SUPPORTING INFORMATION

A comprehensive mechanism of fibrin network formation involving early branching and delayed single- to double-strand transition from coupled time-resolved X-ray/light scattering

Mattia Rocco,1* Matteo Molteni,2 Marco Ponassi,1 Guido Giachi,3 Marco Frediani,3 Alexandros Koutsioubas,4 Aldo Profumo,1 Didier Trevarin,4 Barbara Cardinali,5 Patrice Vachette,6,7 Fabio Ferri,2 and Javier Pérez4

1Biopolimeri e Proteomica, IRCCS AOU San Martino-IST, I-16132 Genova, Italy
2Dip. di Scienza e Alta Tecnologia and To.Sca.Lab, Università dell’Insubria, I-22100 Como, Italy
3Dip. di Chimica "Ugo Schiff", Università di Firenze, I-50019 Sesto Fiorentino (FI), Italy
4SWING beamline, Synchrotron SOLEIL, Saint-Aubin, F-91192 Gif-sur-Yvette, France
5Dept. of Pathology and Laboratory Medicine, University of North Carolina, Chapel Hill, NC 27599-7525, USA
6Université Paris-Sud, IBBM, UMR8619, and 7CNRS, F-91405 Orsay, France

*Email: mattia.rocco@hsanmartino.it; Phone: +39-010555-8310; Fax: +39-010555-8325

CONTENTS:

Sample preparation (Materials and Methods) S2

Figure S1, Schematic representation of the stopped-flow/WA-MALS/SAXS experimental set-up S3

Stopped-flow/WA-MALS/SAXS set-up (Materials and Methods) S4

Figure S2, Chronograph of the computer-controlled events for the SFM 4 syringes and the valves in the stopped-flow/WA MALS/SAXS injection/pulsing operations S6

WA-MALS data analysis (Materials and Methods) S7

Figure S3, WA-MALS data analysis-I S9

Figure S4, WA-MALS data analysis-II S10

Figure S5, $<R_g^2>_c$ vs. $<M>_w$ plots: comparison between theory and simulated WA-MALS data S13

SAXS data analysis (Materials and Methods) S14

YL models definition and scattering properties (Results) S15

Additional References S16

Figure S6, Cross-section Guinier plots comparison of experimental and model SAXS curves S17

Figure S7, Cross section Guinier analysis of a SAXS dataset acquired during a FG polymerization S18

Figure S8, Bifunctional polycondensation polymer distributions S19

YL→DS models generation and parameters computation (Materials and Methods) S20

Figure S9, Parameters setting boxes for the fibrin polymers generation LabView® program S20

Figure S10, YL chain generation scheme S21

Figure S11, YL branched polymer generation scheme S23
Sample preparation.

All chemicals were reagent grade from Merck (VWR International, Milano, Italy) unless otherwise stated, and double-distilled or MilliQ water was used for all the solutions. Lyophilized human FG (FIB1, plasminogen depleted, ERL, South Bend, IN, USA) was reconstituted, dialyzed in tris(hydroxymethyl)aminomethane 50 mM, NaCl 104 mM, aprotinin (Sigma–Aldrich, St. Louis, MO, USA) 10 kallikrein inhibitor units/mL, pH 7.4 (TBS), aliquoted and stored as previously described. A specific absorption coefficient $E_{280} = 1.51 \text{ mL mg}^{-1} \text{ cm}^{-1}$ was used to determine the FG concentration. Ancrod was from Sigma-Aldrich (A-5042, lot 33H9318, nominal 500 National Institutes of Health (NIH) units/mg; no longer available). One 0.1 mg (50 NIH u) vial was reconstituted in 1 mL of water, aliquoted, and stored at -80 °C. One 25 mg vial of GPRP-NH$_2$ (H-1998, MW 424.5, Bachem, Bubendorf, CH) was dissolved in 1 mL of TBS. A DU640 spectrophotometer (Beckman Coulter, Fullerton, CA, USA) was used in the preparative steps, while a Nanodrop ND-1000 (Thermo Scientific, Wilmington, DE, USA) was used at the SOLEIL synchrotron. Before each series of polymerizations, ~2 ml of FG solutions at 16-17 mg/mL were centri-filtered at 12,000 rpm over 0.22 μm pore-size cellulose acetate filters (Costar Spin X, Sigma-Aldrich) and then chromatographed at room temperature over a glass 1.6 cm diameter column filled with Superdex 200 to a 93 cm bed height (GE Healthcare, Philadelphia, PA, USA). The column was connected to an AKTA-Purifier 900 FPLC system (GE Health-care) running with TBS at a 1 mL/min flow rate with UV detection at 280 nm and with 1 mL fraction size collection. FG monomer peak fractions were pooled and diluted with TBS so as to have ~20 mL of solution at ~1.2 mg/mL. Quality control and characterization of the samples by SDS–PAGE and Western blotting was done as previously extensively reported.
Figure S1 - Schematic representation of the stopped-flow/WA-MALS/SAXS experimental set-up. The four panels refer to different combinations of valves position in the various stages of an experiment, as indicated by the text at the bottom of each panel. The fluid lines are color-coded for an easier tracing: red, FG; blue, enzyme Ancrod; magenta, reacting solution; green, buffer; brown, FG plus buffer.
**Stopped-flow/WA-MALS/SAXS set-up.**

In Figure S1, a diagram of the stopped-flow WA-MALS/SAXS set up is shown. Rapid mixing was achieved with a four-syringe stopped-flow mixer (SFM-4, Bio-Logic, Grenoble, FR) equipped with 20 mL PEEK syringes with Teflon-coated plungers. The SFM-4 and the SAXS flow-through capillary were kept at a constant 20 ± 0.1 °C temperature by an external circulating water bath. The SFM-4 inlet ports were fitted with 3-way, Luer-lock nylon valves (Sigma-Aldrich) for loading and priming operations with all-propylene syringes (Norm-Ject, Sigma-Aldrich). Each loading syringe carried a 28 mm diameter, 0.22 µm pore-size filter (Corning 431212, Sigma-Aldrich), to clarify all solutions when filling the SFM-4 syringes (see top-left side of Figure S1A). Several water and buffer cleaning cycles were done until reasonable light scattering baselines were obtained (see below).

The SFM-4 head has an “internal” PEEK piece with a mixer fed by syringes 1 and 2 and two holes fed by the exit channels of syringes 3 and 4. A variety of “external” mixing devices can be mounted on top of this internal piece for mixing the syringes’ 3 and 4 content with the premixed syringes 1-2 flow, and then carry the liquid to observation devices. However, preliminary tests showed that once a first “shot” was made mixing FG and enzyme using any mixer/syringe combination, the relatively long time (up to 10-15 minutes) between shots allowed the enzyme back-diffusion into the syringe holding FG, which then started to polymerize. It was thus mandatory to build a remotely controllable external system of valves and mixers which is schematized in Figure S1 (top-side of each panel).

First, Bio-Logic manufactured a special headpiece having three channels which on the internal face precisely superimpose with the syringes 1-2 mixer and with the syringes 3 and 4 outlets, respectively, and on the external side presents three ports with 1/4”-28 UNF threads. From this point on, 1/16” o.d., 0.020” i.d. PEEK tubing and low-pressure fittings were used for all connections. The syringes 1-2 (holding FG and buffer, respectively) mixer outlet and that of syringe 4 (carrying the enzyme) were each connected with one of the inlet ports of two 3-way solenoid-actuated low-pressure isolation valves (P/N 075T3MP12-32-5M, 0.032” valve orifice, 45 µL internal volume, PEEK body/PTFE membrane; Bio-Chem Valve, msscientific Chromatographie-Handel GmbH, Berlin, DE). The outlet of syringe 3 (carrying buffer) was instead connected to the central port of a PEEK solvent splitter with 0.030” internal channels, whose two side ports were connected to the two solenoid valves other side ports.
Their central exit ports were connected to the lateral inlets of a low-dead volume (3.1 µL) PEEK static mixing tee with a 10 µm UHMWPE filter on the center port (U-466, IDEX Health & Science, Oak Harbor, WA, USA) providing a turbulent mixing. To this point, all the PEEK tubing connection lengths were kept to a minimum. The solenoid valves were powered and driven by a home-made device allowing either their manual control via hard switches, or automated control by computer software via BNC connectors carrying TTL pulses, for remote operation. The mixing tee exit port was connected to an inline, 70 µL dead-volume PEEK/titanium filter holder (#2713, WTC, Santa Barbara, CA, USA) carrying a replaceable 10 mm diameter, 0.1 µm pore-size Durapore membrane (VVLP, Merck Millipore, Billerica, MA, USA).

To allow directing to waste the rapid mixing device efflux and/or the manual cleaning of the WA-MALS and SAXS cells, a 4-way, manually/remotely controlled valve (Rheodyne MXP7900-000 series II, two positions/six ports, 45 µL hold volume, DuraLife™, IDEX Health & Science) was inserted after the inline filter (see Figure S1, middle section of all panels; purging/washing configuration in panels A and C). The other three ports used were equipped with i- a 1/16” o.d., 0.038” i.d. Teflon tubing going to waste; ii- a similar PTFE tubing with integral Luer-lock adapter (A7552, Sigma-Aldrich) for connection to a 20 mL syringe with a 0.2 µm pore-size filter; iii- a ~30 cm-long sample accumulation PEEK loop going toward the observing devices placed sequentially, first WA-MALS and then SAXS.

Before the WA-MALS, a motorized 4-way valve with 1/4”-28 UNF threaded ports and with either manual or remote control via TTL pulses (Bio-Logic, specially made) was inserted (see Figure S1, all panels). This is a crucial step, because the reacting solution cannot be left standing into the SAXS capillary, where successive exposures to the intense X-ray beam will destroy the samples. Thus, the reacting mixture is pushed along the system in small 5 µL steps between each exposure, providing fresh material in front of the X-ray beam. However, if the pulses were also applied to the WA-MALS cell, they would cause spurious long-lasting spikes especially at the lower angles. The 4-way valve was connected via short PEEK tubing to the inlet and outlet of the WA-MALS cell, allowing insulating it after the mixing phase, and to the SAXS capillary inlet. The micropulses are then applied only to the SAXS cell (see Figure S1, lower part of each panel), whose exit was connected to a ~50 cm piece of Teflon 1/16” o.d., 0.038” i.d. tubing going to a waste bottle (placed at a higher level to apply a light
counter-pressure on the entire solution circuit, so that it stops flowing when the syringes stop pushing).

The manual washing/priming operations are carried out initially with valves positioned as shown in Figure S1A, and subsequently as shown in Figure S1B, first with MilliQ grade water, and then with buffer until low-noise baselines at the lower WA-MALS angles are achieved for each SFM-4 syringe. 1.6 ml shots were found to be more than sufficient to completely fill the system with fresh solution and with a large enough volume for the successive pulses required for SAXS data acquisition (see below). After the cleaning step, syringe 1 is primed and filled with the freshly chromatographed FG solution, and all others with TBS. A first series of injections is then made with remote control to acquire WA-MALS and SAXS data for the solvent baselines and for the unreacted FG (see Figure S1B-C). These injections are made exactly as if they were true polymerization reactions, so as to test and reproduce identical conditions. For polymerization reactions, syringe 4 is primed and filled with the enzyme (Ancrod) solution. Polymerizations are initiated with the valves positioned as in Figure S1B, followed by switching to the panel C positions and data are collected with the system in the panel D configuration.

![Figure S2](image)

**Figure S2** – Chronograph of the computer-controlled events for the SFM 4 syringes and the valves in the stopped-flow/WA MALS/SAXS injection/pulsing operations.

The timing of the events is described in Figure S2. The operations start with a 1.6 mL injection over 9 s (flow rate 10.67 mL/min) with the valves configuration as in Figure S1B. At its end, the rising edge of a 10 ms duration TTL pulse triggers SAXS data acquisition and launches another series of events controlled by a pre-running in-house script written in Passerelle language (iSencia Belgium, Gent, BE;
http://www.isencia.be/services/passerelle). After 2 s, both the solenoid valves of the external mixer switch to the washing/pulsing position, excluding cross-contamination between the contents of syringes 1/2 and that of syringe 4, and the semi-automatic 4-way valve diverts the flow to waste (Figure S1C). After another second (3 s from trigger), 250 μL of buffer in 1.5 s are injected via syringe 3, to clean the circuit up to the 4-way semi-automatic valve, and the 4-way rotary valve turns 90° insulating the WA-MALS cell. After 1.5 s (6 s from trigger), the semi-automatic 4-way valve switches back to the injection/pulsing position, and after another 2.5 s (8.5 s from trigger) the first 5 μL of buffer are pulsed in 30 ms from syringe 3 (Figure S1D). The pulses are then repeated every 10 s allowing two SAXS data acquisitions, each lasting 0.5 s, every 5 s, for as long as the programmed total acquisition time of the experiment. The WA-MALS data acquisition (collection time 0.5 s), using the ASTRA® 5.9 software (WTC), is launched before starting any event, as it does not need to be synchronized and allows checking/recording the baselines.

At the end of each run, the motorized 4-way valve is reverted to the inject position (as in Figure S1B) while the solenoid valves are kept as in Figure S1C, and one or more 1.6 mL shots with TBS using syringe 3 are done to clean the WA-MALS and SAXS cells. Cleanliness is judged by the WA-MALS baselines quality; if not satisfactory, the sequence of experiments is interrupted, and the WA-MALS/SAXS cells are manually cleaned sequentially with TBS, water, a 2% in water Hellmanex detergent solution (Hellma Analytics, Müllheim, DE), water, and TBS.

WA-MALS data analysis.

The WA-MALS/QELS data were initially analyzed with the ASTRA® 6.0.3 software (WTC). The TBS viscosity was calculated to be 1.0298 cP at 20 °C. A dn/dc of 0.192 cm³/g was used for FG. The actual concentration in the scattering cells was checked by the fiber-optic spectrophotometer placed on the SAXS capillary. To derive absolute molecular weights from the WA-MALS data, a factory set calibration constant of 3.5575 × 10⁻⁵ V⁻¹ cm⁻¹ was used. A first attempt to normalize the photodiodes was done using the peak fractions (4.3 mg/mL) of a bovine serum albumin (BSA, Sigma-Aldrich) solution in TBS chromatographed over the same Superdex 200 column used for FG, directly injected into the WA-MALS flow cell through a 0.02 μM pore-size filter. However, since we were then
measuring the $<R_g^2>_{\zeta}$ of very elongated polymers, we preferred to normalize the detectors with the SAXS-derived $[<R_g^2>_{\zeta}]^{1/2}$, 14.4 ± 0.2 nm (see Table 1) of a chromatographed FG solution loaded into the SFM-4 syringe and then pushed through the WA-MALS and SAXS detectors.

Data were analyzed using the Zimm method in which $K^* c/R(\theta)$ is plotted vs. $\sin^2(\theta/2)$, $\theta$ being the scattering angle. Here $K^*$ is the optical constant [cm$^2$ g$^{-2}$]:

$$K^* = 4 \pi^2 n_0^2 (dn/dc)^2 / N_A \lambda^2$$  \hspace{1cm} (ES1)

where $n_0$ is the refractive index of the solvent, $dn/dc$ [cm$^3$ g$^{-1}$] is the differential refractive index increment of the sample in respect to the solvent, $N_A$ is Avogadro’s number, and $\lambda$ in the wavelength of the incident light in vacuo. $c$ [g cm$^{-3}$] is the sample concentration, and $R(\theta)$ [cm$^{-1}$] is the Rayleigh ratio, i.e. the excess scattering from a solution above that scattered by the solvent alone into a detector centered at angle $\theta$ per unit solid angle subtended by that detector, per incident power, per unit scattering length. After calibration with a known scatterer (i.e., toluene), the geometric parameters, otherwise very difficult to measure, are removed, and the Rayleigh ratio at the $\theta = 90^\circ$ scattering angle is obtained from the measured voltages. The Rayleigh ratios at other scattering angles are then obtained by determining a normalization factor $N_\theta$ for each scattering angle, usually measured with an isotropic scatterer, while here FG was used (see above).

From the $R(\theta)$, the weight-average molar mass $<M_w>= \sum M_i w_i$ and the $z$-average square radius of gyration $<R_g^2>_{\zeta} = \sum M_i w_i (R_{g_i}^2) / \sum M_i w_i$ (with $M_i$, $(R_g^2)_i$ and $w_i$ the molar mass, square radius of gyration, and weight fraction of the $i^{th}$ species) are obtained from the Zimm plot relation:

$$\frac{K^* c}{R(\theta)} = \frac{1}{<M>_w} \left( 1 + \frac{16 \pi^2 n_0^2}{3 \lambda^2} \frac{<R_g^2>_{\zeta}}{\sin^2 \theta} \right)$$  \hspace{1cm} (ES2)

yielding $1/<M>_w$ and $<R_g^2>_{\zeta}$ from the intercept and from the ratio between the initial slope and the intercept, respectively, after fitting the data with an appropriate polynomial. Note that, owing to the dilute conditions in which we operate, we omitted the second virial coefficient of the osmotic pressure term in equation ES2.
Figure S3 – WA-MALS data analysis - I. Panels A-B, normalized voltage data for all the available detectors (color coded as shown at the right side of panel B) for a FG (0.456 mg mL\(^{-1}\)) polymerization induced by Ancrod 0.05 NIH units/mg FG in TBS; the time axis has been rescaled so that 0.0 indicates the end of mixing. A, full data set; B, the first 100 s shown on an enlarged scale. Panels C-D, baselines definition using ASTRA\textsuperscript{®} 6.0.3 for two scattering angles between the initial buffer and final buffer recorded data. Panels E-F, left-side graph of each panel (“Results graph”), Zimm plots with polynomial fitting of the reduced scattering intensities vs. \(\sin^2(\theta/2)\) at two different times during a FG polymerization run (see A). The times are shown in the right side graph of each panel (“Control graph”), as vertical lines inside the “peaks” defining the analyzed portion of each dataset (note that the times shown are absolute times since the start of data acquisition, which includes baselines before and after each run, and cleaning steps after each run; the end of the mixing time was at 1928 s for “peak” #1, and at 3279 s for “peak” #2). (E), “peak” #2 51 s after mixing. (F), “peak” #2 445 s after mixing.
Figure S4 – WA-MALS data analysis - II. Plots of $<M>_w$ (A) and $[<R_g^2>_z]^{1/2}$ (B-C) values derived from different polynomial degrees (e.g. p2, p3) fits including or excluding the lower three scattering angles (e.g. 2_18, 3_18), vs. time for the “peak” #2 region of the data shown in Figure S3.
A typical dataset processing is reported in Figures S3 and S4. Figure S3A-B shows the normalized scattering intensities of a polymerization run, with the full dataset shown in panel A, and the first ~100 s in panel B. As can be seen, very good quality light scattering curves were obtained at all angles. A snapshot of the baseline setting graph is shown in Figure S3C, for two polymerization runs as seen by the 90° scattering angle detector, with FG concentrations of 0.61 and 0.46 mg/ml, respectively. In Figure S3D a blow-up of the baseline region for the first run is reported, as seen by the 20.9° scattering angle detector, highlighting the very good quality of the data even at this low scattering angle. Note how the baseline noise at the end of the run vanishes as a consequence of repeated washing with buffer. After baseline definition “peaks” were marked within each run and the $K* c / R(\theta) \text{ vs. } \sin^2(\theta/2)$ plots were fitted with several polynomials from the 1st to the 5th degree (Figure S3E-F). When solutions of relatively large macromolecules are analyzed with the Zimm method, a 1st-degree polynomial (straight line) is sufficient to fit well the data if the relation $q^2 < R_g^2 >_z \leq 1$ [where $q = 4 \pi n \sin(\theta/2) / \lambda$] holds for every $q$ value. However, here we are dealing with polymerizing solutions of rod-like macromolecular monomers, quickly giving rise to very polydisperse collections of very elongated particles. Under these circumstances, Zimm plots rapidly deviate from linearity (see Figure S3E-F), and can only be fitted with higher degree polynomials. As previously shown, this procedure can yield reasonable $< M >_w$ and $< R_g^2 >_z$ values, but care should be taken, the main rule being that the lowest degree polynomial still giving a reasonable fit to the data should be used. Therefore, each polymerization “peak” was analyzed with several polynomials, and including/excluding the lowest scattering angles, to produce a series of $< M >_w$ and $< R_g^2 >_z$ vs. time graphs, as shown in Figure S4. As can be seen in Figure S4A, $< M >_w$ is less affected by the choice of polynomial and angles included, which become important approximately halfway through the reaction. In contrast, $< R_g^2 >_z$ (plotted as the square root $[< R_g^2 >_z]^{1/2}$ in Figure S4B-C) is much more sensitive to these choices. In Figure S4C, a blow-up of the first 100 s of the reaction is shown to highlight the issues. Apart from the first ~2-3 s, where mixing noise dominates, it is apparent that a 3rd-degree polynomial is already useful from 6-7 s onwards, without substantial differences if the 3rd lowest angle (corresponding to 20.9°) is excluded (magenta points) or included (yellow points). The transition to a 4th-degree polynomial (with the 20.9° scattering angle still as the lowest included) becomes useful after ~100 s, while the 5th-degree polynomial with the full angular range (lowest
scattering angle: 13.5°) data, having a restricted overlap zone with the lower polynomials data, are then used after ~140 s. A joined dataset is then produced by successive transitions to a higher-degree/lower angles polynomial using common regions for the splice points.

Additional checks were performed a posteriori to verify the validity of this procedure. First, synthetic WA-MALS datasets with RLDS and YL→DS/branched polymers made of cylinders were prepared using the distributions generated by the modified Flory-Janmey bifunctional polycondensation mechanism with \( Q = 90 \), up to polymerization degrees corresponding to the highest \(<M>_w\) values derived from the experimental WA-MALS data (see Materials and Methods-Modeling, in the main text, and the YL models definition and scattering properties and YL→DS models generation and parameters computation sections in this Supporting Information). The synthetic datasets were generated by computing, for each cylinder \( k \) of a polymer composed by \( N \) monomers, the scattered electric field \( E_k(\theta) \) at all the scattering angles, then summing the contributions from all cylinders, and finally recovering the polymer form factor \( P_N(\theta) \) as the normalized squared modulus of the total electric field:

\[
P_N(\theta) = \left| \sum_{k=1}^{N} E_k(\theta) \right|^2 / \left| \sum_{k=1}^{N} E_k(0) \right|^2
\]

(ES3)

For each cylinder \( E_k(\theta) \) was computed by taking into account its orientation with respect to the scattering direction and the Rayleigh-Gans approximation was used. The form factors of polymers with the same \( N \) were then averaged over many configurations and orientations, so as to obtain a statistically robust estimate for \( P_N(\theta) \). Finally, the total scattering Rayleigh ratio was calculated by taking the \( z \)-average of the form factors of all the polymers in solution as \( R(\theta) = K_c < M >_w < P(\theta) >_z \) where \( < P(\theta) >_z = \sum_{i}^{N_{\text{max}}} w_i M_i P_i(\theta) \) and \( N_{\text{max}} \) is the number of monomers of the largest polymer. A statistical level of noise equal to that present in the real experimental data was added at each \( R(\theta) \). Three different, highly polydispersed collections of rod-like polymers were generated, ranging from the rigid, unbranched DS type predicted by the classic mechanism, to the distribution of highly branched species fitting our experimental data (see YL models definition and scattering properties in this Supporting Information), passing through a similar distribution of less branched polymers.
Figure S5. Comparison between the exact (solid lines) and recovered (“simul/fit”, open symbols) $<R_g^2>_z$ vs. $<M>_w$ behaviors for three polymer distributions with different morphologies: YL→DS branched ("br") polymers with all parameters equal to the ones used in Figure 3 of the main text (magenta), YL→DS branched polymers with the same parameters but with $<l_{br}> = 25 \pm 5$ (green), and unbranched rigid RLDS polymers (blue). The recovered $<R_g^2>_z$ vs. $<M>_w$ were obtained by applying the Zimm analysis to the synthetic $R(\theta)$ dataset computed as described in the text. Inset: relative residuals plots between recovered and expected $<R_g^2>_z [((<R_g^2>_z)_{rec} / ((<R_g^2>_z)_{exp} - 1)]$.

The results of these tests are shown in Figure S5, where the expected $<R_g^2>_z$ versus $<M>_w$ curves (solid lines) are compared with the corresponding behaviors retrieved by fitting the synthetic $R(\theta)$ dataset with the Zimm method (open symbols). Various polynomials with different angular ranges were used, as described above for the fitting of the experimental data (but without fully optimizing the early phases, since here we were concerned mainly with the late phases). The three datasets refer to: (a) YL→DS branched polymers with all the parameters equal to the ones used in Figure 3 of the main text (magenta), (b) YL→DS branched polymers as in (a) but with $<l_{br}> = 25 \pm 5$ (green), and (c) unbranched rigid RLDS polymers (blue). The first result of the analysis of these datasets is that the correct $<M>_w$ and $<R_g^2>_z$ values can be recovered up to very high $<R_g^2>_z$ values provided that the relation $q^2<R_g^2>_z \leq 1$.
is obeyed by at least the two lowermost scattering angles available. Secondly, as evidenced by the fractional residuals plots ([<R_g^2>^z]_{rec} / ([<R_g^2>^z]_{exp} − 1) reported in the inset of Figure S5, the agreement between the recovered and expected behaviors is excellent for the YL→DS branched polymers of Figure 3 over the entire <$M>_w$ range [magenta, case (a)], fairly accurate over most of the <$M>_w$ range for the less branched YL→DS polymers [green, case (b)], and becoming inaccurate for the unbranched RLDS polymers when <$M>_w ≥ 8×10^5$ g mol$^{-1}$ [blue, case (c)]. The figure therefore shows that the Zimm method adopted in this work for recovering <$R_g^2>^z$ and <$M>_w$ from scattering data should have provided fairly good estimates of these values for the type of polymers we are likely dealing with.

Therefore, all FG WA-MALS data were processed in this way. The resulting <$M>_w$ and <$R_g^2>^z$ vs. time joined curves for the TBS dataset used here as an example, and for two other datasets are shown in Figure 2A-B, main text. Noteworthy are the apparent smoothness of the joined datasets, and the relatively low standard deviations associated with the points.

**SAXS data analysis**

SAXS data analysis was performed using the SWING in-house software FoxTrot. Individual SAXS images were first reduced to unidimensional curves and normalized by pixel solid angle and transmitted intensity. ~50 frames collected from the buffer injection sequence were then averaged to produce a high statistics baseline curve, which was subsequently subtracted from each sample scattering curve. The resulting curves were then analyzed with the SAS module of US-SOMO.$^{32,55}$ The overall data analysis was based on the cross-section Guinier plot for rod-like particles,$^{27}$ $\ln[q I^*(q)]$ vs. $q^2$, according to which

$$
\ln \frac{q I^*(q)}{c} = \ln \frac{q I^*(0)}{c} - \left(\frac{R_c^2}{2}\right) q^2
$$

(for $2\pi L < q < 1/R_c$)  \hspace{1cm} (ES4)

where $I^*(q)$ is the normalized/reduced scattering intensity (see$^{32}$) as a function of the momentum transfer $q$, $R_c$ is the cross sectional radius of gyration of the rod-like particle ($R_c^2 = r^2/2$, $r$ being the geometrical radius of the particle), and $c$ is the sample concentration. Equation ES4 is valid over a finite $q$-range, delimited by $R_c$ and by the FG monomer length $L = 460$ Å. In the derivation of $I^*(q)$, a partial specific volume of 0.715 cm$^3$/g was used for FG.$^{34}$
**YL model definition and scattering properties**

The Y-Ladder (YL) polymers are built by assembling cylinders connected at single center-to-end binding sites, with a 20°-45° range of θ binding angles (with Gaussian distribution of values) between any two units, and with a limited range (5°-20°) of off-plane ϕ azimuthal angles (see the YL→DS models generation and parameters computation section in this Supplementary Information). It was crucial to check to what extent $R_c^2$ values determined by SAXS for the classic DS and the new YL polymers were trustable. To this purpose, curves for the monomer, and DS and YL 20-mers (an average curve over 20 conformations was used for the latter), were computed and compared with experimental data. To compute SAXS curves for the models, each monomer unit was represented by a linear array of tangent beads whose diameter and number was adjusted to best match the width and length of the hydrated cylinder used to represent FG (10 beads with $d = 4.6$ nm). SAXS curves were then generated using the SAS module of US-SOMO. In Figure S6, two data frames derived from experimental SAXS data collected at the beginning (black) and at the end (red) of a FG (0.456 mg/mL) polymerization process induced by Ancrod 0.05 NIH units/mg FG in TBS are shown, together with a series of calculated curves for various models. While the monomer (green) and DS (blue) curves were linear across a large $q^2$ range, each with its expected slope, the YL 20-mers average curve (magenta) exhibited a broad $q^2$ range having a slope comparable with the monomer curve, and a strong upturn at smaller $q^2$ values, partially deriving from the inter-monomer distance in the Y-ladder. Interestingly, a similar upturn is seen in the experimental data as the polymerization proceeds (Figure S6, open red squares). This upturn constrained us to adopt a $q^2_{\text{min}}$ value for $<R_c^2>_{z}$ derivation significantly higher than that imposed by the Guinier law for a rod (ES4). In contrast, the value of $q^2_{\text{max}}$ was adapted to the calculated $<R_c^2>_{z}$ value for each curve to comply with the relation $q^2_{\text{max}} <R_c^2>_{z} \leq 1$. Using this adaptive $q^2$ interval (marked by the vertical dashed lines in Figure S6, see more below and Figure S7) to analyze SAXS data, we see that the YL polymers, if present, would produce $<R_c^2>_{z}$ values comparable to that of the FG monomer and clearly below that of the DS configuration.

To account for the final formation of DS fibrils, intermediate polymers presenting YL and DS segments were also generated, as well as branched polymers with binding and azimuthal branching angles $\theta_b = 20° \pm 5°$ and $\phi_b = 20° \pm 20°$ (see YL→DS models generation and parameters computation in
The theoretical average SAXS curves for twenty conformers of YL→DS 20-mers (orange) and of thirty conformers of the highly branched YL→DS 300-mers (grey) are also displayed in Figure S6, the latter being linear and practically parallel to the RLDS curve (blue) in the fitting region for $<R_c^2>_z$ calculation. In summary, we can conclude that, over the adaptive $q^2$ range defined above, models and experimental data exhibit a linear profile and yield safe estimates of $<R_c^2>_z$ values. This is further illustrated by Figure S7 in which a subset of the SAXS data corresponding to the WA-MALS data of Figure S3A-B is plotted. The data are of reasonably good quality, considering the low protein concentration employed. Note how all these plots exhibit a linear region over the adaptive $q^2$ interval, from $1.0 - 2.5 \times 10^{-3}$ to $1.0 - 1.5 \times 10^{-3}$ Å$^{-2}$, with a time dependent slope and an isoscattering point around $q^2 \approx 1.75 \times 10^{-3}$ Å$^{-2}$. The linear regression lines in Figure S7 nicely show how the slope in this $q^2$ region smoothly evolves.

Additional References

(53) ASTRA® 6 manual, WTC, CA, USA.
Figure S6. Cross-section Guinier plots comparison of experimental and model SAXS curves. Open squares, two SAXS data frames taken at the beginning (black) and at the end (red) of a FG (0.456 mg/mL) polymerization process induced by Ancrod 0.05 NIH units/mg FG in TBS. Solid lines, model curves, arbitrarily distributed along the y axis for clarity. In each calculation, the FG monomer was described as a linear ensemble of 10 touching beads of $d = 4.6$ nm. Green, FG monomer; magenta, average of 20 conformations for a YL 20-mer; orange, average of 20 conformations for a YL→DS 20-mer; blue, a RLDS 20-mer; grey, average of 30 conformations of a 300-mer (see Figure 4B). The three dashed vertical lines mark the $q^2$ ranges (horizontal arrows) used to extract the $<R_c^2>_{z}$ values from experimental data; the range was automatically adjusted as the $<R_c^2>_{z}$ increased to force $q_{\text{max}}^2 < R_c^2 >_{z} \leq 1$. 
**Figure S7** – Cross section Guinier analysis of a SAXS dataset during a FG (0.456 mg mL$^{-1}$) polymerization induced by 0.05 NIH units Ancrod/mg FG in TBS. Open circles, one every five frames are shown in a rainbow-color code (from red to purple, see side panel; error bars omitted for clarity). The solid lines are the corresponding linear regression analyses. The starting $q^2$ range was $q^2_{\text{min}} = 0.001 - q^2_{\text{max}} = 0.0025$ Å$^{-2}$, restricted to lower $q^2_{\text{max}}$ values as the reaction proceeds to obey the relation $q^2_{\text{max}} < R_c^2 \leq 1$. 
Figure S8 – Bifunctional polycondensation polymer distributions as a function of $Q$ and simulation time. Panel A, polymer distributions at three different $Q = 1, 18, \text{ and } 50$ values for a common very long simulation time. Panel B, polymer distributions for $Q = 90$ at various simulation times giving rise to $<M>_w$ values comparable with those of Figure 3, main text (here the times were an order of magnitude shorter than in panel A). Panel C, $z$-fraction [ $z(i) = n(i)w^2(i)[\bar{R}_g^2(i)]^{1/2} / \sum n(i)w^2(i)$ ] distributions of the root mean square radius of gyration $[\bar{R}_g^2(i)]^{1/2}$ calculated as a function of the number of monomers in a polymer $(i)$ (see top $x$-axis) for $Q = 90$ at the same simulations times and $<M>_w$ values of the Panel B datasets (note that the integral of each distribution curve corresponds to the $<R_g^2>_z$ value at that particular polymerization degree).
**YL→DS models generation and parameters computation**

A Y-Ladder to Double Strand (YL→DS) polymer is a branched polymer made of linear YL and DS chains of different lengths joined at some nodal points whose branching order is either 3 or 4. Depending on monomer position inside the polymer, the local chain conformation can be YL or DS, as described below. The generation of a YL→DS polymer made of $N$ monomers requires the use of three algorithms that define how to: (a) generate a YL linear chain; (b) generate a YL branched polymer made of linear YL chains; and (c) induce the polymer transition from the YL to the DS conformation.

![Figure S9 - Parameters setting boxes for the fibrin polymers generation LabView® program.](image)

All three algorithms have been implemented under LabView® (2010, [http://www.ni.com/labview/](http://www.ni.com/labview/)), constituting a single program whose operating parameters are assigned in the various “setting boxes” (red labels) shown in Figure S9. In particular, the physical characteristics of the monomer (some of
which apply also to end-to-end single stranded polymers, if they are generated) and final DS polymers are set in the \textit{Monomer/fibril input parameters} box: monomer length (“L0”, nm), unhydrated and hydrated diameters (nm), density $\rho_{\text{SS}}$ (“$\rho_{\text{SS}}$, g/cm$^3$”) and Kuhn length $\ell_{k,\text{SS}}$ (nm); DS density $\rho_{\text{DS}}$ (“$\rho_{\text{DS}}$, g/cm$^3$”) and Kuhn length $\ell_{k,\text{DS}}$ (nm). The separation of the unhydrated and hydrated diameters is implemented because the former is used to compute the $R_g$ of the cylinder as seen by WA-MALS, and the latter is the parameter derived from the SAXS data. Thus, the HPLC-SAXS-derived data on monomeric FG were used to determine the hydrated diameter (see main text), and hence, combined with the length, the volume of the hydrated cylinders. After calculating the volume occupied by the theoretically bound hydration water molecules (see Materials and Methods, \textit{Basic Modeling}), an anhydrous cylinder volume can be derived, and therefore for the same length an unhydrated diameter is defined. As reported under the yellow-labeled \textit{Monomer/fibril output parameters} box, the combination of L0, unhydrated diameter, and $\rho_{\text{SS}}$ determines the molecular weight “M0”, the overall radius of gyration (“$R_g$, nm”), mass length ratio “M/L$_0$”, and hydrodynamic radius (“$R_h$, nm”) of the monomer, while its cross-section radius of gyration (“$R_{c^2}$, nm$^2$”) is computed from the hydrated diameter. For the DS polymers, the center-center separation between strands (“$h = \text{cent.dist_DS}$”, nm) and their diameter (diam$_{\text{DS}}$, nm,) are controlled from the unhydrated diameter by varying $\rho_{\text{DS}}$, while their cross-section square radius of gyration (“$r_{c\text{_DS}^2}$, nm$^2$”) is computed using the hydrated monomers diameter.

Figure S10 - YL chain generation scheme representing a dimer (a-1), trimer (a-2), and a chain (a-3), corresponding to the three steps of the algorithm described in the text.
(a) - YL linear chain generation algorithm

A YL linear chain is a linear polymer where the monomers are linked together at single center-to-end binding sites, so that the chain conformation resembles that of a ladder with Y-shaped steps (see Figure S10). In a YL chain, each monomer (a cylinder of length $L$ and diameter $d$) is treated as an oriented (tail to tip) vector of length $L$ and orientation defined by two angles $\theta$ and $\phi$. The chain is generated by adding one monomer at a time according to the following algorithm:

(a-1) the first vector (#1) is oriented at random in a 3D space and is linked to the second vector #2 by creating a bond between its middle point and the tail of vector #2. The tail of vector #2 can be shifted perpendicularly (to the vector) by a distance $h$ so that $h \geq d$. This ensures that in the DS configuration (see below), it is possible to generate strands with a certain degree of separation (the real linkages between the two molecules cannot be accounted for with this simple geometrical representation). The bonding angle $\theta$ between the two vectors is drawn randomly from a Gaussian distribution with average and standard deviation $<\theta>$ and $\sigma_\theta$, defined in the YL angles box of Figure S9. The angle $\theta$ can assume values from $0^\circ$ (double-strand configuration) to $180^\circ$ (negative-double strand configuration). The positive versus of the angle $\theta$ is defined according to the vector product $\mathbf{2} \times \mathbf{1}$ [see Figure S10-(a-1)].

(a-2) the dimer built at point (a-1) has two binding sites, namely one end-site corresponding to the tail of vector #1, and one center-site corresponding to the middle point of vector #2. Therefore, vector #3 is joined (randomly) to one of these two sites, either binding its tail to the center-site or its middle point to the end-site. The orientation of the third vector is defined by the angle $\theta$ formed with the vector to which is linked and by the azimuthal angle $\phi$ between the planes defined by the vectors 1-2 and 2-3 (or 3-1 and 1-2). The angle $\phi$, which is drawn at random from a Gaussian distribution with average and standard deviation $<\phi>$ and $\sigma_\phi$, (see YL angles box of Figure S9), can range from $-180^\circ$ to $+180^\circ$ and defines the spatial twisting along the chain [see Figure S10-(a-2)].

(a-3) generalizing, if $k$ vectors have already been linked sequentially to form a chain of $k$ monomers, the next $k+1$ vector is added either by binding its tail to the center-site of the chain or by binding its middle point to the end-site of the chain. The orientation of the $k+1$ vector is
defined by the binding angle $\theta$ between the $k+1$ vector and the vector to which is bound (called $p$ for simplicity), and by the azimuthal angle $\varphi$ between the planes defined by the couples of vectors $(k+1)-p$ and the vectors $p-p'$, where $p'$ is the vector bound to $p$ [see Figure S10-(a-3)].

Note that the definition of the positive versus of the bonding angle $\theta$ (see point a-1) guarantees the correct reciprocal orientation between the binding sites on two adjacent vectors, so that, when $\theta \to 0$, a proper DS configuration is obtained.

Figure S11 - YL branched polymer generation scheme representing: (b-2), the incipient linear chain of length $d_1$ between its first two branching points, represented as light blue dots; (b-5), a growing polymer made of two linear YL stretches of lengths $d_1$ and $d_2$ between 3 branching points; (b-6), a general growing YL polymer containing single a double branching points. Labels (b-2), (b-5), and (b-6) refer to the corresponding steps of the algorithm described in the text. The dotted segments represent the two potential binding monomers at each branching point, the black one being aligned along the original chain, and the red one being the actual branched monomer.
(b) - YL branched polymer generation algorithm.

The generation of a YL polymer made of \( N \) monomers forming linear YL chains that can branch, is carried out by using the following algorithm.

(b-1) A set of \( M \) (\( M > N \)) integer numbers representing the distances \( d_j \), (expressed in monomer units (m.u.), \( j = 1, 2, \ldots, M \)) of the branching points inside the polymer is generated at random according to a Gaussian distribution with average and standard deviation \( <d> \) and \( \sigma_d \) (see Branching probabilities box in Figure S9).

(b-2) The first linear chain containing \( d_1 \) monomers is generated as described above throughout points (a-1/a-3). The two end-monomers of the chain are enabled to be sources of possible branching points [see Figure S11, panel (b-2)].

(b-3) The polymer is let to grow by adding a new monomer at one of the four available binding sites. Here we assume that at the branching location both FpAs have been released from that particular unit in the YL configuration, enabling a second monomer to bind with its end to the center of the branching unit. The site is chosen at random (between the four) and the added monomer is aligned either along the original chain (as described at point (a-2)), or along a new direction defined by the branching angles \( \theta_b \) and \( \phi_b \). The latter are chosen at random according to Gaussian distributions characterized by averages and standard deviations \( <\theta_b>, \sigma_{\theta_b}, <\phi_b> \) and \( \sigma_{\phi_b} \) (see Branching angles box in Figure S9).

(b-4) Point (b-3) is iteratively repeated until the longest of the four dangling ends of the polymer attains a length equal to \( d_2 \). The corresponding end-monomer is enabled for branching.

(b-5) The polymer is now characterized by 3 branching points, 2 linear stretches of fixed lengths, and can grow along 5 dangling ends, until the longest one attains a length equal to \( d_3 \), after which the procedure is repeated.

(b-6) Generalizing, a polymer with \( p \) branching points is characterized by \( p-1 \) linear stretches and is let to grow along \( p+2 \) dangling ends, until the longest dangling end attains a length equal to \( d_p \).

(b-7) Step (b-6) is repeated until the final polymer is made of \( N \) monomers, \( i.e. \) when \( \sum_{j=1}^{p-1} d_j + \sum_{j=1}^{p+2} h_j = N \), where \( h_j \) is the length of the \( j \)-th dangling end.
It should be mentioned that it is possible to define a double branching probability \((d/s)\) as well. This parameter sets the probability that two secondary chain start at the same branching position and are added to the collection of the other terminal points. Finally, it is possible to tune the ratio of the probability that a monomer is added to one of the secondary branched chains over the probability that the monomer is added to the primary main chain by a growth probability ratio \((b/m)\). All these parameters can be set in the Branching probabilities box shown in Figure S9.

\((c)\) Algorithm for tuning the YL\(\rightarrow\)DS transition

Once the overall conformation of a YL branched polymer has been obtained, the YL\(\rightarrow\)DS transition of the binding sites of all its monomers is controlled by two independent probabilities. For each monomer, we assign:

\((c-1)\) the probability \(P_1\) that a monomer located at a distance \(x\) from the nearest terminal point of the structure does not undergo a YL\(\rightarrow\)DS transition as

\[
P_1(x) = 0.5 \left\{1 - \tanh\left(\frac{(x-t_1)}{w_1}\right)\right\} \quad \text{(ES5)}
\]

where “tanh” is the hyperbolic tangent, and \(t_1\) and \(w_1\) are two parameters that control the transition location and the width of the sigmoid-like YL\(\rightarrow\)DS transition.

\((c-2)\) the probability \(P_2\) that a monomer belonging to a polymer with a total number \(y\) of branching points does not undergo a YL\(\rightarrow\)DS transition as

\[
P_2(y) = 0.5 \left\{1 - \tanh\left(\frac{(y-t_2)}{w_2}\right)\right\} \quad \text{(ES6)}
\]

where, as in (ES5), \(t_2\) and \(w_2\) control the transition decay and its width. This correction was found to be necessary when many branches are added, whose ends will otherwise add a too large number of YL segments, making difficult to match the experimental \(<R_z^2>_z\) vs. \(<M>_w\) data.
Therefore, the probability that a monomer does not undergo a YL→DS transition is $P_1 \times P_2$, or, equivalently, the probability $P_{\theta=0}$ that a monomer does undergo a YL→DS transition is

$$P_{\theta=0} = 1 - P_1 \times P_2 \quad (ES7)$$

All the parameters appearing in (ES5) and (ES6) can be set in the \textit{YL-DS conversion} box of Figure S9.

Once the skeleton of the structure is completely built, its squared radius of gyration is computed as $R_g^2 = \Sigma r_{0i}^2/N + r_g^2$, where $r_{0i}$ are the distances from the global center of mass of the centers of mass of each cylinder in the chain, with $i = 1, \ldots, N$, and $r_g$ is the radius of gyration of a single cylinder given by $r_g^2 = L^2/12 + d_u^2/8$, where $L$ is the length of the cylinder and $d_u$ its unhydrated diameter.

The squared cross-section radius of gyration ($R_c^2$) is also computed. For a single cylinder of hydrated diameter $d_h$ this is $R_c^2 = (d_h/2)^2/2$. For a DS chain the squared radius of gyration of the cross-section is $R_{c(DS)}^2 = R_c^2 + (d_{DS}/2)^2$, where $d_{DS}$ is the distance between the long axes of the cylinders. In the general case of YL→DS polymers generated by the above defined algorithms, there are regions with double-strand configuration ($\theta = 0$) and regions with single strand (YL) configurations. Since we found out that in the scattering vector $q$ range where we analyze the SAXS data only the monomers’ $R_c^2$ are recovered for the YL polymers (see \textit{YL models definition and scattering properties} in this Supporting Information), a weight average is made between the two analytic values of $R_c^2$ for the mixed YL/DS polymers.\textsuperscript{33}

While a pure DS configuration without branching for each polymer made of $N$ monomers is precisely defined, the corresponding YL and YL→DS polymers without or with branching can be generated in many conformations, whose numbers grows as a function of $N$. To attain a satisfactory statistical representation for each kind of polymer made of $N$ monomers, the program will generate a number of conformations proportional to the number $N$ of monomers in the polymer, and average properties are then computed. However, to avoid very long computational times, only a subset of polymers is generated, and the properties of the intermediate polymers are then interpolated. All these parameters can be set in the \textit{Configuration settings} in Figure S9. After a generation run is completed, the program will save two matrices containing the average $<R_g^2>$ and $<R_c^2>$ values for each kind of polymer. These
matrices must be regenerated each time one of the monomer/DS physical parameters or one the configuration parameters are changed. Polymers configurations can also be saved as initial and final coordinates of each monomer, or as a corresponding set of touching beads, for visualization or further computations. These operations are set in a “Y-ladder save structure” box (not shown).

Finally, the program can generate time-dependent polymers distributions using the Flory-Janmey bifunctional polycondensation scheme controlled by the Janmey parameters box, where “F0” is the monomers starting molar concentration, “Thr” is the enzyme concentration in NIH units/ml, “K2” is the enzyme catalytic constant (sec\(^{-1}\) ml/NIH units), “Km” is the enzyme Michaelis-Menten constant (M), and “Q” is the ratio of release between the first and the second FpA from the same FG molecule (see Figure S9). The length and time-step of the simulation can be controlled from a Numerical integration box (not shown). See the Materials and Methods Basic modeling in the main text for more details and references to the functionality and implementation of the Flory-Janmey bifunctional polycondensation scheme. It is important to note that the Janmey parameters (in practice only the Q value) can be changed and the simulation re-run without to have to re-generate the polymer properties matrices.

Once a polymer distribution has been generated, the \(z\)-averages \(<R_g^2>_z\) and \(<R_c^2>_z\), and the \(w\)-average \(<M>_w\) can be computed at each time step, and saved in a file (controlled under a “save” box; not shown). The program then allows comparing the synthetic \(<R_g^2>_z\) vs. \(<M>_w\) and \(<R_c^2>_z\) vs. \(<M>_w\) curves for each type of polymerization model generated (e.g. RLDS, WLDS, YL→DS) with experimental data (pre-loaded from an “exp. data” form; not shown), in a “Rg, Rc vs. Mw plots” screen (not shown). SD weighted or un-weighted residuals are also point-wise computed and can be saved in separate files.