Supplementary Information

Caught in the Act: Mechanistic insight into Supramolecular Polymerization driven Self-Replication from Real Time Visualization.

Sourav Maity,[1] Jim Ottelé,[2] Guillermo Monreal Santiago,[2] Pim W. J. M. Frederix,[3] Peter Kroon,[3] Omer Markovitch,[2,4] Marc C. A. Stuart,[2] Siewert J. Marrink,[3]* Sijbren Otto[2]* and Wouter H. Roos[1]*

Affiliations:
[1] Molecular Biophysics, Zernike Institute for Advanced Materials, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands
[2] Centre for Systems Chemistry, Stratingh Institute, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands
[3] Groningen Biomolecular Sciences and Biotechnology Institute & Zernike Institute for Advanced Materials, University of Groningen, Nijenborgh 7, 9747 AG Groningen, The Netherlands
[4] Origins Center, Nijenborgh 7, 9747 AG, Groningen, The Netherlands

*Correspondence to: s.j.marrink@rug.nl, s.otto@rug.nl, w.h.roos@rug.nl

This supplementary information contains:

Supplementary Video description:

Supplementary Video 1. Growth of a fibre captured by HS-AFM. The video shows the growth of a fibre weakly immobilized on a membrane bilayer. Imaging rate is 0.5 frame/second. For detailed description see **Figure S2 A-F** and the main text.

Supplementary Video 2. No observable growth of fibre with precursors attached in the middle as captured by HS-AFM. The video captured a fibre immobilized on a membrane bilayer with precursors attached in the middle. Imaging rate is 2 frame/second. For detailed description see **Figure S2 G-I** and the main text

Supplementary Video 3. Growth of a fibre, accumulation and diffusion of precursor aggregate captured by HS-AFM. The video shows the pathway of fibre growth capture for a long period of time (1930 s). Imaging rate 0.5 frame/second. For detailed description see **Figure 3** and the main text.

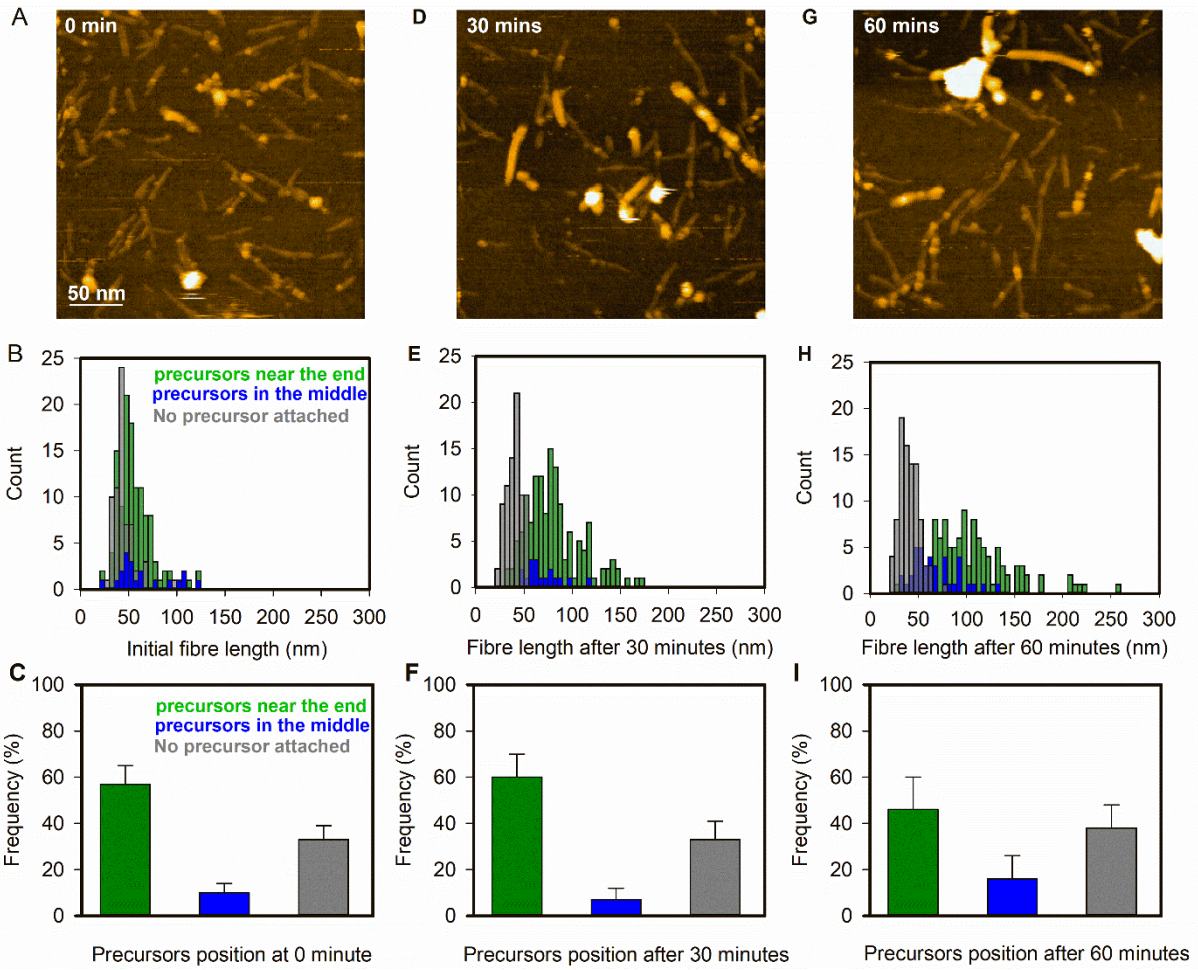
Supplementary Video 4. Precursors diffusivity on a fibre surface captured by HS-AFM. The video shows growth of a fibre without mask (left), and with mask (right) used for the height of a fibre (3.5 nm). The black dots represent all protrusions above 3.5 nm. Imaging rate 0.5 frame/second. For more detailed description see **Figure S3**.

Supplementary Video 5. Growth of a fibre captured by HS-AFM at an extended time. The video shows the growth of a fibre weakly immobilized on a membrane bilayer at a longer time frame (after 32 minutes). Imaging rate is 0.5 frame/second. For detailed description see **Figure S8** and the main text.

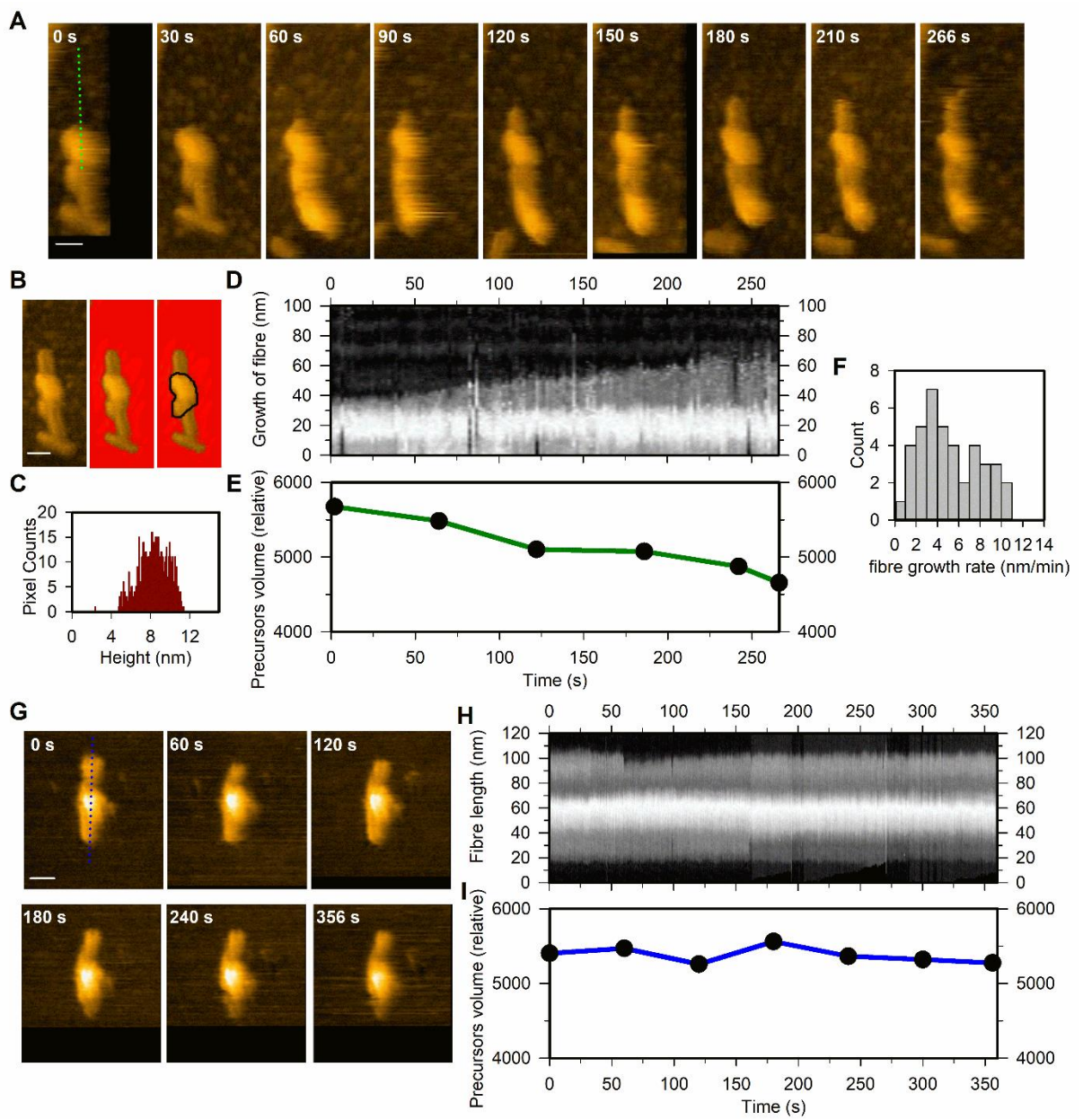
Supplementary Video 6. GROMOS all-atom simulation results of the diffusion of a precursor on a fibre. The video shows the normalized density profile of a **1**$_3$ precursor averaged over 200 independent 60 ns simulations. The fibre outline is represented by the black line.

Supplementary Video 7. Martini coarse grain simulation results of the diffusion of a precursor on a fibre. The video shows the normalized density profile of a **1**$_3$ precursor averaged over 400 independent 500 ns simulations. The fibre outline is represented by the black line.

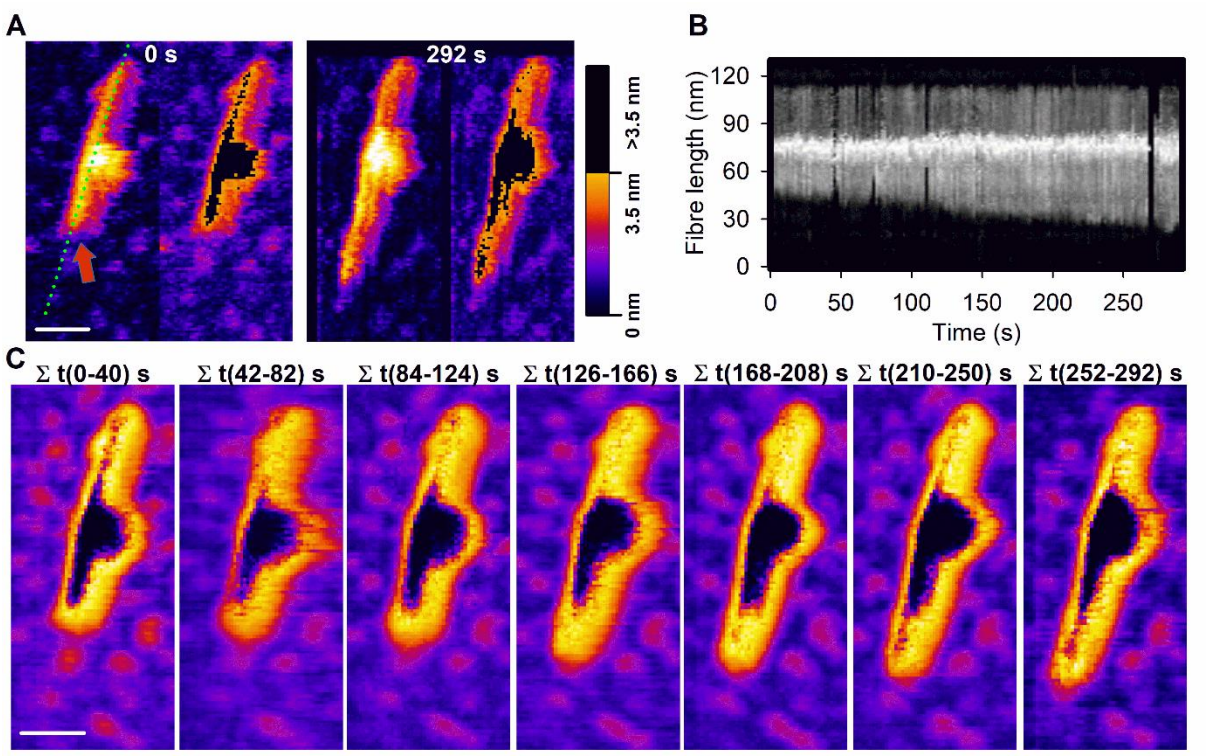
Supplementary Figures:



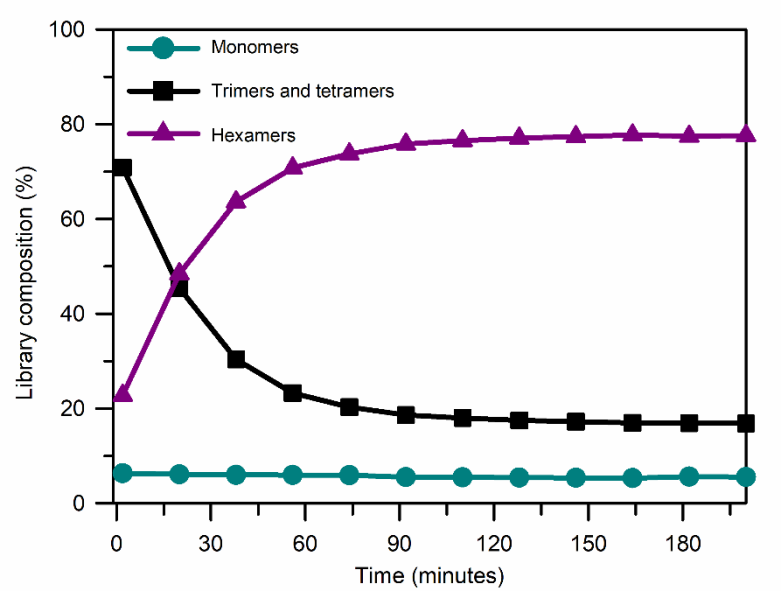
Supplementary Figure 1 | Fibres having precursors attached near the fibre end have the highest frequency of occurrence and grow the most. (A) Example of an AFM image of fibres immobilized on a membrane surface at 0 min, in presence of 2.31 mM precursors. **(B)** Histogram of fibre length at 0 min. N= 210 (fibres) from 10 different images. **(C)** Relative frequencies of fibres having precursors attached near the end (green), in the middle (blue) or no precursor attachment (grey). **(D)** Same as A, but after 30 min of incubation on the membrane surface with 2.31 mM precursors in solution. **(E, F)** Same as B and C, respectively, but after 30 min. N= 234 (fibres) from 10 different images. **(G, H, I)** Same as A, B and C, respectively, but after 60 min of incubation on the membrane surface with precursors in solution. N= 238 (fibres) from 10 different images. The error bar in panels C, F and I are representing the standard deviation of the distribution. Statistical analysis using the two-sample Kolmogorov-Smirnov test confirmed that at t=60 min the lengths of fibres with precursors attached to their ends are significantly different from those with precursors attached to their middle (highest p-value=4.39×10^{-7}, **Supplementary Table 2**). Evaluating the elongation of fibres between t=0, 30 and 60 min confirmed that only the group of fibres with precursors attached at their ends grew significantly (highest p-value=2.41×10^{-7}, **Supplementary Table 3**), while the group of fibres with precursors attached at the middle and that without precursors did not significantly grow in length (lowest p-value= 5.22×10^{-2} and 7.07×10^{-2}, respectively).



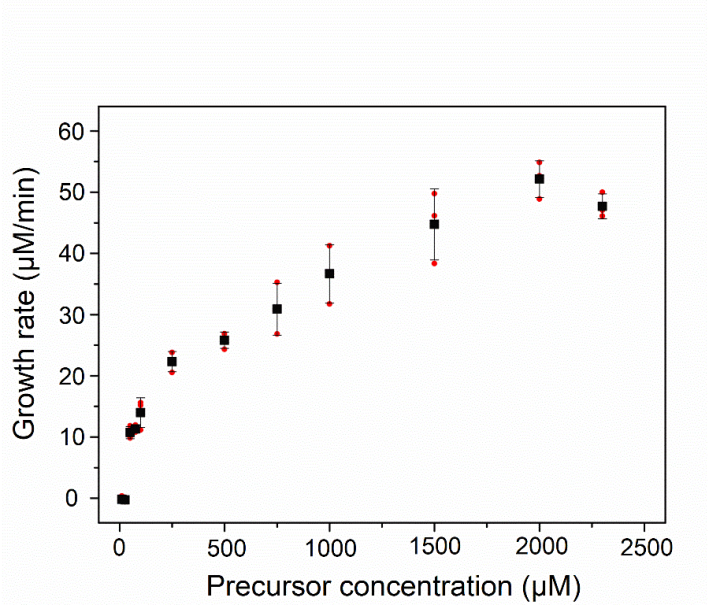
Supplementary Figure 2 | The attached precursors participate actively in the growth process. (A) Snapshots of fibre growth on a membrane surface at different times. Scale bar 20 nm. **(B)** Image processing for the calculation of relative volume. The first frame (left) represents an example of a raw image. In the second frame (middle) the background is masked. In the third frame (right) the black line indicates the area occupied by the precursors. The selected area was fixed throughout image acquisition. **(C)** Example of a height distribution histogram (bin size of 0.08 nm) of the selected area in panel B (right). **(D)** Kymograph of relative height (in terms of intensity) along the green dotted line in panel A for the duration of the imaging (266 s). **(E)** Change in precursor volume with time from the experiment in panel A. **(F)** Histogram of fibre growth rate for all the growing fibres, calculated every 60 s (N=40). **(G)** Snapshot of AFM images of a fibre with precursors attached in the middle. Scale bar 20 nm. **(H)** Kymograph of relative height along the blue dashed line in panel **G**. **(I)** Measured relative volume of the attached precursors over time from panel **G**.



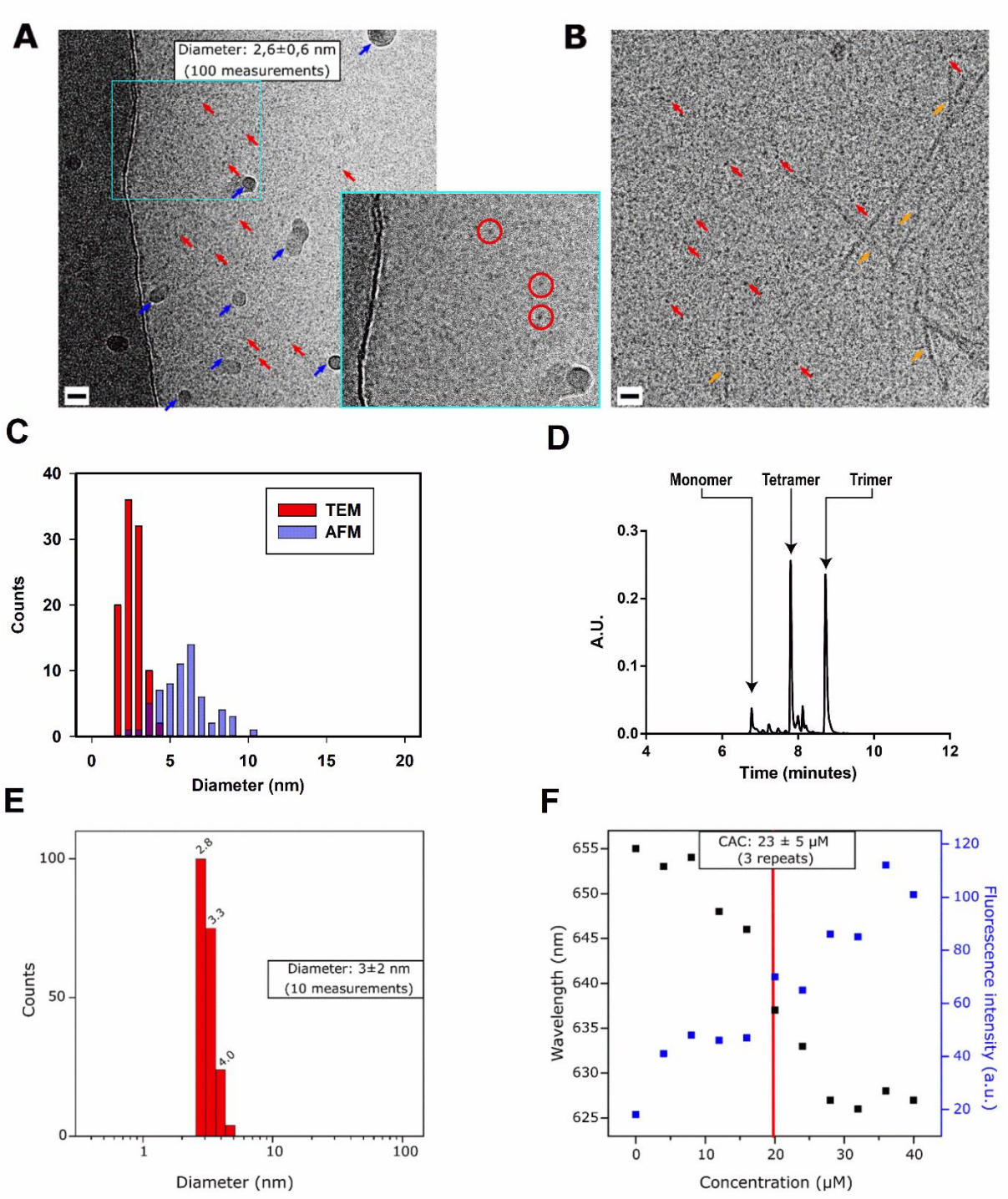
Supplementary Figure 3 | Precursor diffusion on a fibre surface favouring directionality in a fibre growth. (A) Snapshots of AFM images of a growing fibre on a membrane surface taken at initial time, and after 292 seconds. The left image shows the unmasked fibre, and right image after masking for the fibre height of 3.5 nm. The black dots on the fibre surface represent any protrusion above the fibre surface (>3.5 nm). **(B)** kymograph along the dotted green line in panel A for the period of 292 s. **(C)** Superimposed masked images of the fibre for every 40 s intervals, showing overall diffusion of precursors in every 40 seconds. The red arrow indicates the direction of growth. It can be observed that there is a higher number of black dots (i.e. precursors) at the growing end. Scale bars 20 nm.



Supplementary Figure 4 | UPLC analysis of library composition over time with 0.154 mM precursors in presence of 5 % seed, and without mechanical agitation. The data (monomers as circles, trimers and tetramers as squares, and hexamers as triangles) shown is the average of four separate experiments. The calculated growth rate for hexamers from the initial linear part of purple curve is ~1 nm/min.

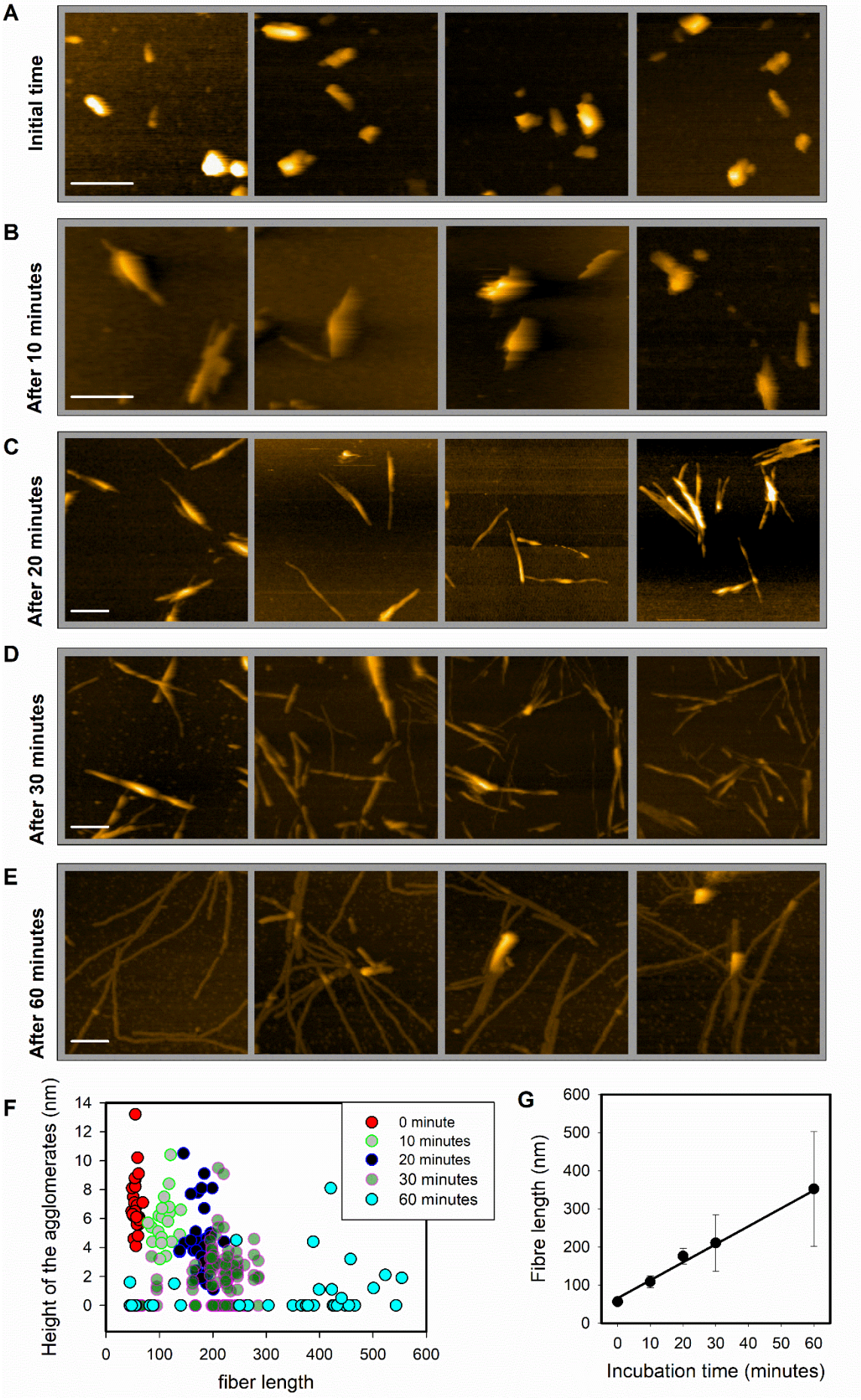


Supplementary Figure 5 | Initial growth rate of fibres in presence of different concentrations of precursors, determined by fluorescence spectroscopy (See Methods). The black squares represent the average of three experiments and the error bars their standard deviation, with the red circles representing each of the individual experiments. The fact that the fibre growth rate levels off at higher precursor concentrations suggests that precursor aggregates form which are off-pathway: i.e. these aggregates do not contribute to fibre growth.

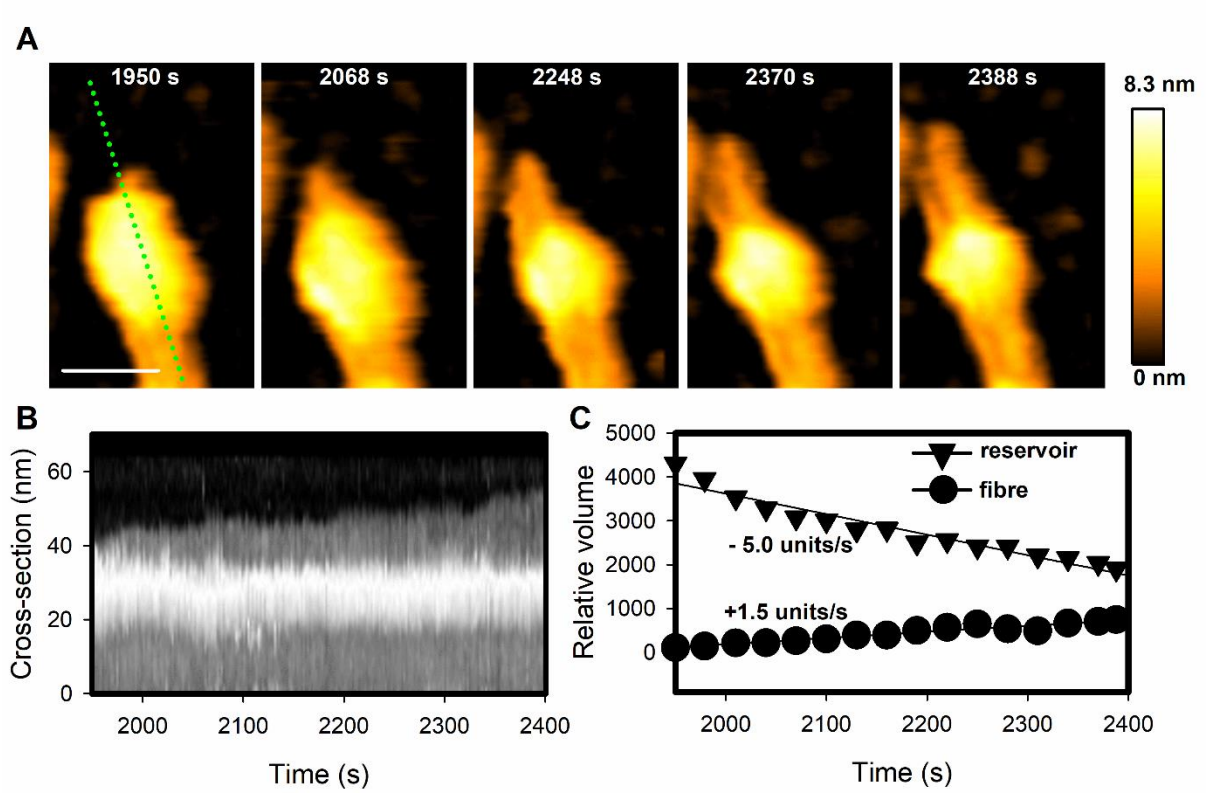


Supplementary Figure 6 | Characterization of the aggregates formed by precursors in solution. Cryo-TEM micrographs of samples containing **(A)** precursors and **(B)** 6mer fibres and precursors. To aid interpretation, arrows have been added indicating the location of aggregates (red), fibres (orange), and ice contamination (blue).[1–3] Scale bars are 20 nm. The insert in A highlights (red circles) typical examples of precursors aggregates. **(C)** A comparative histogram of measured aggregate diameter by TEM, and aggregate height by AFM. The diameter of 100 different aggregates was measured by TEM in different micrographs, giving an average value of 2.6 ± 0.6 nm. By AFM, the aggregates were only visible when they were attached on a fibre surface. The measured height of 63 different aggregates is 6 ± 2 nm, comparatively higher than that measured by TEM. **(D)** Chromatogram

representation of the sample from A. The results depict the major composition of molecules present in the precursors solution. **(E)** Size distribution of a precursor sample, as measured by DLS. The accuracy of DLS is relatively low for the size range measured, so 10 individual measurements were recorded (from 10 repeats each) and averaged, obtaining a diameter of 3 ± 2 nm. **(F)** Determination of the critical aggregation concentration of precursors using Nile Red as a fluorescent probe. The blue shift and increase in intensity of the fluorescence band of Nile Red above a certain concentration corresponds to the formation of a supramolecular structure with a hydrophobic environment.[4,5]


A
Initial time
B
After 10 minutes
C
After 20 minutes
D
After 30 minutes
E
After 60 minutes
F
Height of the agglomerates (nm)
fiber length
0 minute
10 minutes
20 minutes
30 minutes
60 minutes
G
Fibre length (nm)
Incubation time (minutes)

Supplementary Figure 7 | Growth of fibres free in solution in presence of 2.31 mM precursors as periodically observed by AFM. (A-E) Example of AFM images of fibres immobilized on mica after 0, 10, 35, and 60 min of incubation, respectively. The relative position of the precursors on the fibres changes with time. In the 0-10 min period, the aggregates do not show significant differences in position; while after 20 min, a spread in the precursor density can be observed. After 30 min, an accumulation of the precursors towards the fibre end can be seen. After 60 min, the precursors had disappeared as the fibres grew to the full extent. Scale bars are 100 nm. (**F**) Plots of aggregates height vs fibre length as observed at different times. The aggregates height decreases as the fibre elongates and the precursors spread over the fibre. (**G**) Fibre growth upon incubation with 2.31 mM (final 2.19 mM) precursor solution. Growth rate is ~5 nm/min. This number is likely an underestimate as the longer fibres tend to break in smaller pieces upon attachment to the surface.



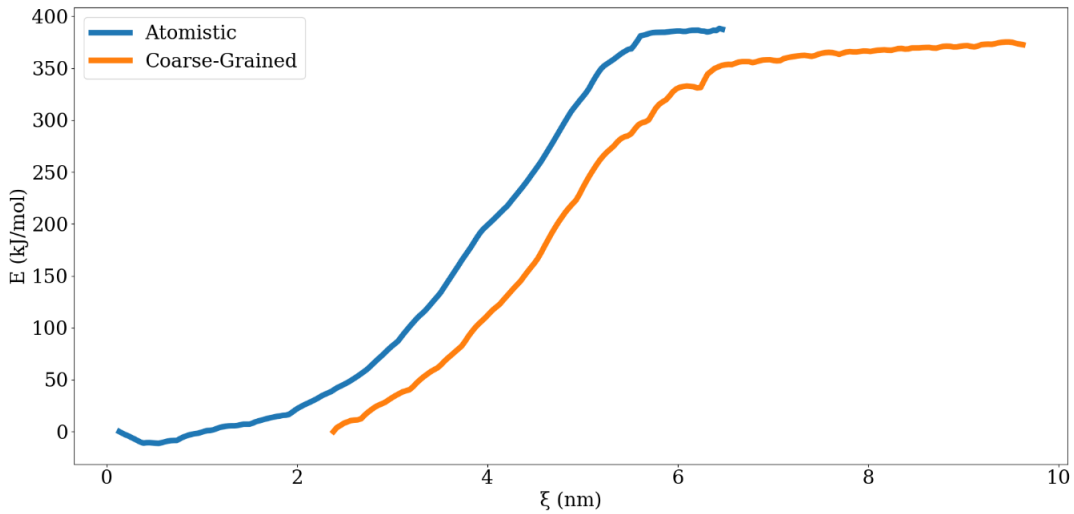
Supplementary Figure 8. Correlation of diffusivity with the fibre growth. (A) Snapshots of HS-AFM images of a growing fibre on a membrane surface after 32 min incubation period. Scale bar 20 nm. **(B)** kymograph of the growth along the dotted green line in panel **A**. An average growth rate of 3.6 nm/min is estimated from this kymograph. **(C)** Comparative measurements of relative volume over time for the precursors aggregate (reservoir), and the fibre. The slope estimated for the reservoir indicates a decrease in volume at least at a rate of 5 relative volume units/s. It is important to mention that this rate may be an underestimation, as free precursors in the solution can still be absorbed into the existing aggregates during this time period. The slope estimated for the fibre shows an increase in volume at a rate of 1.5 relative volume units/s.

Supplementary table 1. Summary of experimentally measured elongation rate of the fibre using different techniques and sampling.

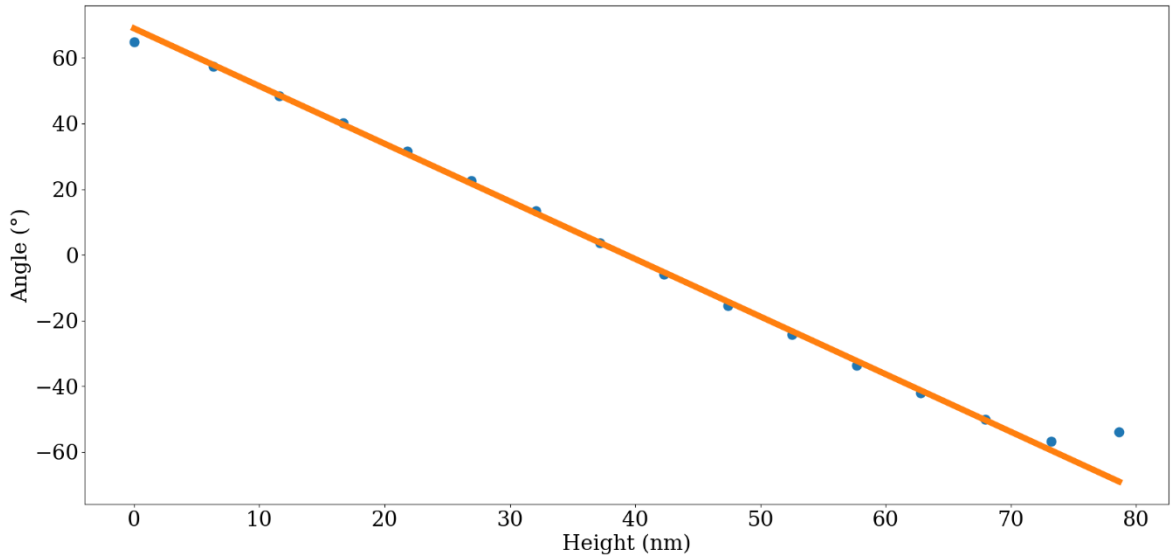
AFM measurements.		UPLC measurements.		Measured by fluorescence spectroscopy.	
Fibres grown on a surface (2.31 mM precursors)	Fibres grown free in solution (2.31 mM precursors)	In presence of 2.31 mM precursors	In presence of 0.154 mM precursors	In presence of 2.31 mM precursors	In presence of 0.100 mM precursors
~5 nm/min	~5 nm/min	~4 nm/min	~1 nm/min	~9 nm/min	~1 nm/min

Supplementary Discussion 1.

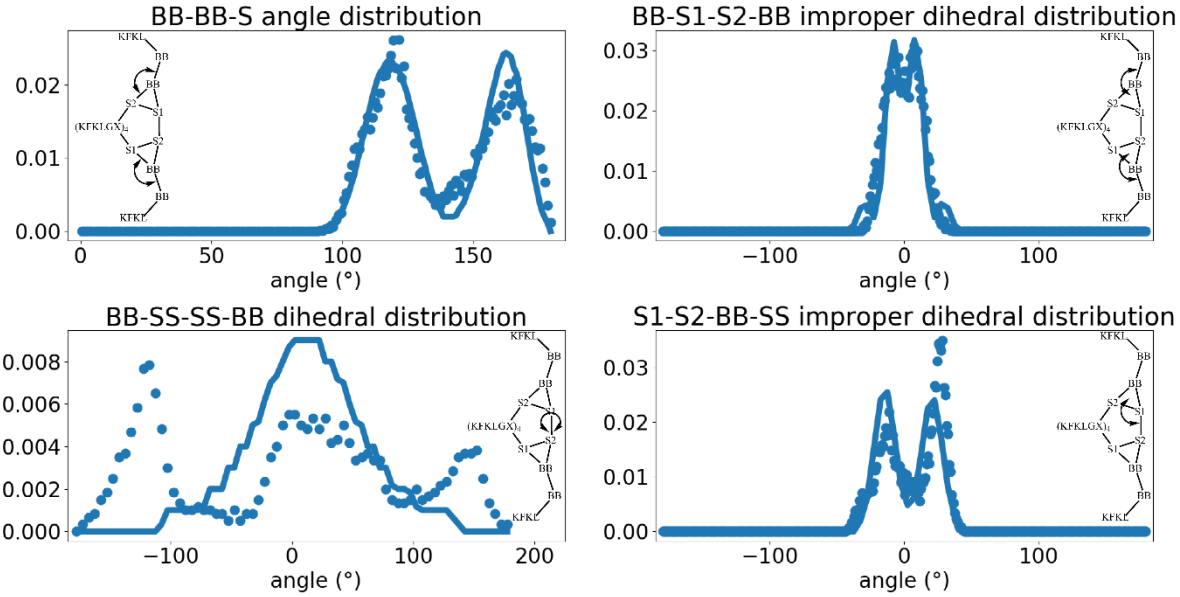
The growth rate values measured by fluorescence and UPLC should only be taken as an approximation because of the assumptions made in the conversion from µM/minute to nm/minute/fibre end (i.e. the calculated number of fibre ends per mM or the width of each hexamer). The experiments shown in **Supplementary Figure 5**, however, were performed simultaneously and under exactly the same conditions, and the above mentioned conversion factor is the same for all results.



Supplementary Figure 9 | Free energy profiles of binding comparing the atomistic and coarse-grained models. Both profiles were obtained using umbrella sampling by pulling one hexamer from a fibre of 8 hexamers. The reaction coordinate for the atomistic simulation was defined as the center of mass (com) distance between neighbouring C-terminal lysines of hexamers 1 and 2, simulating for 10 ns per window. The reaction coordinate for the coarse-grained simulation was defined as the com distance between one of the C-terminal lysines of hexamer 1 and the com of full hexamer 2 (hence the horizontal shift of the curve), simulating for 6 ns per window.



Supplementary Figure 10 | Angle of the hexamer as function of position in the fibre in a 1 µs coarse-grained simulation. To obtain the period of the twist of the fibre a line was fit (orange line) to the found angles (blue dots), excluding the first and last hexamers. This resulted in a slope of -1.75 ± 0.01 °/nm (R^2=0.9991). The fit was performed using SymFit.[6] This amounts to 205 ± 3 nm for a full twist, somewhat overestimating the experimental and atomistic findings.[7]



Supplementary Figure 11 | Distributions of angles and (improper) dihedrals in the dimercaptobenzene headgroup. Evidence of CG model accuracy vs atomistic data. Distributions determined based on simulations of a single hexamer in solution. Atomistic reference data is plotted as dots, the corresponding coarse-grained data is plotted as solid lines. Bead notation as in insets, S/SS denote either S1 or S2. The observable mismatch between atomistic and CG model in the BB-SS-SS-BB dihedral distribution at angles below -100° and above 100° is considered acceptable as the main mode is still captured, especially since the macroscopic CG fibre behaviour matches the AA fibre behaviour.

Supplementary Note

Supplementary Table 2: p-values of statistical tests comparing the observed fibre lengths at specific times. The two-sample Kolmogorov-Smirnov test was used, with null hypothesis that two datasets are from the same continuous distribution and the alternative hypothesis is that they are from different continuous distributions.

	0 min.	30 min.	60 min.
End vs. Middle	4.54×10^{-1}	3.22×10^{-1}	4.39×10^{-7}
End vs. None	3.16×10^{-8}	1.14×10^{-30}	5.71×10^{-37}
Middle vs. None	7.41×10^{-4}	5.08×10^{-10}	2.73×10^{-10}

Supplementary Table 3: p-values when comparing the same group of fibres at different times. The analysis is identical to that in **Supplementary Table 2**.

	End	Middle	None
0 vs. 30 min.	2.21×10^{-13}	5.22×10^{-2}	2.03×10^{-1}
0 vs. 60 min.	5.05×10^{-26}	5.04×10^{-1}	7.07×10^{-2}
30 vs. 60 min	2.41×10^{-7}	3.16×10^{-1}	9.41×10^{-1}

Supplementary Discussion 2.

Mass-action kinetic modelling and data-fitting

Methods

Reactions in the model are given in the form of ordinary differential equations (ODEs), with free parameters representing rate-constants. For a set of rate-constants, ODEs are solved numerically via MATLAB's *ode45* routine. Data-fitting between model and experiment used a derivative-free routine *fminsearch* employing the Nelder-Mead simplex algorithm search method.[8] The objective function that *fminsearch* minimized was the root-mean-square-difference (RMSD) between the model and the experiment:

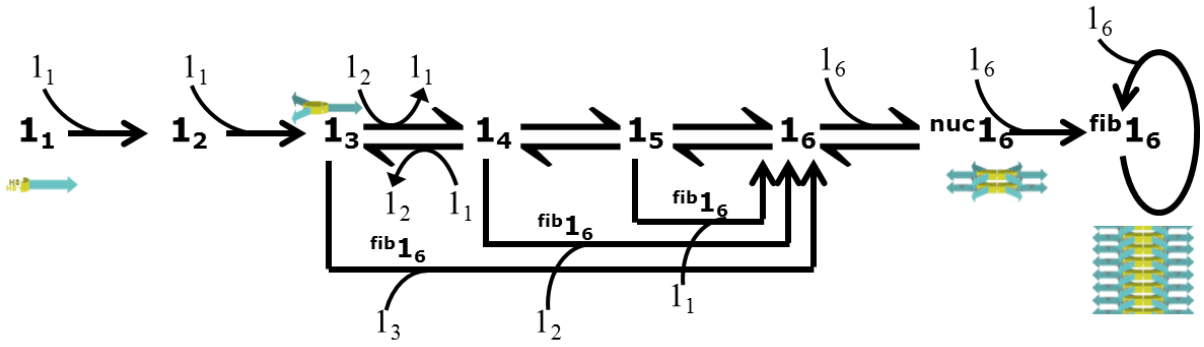
$RMSD(k) = \sqrt{\langle \sum_i \left(\sum_t \left[Y_t^i - S(k)_t^i \right]^2 \right) \rangle}$, where k is set of rate-constants values (**Supplementary Table 4**), and for species i (i=monomer to hexamer) Y_i^t is its experimentally measured concentration at time t and S_i^t is its modeled concentration at that time. <∘> denotes an average. All MATLAB's functions are used with their default parameters unless stated otherwise. Because the UPLC measurements do not differentiate between any state of the hexamers (i.e. free or in form of a fibre), when the model was compared with an experiment the total concentration of modelled hexamer in all its states was considered. Fitting was constrained such that the rate constant for the nucleation stage (k^2_f, **Supplementary Table 4**) is the lowest, by penalizing those cases.

Simulated seeding at time t was done as follows: first, the model was simulated until time t (with initial concentrations as in the experiment[4]: $[1_1]$=966, $[1_2]$=94, $[1_3]$=437, $[1_4]$=327 and the rest at 0 μM) and the last concentrations of all species were recorded, then concentration of $^{fib}1_6$ was increased by the seed amount and the simulation was continued.

Model Description

The mass-action model of the system is given in the form of ODEs representing the reactions in the system (**Supplementary Figure 12** and **Supplementary Table 4**). The different types of reactions occurring in the system are:

1. Oxidation: 2*RSH → RSSR
2. Exchange: while the complete exchange of thiols is likely to involve multiple steps and possibly also contributions from a radical mechanism, here it is considered in a simplified form while maintaining mass-balance, including the number of thiols and thiolates (–SH and $–S^-$): $c_n+l_m \leftrightharpoons c_{n+1}+l_{m-1}$, where c and l represent cyclic and linear species and n and m represent the number of building blocks in each species (n=3..6, m=1,2).
3. Self-assembly: the existence of fibres is revealed by imaging. Such fibres are homologue to those often studied in the field of amyloid fibrillation[5], and therefore analogue stacking/assembling pathways are considered here. In order to reduce the model's complexity and number of parameters, only a single nucleation stage is assumed ($^{nuc}1_6$) before the stack is considered to be a full-fledged fibre ($^{fib}1_6$). The model does not assign a specific structure/length to fibres, and therefore the concentration of $^{fib}1_6$ is defined as the concentration of hexamers in this state (and similarly for $^{nuc}1_6$).
4. Fibre-mediated-reaction: in line with the high-speed AFM observations of the present work, a tri-molecular reaction is considered, whereby cyclic species react with linear species in the vicinity of fibres in a non-reversible manner, resulting in hexamers. For the reaction in which hexamer is produced from two trimers, the latter means that one of the trimers is assumed to be linear.



Supplementary Figure 12. An overview of the self-assembly and replication model, that provides the best fit between model and experiment. Further information is Supplementary Table 4 and Supplementary Equation 1.

Supplementary Table 4. Listing of the model's reactions and their rate-laws (see also Supplementary Figure 12).

	Reaction	Rate law		Units
1	$1_1+1_1 \rightarrow 1_2$	$r_{112} = k_{112}[1_1]^2$	k_{112}	Bi-molecular
2	$1_1+1_2 \rightarrow 1_3$	$r_{123} = k_{123}[1_1][1_2]$	k_{123}	Bi-molecular
3	$1_3+1_2 \rightarrow 1_4+1_1$	$r_{324} = k_{324}[1_3][1_2]$	k_{324}	Bi-molecular
4	$1_3+1_2 \leftarrow 1_4+1_1$	$r_{413} = k_{413}[1_4][1_1]$	k_{413}	Bi-molecular
5	$1_4+1_2 \rightarrow 1_5+1_1$	$r_{425} = k_{425}[1_4][1_2]$	k_{425}	Bi-molecular
6	$1_4+1_2 \leftarrow 1_5+1_1$	$r_{514} = k_{514}[1_5][1_1]$	k_{514}	Bi-molecular
7	$1_5+1_2 \rightarrow 1_6+1_1$	$r_{526} = k_{526}[1_5][1_2]$	k_{526}	Bi-molecular
8	$1_5+1_2 \leftarrow 1_6+1_1$	$r_{615} = k_{615}[1_6][1_1]$	k_{615}	Bi-molecular
9	$1_3+1_3+{}^{fib}1_6 \rightarrow 1_6+{}^{fib}1_6$	$r_{cat3} = k_{cat3}[{}^{fib}1_6][1_3][1_3]$	k_{cat3}	Tri-molecular
10	$1_4+1_2+{}^{fib}1_6 \rightarrow 1_6+{}^{fib}1_6$	$r_{cat4} = k_{cat4}[{}^{fib}1_6][1_4][1_2]$	k_{cat4}	Tri-molecular
11	$1_5+1_1+{}^{fib}1_6 \rightarrow 1_6+{}^{fib}1_6$	$r_{cat5} = k_{cat5}[{}^{fib}1_6][1_5][1_1]$	k_{cat5}	Tri-molecular
12	$1_6+1_6 \rightarrow 2\cdot{}^{nuc}1_6$	$r_f^{62} = k_f^2[1_6]^2$	k^2_f	Bi-molecular
13	$1_6+1_6 \leftarrow 2\cdot{}^{nuc}1_6$	$r_b^{62} = k_b^2[{}^{nuc}1_6]$	k^2_b	Uni-molecular
14	$2\cdot{}^{nuc}1_6+1_6 \rightarrow 3\cdot{}^{fib}1_6$	$r_{fiber} = k_{fiber}[{}^{nuc}1_6][1_6]$	k_{fibre}	Bi-molecular
15	${}^{fib}1_6+1_6 \rightarrow 2\cdot{}^{fib}1_6$	$r_{elong} = k_{elong}[{}^{fib}1_6][1_6]$	k_{elong}	Bi-molecular

Supplementary Equation 1. collects all the reactions that are listed in **Supplementary Table 4**, the model is now given in the form of the rates at which each species' concentration change:

a) $\frac{d}{dt}[1_1] = +r_{324} + r_{425} + r_{526} - 2r_{112} - r_{123} - r_{413} - r_{514} - r_{615} - r_{cat5}$

b) $\frac{d}{dt}[1_2] = +r_{112} + r_{413} + r_{514} + r_{615} - r_{123} - r_{324} - r_{425} - r_{526} - r_{cat4}$

c) $\frac{d}{dt}[1_3] = +r_{123} + r_{413} - r_{324} - 2r_{cat3}$

d) $\frac{d}{dt}[1_4] = +r_{324} + r_{514} - r_{413} - r_{425} - r_{cat4}$

e) $\frac{d}{dt}[1_5] = +r_{425} + r_{615} - r_{514} - r_{526} - r_{cat5}$

f) $\frac{d}{dt}[1_6] = +r_{526} + 2r_b^{62} + r_{cat3} + r_{cat4} + r_{cat5} - r_{615} - 2r_f^{62} - r_{fiber} - r_{elong}$

g) $\frac{d}{dt}[{}^{nuc}1_6] = +2r_f^{62} - 2r_b^{62} - 2r_{fiber}$

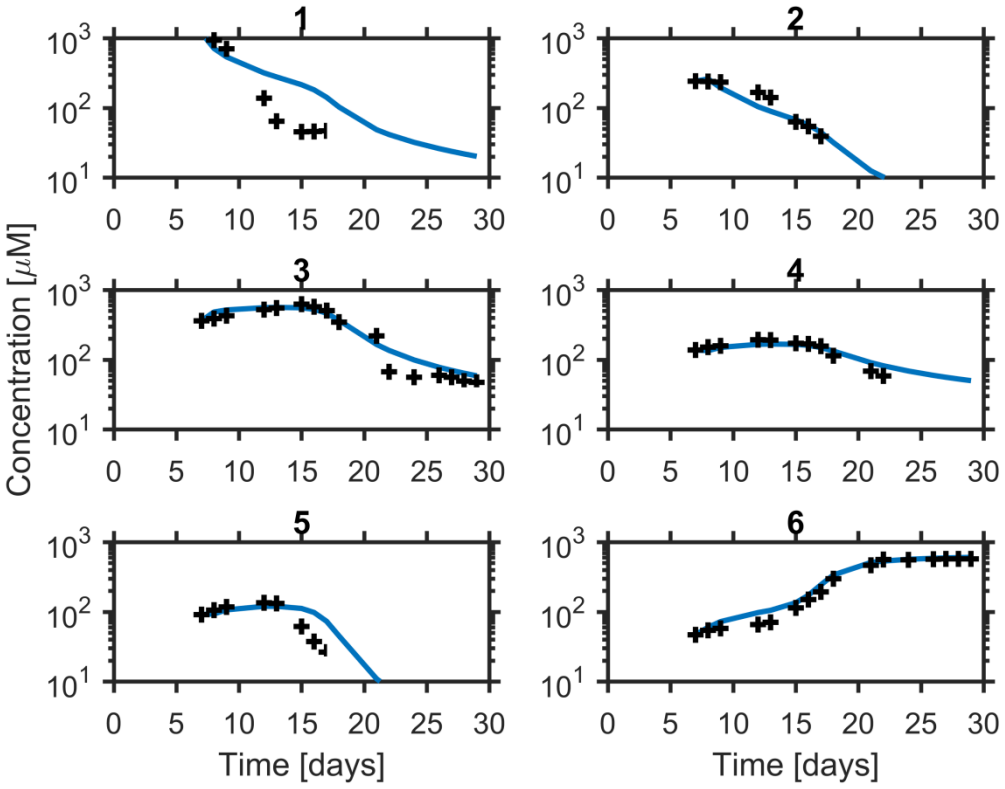
h) $\frac{d}{dt}\left[{}^{fib}1_6\right] = +3r_{fiber} + r_{elong}$

Data fitting

Experimental data is taken from a previous publication[4]. After data fitting, the modelled concentrations of all the species are now in good agreement with the experimentally measured ones (**Supplementary Figure 13**, note the log-scale). In order to test if other parameter sets could result in lower RMSD value, 5000 data-fitting trials were executed, each with a randomly picked set of initial parameter guesses distributed in the range 10^{-4}-10^4. The final parameter values that gave the lowest RMSD are shown in **Supplementary Table 5**.

While the model fits the experiment, differences could arise from experimental noise in concentration measurements and temperature fluctuations during the experiment. Another source of potential differences is that the model assumes a single nucleation stage and that all fibres are identical.

Furthermore, many of the reactions modelled (**Supplementary Table 5**) consist of multiple fundamental steps (for example, disulfide exchange involves a ring-opening step and a ring closing step), that are combined into a single apparent rate constant in the model. As solubility of oxygen in water is reported[9] to be in the order of 10^2-10^3 μM which is lower than the initial, un-oxidized, building block concentration in the experiment, the data for the first 6 days was not included in the fitting process to avoid introducing biases due to the low relative oxygen concentration at the early stage of the experiment.



Supplementary Figure 13. Experimentally measured concentrations (+ symbol) and modelled values (lines) for all of the system's species. Model is based on the best fitted rate-constants (**Supplementary Table 5**) that give the lowest RMSD of 57.47.

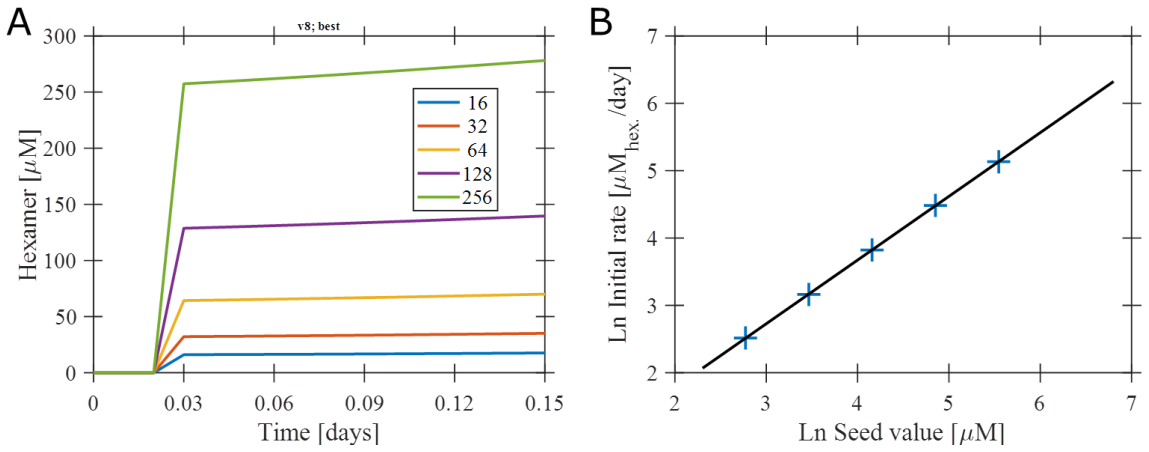
Supplementary Table 5. Values of fitted parameters.

	Fitted value	**Units**
k_{112}	2.34×10^{-4}	1/(μM×Day)
k_{123}	5.41×10^{-4}	1/(μM×Day)
k_{324}	1.53×10^{-3}	1/(μM×Day)
k_{413}	1.50×10^{-3}	1/(μM×Day)
k_{425}	3.306×10^{-3}	1/(μM×Day)
k_{514}	1.30×10^{-3}	1/(μM×Day)
k_{526}	9.66×10^{-4}	1/(μM×Day)
k_{615}	2.27×10^{-4}	1/(μM×Day)
k_{cat3}	1.44×10^{-6}	$1/(\mu M^2\times Day)$
k_{cat4}	6.07×10^{-6}	$1/(\mu M^2\times Day)$

k_{cat5}	2.54×10^{-5}	$1/(\mu M^2\times Day)$
k^2_f	1.35×10^{-5}	$1/(\mu M\times Day)$
k^2_b	2.70×10^{-5}	$1/Day$
k_{fibre}	1.12×10^{-3}	$1/(\mu M\times Day)$
k_{elong}	8.43×10^{-3}	$1/(\mu M\times Day)$

Simulated seeding

Previously, when a fresh batch of building block was prepared and seeded with pre-existing fibres it was experimentally observed to immediately give rise to the exponential growth, thus providing firm evidence for the templating ability of fibres/hexamers.[4] Seeding with different amounts of pre-formed fibres and following the immediate change exponential growth rate allowed for the determination of the order of replicator to be 0.996±0.166. Performing an identical seeding procedure in the model gave similar value of 0.946 (**Supplementary Figure 14**).



Supplementary Figure 14. Simulated seeding. (A) Initial hexamer growth for seeded fibre concentrations (in μM). Initial replication rate is obtained via linear fitting to hexamer concentration in the range: t=0.03..0.12, in line with the seeding experiment.[4] (B) Initial replication rates against initial seed concentration. Black line is a fit: ln(rate)=ln(seed)*0.946-0.109.

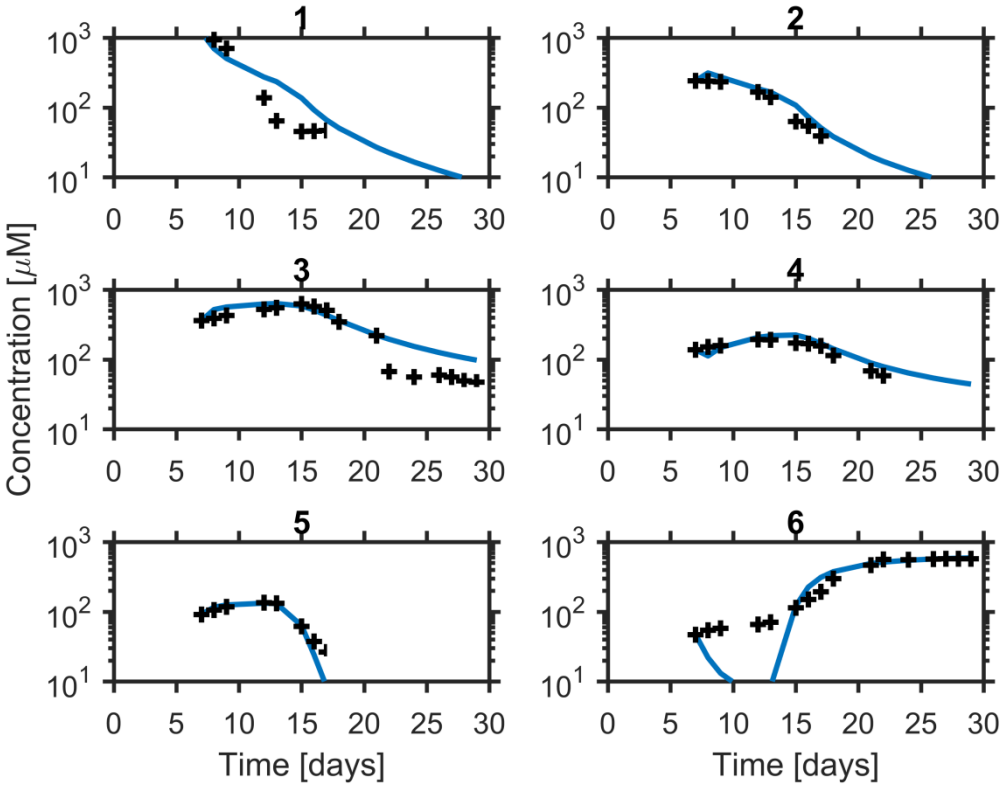
Alternative models

While both the fit quality and the simulated seeding results support the validity of the model, it is important to assess whether alternative models can reproduce the experimentally observed dynamics of the system. To this end, several model variants were constructed, representing alternative fibre-mediated-reactions (including non-catalytic ones) (**Supplementary Table 6**). These variants were fitted in the same way as described above and the figures below show the best fits (i.e. that gave lowest RMSD) for each one.

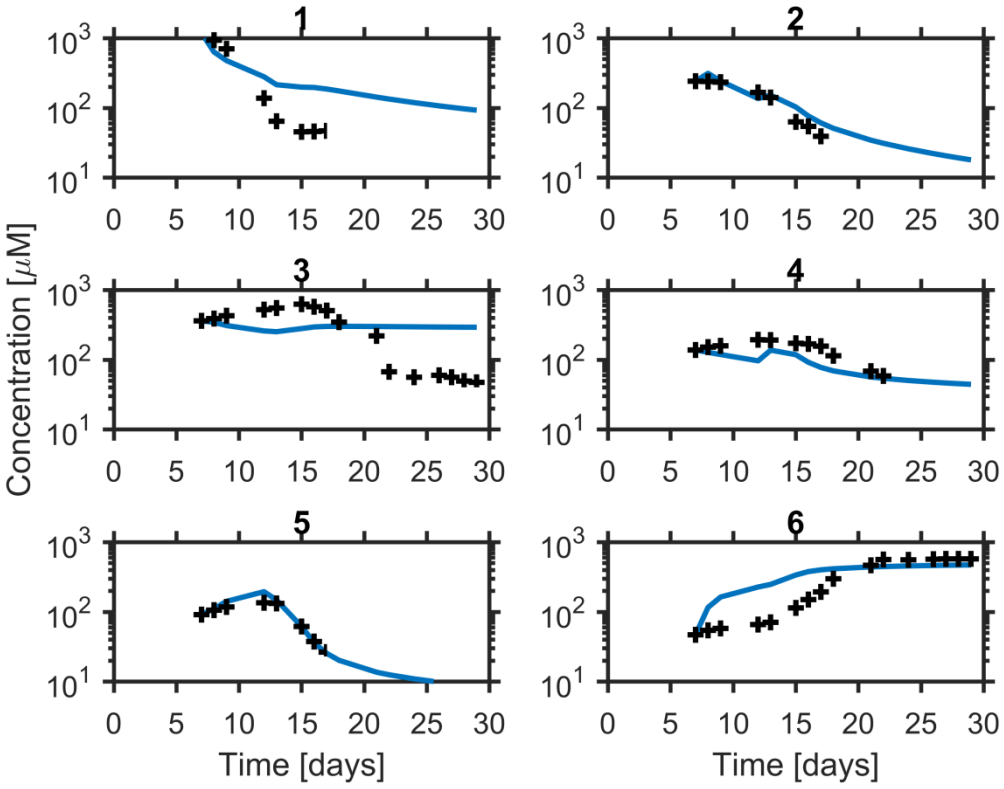
Variants B-D gave significantly worse RMSD values compared to the original model. Only variant A gave a RMSD value that is comparable to that of the original model. However, variant A performs poorly on fitting the growth of the species of interest: the replicating hexamers.

Supplementary Table 6. Description of alternative models tested.

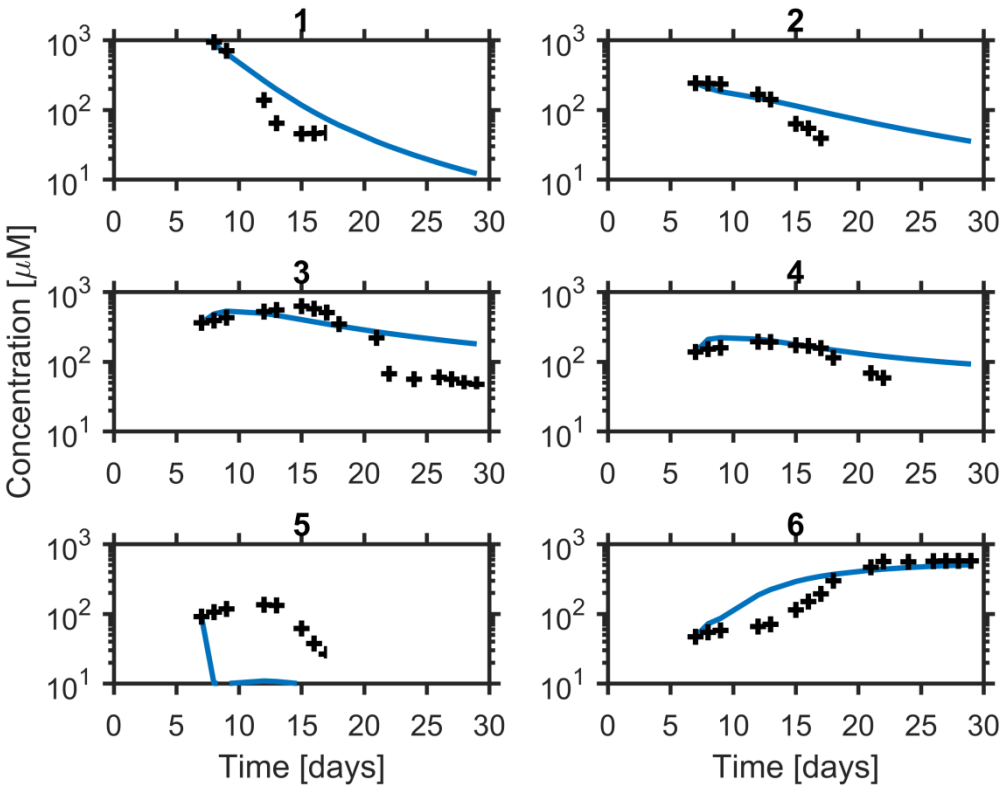
Variant	Change reactions 9, 10 and 11:	Description	Best RMSD
A	$1_3+1_3+{}^{fib}1_6\rightarrow {}^{fib}1_6+{}^{fib}1_6$ $1_4+1_2+{}^{fib}1_6\rightarrow {}^{fib}1_6+{}^{fib}1_6$ $1_5+1_1+{}^{fib}1_6\rightarrow {}^{fib}1_6+{}^{fib}1_6$	Formed hexamers are inserted into fibre unassisted $r_3 = k_3[{}^{fib}1_6][1_3][1_3]$ $r_4 = k_4[{}^{fib}1_6][1_4][1_2]$ $r_5 = k_5[{}^{fib}1_6][1_5][1_1]$	59.07
B		Fibres are allowed to catalyse the disulfide exchange reactions (3-8 in **Supplementary Table 4**). Example: $r_{324} = (k_{324} + k_{cat3}[{}^{fib}1_6])[1_3][1_2]$	124.7
C	$1_3+1_3\rightarrow 1_6$ $1_4+1_2\rightarrow 1_6$ $1_5+1_1\rightarrow 1_6$	Hexamers are formed, but not catalysed by fibres $r_3 = k_3[1_3][1_3]$ $r_4 = k_4[1_4][1_2]$ $r_5 = k_5[1_5][1_1]$	83.90
D	$1_3+1_3\rightarrow {}^{fib}1_6$ $1_4+1_2\rightarrow {}^{fib}1_6$ $1_5+1_1\rightarrow {}^{fib}1_6$	Same as C, hexamers are inserted into fibre unassisted $r_3 = k_3[1_3][1_3]$ $r_4 = k_4[1_4][1_2]$ $r_5 = k_5[1_5][1_1]$	97.10



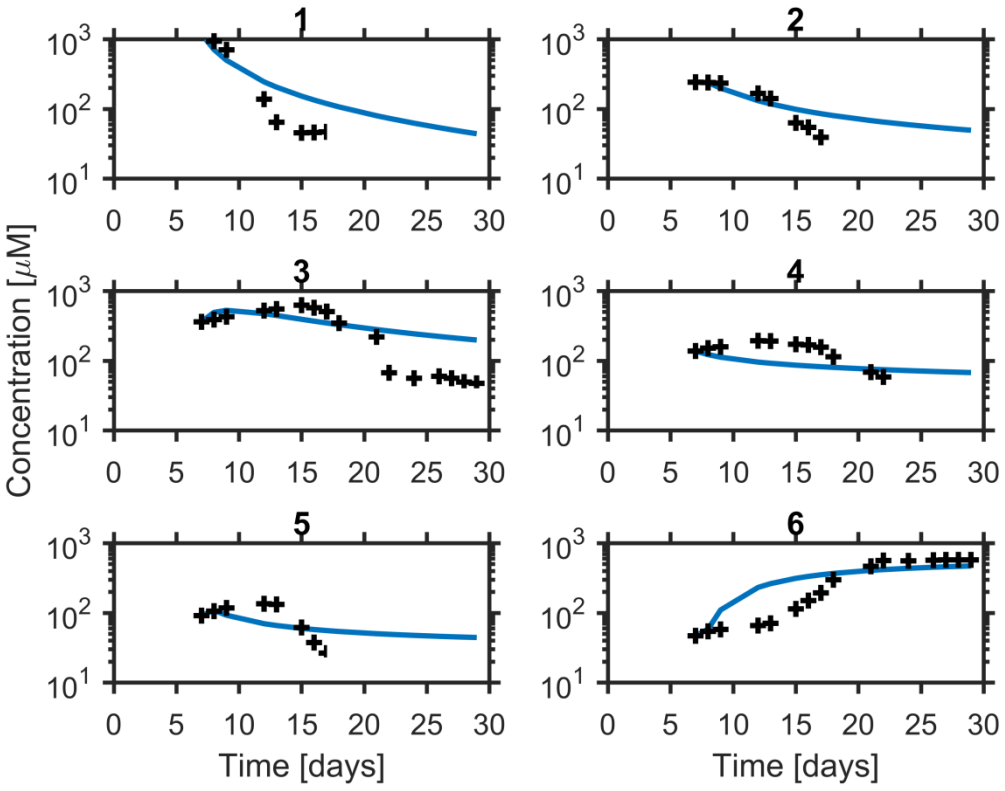
Supplementary Figure 15. Fit to model variant A (see **Supplementary Table 6**).



Supplementary Figure 16. Fit to model variant B (see **Supplementary Table 6**).



Supplementary Figure 17. Fit to model variant C (see **Supplementary Table 6**).



Supplementary Figure 18. Fit to model variant D (see **Supplementary Table 6**).

SI References:

(1) Hahakura, S.; Isoda, S.; Ogawa, T.; Moriguchi, S.; Kobayashi, T. Formation of Ultrafine Platinum Particles in an Aqueous Solution with a Surfactant. *J. Cryst. Growth* **2002**, *237–239* (1–4), 1942–1945.

(2) Liu, Y.; Stuart, M. C. A.; Buhler, E.; Lehn, J. M.; Hirsch, A. K. H. Proteoid Dynamers with Tunable Properties. *Adv. Funct. Mater.* **2016**, *26* (34), 6297–6305.

(3) Franken, L. E.; Boekema, E. J.; Stuart, M. C. A. Transmission Electron Microscopy as a Tool for the Characterization of Soft Materials: Application and Interpretation. *Adv. Sci. (Weinheim, Ger.).* **2017**, *4* (5), 1600476.

(4) Colomb-Delsuc, M.; Mattia, E.; Sadownik, J. W.; Otto, S. Exponential Self-Replication Enabled through a Fibre Elongation/Breakage Mechanism. *Nat. Commun.* **2015**, *6* (May), 7427.

(5) Lee, C. C.; Nayak, A.; Sethuraman, A.; Belfort, G.; McRae, G. J. A Three-Stage Kinetic Model of Amyloid Fibrillation. *Biophys. J.* **2007**, *92* (10), 3448–3458.

(6) Roelfs M., and Kroon P. C. "symfit". In: Zenodo 24/10/2018. http://doi.org/10.5281/zenodo.1133336

(7) Frederix, P. W. J. M.; Idé, J.; Altay, Y.; Schaeffer, G.; Surin, M.; Beljonne, D.; Bondarenko, A. S.; Jansen, T. L. C.; Otto, S.; Marrink, S. J. Structural and Spectroscopic Properties of Assemblies of Self-Replicating Peptide Macrocycles. *ACS Nano* **2017**, *11* (8), 7858–7868.

(8) Lagarias, J. C.; Reeds, J. A.; Wright, M. H.; Wright, P. E. Convergence Properties of the Nelder-Mead Simplex Method in Low Dimensions. *SIAM J. Optim.* **1998**, *9* (1), 112–147.

(9) Reynafarje, B.; Costa, L. E.; Lehninger, A. L. O2 Solubility in Aqueous Media Determined by a Kinetic Method. *Anal. Biochem.* **1985**, *145* (2), 406–418.