{-1} h^{-1}, \quad k_f = 2.76 \times 10^{-1} h^{-1}, \quad k_b = 8.99 \times 10^{-2} h^{-1}, \quad k_b = 1.24 h^{-1}, \quad r_c = 1.8 \text{ nm}, \quad r_1 = 2.26 \text{ nm}, \quad r_{sc} = 2.0 r_1, \quad A_1 = 2.35 \times 10^{-8} \text{ nm}^\chi, \quad \text{and} \quad \chi = 0.67. \quad \text{See Fig. 7(b).}
\]

I. \( \alpha \beta_2 (1-42) \)-low from Meisl, et al.\(^10\)

\[
n_c = n_2 = 3, \quad c_0 = 5.0 \mu M, \quad k_0^0 = 8.34 \mu M^{-1} h^{-1}, \quad k_0^* = 1.43 \times 10^{-3} h^{-1}, \quad k_0^* = 9.91 \times 10^{-8} \mu M^{-1} h^{-1}, \quad \tilde{k}_+ = 1.74 \times 10^{-3} h^{-1}, \quad k_f = 1.72 \times 10^{-2} h^{-1}, \quad k_b = 1.78 \times 10^{9} h^{-1}, \quad r_c = 1.8 \text{ nm}, \quad r_1 = 2.26 \text{ nm}, \quad r_{sc} = 2.5 r_1, \quad A_1 = 2.35 \times 10^{-8} \text{ nm}^\chi, \quad \text{and} \quad \chi = 0.67. \quad \text{See Fig. 7(c).}
\]

J. \( \alpha \beta_2 \) from Seeliger, et al.\(^11\)

\[
n_c = 1, \quad c_0 = 10.0 \mu M, \quad k_0^0 = 5.11 \times 10^{-5} h^{-1}, \quad \tilde{k}_+ = 2.28 \times 10^{-3} h^{-1}, \quad k_0^* = 4.86 \times 10^{-4} \mu M^{-1} h^{-1}, \quad r_c = 1.8 \text{ nm}, \quad r_1 = r_{sc} = 2.74 \text{ nm}, \quad A_1 = 2.92 \times 10^{-9} \text{ nm}^\chi, \quad \text{and} \quad \chi = 1.82. \quad \text{See Fig. S6.}
\]

K. Rabbit PrP from Zhou, et al.\(^3\)

\[
n_c = 1, \quad c_0 = 10.0 \mu M, \quad k_0^0 = 36.83 \mu M^{-1} h^{-1}, \quad \tilde{k}_+ = 4.41 \times 10^{-2} h^{-1}, \quad k_0^* = 6.56 \times 10^{-4} \mu M^{-1} h^{-1}, \quad \tilde{k}_+ = 2.32 \times 10^{-3} h^{-1}, \quad k_0^* = 1.86 \times 10^{-6} \mu M^{-1} h^{-1}, \quad r_c = 1.8 \text{ nm}, \quad r_1 = 3.34 \text{ nm}, \quad r_{sc} = 2.21 \text{ nm}, \quad A_1 = 4.65 \times 10^{-11} \text{ nm}^\chi, \quad \text{and} \quad \chi = 2.04. \quad \text{See Fig. S7.}
\]
L. **BCA from Mittal, et al.**

\[ n_c = 1, \ c_0 = 7.0 \, \mu M, \ k_0^0 = 1161.56 \, \mu M^{-1}h^{-1}, \ k_0^- = 48.31 \, h^{-1}, \ k_0^+ = 9.45 \times 10^{-2} \, h^{-1}, \ k_0^n = 4.14 \times 10^{-3} \, \mu M^4 h^{-1}, \]
\[ r_+^0 = 1.8 \, \text{nm}, \ r_1 = 3.16 \, \text{nm}, \ r_{sc} = 2.27 \, \text{nm}, \ A_1 = 4.72 \times 10^{-10} \, \text{nm}^{-x}, \text{ and } \chi = 1.84. \text{ See Fig. S8.} \]

M. **Aβ(1-40) from Lee, et al.**

\[ n_c = n_2 = 2, \ c_0 = 30.0 \, \mu M, \ k_+^0 = 7.50 \, \mu M^{-1}h^{-1}, \]
\[ k_-^0 = 6.59 \times 10^{-11} \, h^{-1}, \ \tilde{k}_+^0 = 1.58 \times 10^{-9} \, h^{-1}, \ k_-^n = 8.39 \times 10^{-9} \, \mu M h^{-1}, \ \tilde{k}_+^n = 1.28 \times 10^{-5} \, h^{-1}, \ k_f = 5.09 \times 10^8 \, \mu M h^{-1}, \]
\[ k_b = 2.6 \times 10^{10} \, h^{-1}, \ r_+^0 = 1.8 \, \text{nm}, \ r_1 = 2.41 \, \text{nm}, \ r_{sc} = 1.56 \, \text{nm}, \ A_1 = 1.26 \times 10^{-9} \, \text{nm}^{-x}, \text{ and } \chi = 0.50. \text{ See Fig. S9.} \]


FIG. S1. Human PrP. In (a) the mass concentration of fibrils $M(t)$ is shown and was fit to ThT data from Zhou, et al. In (b) the predicted $L(t)$ is shown for different crowder concentrations $\phi$. In (c) the half-time ($t_{1/2}$) obtained from the curves in (a) is shown versus $\phi$. The different symbol shapes indicate different values for $\phi$. In (a) they also refer to experimental data.

FIG. S2. ApoC2. In (a) the mass concentration of fibrils $M(t)$ is shown and was fit to ThT data from Binger, et al. In (b) the predicted $L(t)$ is shown for different crowder concentrations $\phi$. In (c) the half-time ($t_{1/2}$) obtained from the curves in (a) is shown versus $\phi$. The different symbol shapes indicate different values for $\phi$. In (a) they also refer to experimental data.

FIG. S3. Aβ(1-40) in agitated (shaken) conditions. In (a) the mass concentration of fibrils $M(t)$ is shown and was fit to ThT data from Lee, et al. In (b) the predicted $L(t)$ is shown for different crowder concentrations $\phi$. In (c) the half-time ($t_{1/2}$) obtained from the curves in (a) is shown versus $\phi$. The different symbol shapes indicate different values for $\phi$. In (a) they also refer to experimental data.
FIG. S4. β-LAC. In (a) the mass concentration of fibrils $M(t)$ is shown and was fit to ThT data from Ma, et al. In (b) the predicted $L(t)$ is shown for different crowder concentrations $\phi$. In (c) the half-time ($t_{1/2}$) obtained from the curves in (a) is shown versus $\phi$. The different symbol shapes indicate different values for $\phi$. In (a) they also refer to experimental data.

FIG. S5. β2m. In (a) the mass concentration of fibrils $M(t)$ is shown and was fit to ThT data from Luo, et al. In (b) the predicted $L(t)$ is shown for different crowder concentrations $\phi$. In (c) the half-time ($t_{1/2}$) obtained from the curves in (a) is shown versus $\phi$. The different symbol shapes indicate different values for $\phi$. In (a) they also refer to experimental data.

FIG. S6. IAPP. In (a) the mass concentration of fibrils $M(t)$ is shown and was fit to ThT data from Seeliger, et al. In (b) the predicted $L(t)$ is shown for different crowder concentrations $\phi$. In (c) the half-time ($t_{1/2}$) obtained from the curves in (a) is shown versus $\phi$. The different symbol shapes indicate different values for $\phi$. In (a) they also refer to experimental data.
FIG. S7. Rabbit PrP. In (a) the mass concentration of fibrils $M(t)$ is shown and was fit to ThT data from Zhou, et al.\textsuperscript{3} In (b) the predicted $L(t)$ is shown for different crowder concentrations $\phi$. In (c) the half-time ($t_{1/2}$) obtained from the curves in (a) is shown versus $\phi$. The different symbol shapes indicate different values for $\phi$. In (a) they also refer to experimental data.

FIG. S8. BCA. In (a) the mass concentration of fibrils $M(t)$ is shown and was fit to ThT data from Mittal, et al.\textsuperscript{12} We could not obtain the half-time for the data at $\phi = 0.26$ (purple right-facing triangles) using Eq. (S4). In (b) the predicted $L(t)$ is shown for different crowder concentrations $\phi$. In (c) the half-time ($t_{1/2}$) obtained from the curves in (a) is shown versus $\phi$. The different symbol shapes indicate different values for $\phi$. In (a) they also refer to experimental data.

FIG. S9. $A\beta(1-40)$. In (a) the mass concentration of fibrils $M(t)$ is shown and was fit to ThT data from Lee, et al.\textsuperscript{5} In (b) the predicted $L(t)$ is shown for different crowder concentrations $\phi$. In (c) the half-time ($t_{1/2}$) obtained from the curves in (a) is shown versus $\phi$. The different symbol shapes indicate different values for $\phi$. In (a) they also refer to experimental data.