Electronic Supplementary Information (ESI)

Mejias et al. "Controlled nanometric fibers of self-assembled designed protein scaffolds"

Methods:

Protein design and purification

Consensus CTPR20 protein without solvating helix was generated by the introduction of a stop codon after the last consensus repeat of CTPR20 gene cloned into pPro-EX-HTb.[1] Two single Cys residues at each end of the protein were introduced then by quick change site directed mutagenesis. Cys-CTPR20-Cys protein was expressed as his-tagged fusion and purified using standard affinity chromatography methods as previously described.[1, 2] The his-tag was cleaved by TEV protease digestion and the His-tag and the tagged protease removed by a second Ni-NTA column. The protein sample was dialyzed into PBS (150 mM NaCl, 50 mM phosphate buffer pH 7.4) and stored frozen at -20°C. The protein concentration was determined by *uv* absorbance at 280 nm using the extinction coefficient calculated from the amino acid composition.[3]

Protein polymerization in solution by DLS

Purified Cys-CTPR20-Cys samples at a protein concentration between 40 -100 μM were freshly reduced in the presence of 5 mM dithiothreitol (DTT) during 20 minutes at room temperature. DTT was then removed from the protein sample using a NAP-5 (GE Healthcare Life Science) size exclusion column. Protein fractions without DTT were collected at 4ºC and the protein concentration determined. The samples were filtered through a 0.22 μm size pore to remove any potential large particles that would

interfere in the DLS measurements. The polymerization is initiated immediately by placing the protein in the DLS instrument at constant temperature (25ºC or 42ºC).

Dynamic light scattering measurements

To monitor the increase in size upon protein polymerization by DLS was used a Zetasizer NanoZS instrument (Malvern Instruments, Malvern, UK), which measures scattering in small volumes at low concentration (from 0.1mg/mL) and sizes from 0.3nm (diameter). An incident light wavelength of 532 nm is used and the scattered light is collected at a fix angle of 173º, which is optimal for low sample concentration. The distributions of the hydrodynamic sizes in the sample were collected periodically from time 0 at the different starting protein concentrations and polymerization temperatures. Three independent measurements were taken at each data point to calculate the error.

Polymerization kinetics and thermodynamics

Protein polymerization kinetics as given by equation 1 in the main text is expected to follow a linear dependence with time.[4] Polymerization rates can be determined by applying linear fit to the initial reaction times. When monitored by DLS, polymerization rates can be estimated by a fit to the following equation at short reaction times:

$$D = k*t + D_0 \quad \text{(S1)}$$

where D is the average hydrodynamic diameter of the polymeric sample in nm; k is the rate of polymerization; t is the time in minutes and D_0 is the hydrodynamic diameter at time 0.

The polymerization speed dependence on the protein concentration is fitted to the following equation, which assumes Michaelis-Menten-like kinetics:

$$V = \frac{V_{\max}[P]}{K + [P]} \tag{S2}$$

where V is the velocity of the polymerization, V_{max} is the maximum velocity, [P] is the protein concentration and K is the concentration at which the polymerization velocity is half of the V_{max}.

The Arrhenius equation for two reaction rates (k_1 and k_2) at two temperatures (T_1 and T_2) is used to estimate the activation energy (E_a) for the CTPR polymerization process:

$$\frac{k_2}{k_1} = \exp\left[\frac{-E_a}{R}\left(\frac{1}{T_2} - \frac{1}{T_1}\right)\right] \tag{S3}$$

The equilibrium constant can be related with the Gibbs free energy change as follows:

$$\Delta G = -RT \ln (K_D) \tag{S4}$$

where R is the ideal gas constant (1.98 10^{-3} kcal $K^{-1}mol^{-1}$), T is the temperature in Kelvin and K_D the dissociation constant.

Hydrodynamic size calculations of CTPR proteins

The experimental diffusion coefficients, and therefore, hydrodynamic sizes of CTPR proteins with different number of repeats (2-20) were previously calculated by fluorescence correlation spectroscopy (FCS).[5] The theoretical diffusion coefficients and the hydrodynamic radius for CTPR proteins of different lengths were calculated from the x-ray crystal structure coordinates[1, 6] using the program Hydropro version

7c19 (http://Leonardo.fcu.um.es/macromol/programs/hydropro/hydropro.htm).[7] The correlation between number of repeats and the hydrodynamic size can be fitted to the equation $R_h=R_0 N^v$, where R_h is the hydrodynamic radius and N the number of repeats per protein.

Transmission electron microscopy (TEM)

Monomeric or polymerized CTPR protein samples at 10 μM protein concentration were deposited on glow discharged Cu/Rh grids coated with carbon, and negatively stained with 2% uranyl acetate. Micrographs were recorded using Kodak SO-163 film, in a JEOL JEM1200EXII electron microscope with a tungsten filament operated at 100 kV and at 60 K magnifications. The particle sizes were the average size of 5 molecules measured using ImageJ program in order to obtain SD.

Quantification of free thiols

For the quantification of the number of cysteins that have not formed a disulfide bond was used Ellman's reagent (5,5'-dithiobis-(2-nitrobenzoic acid) or DTNB).[8] At different polymerization times protein samples were taken to measure the concentration of free thiol groups. A standard curve using freshly reduced CTPR20-Cys was used for quantification. 10 μl sample were mixed in 0.1 M Tris pH 8.0, 0.1 mM DTNB and 2.5 mM sodium acetate and let it react for 10 minutes at room temperature. NTB^{2-} formation was quantified by measuring the absorbance at 412 nm using an extinction coefficient of 13,600 M^{-1} cm^{-1}.

Step growth polymerization model description

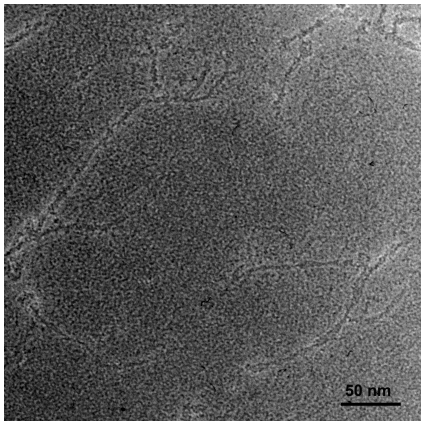
In a step growth model the polymerization process is defined by the average degree of polymerization (X_n), and the fraction of functional groups forming a bond *(p)* at a defined polymerization time.

The relation between p and X_n for an ideal step-growth polymerization is given by equation S5.[4]

$$X_n = 1/(1-p) \quad \text{(S5)}$$

As p grows, X_n becames larger, because more functional groups are bonded within the polymers.

Supporting Figure 1. TEM imaging of protein polymerization. CTPR20 sample after polymerization saturation, where several linear polymeric structures are observed.



References

1 T. Kajander, A. L. Cortajarena, S. G. Mochrie and L. Regan, *Acta Crystallographica*, 2007, **D63**, 800-811.

2 T. Kajander, A. L. Cortajarena and L. Regan, *Methods Mol. Biol.*, 2006, **340**, 151-170.

3 C. N. Pace, F. Vajdos, L. Fee, G. Grimsley and T. Gray, *Protein Sci.*, 1995, **4**, 2411-2423.

4 G. Odian, *Principles of polymerization*, Wiley-Interscience2004.

5 A. L. Cortajarena, G. Lois, E. Sherman, C. S. O'Hern, L. Regan and G. Haran, *J. Mol. Biol.*, 2008, **382**, 203-212.

6 E. R. G. Main, Y. Xiong, M. J. Cocco, L. D'Andrea and L. Regan, *Structure*, 2003, **11**, 497-508.

7 A. Ortega, D. Amoros and J. Garcia de la Torre, *Biophys J.*, 2011, **101**, 892-898.

8 G. L. Ellman, *Arch. Biochem. Biophys.*, 1959, **82**, 70-77.