Electronic Supplementary Material (ESI) for Chemical Science.

Supporting Information:

Reaction network connectivity governs scaling behaviour and rate-determining steps in filamentous self-assembly.

Georg Meisl, Luke Rajah, Samuel I. A. Cohen, Manuela Pfammatter, Andela Saric, Erik Hellstrand, Alexander K. Buell, Adriano Aguzzi, Sara Linse, Michele Vendruscolo, Christopher M. Dobson and Tuomas P. J. Knowles

1 Approximations in the derivation of the moment equations

The original differential equations for M and P are obtained by calculating the moments of the master equation [1], which in turn is derived using mass action and considering how the aggregates of all sizes may inter-convert (as shown in the reaction network in Fig. 2). To obtain the differential equations presented here, several approximations are made, namely the fibril mass produced directly by nucleation, as well as the mass lost through fragmentation into pieces smaller than the critical nucleus, is negligible, hence the corresponding terms were not included in the above equations. These approximations are generally valid for systems of linear self-assembly that produce elongated fibrils, that are on average significantly larger than the critical nucleus size. For more details see Cohen et al. [2]. In addition we also neglect the back rates for all processes except for the formation of the intermediate species during elongation and secondary nucleation, denoted by M^* and P^* in the Petri net, as these rates usually represent only a minor contribution to the kinetics. These neglected processes include the dissociation of from the ends of fibrils, the reverse of fragmentation (i.e. end to end joining of fibrils) and the reverse of the nucleation processes. Descriptions that explicitly include the monomer dissociation can be found in Cohen et al. [2] and are very similar to the results presented here.

2 Primary nucleation in the presence of dominant secondary processes

Whatever the dominant process in a particular system might be, to aggregate from soluble monomer it must always be preceded by primary nucleation. Thus primary nucleation can always be expected to limit the rate of growth to a certain degree.

However, it has been shown [1, 2] that the effect of primary nucleation on a system dominated by a secondary process is manifested in the scaling laws only through a logarithmic correction, which was therefore neglected in the main text. To illustrate, consider the case of an aggregation reaction proceeding via unsaturated elongation and secondary nucleation:

$$\tau_{1/2} \sim \log(2\kappa^2/\lambda^2)\kappa^{-1} \tag{S1}$$

where

$$\kappa = \sqrt{m(0)k_+ m(0)^{n_2} k_2} \tag{S2}$$

$$\lambda = \sqrt{m(0)k_+ m(0)^{n_c-1} k_n} \tag{S3}$$

This results in a correction to the scaling exponent due to primary nucleation of the form:

$$\gamma = -\frac{1+n_2}{2} - \frac{1+n_2-n_c}{\log(\lambda^2/[2\kappa^2])} \tag{S4}$$

It is interesting to note from the functional form of the correction in Eq. (S4) that the scaling exponent is, therefore, affected by primary nucleation even at small values of the ratio λ^2/κ^2.

This is in contrast to the result obtained in the case of 2 secondary mechanisms competing in parallel (Eq. (5) of the main text), which shows that the presence of a weak additional secondary process, active in addition to a more dominant secondary process, is not reported to a significant level in the scaling exponent. Fundamentally, this difference emerges since in a polymerisation reaction where all of the peptide is initially monomeric, there must always be a timescale within which primary nucleation is the dominant mechanism, and is thus rate determining. This occurs before enough aggregated material is formed for secondary nucleation to become dominant.

If the primary process competes with the secondary process,

3 The effect of seeding

The network and scalings discussed in the main paper all focused on reactions of a solution of monomeric proteins alone, in the absence of preformed fibrils. The reason for this is two-fold: scalings are more informative in the absence of seeds and seeded experiments are inherently more difficult as more parameters need to be controlled,

such as the exact composition of the seed stock in terms of concentration and length distribution, as well as the time of addition of seeds.

The addition of such preformed seed fibrils at the beginning of the reaction does not alter the reaction network, i.e. a fibril that is prone to display surface catalysed secondary nucleation will still do so in the presence of seeds. However, the importance of some of the inherent processes of aggregation can decrease compared to the simple growth of the added seed fibrils. In particular, primary nucleation can sometimes be bypassed already at nM concentrations of seeds for μM concentrations of monomer (Aβ42) and the large effect on the kinetics of a small amount of seeds can be used as evidence for the presence of secondary processes [3].

By adding a large amount of preformed fibrils (i.e. a protein mass comparable to the mass of free monomers), one may be able to acquire kinetics that are dominated only by the elongation of these seed fibrils and are not affected by nucleation processes. Such experiments can be used to investigate the behaviour of the elongation process separately, without interference from the nucleation processes, for example by looking at the early time gradients of strongly seeded aggregation curves [4].

4 Parallel versus serial processes in the high and low concentration limits

Here we will present a detailed mathematical analysis of the example system in Fig. 4b of the main text. To recapitulate, we identified an increase in the magnitude of scaling with increasing monomer concentration as characteristic of a parallel pathway. By contrast a decrease in the magnitude of scaling with increasing monomer concentration is characteristic of a serial (i.e. saturating) pathway. For simplicity we will now refer to the reactants / intermediates as A and B respectively and the product as C. Now consider the kinetic equations describing the reactions in both cases.

4.1 Parallel

The parallel process is described by the equations

$$\frac{dA}{dt} = -2k_{\mathrm{P1}}A^2 \tag{S5}$$

$$\frac{dB}{dt} = -k_{\mathrm{P2}}B \tag{S6}$$

$$\frac{dC}{dt} = k_{\mathrm{P1}}A^2 + k_{\mathrm{P2}}B \tag{S7}$$

From equation S7 we see that the process are additive, the faster one will dominate: If A is small (where by small we mean $A^2 \ll k_{\mathrm{P2}}B/k_{\mathrm{P1}}$) the rate of formation of C is

proportional to B and if A is large the rate of formation of C is proportional to A^2. Hence the scaling increases as A increases.

The differential equations for A and B can be solved directly to yield:

$$A(t) = \frac{A_0}{1 + A_0 2k_{\mathrm{P1}} t} \tag{S8}$$

and

$$B(t) = B_0 e^{-k_{\mathrm{P2}} t} \tag{S9}$$

where A_0 and B_0 are the initial reactant concentrations. The time evolution of the product concentration is then simply given by

$$C(t) = B_0(1 - e^{-k_{\mathrm{P2}} t}) + A_0(1 - \frac{1}{1 + A_0 2k_{\mathrm{P1}} t}) \tag{S10}$$

assuming that no product is present at time $t = 0$. The half time is given by

$$C(t_{1/2}) = B_0(1 - e^{-k_{\mathrm{P2}} t_{1/2}}) + A_0(1 - \frac{1}{1 + A_0 2k_{\mathrm{P1}} t_{1/2}}) = \frac{A_0/2 + B_0}{2} \tag{S11}$$

In the limit of large A_0 this simplifies to $t_{1/2} = \frac{1}{6A_0 k_{\mathrm{P1}}}$, i.e. a scaling exponent of -1. In the limit of small A_0 this simplifies to $t_{1/2} = \frac{\log(2)}{k_{\mathrm{P2}}}$, i.e. a scaling exponent of 0. The plot in the main text was obtained by numerically solving equation S11 and agrees with these predicted approximate scaling exponents.

4.2 Serial with saturation (intermediate conversion through catalyst)

The case of the saturating serial process, i.e. one where conversion of the intermediate is through action of a catalyst, is similar to Michaelis-Menten kinetics, which is also similar to the mechanism of secondary nucleation for Aβ40 [4]. Therefore this is the process discussed in the main text as it is most relevant in the context of the kinetic models presented here. However, as shown below the same general rules about scaling exponents hold for a simpler case of serial reactions. The relevant kinetic equations are:

$$\frac{dA}{dt} = -k_{\mathrm{S1}} A\mathrm{Cat(free)} \tag{S12}$$

$$\frac{dB}{dt} = k_{\mathrm{S1}} A\mathrm{Cat(free)} - 2k_{\mathrm{S2}} B^2 \tag{S13}$$

$$\frac{dC}{dt} = k_{\mathrm{S2}} B^2 \tag{S14}$$

One could equally well consider an alternative description, with the dimerisation step before attachment, analogous to the two descriptions of secondary nucleation in section 5.3. The qualitative effect on the scaling exponent remains the same.

The fact that the catalyst is recovered upon formation of C from B gives us the condition

$$\mathrm{Cat(free)} + \mathrm{Cat(bound)} = \mathrm{Cat(free)} + B = \mathrm{Cat(total)} \tag{S15}$$

where we used the fact that the concentration of B is equivalent to the concentration of bound catalyst.

The differential equations are now coupled and cannot be solved as simply as the ones in the parallel network. Assuming steady state for B yields the equations below (see the part on the fast initial conversion limit in Sec. 4.3 for a discussion of the behaviour when steady state does not apply).

$$B = \frac{-A(t)k_{\mathrm{S1}} + \sqrt{A(t)^2 k_{\mathrm{S1}}^2 + 8A(t)k_{\mathrm{S1}}\mathrm{Cat(total)}k_{\mathrm{S2}}}}{4k_{\mathrm{S2}}} \tag{S16}$$

For large A ($A \gg \mathrm{Cat(total)}k_{\mathrm{S2}}/k_{\mathrm{S1}}$) this expression converges to $B = \mathrm{Cat(total)}$, i.e. all the catalyst is bound. This gives a constant rate of product formation hence

$$C(t) = \mathrm{Cat(total)}^2 k_{\mathrm{S2}} t \tag{S17}$$

and therefore $t_{1/2} = \frac{A_0}{4\mathrm{Cat(total)}^2 k_{\mathrm{S2}}}$, i.e. a scaling of +1. We cut the plots in the main text when the scaling reaches 0, to avoid confusion that may arise from a positive scaling, which can only rarely be found in filamentous aggregation, requiring a complete monomer independence of all processes.

For low concentrations of A ($A \ll \mathrm{Cat(total)}k_{\mathrm{S2}}/k_{\mathrm{S1}}$) the steady state approximation is poor, in particular for slow conversion of intermediate.

We can obtain an expression in this limit by making a number of approximations. First we neglect the concentration of bound catalyst, replacing Cat(free) with Cat(total) in equation S14. We then obtain the concentration of reactant simply as

$$A(t) = A_0 e^{-\mathrm{Cat(total)}k_{\mathrm{S1}}t} \tag{S18}$$

and the differential equation for the intermediate becomes

$$\frac{dB}{dt} = k_{\mathrm{S1}} A_0 e^{-\mathrm{Cat(total)}k_{\mathrm{S1}}t} - 2k_{\mathrm{S2}}B^2 \tag{S19}$$

an analytical solution to this equation exists in terms of Bessel functions, however, for the purpose of obtaining an approximate scaling exponent this is not very suitable and we attempt to construct an approximate solution to B instead. Keep in mind that the aim of this analysis is not to accurately determine the half times themselves, but only estimate the scaling exponent, which will not be sensitive to approximations that do

not affect the concentration dependence. Consider the following approximation: For slow conversion rates k_2 and low concentrations of reactant A_0, most of A will initially convert to B, which then decays as governed by the rate k_2. Hence we can approximate $B(t)$ at late times as

$$\frac{dB}{dt} \approx -2k_{S2}B^2 \tag{S20}$$

assuming $B(t=0) = A_0$, which gives

$$B_{\text{approx}}(t) = \frac{A_0}{1 + 2A_0k_2t} \tag{S21}$$

Because most of A is converted to B immediately, the half time for B, i.e. the point at which $B = A_0/2$, will also approximately correspond to the half time for the product C. The half time will occur at late times, so the approximation made for B will be valid. A comparison of this approximation with the numerically integrated solution to equation S14 is shown in Fig. S1.

$$B_{\text{approx}}(t_{1/2}) = \frac{A_0}{(1 + 2A_0k_2t)} = \frac{A_0}{2} \tag{S22}$$

which yields the half time as $t_{1/2} = \frac{1}{2A_0k_2}$, i.e. a scaling exponent of -1. The plot in the main text was obtained from numerical integration of equations S14 and the scaling exponents agree with the approximate ones obtained here.

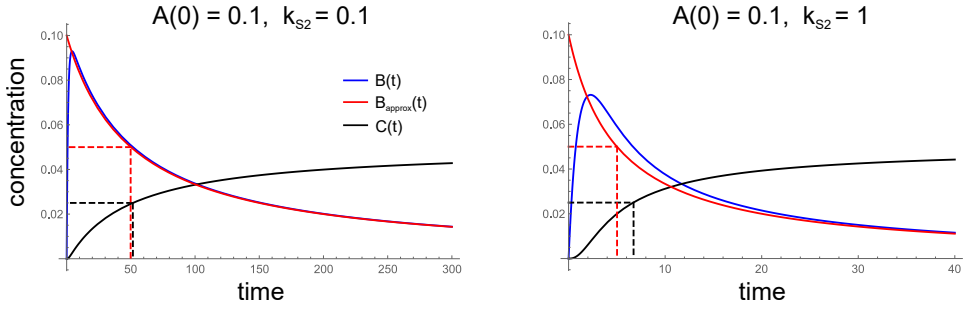

Figure S1: **Accurate numerical solution and approximation.** The accurate numerically integrated solution of equation S14 for $B(t)$ (blue) is compared to the approximate solution given in equation S21 (red). The time at which the approximation of $B(t)$ reaches half the reactant concentration (this is during the decay of B, not during its production) is marked with a dashed red line. The half time of $C(t)$, which determines the scaling exponent, is marked with a dashed black line. Without loss of generality, the initial attachment rate and the catalyst concentration are fixed at $k_{S1} = 1$ and Cat(total) = 1. We then look at the low reactant regime by setting $A_0 = 0.1$ and consider the case of a slow conversion, $k_{S2} = 0.1$ (left), and a faster conversion, $k_{S2} = 1$ (right).

4.3 Serial processes without saturation

Although this process is not discussed in the main text in any detail, we include it here for completeness, to illustrate that the decrease in scaling upon an increase in concentration is a general feature of serial processes not just the specific one chosen in the main text. We generalize the reaction orders, with the initial reactant A having reaction order n_1 and the intermediate B having reaction order n_2, and we will show that the higher reaction order always dominates the scaling at low concentrations, whereas the lower reaction order dominates at high concentrations.

$$\frac{dA}{dt} = -n_1 k_{S1} A^{n_1} \tag{S23}$$

$$\frac{dB}{dt} = k_{S1} A^{n_1} - n_2 k_{S2} B^{n_2} \tag{S24}$$

$$\frac{dC}{dt} = k_{S2} B^{n_2} \tag{S25}$$

At this stage, we will assume that the reaction orders are different, i.e. $n_1 \neq n_2$ (the case when they are equivalent will be discussed at the end), and hence we can set $k_{S1} = 1$ and $k_{S2} = 1$ without limitation of generality as these rate constants now define the units of concentration and time.

The differential equation for $A(t)$ can be solved to give for $n_1 \neq 1$:

$$A(t) = (A_0^{1-n_1} + n_1(n_1 - 1)t)^{\frac{1}{1-n_1}} \tag{S26}$$

where A_0 is the initial concentration of A.

To determine the dependence of the half time of product formation on the initial concentration A_0, first consider the steady state limit, i.e. $\frac{dB}{dt} = 0$, which will be valid when the rate of the second process is comparable to that of the first process. The differential equations are now:

$$\frac{dA}{dt} = -n_1 A^{n_1} \tag{S27}$$

$$\frac{dC}{dt} = A^{n_1} \tag{S28}$$

The half time for product formation is simply given by the half time for A, which can be obtained by solving $A(t_{1/2}) = A_0/2$ as $t_{1/2} = A_0^{1-n_1} \frac{2^{n_1-1}-1}{n1(n1-1)}$. Therefore in the steady state limit the scaling is given by $\gamma_{ss} = -(n_1 - 1)$. This limit is analogous to the limit of low concentration of A in the above case with catalyst ($A \ll \mathrm{Cat(total)} k_{S2}/k_{S1}$).

Now consider the other limit where the initial conversion of A to B is much faster than conversion of B to C. In this case the majority of A is immediately converted to

B, and we can instead consider the differential equations:

$$\frac{dB}{dt} = -n_2 B^{n_2} \tag{S29}$$

$$\frac{dC}{dt} = B^{n_2} \tag{S30}$$

with the initial condition $B(t=0) = A(t=0)/n$. These equations are the same as in the previous equation with renamed variables $A \to B$, $n_1 \to n_2$ and $A_0 \to A_0/n_1$, hence giving for the half time $t_{1/2} = A_0^{1-n_2} \frac{2^{n_2-1}-1}{n2(n2-1)n1^{1-n_2}}$. Therefore in the limit of fast initial conversion, the scaling is given by $\gamma_{\text{init}} = -(n_2 - 1)$. Note that for the saturating serial case above we did not consider this fast initial conversion limit. There the implicit assumption was that the total amount of catalyst is much smaller than the amount of reactant A present, therefore, for a fast initial conversion of A to the catalyst-bound intermediate state, the system would saturate fast at the beginning of the reaction and then proceed in steady state. If the amount of catalyst is larger than the total amount of A, the case of serial processes without saturation discussed here is recovered.

What remains is to establish which limit is applicable at low and high concentrations of A, for given values of n_1 and n_2. In order to do so we need to estimate the rates of production and loss of B, in order to determine which limit is valid. Initially the concentration of B will always be increasing, followed by a decrease in concentration of B as it is converted to C. If the concentration of B increases until most of A is used up, we are in the limit of fast initial conversion, if by contrast the concentration of B already decreases when only a small amount of A has been used up, we are in the steady state limit. We will now estimate the fraction, α, of A left when the rates of production and loss of B equal each other. So $\alpha = A(t_c)/A_0$ where t_c is given by $A(t_c)^{n_1} = n_2 B(t_c)^{n_2}$. Whilst the rate of production of B is easily determined, we need to estimate the rate of loss. We can give an upper bound on the rate of loss of B by using the amount of A converted at t_c, which is $(1-\alpha)A_0$ as an upper bound for the amount of B present. This gives as the loss rate $n_2 \left(\frac{(1-\alpha)A_0}{n_1}\right)^{n_2}$. As we use an upper bound for the rate of loss, we will overestimate α. Equating the two rates we get

$$(\alpha A_0)^{n_1} = n_2 \left(\frac{(1-\alpha)A_0}{n_1}\right)^{n_2} \tag{S31}$$

which can be rearranged to give

$$(A_0)^{n_1-n_2} = \frac{(1-\alpha)^{n_2}}{\alpha^{n_1}} \frac{n_2}{n_1^{n_2}} \tag{S32}$$

Given that the reaction orders n_1 and n_2 will be small integers larger or equal to 1, it can be assumed that the term $\frac{n_2}{n_1^{n_2}}$ is close to unity for the purposes of estimating α. There is four cases to consider, $n_1 < n_2$ and $n_1 > n_2$ at both high and low concentrations of A.

Fist consider $n_1 > n_2$ and start with the low concentration limit, $A \ll 1$. The left hand side in equation S32 is much smaller than 1, therefore we require a small numerator on the right hand side. Again noting that n_1 and n_2 will be small integers larger or equal to 1, we require that α is close to 1. In other words, the concentration of B already decreasing when most of A has not been converted yet, putting us in the steady state limit and giving a scaling of $\gamma = -(n_1 - 1)$. By contrast, in the limit of high concentrations, $A \gg 1$, we require a small denominator on the right hand side, giving a small α and putting us in the fast initial conversion limit with $\gamma = -(n_2 - 1)$.

If $n1 < n_2$ the left hand side is now small for large concentrations of a and large for small concentrations, therefore inverting the above results and giving a scaling of $\gamma = -(n_2 - 1)$ at low concentrations and $\gamma = -(n_1 - 1)$ at high concentrations. Finally, it can be shown that in the case of $n_1 = n_2$ there is no change in the scaling upon a change in the initial concentration of A.

In summary, the scaling changes from $\gamma = -(n_{\text{larger}} - 1)$ at low concentrations to $\gamma = -(n_{\text{smaller}} - 1)$ at high concentrations, where n_{larger} is the larger reaction order out of (n_1, n_2) and n_{smaller} the smaller one. Therefore the feature of positive curvature in half time plots is indeed general to serial processes and not a result of the choice of reaction orders or of the inclusion of a catalyst for the intermediate conversion.

4.4 Mapping of the elongation reaction to a simpler serial reaction

Here we discuss what modifications of a simple serial process are necessary to reproduce the behaviour observed for the overall conversion of monomer to fibrils. In the above simple serial case, the reaction was dominated by either of the steps in the two limits, i.e. the faster step became kinetically invisible. The combination of a nucleation and an elongation reaction as encountered in the linear aggregation reactions also constitutes a serial reaction, however, in that case the elongation step does contribute even when it is much faster than nucleation. The crucial difference is that the intermediate species in this case, the fibril number concentration P, is not being destroyed in the production of new fibril mass M. To map this to a simple chemical reaction, the free ends P can be thought of as catalysts converting protein from its monomeric state m to its fibrillar state M. The fact that P is not being lost means P can never reach steady state and is therefore always kinetically visible. Mathematically, this effect can be easily incorporated by simply removing the loss term in equation S24:

$$\frac{dA}{dt} = -n_1 k_{\text{S1}} A^{n_1} \tag{S33}$$

$$\frac{dB}{dt} = k_{\text{S1}} A^{n_1} \tag{S34}$$

$$\frac{dC}{dt} = k_{\text{S2}} B^{n_2} \tag{S35}$$

Although these equations now violate conservation of mass, and C diverges in the long time limit, they are sufficient to highlight the crucial differences compared to the simple serial example. We simply consider the half time to still be given by the time at which half of the initial concentration of A is present as C.

The limit of fast initial conversion is not affected, the overall scaling depends only on the second step, the conversion of the intermediate B to product C. C is formed with constant rate, giving $C(t) = k_{\mathrm{S2}} B_0^{n_2} t$ where $B_0 = A_0/n_1$ and A_0 is the initial concentration of A. Therefore the scaling is given by $\gamma_{\mathrm{init\_fast}} = -(n_2 - 1)$ as expected.

The steady state limit however no longer exists, we instead consider the reaction when A is not significantly depleted, i.e. $A(t) \approx A_0$. This simply gives $B(t) = k_{\mathrm{S1}} A_0^{n_1} t$ and therefore

$$C(t) = A_0^{n_1 n_2} \frac{k_{\mathrm{S1}}^{n_2} k_{\mathrm{S2}}}{n_2 + 1} t^{n_2+1} \tag{S36}$$

The scaling exponent is given as $\gamma_{\mathrm{init\_slow}} = -\frac{n_1 n_2 - 1}{n_2 + 1}$, i.e. it is affected by the reaction orders of both the first and the second process.

In this simplified model, if we set $n_2 = 1$ and make the rate of formation of C dependent on A, we recover the equations for nucleated linear polymerization without secondary processes [5]. Usually only the latter limit is relevant because nucleation, the first step, is much slower than elongation. The case where nucleation is faster than elongation is not usually encountered in the systems studied in the context of fibrillar assembly because it will not produce long fibrils but only species of a size similar to the nucleus. A scenario where elongation dominates as in the fast initial conversion case does occur when large amounts of preformed seeds are added at the start of the aggregation reaction (analogous to the late time limit).

5 Derivation of integrated rate laws

Although the non-linear differential equations describing the time evolution of aggregate number and aggregate mass are not readily integrable, it is possible to derive closed-form solutions using a self-consistent scheme, as we have used previously for related growth problems [1,2,4]. This approach involves reformulating the equations in terms of an integral operator, and then applying this integral operator (repeatedly) to an initial guess. This results in a closed-form expression for the fibril mass concentration.

5.1 Approximate form of half time and scaling

In the main text we stated 'half times are approximately given by $t_{1/2} \approx 1/\kappa$, where κ is of the form $\sqrt{\text{elongation}} \cdot \sqrt{(\text{sum } 2^{\text{o}} \text{ processes})}$'. To illustrate, consider the generalised

moment equations:

$$\frac{dP}{dt} = \alpha_P M(t) + \beta_P \tag{S37}$$

$$\frac{dM}{dt} = \alpha_M P(t) \tag{S38}$$

where α_P is the rate of nucleus formation from secondary processes, per fibril mass, β_P is the rate of nucleus formation from primary nuclei and α_M is the rate of elongation, per fibril. For example, in the case of a single-step secondary nucleation and fragmentation model, we would have $\alpha_P = k_2 m(t)^{n_2} + k_-$, $\beta_P = k_n m(t)^{n_c}$ and $\alpha_M = k_+ m(t)$.

If we consider the early time linearised solution to this system, where $m(t) \approx m_0$, then α_P, β_P and α_M all remain constant and the early time solution simply is:

$$M_{\text{lin}}(t) \propto e^{\sqrt{\alpha_P \alpha_M} t} \tag{S39}$$

where we neglected the contribution from primary nucleation. We identify $\sqrt{\alpha_P \alpha_M}$ with κ of the main text, which is indeed of the form $\sqrt{\text{elongation}} \cdot \sqrt{(\text{sum } 2^{\text{o}} \text{ processes})}$. It remains to show that $t_{1/2} \approx 1/\kappa$.

In order to get an approximate half time we solve equation (S39) for a specific aggregate mass, i.e. $M(t_{1/2}) = c$, to obtain:

$$t_{1/2} = \frac{1}{\sqrt{\alpha_P \alpha_M}} + \text{logarithmic terms} \tag{S40}$$

hence, neglecting the logarithmic terms, $t_{1/2} \approx 1/\kappa$.

5.2 Saturating elongation

We first explicitly include the concentration of bound ends, P^*, into the kinetics and the use a steady state approximation to obtain the Michaelmas Menten-like kinetics.

$$\frac{dP_f}{dt} = k_n m(t)^{nc} + k_- M(t) - k_+^* m(t) P_f(t) + \frac{1}{\tau_r} P^*(t) \tag{S41}$$

$$\frac{dP^*}{dt} = k_+^* m(t) P_f(t) - \frac{1}{\tau_r} P^*(t) \tag{S42}$$

$$\frac{dM}{dt} = \frac{1}{\tau_r} P^*(t) \tag{S43}$$

where k_+^* is the attachment rate of monomer to a free end, $1/\tau_r$ is the rate at which the monomer on a bound end rearranges to be incorporated into the fibril, recovering a free end, $P_f(t)$ is the concentration of free ends and $P^*(t)$ is the concentration of bound ends. We assume steady state, $dP^*/dt = 0$. This assumption is valid as long as the steady state

concentration of free ends is established quickly as soon as the fibril number changes. Using the fact that these processes do not change the fibril number $P_f(t)+P^*(t) = 2P(t)$ where $P(t)$ is the total fibril number concentration (the factor of 2 comes from the fact that the original definition of $P(t)$ is in terms of number of fibrils, whereas when we consider bound and free ends, there are 2 ends per fibril), equation (S42) becomes:

$$k_+^* m(t)(2P(t) - P^*(t)) = \frac{1}{\tau_r} P^*(t) \tag{S44}$$

Solving for $P^*(t)$ and inserting into (S43) yields

$$\frac{dM}{dt} = \frac{1}{\tau_r} \frac{2\tau_r k_+^* m(t) P(t)}{\tau_r k_+^* m(t) + 1} = 2k_+ \frac{m(t)}{1 + m(t)K_E^{-1}} P(t) \tag{S45}$$

where we have renamed $k_+ = k_+^*$ and $K_E = 1/(k_+^* \tau_r)$.

The full equations describing the fibril number and mass concentrations for a system with fragmentation and saturating elongation are hence given by:

$$\frac{dP}{dt} = k_- M(t) + k_n m(t)^{n_c} \tag{S46}$$

$$\frac{dM}{dt} = 2k_+ \frac{m(t)}{1 + m(t)K_E^{-1}} P(t) \tag{S47}$$

where we assumed the increase in mass from nucleation processes and the loss in aggregates due to fragmentation into pieces smaller than the critical nucleus size is negligible. In the limit $m \ll K_E$ the monomer concentration is so low that no saturation effects are evident because the fraction of bound ends, P^*, is negligible. Equation (S47) becomes $dM/dt = 2k_+ m(t) P(t)$, recovering the single-step kinetics. By contrast, in the highly saturated limit, $m \gg K_E$, equation (S47) becomes $dM/dt = 2k_+ K_E P(t)$, corresponding to a growth process whose rate is independent of the monomer concentration. Whilst the single-step description can be valid for the full time-course of a reaction, the monomer concentration-independent expression inevitably fails at late times when monomer is depleted below the saturation concentration.

The above equation can then be integrated to yield:

$$M(t) = m(0) \frac{W\left(K_E^{-1} m(0) \exp\left(m(0)K_E^{-1} - 2k_+ \int_0^t P(t)dt\right)\right)}{K_E^{-1}} \tag{S48}$$

where $W(x)$ is the Lambert W (product log) function defined by $W(x\exp(x)) = x$.

Now we obtain the linearised solutions, $P_{\text{lin}}(t)$ and $M_{\text{lin}}(t)$, by assuming a constant concentration of monomer and solving Eqs. (S46) and (S47) with $m(t) \to m(0)$. Inserting this linearised solution for the aggregate number concentration, $P_{\text{lin}}(t)$, into Eq. (S48) yields the first-order self-consistent solution for $M(t)$:

$$\frac{M}{m_{\rm tot}} = 1 - \frac{K_E}{m_{\rm tot}} \mathrm{W}\left(\frac{m_{\rm tot} - M_0}{K_E} \exp\left[\frac{(m_{\rm tot} - M0)}{K_E} - \frac{2k_+(P_0\kappa\sinh(\kappa t) + \alpha(\cosh(\kappa t) - 1))}{\kappa^2}\right]\right) \quad \text{(S49)}$$

where the definitions of the parameters are

$$\kappa = \sqrt{2\frac{m_0 k_+}{1 + m_0/K_E} k_-} \quad \text{(S50)}$$

$$\alpha = k_n m_0^n c + k_- M0 \quad \text{(S51)}$$

and P_0 and M_0 denote the aggregate number and mass concentrations at the beginning of the reaction respectively. In the case of unseeded data $P_0 = M_0 = 0$ and these equations will simplify significantly.

The half time of the reaction is determined by the argument in the hyperbolic function:

$$\tau_{\rm lag} \sim \kappa^{-1} \quad \text{(S52)}$$

The scaling exponent is therefore given by

$$\gamma \approx -\frac{1}{2(1 + m(0)/K_E)} \quad \text{(S53)}$$

which interpolates between the limits $\gamma = -1/2$ when $m(0)K_E^{-1} \ll 1$ and $\gamma = 0$ in the opposite limit, i.e. the scaling exponent increases with increasing monomer and there is positive curvature in the double logarithmic plots.

As expected, Eq. (S49) recovers the result known previously for cases where the elongation rate does not saturate, in the limit $K_E^{-1} \to 0$.

In the opposing limit, $K_E^{-1} \to \infty$, Eq. (S49) results in

$$M(t) = m(0) - W(K_E^{-1} m(0) \exp(K_E^{-1} m(0)))/K_E^{-1} = 0 \quad \text{(S54)}$$

Physically this corresponds to the equilibrium constant of monomer binding to the fibril ends being extremely in favour of unbound monomer, hence no elongation reaction takes place and the fibril mass is zero.

5.3 Saturating secondary nucleation

The derivation of the Michaelis Menten-type kinetics follows the same principle as for elongation above and the full details can be found in the supplementary of Meisl *et al.* [4]. The rates in the Petri nets in Fig. 2 and Fig. 5 of the main text are linked to the rate equations as follows: $k_2 = k_2$ and $K_M = k_d/k_2$.

The differential equations for this system are given by:

$$\frac{dP}{dt} = k_n m(t)^{n_c} + k_2 \frac{m(t)^{n_2}}{1 + m(t)^{n_2}/K_M} M(t) \tag{S55}$$

$$\frac{dM}{dt} = 2m(t)k_+ P(t) \tag{S56}$$

In the limit $m(t)^{n_2} \ll K_M$ the rate limiting step is the binding to the fibril surface and equation (S55) becomes, $dP/dt = k_2 m(t)^{n_2} M(t) + k_n m(t)^{n_c}$, recovering the single-step description [2]. In the opposite limit of $m(t)^{n_2} \gg K_M$, all binding sites on the fibril surface are fully covered and the rate-determining step is the rearrangement and detachment of new nuclei, leading to monomer-independent kinetics. In this limit equation (S55) becomes formally equivalent to that for a fragmentation dominated system (see equation (S46)) $dP/dt = k_2 K_M M(t) + k_n m(t)^{n_c}$, where replacing $k_2 K_M$ by k_- yields the equation for fragmentation. This equivalence has significant practical consequences, showing that it is not possible to distinguish between a mechanism dominated by fragmentation and one dominated by a fully saturated secondary nucleation mechanism from measurement of $P(t)$ and $M(t)$ alone.

The detailed derivation of an approximate analytical solution can be found in Meisl *et al.* [4] and yields:

$$\begin{aligned} \frac{M}{M_\infty} &= 1 - \left(1 - \frac{M_0}{M_\infty}\right) e^{-k_\infty t} \\ &\quad \cdot \left(\frac{B_- + C_+ e^{\kappa t}}{B_+ + C_+ e^{\kappa t}} \cdot \frac{B_+ + C_+}{B_- + C_+}\right)^{\frac{k_\infty}{\kappa \bar{k}_\infty}} \end{aligned} \tag{S57}$$

where the definitions of the parameters are

$$\kappa = \sqrt{2m_0 k_+ \frac{m_0^{n_2} k_2}{1 + m_0^{n_2}/K_M}} \tag{S58}$$

$$\lambda = \sqrt{2k_+ k_n m_0^{n_c}} \tag{S59}$$

$$C_\pm = \frac{k_+ P_0}{\kappa} \pm \frac{k_+ M_0}{2m_0 k_+} \pm \frac{\lambda^2}{2\kappa^2} \tag{S60}$$

$$k_\infty = \sqrt{(2k_+ P(0))^2 - 2A - 4k_+ k_2 m_{\text{tot}} K_M \frac{\log\left[K_M\right]}{n_2}} \tag{S61}$$

$$\begin{aligned} A &= -\frac{2k_+ k_n m_0^{n_c}}{n_c} - 2k_+ k_2 m_{\text{tot}} K_M \frac{\log\left[K_M + m_0^{n_2}\right]}{n_2} \\ &\quad - 2k_+ k_2 K_M m_0 \left({}_2F_1\left[\frac{1}{n_2}, 1, 1 + \frac{1}{n_2}, -\frac{m_0^{n_2}}{K_M}\right] - 1\right) \end{aligned} \tag{S62}$$

$$\bar{k}_\infty = \sqrt{k_\infty^2 - 2C_+ C_- \kappa^2} \tag{S63}$$

$$B_\pm = \frac{k_\infty \pm \bar{k}_\infty}{2\kappa} \tag{S64}$$

where M_∞ is the aggregate mass concentration at long times and ${}_2F_1(a,b,c,d)$ is the ordinary hyper-geometric function.

The approximate scaling exponent is:

$$\gamma \approx -\frac{1}{2}\left(\frac{n_2}{1+m(0)^{n_2}/K_M}+1\right) \tag{S65}$$

In the limit of low and high monomer concentration this becomes $\gamma = -(n_2+1)/2$ and $\gamma = -1/2$ respectively, i.e. the scaling exponent increases with increasing monomer and there is positive curvature in the double logarithmic plots.

5.3.1 Alternate model for saturating secondary nucleation

For completeness we add a note here on the use of two slightly different, but so far equally valid, models for secondary nucleation. In all of the main text the model discussed above was used. A physical interpretation of the above equations would be as follows: Monomers in solution are in equilibrium with pre-nucleus species, dimers, trimers, etc. These pre-nuclei need to undergo structural rearrangement to become growth competent nuclei and this process can be catalysed on the surface of existing fibrils. In short, the above equations describe monomers meeting before they attach to the surface and rearrange. A physically equally valid model is that monomeric species attach to the fibril surface and formation of the pre-nucleus cluster occurs on the surface. Mathematically this corresponds to replacing the term $\frac{m_0^{n_2}}{1+m_0^{n_2}/K_M}$ by $\left(\frac{m_0}{1+m_0/K_M}\right)^{n_2}$ and the solutions obtained in the two cases behave very similarly (see Saric *et al.* [6] for the detailed solution). Indeed we found the two models to yield equally good results in the fitting of experimental data so far, although the values of the fitted parameters will differ, preventing a determination of the correct model based on the aggregation data alone. In a recent *in silico* study [6] we observed direct attachment of monomers to the surface of fibrils and hence used the latter model to analyse this data. In all previous work we have opted to use the model assuming an attachment of pre-nuclei and therefore also use this description here.

5.4 Competing secondary processes

The relevant differential equations for the rate of new aggregate formation are:

$$\frac{dP}{dt} = k_n m(t)^{n_c} + k_- M(t) + k_2 m(t)^{n_2} M(t) \tag{S66}$$

$$\frac{dM}{dt} = 2(m(t)k_+ - k_{\text{off}})P(t) \tag{S67}$$

Again the linearised solution is obtained by setting $m(t) = m_0$ (see also see Meisl *et al.* [3]). This is then used in a fixed point iteration to give the first order fixed point solution:

$$
\begin{aligned}
M(t) &= M_\infty + \mathrm{Exp}\left[-\frac{k_+(4c\kappa\mathrm{Cosh}(\kappa t) + 4P_0\kappa^2\mathrm{Sinh}(\kappa t))}{2\kappa^3}\right] \\
&\quad \left((M_0 - M_\infty)e^{\frac{2k_+c}{\kappa^2}}\right) \qquad \text{(S68)}
\end{aligned}
$$

where

$$
\begin{aligned}
a &= k_2 m_0^{n_2} + k_- \\
c &= k_n m_0^{n_c} + aM_0 \\
\kappa &= \sqrt{2(k_+ m_0 - k_{\mathrm{off}})(k_2 m_0^{n_2} + k_-)} \\
M_\infty &= m_{\mathrm{tot}} - k_{\mathrm{off}}/k_+ \qquad \text{(S69)}
\end{aligned}
$$

The approximate scaling exponent is:

$$
\gamma = \frac{d\log(t_{1/2})}{d\log(m(0))} \approx -\frac{1}{2}\left(\frac{n_2}{1 + K/m(0)^{n_2}} + 1\right) \qquad \text{(S70)}
$$

where $K = k_-/k_2$. In the limit of low and high monomer concentration this becomes $\gamma = -1/2$ and $\gamma = -(n_2 + 1)/2$ respectively, i.e. the scaling exponent decreases with increasing monomer and there is negative curvature in the double logarithmic plots.

References

[1] Tuomas P J Knowles, Christopher A Waudby, Glyn L Devlin, Samuel I A Cohen, Adriano Aguzzi, Michele Vendruscolo, Eugene M Terentjev, Mark E Welland, and Christopher M Dobson. An analytical solution to the kinetics of breakable filament assembly. *Science*, 326(5959):1533–1537, 2009.

[2] Samuel I A Cohen, Michele Vendruscolo, Mark E Welland, Christopher M Dobson, Eugene M Terentjev, and Tuomas P J Knowles. Nucleated polymerization with secondary pathways. i. time evolution of the principal moments. *J Chem Phys*, 135(6):065105, 2011.

[3] Georg Meisl, Julius B. Kirkegaard, Paolo Arosio, Thomas T. C. Michaels, Michele Vendruscolo, Christopher M. Dobson, Sara Linse, and Tuomas P. J. Knowles. Molecular mechanisms of protein aggregation from global fitting of kinetic models. *Nature Protocols*, 11(2):252–272, 2016.

[4] Georg Meisl, Xiaoting Yang, Erik Hellstrand, Birgitta Frohm, Julius B. Kirkegaard, Samuel I. A. Cohen, Christopher M. Dobson, Sara Linse, and Tuomas P. J. Knowles. Differences in nucleation behavior underlie the contrasting aggregation kinetics of the aβ40 and aβ42 peptides. *Proceedings of the National Academy of Sciences*, 111:9384–9389, 2014.

[5] Fumio Oosawa and Sho Asakura. *Thermodynamics of the Polymerization of Protein*. Academic Press, 1975.

[6] Andela Saric, Alexander Buell, Georg Meisl, Thomas C. T. Michaels, Christopher M. Dobson, Sara Linse, Tuomas P. J. Knowles, and Daan Frenkel. Physical determinants of the self-replication of protein fibrils. *Nature Physics*, 12:874–880, 2016.