Supporting Information

to

"What is Responsible for Atypical Dependence of the Rate of Amyloid Formation on Protein Concentration: Fibril-Catalyzed Initiation of New Fibrils or Competition with Oligomers?"

by

Alexey V. Finkelstein, Nikita V. Dovidchenko and Oxana V. Galzitskaya

Institute of Protein Research, Russian Academy of Sciences, Pushchino, 142290, Moscow Region, Russian Federation

1. Free monomers, oligomers and polymers in solution

When a protein solution contains free monomers, oligomers and polymers, their momentary concentrations obey the equations

$$[\mathrm{P_n}]+[C] = [C_\Sigma], \tag{S.1}$$

$$[C_1]+m[C_\mathrm{m}] = [C], \tag{S.2}$$

where:

$[C_\Sigma]$ is the total concentration of protein molecules in solution,

$[\mathrm{P_n}]$ is the concentration of polymerized monomers,

$[C]$ is the concentration of non-polymerized monomers,

$[C_1]$ is the concentration of free monomers in solution,

$[C_\mathrm{m}]$ is the concentration of oligomers that contain m monomers ("m-mers") in solution,

When momentary concentrations of free monomers and m-mers are in equilibrium, they obey the law of mass action[1]

$$[C_1]^m = K \times [C_\mathrm{m}], \tag{S.3}$$

where K is the equilibrium constant. Equations (S.2), (S.3) determine $[C_1]$ (and then $[C_\mathrm{m}]$ as well) as functions of $[C]$, m and K from the following equation:

$$[C_1]^m = (K/m) \times ([C] - [C_1]). \tag{S.3a}$$

At low $[C]$, when $K >> m[C]^{m-1}$, we have $[C_1] \approx [C] \times (1 - m[C]^{m-1}/K) \approx [C]$ to a first approximation. At very high $[C]$, when $K << m[C]^{m-1}$, we have $[C_1] \approx (K[C]/m)^{1/m} << [C]$ to a first approximation.

2. Linear regime of polymer growth

This scenario of fibril growth from free monomers (Fig. 1S) is described by the following system of equations:[2]

$$\begin{cases} \dfrac{d[\mathrm{P^E}]}{dt} = 2k_+[C_1]^{n^*} \\ \dfrac{d[\mathrm{P}_n]}{dt} = n_s k_+[C_1]^{n^*} + k_2[C_1][\mathrm{P^E}]; \end{cases} \tag{S.4}$$

the first of them describes an increase of fibril ends, the last – an increase of monomers involved in fibrils. Here $[\mathrm{P^E}]$ is concentration of the ends of all polymers made of monomers, n^* is the effective number of the former free monomers in the unstable nucleus of polymerization (Fig. 1S), n_s is the number of monomers in the smallest stable polymer (a "seed"), and k_+, k_2 are the effective rate constants for initiation and elongation of polymers ("2" in "$2k_+$" takes into account that creation of one new polymer creates two new polymer ends). In the following equations, we assume that k_2, as well as

k_+, are the concentration-independent constants. Analyses of these assumptions see in Chapter 4 of this Supplement.

The term $k_2[C_1][\mathrm{P}^{\mathrm{E}}]$ in the last equation stands for involvement of monomers in the seeded polymerization, and the term $n_s k_+[C_1]^{n^*}$ in this equation stands for involvement of monomers in "seeds" (i.e., the smallest stable polymers made of n_s monomers); one can neglect this term because the rate of seeding (associated with k_+) much smaller than the rate of seeded polymerization (associated with k_2).

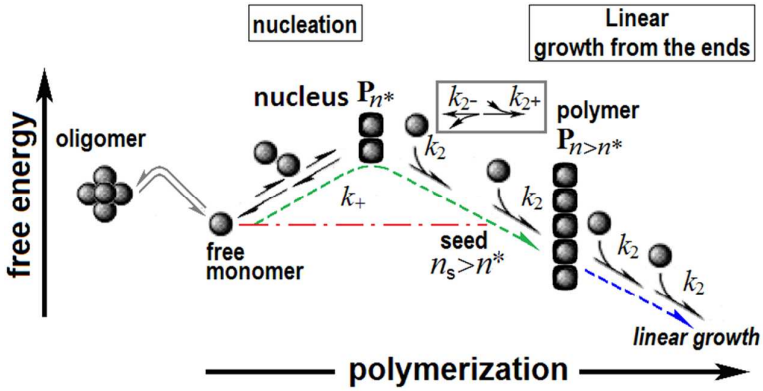

Figure 1S. Free energy change during linear growth of a polymer and formation of competing oligomers. The "polymerization seed" of n_s monomers is the smallest stable polymer; all polymers with $n < n_s$ are unstable. The effective rate of fibril growth is $k_{2+}[C_1] - k_{2-}$, where $k_{2+}[C_1]$ is the true rate of polymerization, and k_{2-} is the true rate of depolymerization; it is assumed that the free monomer concentration $[C_1]$ is high, so that the polymerization is irreversible; then k_{2-} is negligible as compared to $k_{2+}[C_1]$, and the effective rate of polymerization of a fibril can be presented as $k_2[C_1] \approx k_{2+}[C_1]$.

Then Eq. (S.4) takes a simpler form[2]

$$\begin{cases} \dfrac{d[\mathrm{P}^{\mathrm{E}}]}{dt} = 2k_+[C_1]^{n^*} \\ \dfrac{d[\mathrm{P}_n]}{dt} \approx k_2[C_1][\mathrm{P}^{\mathrm{E}}], \end{cases} \tag{S.5}$$

where t is time. Using conditions $[\mathrm{P_n}]+[C] = [C_\Sigma] = const$ and $[C_1] \equiv [C]\times([C_1]/[C])$, this system can be presented as

$$\begin{cases} \dfrac{d[\mathrm{P}^{\mathrm{E}}]}{dt} = 2k_+^1[C]^{n^*} \\ \dfrac{d[C]}{dt} \approx -k_2^1[C][\mathrm{P}^{\mathrm{E}}], \end{cases} \tag{S.6}$$

where

$$k_+^1 = k_+\left(\frac{[C_1]}{[C]}\right)^{n^*}, \quad k_2^1 = k_2\frac{[C_1]}{[C]} \tag{S.7}$$

depend on relative concentrations of free monomers and oligomers. However, to a first approximation, one can take k_+^1 and k_2^1 as constants

$$k_+^1 \approx k_+\left(\frac{[C_1]_{t=0}}{[C_\Sigma]}\right)^{n^*}, \quad k_2^1 \approx k_2\frac{[C_1]_{t=0}}{[C_\Sigma]} \tag{S.7a}$$

during the first half period of amyloid aggregation (when it proceeds without pre-formed seeds), because then $[P_n]_{t=0} = 0$, $[P^E]_{t=0} = 0$, $d[P^E]/dt|_{t=0} \neq 0$, while $[C_1]_t \approx [C_1]_{t=0}$, $[C]_t \approx [C]_{t=0}$ (and $[C_1]_{t=0} \sim [C]_{t=0} \approx [C_\Sigma]$ when the fraction of oligomers is relatively small); thus, $[P^E]$ changes with time t much faster than $[C]$, $[C_1]$, k_+^1 and k_2^1. In this case, the system (S.6) can be solved (with constant k_+^1, k_2^1) analytically using a substitution $\xi_t = ([C]_t/[C_\Sigma])^{n^*}$ (see[2] item 1 in Supporting Information). The time-dependent degree of involvement of monomers in polymers is[2]

$$\mu_t \equiv \frac{[P_n]_t}{[C_\Sigma]} = 1 - \frac{[C]_t}{[C_\Sigma]} = 1 - \frac{4^{1/n^*}\exp(-t/T)}{\left[1+\exp(-n^* t/T)\right]^{2/n^*}}, \tag{S.8}$$

where

$$T = \sqrt{\frac{n^*}{4k_+^1 k_2^1}}[C_\Sigma]^{-\frac{n^*}{2}} \approx \sqrt{\frac{n^*}{4k_+ k_2}}[C_\Sigma]^{-\frac{n^*}{2}} \times \left(\frac{[C_1]_{t=0}}{[C_\Sigma]}\right)^{-\frac{n^*+1}{2}} \equiv \sqrt{\frac{n^*}{4k_+ k_2}}[C_1]_{t=0}^{\frac{-n^*}{2}} \times \left(\frac{[C_1]_{t=0}}{[C_\Sigma]}\right)^{-\frac{1}{2}} \tag{S.9}$$

is the single time parameter for the linear regime of the fibril growth process.

Thus, for this regime, logarithms of all characteristic times of fibril growth can be presented as

$$\ln(\text{time}) = const(n^*, k_+, k_2) - \frac{n^*}{2}\ln([C_\Sigma]) - \frac{n^*+1}{2}\ln\left(\frac{[C_1]_{t=0}}{[C_\Sigma]}\right) \tag{S.10}$$

These characteristic times include: T_{lag}, duration of the lag-period; T_2, duration of the fast transition period; and $t_{1/2}$, the transition half-time (i.e., the time required for 50% protein to aggregate); these T_{lag}, T_2, and $t_{1/2}$ have the same dependence on $\ln[C_\Sigma]$ and $[C_1]_{t=0}/[C_\Sigma]$, but different $const(n^*, k_+, k_2)$.

3. Exponential regimes of polymer growth: "fragmentation" and "bifurcation" scenarios

The "fragmentation" scenario (Fig. 2S a, b) can be described by the following system of equations:[2]

$$\begin{cases} \dfrac{d[P^E]}{dt} = 2k_+[C_1]^{n^*} + 2\lambda_+[C_1]^{n_2}\{[P_n] - n_s[P^E]\} - \lambda_-[P^E]^2 \\ \dfrac{d[P_n]}{dt} = n_s k_+[C_1]^{n^*} + n_2\lambda_+[C_1]^{n_2}\{[P_n] - n_s[P^E]\} + k_2[C_1][P^E] \end{cases} \tag{S.11}$$

where the rate constant λ_+ describes the rate of "secondary nucleation" by fragmentation, i.e., emergence of a pair of stable fibril fragments (with two "new" ends) from an "old" fibril *and* n_2 monomers forming the "secondary nucleus" of fragmentation (Fig. 2S b), or from an "old" fibril alone (Fig. 2S a) (in this case, the "secondary nucleus" consists of n_2=0 monomers), while the rate constant λ_- describes the rate of fusion of the ends of separate fibrils. In the both cases (a and b), the fragmentation creates m=2 new ends of the polymer. As above, we assume that λ_+ is a concentration-independent constant. For analysis of such an assumption, see Chapter 4 of this Supplement.

The "bifurcation" scenario (Fig. 2S c, d) can be described by the following system of equations:[2]

$$\begin{cases} \dfrac{d[P^E]}{dt} \approx 2k_+[C_1]^{n^*} + m\lambda_+[C_1]^{n_2}[P_n] \\ \dfrac{d[P_n]}{dt} = n_s k_+[C_1]^{n^*} + n_2\lambda_+[C_1]^{n_2}[P_n] + k_2[C_1][P^E], \end{cases} \tag{S.12}$$

where λ_+ is the "secondary nucleation" rate constant by bifurcation, m is the number of new polymer ends created by one bifurcation, which can be m=1 when a new fibril forms as a branch of the existing one, thus producing one new end (Fig. 2S c), or m=2 in the case of lateral formation of a new fibril, thus producing two new ends (Fig. 2S d).

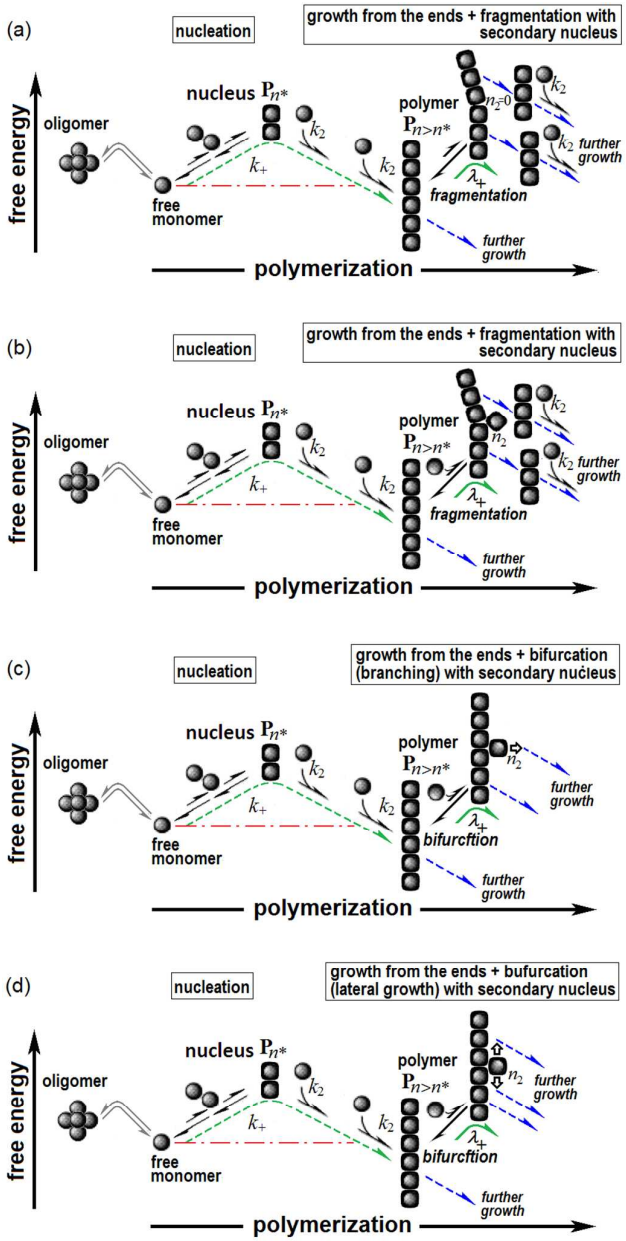

Figure 2S. Exponential regimes of polymer growth: (a) growth with a zero-size (n_2=0) "secondary nucleus" of fragmentation; (b) growth with a fragmentation "secondary nucleus" of n_2>0 monomers; (c) growth with the bifurcation "secondary nucleus" creating the fibril's new branch with one new growing end; (d) growth with the bifurcation "secondary nucleus" leading to the further lateral growth of the new fibril from its both ends.

Neglecting relatively small terms in systems (S.11), (S.12) (in the last equations of these systems: those with initiation rate constants k_+, λ_+, as compared to those with the elongation rate constant k_2; in the first equations of the systems: those with concentration of ends $[P^E]$, as compared to that of free monomers $[C_1]$), one obtains a generalized system in the form

$$\begin{cases} \dfrac{d[P^E]}{dt} \approx 2k_+[C_1]^{n^*} + \lambda_+^0[C_1]^{n_2}[P_n] \\ \dfrac{d[P_n]}{dt} \approx k_2[C_1][P^E], \end{cases} \qquad \text{(S.13)}$$

which describes both fragmentation (with $\lambda_+^0 = 2\lambda_+$) and bifurcation (with $\lambda_+^0 = m\lambda_+$) scenarios of fibril growth.

Using conditions $[P_n]+[C] = [C_\Sigma] = const$ and $[C_1] \equiv [C]\times([C_1]/[C])$, this system can be presented as

$$\begin{cases} \dfrac{d[P^E]}{dt} = 2k_+^1[C]^{n^*} + \lambda_+^1[C]^{n_2}[P_n] \\ \dfrac{d[C]}{dt} \approx -k_2^1[C][P^E], \end{cases} \tag{S.13a}$$

where

$$k_+^1 = k_+\left(\frac{[C_1]}{[C]}\right)^{n^*}, \quad k_2^1 = k_2\frac{[C_1]}{[C]}, \quad \lambda_+^1 = \lambda_+^0\left(\frac{[C_1]}{[C]}\right)^{n_2} \tag{S.14}$$

depend on relative concentrations of free monomers and oligomers. However, to a first approximation, one can take k_+^1, k_2^1 and λ_+^1 as constants,

$$k_+^1 \approx k_+\left(\frac{[C_1]_{t=0}}{[C_\Sigma]}\right)^{n^*}, \quad k_2^1 \approx k_2\frac{[C_1]_{t=0}}{[C_\Sigma]}, \quad \lambda_+^1 \approx \lambda_+^0\left(\frac{[C_1]_{t=0}}{[C_\Sigma]}\right)^{n_2}, \tag{S.14a}$$

during the first period of amyloid aggregation (when it proceeds without pre-formed seeds), because then $[P_n]_{t=0} = 0$, $[P^E]_{t=0} = 0$, $d[P^E]/dt|_{t=0} \neq 0$, while $[C_1]_t \sim [C_1]_{t=0}$, $[C]_t \sim [C]_{t=0}$ (and $[C_1]_{t=0} \sim [C]_{t=0} \approx [C_\Sigma]$ when the fraction of oligomers is relatively small). Thus, $[P^E]_t$ and $[P_n]_t = [C_\Sigma] - [C]_t$ change with time t much faster than $[C]_t$. As a result, the system (S.13a) with constant k_+^1, k_2^1, λ_+^1 and $[C] = [C_\Sigma]$ takes the form

$$\begin{cases} \dfrac{d[P^E]}{dt} \approx 2k_+^1[C_\Sigma]^{n^*} + \lambda_+^1[C_\Sigma]^{n_2}\{[C_\Sigma]-[C]\} \\ \dfrac{d[C]}{dt} \approx -k_2^1[C_\Sigma][P^E], \end{cases} \tag{S.13b}$$

(see[2] items 3, 4 in Supporting Information), which is sufficiently accurate for the first part of the process.

Now, using the time-dependent degree of involvement of monomers in polymers, $\mu_t \equiv 1-[C]_t/[C_\Sigma]$ (cf. Eq. (S.8)), one can present the system (S.13b) in a simple form

$$\begin{cases} \dfrac{d[P^E]}{dt} \approx 2k_+^1[C_\Sigma]^{n^*} + \lambda_+^1[C_\Sigma]^{n_2+1}\mu_t \\ \dfrac{d\mu_t}{dt} \approx k_2^1[P^E], \end{cases} \tag{S.13c}$$

and one can see that

$$\frac{d^2\mu_t}{dt^2} \approx k_2^1\frac{d[P^E]}{dt} = (\mu_t + A)/T_2^2, \tag{S.15}$$

where

$$A = \frac{2k_+^1[C_\Sigma]^{n^*}}{\lambda_+^1[C_\Sigma]^{n_2+1}} = \frac{2k_+[C_\Sigma]^{n^*}}{\lambda_+^0[C_\Sigma]^{n_2+1}}\times\left(\frac{[C_1]_{t=0}}{[C_\Sigma]}\right)^{n^*-n_2} \equiv \frac{2k_+[C_1]_{t=0}^{n^*-n_2}}{\lambda_+^0[C_\Sigma]} \tag{S.16}$$

and

$$T_2 = \frac{1}{\sqrt{\lambda_+^1k_2^1}}[C_\Sigma]^{-\frac{n_2+1}{2}} = \frac{1}{\sqrt{\lambda_+^0k_2}}[C_\Sigma]^{-\frac{n_2+1}{2}}\times\left(\frac{[C_1]_{t=0}}{[C_\Sigma]}\right)^{-\frac{n_2+1}{2}} \equiv \frac{1}{\sqrt{\lambda_+^0k_2}}\times[C_1]_{t=0}^{-\frac{n_2+1}{2}}, \tag{S.17}$$

so that

$$\mu_t = A \times \left[\frac{\exp(t/T_2) + \exp(-t/T_2)}{2} - 1 \right] \tag{S.18}$$

is a solution[2] of the system (S.14), which is valid for the first part of the process. Note that the characteristic time T_2 is now (in the case of exponential growth) independent of the primary fibril initiation rate constant k_+.

An analysis of this solution shows two main regimes of fibril growth with possible fragmentation or bifurcation:

1) With $A >> 1$ (that is, with $k_+ >> \lambda_+^0 [C_\Sigma]/[C_1]_{t=0}^{n^*-n_2}$), the value in parentheses in Eq. (S.18) is small (because μ_t cannot exceed unity), which means that 50% of the protein will be aggregated during the time $t_{1/2} << T_2$, where

$$t_{1/2} = \frac{T_2}{\sqrt{A}} = \frac{1}{\sqrt{2k_+^1 k_2^1}} [C_\Sigma]^{-\frac{n^*}{2}} \approx \sqrt{\frac{1}{2k_+ k_2}} [C_\Sigma]^{-\frac{n^*}{2}} \times \left(\frac{[C_1]_{t=0}}{[C_\Sigma]} \right)^{-\frac{n^*+1}{2}} \equiv \sqrt{\frac{1}{2k_+ k_2}} [C_1]_{t=0}^{\frac{-n^*}{2}} \times \left(\frac{[C_1]_{t=0}}{[C_\Sigma]} \right)^{-\frac{1}{2}}. \tag{S.19}$$

This $t_{1/2}$ does not contain λ_+^1 or λ_+^0 (showing that neither fragmentation nor bifurcation is important when $A >> 1$) and the same dependence on $[C_\Sigma]$ and $[C_1]_{t=0}/[C_\Sigma]$ as the T value given by Eq. (S.9) for the linear regime of the fibril growth process.

2) With $A << 1$, the value in parentheses in Eq. (S.18) is large at the time $t_{1/2}$ when 50% of the protein is aggregated, which means that $t_{1/2} >> T_2$, and $t_{1/2}$ can be estimated from the equation

$$(A/2) \times \exp(t_{1/2}/T_2) \approx 1/2, \tag{S.20}$$

that is

$$t_{1/2} \approx T_2 \times \ln(1/A) = [C_\Sigma]^{-\frac{n_2+1}{2}} \times \frac{\ln(1/A)}{\sqrt{\lambda_+^1 k_2^1}} = \left([C_\Sigma] \times \frac{[C_1]_{t=0}}{[C_\Sigma]} \right)^{-\frac{n_2+1}{2}} \times \frac{\ln(1/A)}{\sqrt{\lambda_+^0 k_2}} \equiv [C_1]_{t=0}^{-\frac{n_2+1}{2}} \times \frac{\ln(1/A)}{\sqrt{\lambda_+^0 k_2}}, \tag{S.21}$$

or

$$\ln(t_{1/2}) \approx -\frac{n_2+1}{2} \ln[C_\Sigma] + \ln\left(\frac{\ln(1/A)}{\sqrt{\lambda_+^1 k_2^1}} \right) = -\frac{n_2+1}{2} \left(\ln[C_\Sigma] + \ln\left(\frac{[C_1]_{t=0}}{[C_\Sigma]} \right) \right) + \ln\left(\frac{\ln(1/A)}{\sqrt{\lambda_+^0 k_2}} \right)$$

$$\equiv -\frac{n_2+1}{2} \ln[C_1]_{t=0} + \ln(\ln(1/A)) - \frac{1}{2} \ln(\lambda_+^0 k). \tag{S.22}$$

Note that this characteristic time $t_{1/2}$ is now (in the case of exponential growth with $A << 1$, i.e., when $k_+ << \lambda_+^0 [C_\Sigma]/[C_1]_{t=0}^{n^*-n_2}$) has a very weak (logarithmic) dependence on the primary fibril initiation rate constant k_+ participating in A.

The dependence on $\ln[C_\Sigma]$ (or $\ln[C_1]_{t=0}$) comes mostly from the first term of the sum; the value of $\ln(\ln(1/A))$ also depends on $\ln[C_\Sigma]$, but this dependence is rather weak, because $\ln(1/A) = \ln(\lambda_+^0/2k_+) + \ln[C_\Sigma] + \ln[C_1]_{t=0}^{n^*-n_2} >> 1$.

If the difference between $[C_1]$ and $[C_\Sigma]$ is ignored (or the monomer fraction $[C_1]_{t=0}/[C_\Sigma]$ does not change with $[C_\Sigma]$), the first derivative over $\ln[C_\Sigma]$ is

$$\frac{d \ln(t_{1/2})}{d \ln[C_\Sigma]} \approx -\frac{n_2+1}{2} + \frac{n_2+1-n^*}{\ln(1/A)}, \tag{S.23}$$

where the second summand is small because $\ln(1/A)$ is large, and the second derivative

$$\frac{d^2 \ln(t_{1/2})}{d(\ln[C_\Sigma])^2} \approx -\frac{(n_2+1-n^*)^2}{(\ln(1/A))^2}, \tag{S.24}$$

is small too (for the same reason) and, importantly, it is *never* positive for all values of n^*, n_2 and A.

Thus, the strong *positive* curvature of the experimental $\ln(t_{1/2})$ on $\ln[C_\Sigma]$ dependence observed by Meisl *et al.*[3] can in no way be explained by the non-linear (with respect to $\ln[C_\Sigma]$) summand $\ln(\ln(1/A))$ present in $\ln(t_{1/2})$, and this positive curvature has to be attributed to a non-linear dependence of $[C_1]$ on $[C_\Sigma]$.

4. Kinetic constants for the polymerization process

The effective rate of polymerization of a fibril is $k_{2+}[C_1] - k_{2-}$, where $k_{2+}[C_1]$ is the true rate of polymerization, and k_{2-} is the true rate of depolymerization (Fig. 1S). If the free monomer concentration $[C_1]$ is much higher than the critical concentration of aggregation

$$[\mathrm{C}_{1*}] \equiv k_{2-}/k_{2+}, \tag{S.25}$$

the polymerization process is irreversible, and its effective rate, $k_2[C_1]$, is $k_{2+}([C_1] - [\mathrm{C}_{1*}]) \approx k_{2+}[C_1]$. This approximation ($k_2[C_1] \approx k_{2+}[C_1]$) is not valid only when $[\mathrm{C}_1]$ is close to $[\mathrm{C}_{1*}]$ (say, when $[\mathrm{C}_{1*}] < [\mathrm{C}_1] \lesssim 3[\mathrm{C}_{1*}]$). However, the concentration range examined in Ref. 3 covers almost two orders of magnitude (Fig. 1), so for the whole or at least the best part of this range, and especially for the high concentrations of interest, the approximation $k_2[C_1] \approx k_{2+}[C_1]$ (with the concentration-independent elongation constant k_2) is valid.

The rate of primary fibril initiation (estimated in the "steady-state approximation",[4] widely used in chemical kinetics) has a general form $k_+[C_1]^{n^*-1}$, which follows from the kinetic equation (S.2),

$$d[\frac{1}{2}\mathrm{P}^{\mathrm{E}}]/dt = k_+[C_1]^{n^*} \equiv [C_1]\cdot\left\{k_+[C_1]^{n^*-1}\right\}, \tag{S.26}$$

Where $[\frac{1}{2}\mathrm{P}^{\mathrm{E}}]$ is the concentration of fibrils (because $[\mathrm{P}^{\mathrm{E}}]$ is the concentration of fibril ends, and a fibril has two ends), $[C_1]$ is the concentration of free monomers, n^* is the effective number of monomers absorbed by a primary nucleus of polymerization (that is, the size of a polymer corresponding to the free-energy barrier, i.e., having the highest free energy over the polymerization pathway), and k_+ is the rate constant.

It should be noted that the steady-state approximation usually implies that the high free-energy barrier corresponds to one step of the reaction; however, this approximation is also applicable when the free-energy barrier includes several steps[5-7] (which is just the case for the fibril initiation, see Fig. 1S).

The rate constant k_+ is usually assumed to be concentration-independent; this assumption is strictly valid for the barrier corresponding to one step of reaction,[4] but the case of a multi-step barrier requires more attentive consideration.[5-7] Figure 3S presents a multi-state process with the free-energy barrier (corresponding to the true activated complex) made of n^{**} monomers.

The rate $æ_{1\to2\to3\to\ldots}$ of overcoming of a long free-energy barrier (leading to the n_s-monomer seed, see Fig. 3S) can be presented[5-7] as

$$\frac{1}{æ_{1\to2\to3\ldots}} \cong \sum_{j=2}^{n_s} \exp\left[\frac{G_{j-1}-G_1}{k_BT}\right]\frac{1}{æ_{j-1\to j}}, \tag{S.26}$$

where $æ_{j\text{-}1\to j}$ is the rate of passage from the state j-1 to j; that is, in the polymerization reaction shown in Fig. 3S, $æ_{1\to2} = k_{+1}[C_1]^{n^{**}\text{-}1}$ (where n^{**} is the number of monomers in the activated, highest-free-energy complex), and $æ_{2\to3} = æ_{3\to4} = \ldots = k_{2+}[C_1]$.

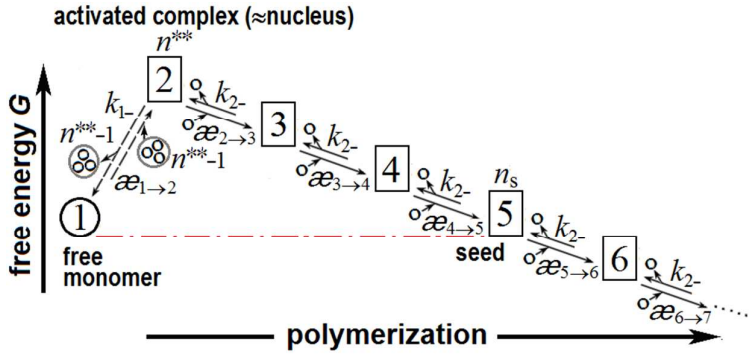

Figure 3S. Free energy change during linear growth of a polymer. The activated complex consists of n^{**} monomers (in the picture, n^{**}=2 for simplicity, but between states "1" and "2" there can be a few intermediate states with free energies between G_1 and G_2). All polymers shorter than the "seed" are unstable. $æ_{j\text{-}1\to j}$ (j>1) is the rate of passage from the state j-1 to j. The true rate of formation of the activated complex is $æ_{1\to n^{**}=2} = k_{1+}[C_1]^{n^{**}\text{-}1}$. The true rate of decay of this complex into monomers is $k_{1\text{-}}$. The true rates of polymerization after formation of the activated complex are assumed to be equal to $k_{2+}[C_1]$ (i.e., $æ_{n^*=2\to 3} = æ_{3\to 4} = \ldots = k_{2+}[C_1]$). The corresponding depolymerization rates are all equal to $k_{2\text{-}}$. Thus, the free energy change at each polymerization step, i.e., G_j-$G_{j\text{-}1}$ at $j > 2$, is $k_B T \cdot \ln\left(\frac{k_{2-}}{k_{2+}[C_1]}\right) = k_B T \cdot \ln\left(\frac{[C_{1*}]}{[C_1]}\right)$, and the free energy change at the initiation step, G_2-G_1, is $k_B T \cdot \ln\left(\frac{k_{1-}}{k_{1+}[C_1]^{n^{**}-1}}\right)$.

In the steady-state approximation, after passing the barrier state "2" (i.e., at $j > 2$) we have $\frac{G_{j-1}-G_1}{k_B T} = \ln\left(\frac{k_{2-}}{k_{2+}[C_1]}\right) \equiv \ln\left(\frac{[C_{1*}]}{[C_1]}\right)$ (that is, $\frac{G_{j-1}-G_1}{k_B T} = \frac{G_2-G_1}{k_B T} + (j-2)\cdot\ln\left(\frac{[C_{1*}]}{[C_1]}\right)$ at $j > 2$), and for the first, "barrier-climbing" step we have $\frac{G_2-G_1}{k_B T} = \ln\left(\frac{k_{1-}}{k_{1+}[C_1]^{n^{**}-1}}\right)$. Therefore,

$$\frac{1}{æ_{1\to 2\to 3\ldots}} \cong \frac{1}{k_{1+}[C_1]^{n^{**}-1}} + \frac{1}{k_{2+}[C_1]} \cdot \frac{k_{1-}}{k_{1+}[C_1]^{n^{**}-1}} \sum_{j=3}^{n_s} \left(\frac{[C_{1*}]}{[C_1]}\right)^{j-3}. \quad \text{(S.27)}$$

Now, because $\sum_{j=3}^{n_s}\left(\frac{[C_{1*}]}{[C_1]}\right)^{j-3} = \frac{1-\left(\frac{[C_{1*}]}{[C_1]}\right)^{n_s-2}}{1-\frac{[C_{1*}]}{[C_1]}} \approx \frac{[C_1]}{[C_1]-[C_{1*}]} \approx 1$ when $[C_1] \gg [C_{1*}]$ (which is true for the case examined in Ref. 3, see above), we have

$$\frac{1}{æ_{1\to 2\to 3\ldots}} \approx \frac{k_{1-}+k_{2+}[C_1]}{k_{2+}\cdot k_{1+}[C_1]^{n^{**}}}. \quad \text{(S.27a)}$$

Equation (S.26a) means that $æ_{1\to 2\to 3\ldots} \approx k_{1+}[C_1]^{n^{**}-1} \times \left(\frac{k_{2+}[C_1]}{k_{1-}+k_{2+}[C_1]}\right)$ is the rate of formation of the effective nucleus.

The multiplier $\left(\frac{k_{2+}[C_1]}{k_{1-}+k_{2+}[C_1]}\right)$ is close to $\frac{k_{2+}[C_1]}{k_{1-}}$ when $k_{2+}[C_1] << k_{1-}$, but when $[C_1]$ grows, the multiplier grows up to unity, and then (with $k_{2+}[C_1] >> k_{1-}$) remains constant. That is,

$$æ_{1\to 2\to 3\ldots} \approx \frac{k_{2+}}{k_{1-}} k_{1+}[C_1]^{n^{**}} \text{ when } [C_1] \le k_{1-}/k_{2+}, \quad \text{(S.28)}$$

and

$$æ_{1\to 2\to 3\ldots} \approx k_{1+}[C_1]^{n^{**}-1} \text{ when } [C_1] \ge k_{1-}/k_{2+}. \quad \text{(S.29)}$$

As a result, in all conditions the rate has a form "*constant*×[concentration]$^{\text{some\_power}}$", but these "*constants*" and "powers" are different.

With large $[C_1] \geq k_{1-}/k_{2+}$, the power is n^{**}-1, which means that the effective nucleus corresponds to n^{**} monomers, i.e., just to the n^{**}-monomer activated complex (the free energy maximum over the reaction pathway).

With small $[C_1] \leq k_{1-}/k_{2+}$, the power is n^{**}, which means that the effective nucleus is larger by one monomer: it corresponds to n^{**}+1 monomers, because it includes the n^{**}-monomer activated complex *plus* one additional, absorbed monomer (deficient in solution).

The transition from one regime to another occurs in vicinity of the specific free monomer concentration

$$[C_{1**}] = k_{1-}/k_{2+}; \tag{S.30}$$

note that $[C_{1**}] \gg [C_{1*}]$ because k_{1-}, the rate of monomer loss by the unstable nucleus is much greater than k_{2-}, the rate of monomer loss by a stable fibril.

The same considerations refer to overcoming the free-energy barriers associated with the secondary nucleation (where the rate constant is λ_+).

Recurring to the slope of the $\ln(t_{1/2})$ vs. $\ln([C_\Sigma])$ dependence considered in the main article, we see from Eq. (S.22) and the above results that this slope can change from $0.5\times n^{**}$ at low (below $[C_{1**}]$) concentration of monomers to $0.5\times(n^{**}\text{-}1)$ at their high (above $[C_{1**}]$) concentration. The difference in slopes for the low and high monomer concentrations (obtained here from analysis of multi-step reactions of fibril initiation) is similar to that suggested by Meisl *et al.*[3] from their model of fibril-catalyzed initiation with saturation of catalytic sites; both of these produce effects which appear insufficient for the observed[3] deceleration of the amyloid formation rate at very high protein concentrations if competition between oligomers and fibrils is not taken into account.

5. Additional electron microscopy images of the growing fibrils and oligomers

One can find in literature numerous electron microscopy (EM) and atomic force microscopy (AFM) images of amyloid fibrils; many of these images present aggregates and oligomers as well (see, e.g., Refs. 8-16).

Figure 4S(A) shows that amorphous aggregates and various oligomers disappear with the growth of amyloid fibrils; however, the resolution of these EM images does not allow seeing the monomeric structure of these aggregates and oligomers.

Figure 4S(B), made in a higher resolution, allows seeing monomeric structure of Aβ peptide oligomers and shows that they apparently consist of 4-6 monomers.

The EM studies are performed at protein concentration of about 0.2 mg/ml = 50 μM (for Aβ peptides); but, unfortunately, this is a nominal (i.e., initial) concentration only, because during the sample drying (which is a necessary step of EM studies) the actual protein concentration increases manifold, and the structures that we see may pertain a concentration which is much higher than the nominal one, if the oligomer restructuring is faster that drying.

The AFM studies can be done without drying (see, e.g. Refs. 12, 14), but, unfortunately, the AFM resolution does not allow seeing the monomeric structures of oligomers.

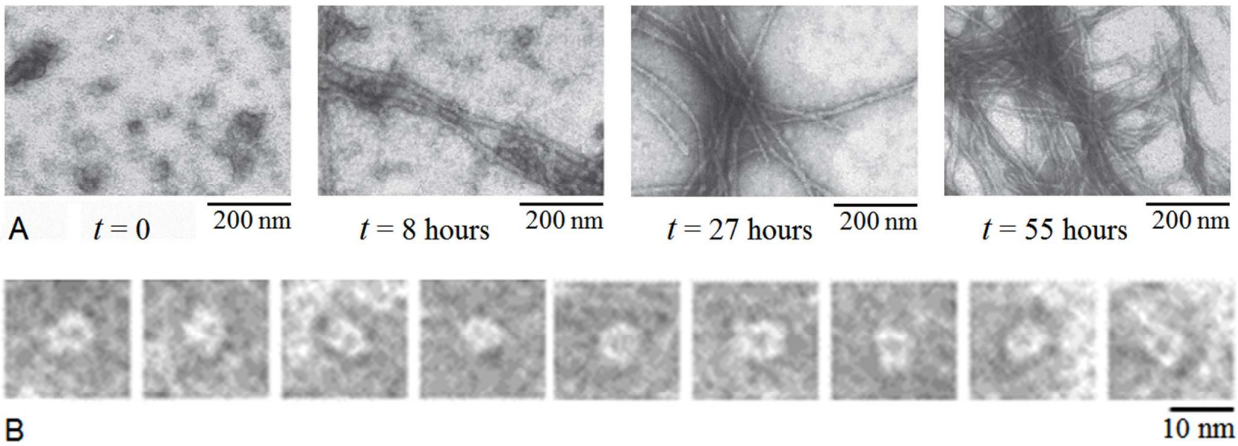

Figure 4S. (A) Kinetics of amyloid formation by recombinant Aβ40 peptide (50mMTris–HCl, pH7.5, 25°C, 5%DMSO, $[C_{\Sigma}]_{A\beta 40}$ = 0.2 mg/ml = 50 μM): Fibrils arise and grow with time, oligomers and amorphous aggregates (including those covering the fibrils) gradually disappear. Adapted from Ref. 15. (B) Additional electron microscopy images of oligomers formed by Aβ42) peptides. Adapted from Ref. 16.

Acknowledgements

We are grateful to T.P.J. Knowles for valuable discussions and to E.V. Serebrova for assistance in manuscript preparation.

This part of the study was supported in part by the Russian Academy of Sciences Program "Molecular and Cell Biology" (Grant No. 01201358029).

References

(1) Pauling, L., **1970**. *General Chemistry*. W.H. Freeman & Co, New York (Chapter 11).

(2) Dovidchenko, N. V.; Finkelstein, A. V.; Galzitskaya, O. V. How to Determine the Size of Folding Nuclei of Protofibrils from the Concentration Dependence of the Rate and Lag-Time of Aggregation. I. Modeling the Amyloid Protofibril Formation. *J. Phys. Chem*. B **2014**, *118*, 1189-1197.

(3) 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. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 9384–9389.

(4) Eyring, H.J. The Activated Complex in Chemical Reactions. *Chem. Phys*. **1935**, *3*, 107−115.

(5) Becker, R.; Döring, W. Kinetische Behandlung der Keimbildung in Übersättigten Dämpfen. *Ann. Phys.* **1935***, 24*, 719−749.

(6) Goldstein, R. F.; Stryer, L. Cooperative Polymerization Reactions. *Biophys. J*. **1986**, *50*, 583−599.

(7) Finkelstein, A.V. Time to Overcome the High, Long and Bumpy Free-Energy Barrier in a Multi-Stage Process: The Generalized Steady-State Approach. *J. Phys. Chem. B* **2015**, *119*, 158-163.

(8) Shirahama, T.; Cohen, A.S. High-Resolution Electron Microscopic Analysis of the Amyloid Fibril. *J. Cell Biol*. **1967**, *33*, 679-708.

(9) Nielsen E.H.; Nybo, M.; Svehag, S.E. Electron Microscopy of Prefibrilar Structures and Amyloid Fibrils. *Methods in Enzymology* **1999**, *309*, 491-496.

(10) Lashuel, H.A.; Hartley, D.M.; Petre, B.M.; Wall, J.S.; Simon, M.N.; Walz, T.; Lansbury, P.T, Jr. Mixtures of Wild-Type and a Pathogenic (E22G) Form of Abeta40 in Vitro Accumulate Protofibrils, Including Amyloid Pores. *J. Mol. Biol.* **2003**, *332*, 795–808.

(11) Myers, S.L.; Jones, S.; Jahn, T.N.; Morten, I.J.; Tennent, G.A.; Hewitt, E.W., Radford, S.E. A Systematic Study of the Effect of Physiological Factors on Beta2-Microglobulin Amyloid Formation at Neutral Ph. *Biochemistry* **2006**, *45*, 2311-2321.

(12) Goldsbury, C.; Frey P.; Olivieri, V.; Aebi, U.; Müller, S.A. Multiple Assembly Pathways Underlie Amyloid-Beta Fibril Polymorphisms. *J. Mol. Biol*. **2005**, *352*, 282-298.

(13) Lu, J.X.; Qiang, W.; Yau, W.M.; Schweiters, C.D.; Meredith, S.C.; Tycko, R. Molecular Structure of β-Amyloid Fibrils in Alzheimer's Disease Brain Tissue. *Cell* **2013**, *154*, 1257–1268.

(14) Watanabe-Nakayama, T.; Ono, K.; Itami, M.; Takahashi, R.; Teplow, D.B.; Yamada, M. High-Speed Atomic Force Microscopy Reveals Structural Dynamics of Amyloid B1-42 Aggregates. *Proc. Natl. Acad. Sci. USA*, **2016**, *113*, 5835-5840

(15) Selivanova, O.M.; Grigorashvili, E.I.; Suvorina, M.Yu.; Dzhus, U.F.; Nikulin, A.D.; Marchenkov, V.V.; Surin, A.K.; Galzitskaya, O.V. X-ray Diffraction and Electron Microscopy Data for Amyloid Formation of Aβ40 and Aβ42. *Data Brief*. **2016**, *8*, 108-113.

(16) Suvorina, M.Y.; Selivanova, O.M.; Grigorashvili, E.I.; Nikulin, A.D.; Marchenkov, V.V.; Surin, A.K.; Galzitskaya, O.V. Studies of Polymorphism of Amyloid-β42 Peptide from Different Suppliers. *J. Alzheimers Dis*. **2015**, *47*, 583-593.