Supporting Information

Arrested Phase Separation of Elastin-Like Polypeptide Solutions Yields Stiff, Thermoresponsive Gels

Matthew J. Glassman,[a] *and Bradley D. Olsen*[a]*

[a]Department of Chemical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Ave, Room 66-153, Cambridge, MA 02139. Email: bdolsen@mit.edu.

Genetic Engineering and Biosynthesis

Oligonucleotides 75 nt in length coding for the desired repetitive polypeptides were designed containing flanking recognition sequences for a Type IIS restriction enzyme, BsaI (NEB). The directionality of the asymmetric recognition sequences was chosen so that digestion produced fragments were flanked with TTGG overhangs, which are non-palindromic to ensure robust head-to-tail ligation, and so that the enzyme releases fragments that do not contain BsaI sites.. More specifically, the following cassette was inserted between the NcoI and XhoI sites in pET-28b:

A_1 insert:
CCATGGGCGGATCCGCTAGCGGTCTCGTTGGT**ATTCCTGCTGTTGGTGTGC**
CGGCTGTTGGTATCCCAGCTGTTGGCGTTCCGGCTGTAGGTATTCCGGC
TGTTGGTGAGACCACTAGTTAAATGAATAAGCTTTAACTCGAG

A_1':
CCATGGGCGGATCCGCTAGCGGTCTCGTTGGT**GTTCCTGCTGTCGGTGTGC**
CGGCTGTTGGTATTCCAGCTGTTGGCGTGCCGGCTGTAGGTGTCCCGGC
TGTTGGCGAGACCACTAGTTAAATGAATAAGCTTTAACTCGAG

G_1:
CCATGGGCGGATCCGCTAGCGGTCTCGTTGGT**GTACCTGGTGTTGGCGTCC**
CGGGTGTAGGTATCCCAGGCGTTGGTGTACCGGGTGTAGGCGTTCCAGG
CGTTGGCGAGACCACTAGTTAAATGAATAAGCTTTAACTCGAG

The 75 nt sequences that code a set of five pentapeptide repeats for each ELP are highlighted in bold, and the asymmetric BsaI recognition sequences are underlined.

These fragments were ligated at 100-fold molar excess into modified pET-28b vectors (pETA) with the same arrangement of BsaI recognition sequences in the multiple cloning site (MCS). Because concatemerized genes did not have internal BsaI sites, multiple rounds of digestion and concatemerization were possible. Up to three rounds were typically required to produce the gene sizes used in this study, coding for up to 120 pentapeptide repeats.

Following expression and harvesting, cell pellets were resuspended in non-denaturing lysis buffer (MENT buffer: 10 mM Tris, 1 mM EDTA, 100 mM NaCl, 5 mM $MgCl_2$, pH = 7.5) at a concentration of approximately 30 g wet cell mass (WCM) per 100 mL buffer. Resuspensions were frozen at -20°C. After thawing on ice, lysozyme (100 mg per 100 mL resuspension) was added, and after approximately 1 hour the suspension was sonicated. Cell debris was removed by centrifugation, and DNAse I and RNAse A (2 mg each) were added to the clarified supernatant and incubated for 2-3 hours at 37 ^{o}C. The turbid lysates were then centrifuged at 37 ^{o}C, and the pellets were redissolved in MENT buffer at 5 ^{o}C, typically overnight. The protein solutions were thermally cycled between 5°C and 37°C in MENT buffer for two additional cycles. The solutions were then dialyzed against MilliQ water and purified in a final step by passing over anion exchange resin in 6 M urea, 20 mM Tris, pH = 8.0, using HiTrap Q pre-packed columns (GE Healthcare) by automated chromotography. Bound contaminants were discarded, and the target proteins were collected in the flow-through step, dialyzed against MilliQ water, and lyophilized.

Protein Sequences:

$(X^1PGVG)_{50}$: $[(VPGVG)_2(IPGVG)(VPGVG)_2]_{10}$

MGWGSASGLVGVPGVGVPGVGIPGVGVPGVGVPGVGVPGVGVPGVGIPGVGVPGVGV
PGVGVPGVGVPGVGIPGVGVPGVGVPGVGVPGVGVPGVGIPGVGVPGVGVPGVGVPG
VGVPGVGIPGVGVPGVGVPGVGVPGVGVPGVGIPGVGVPGVGVPGVGVPGVGVPGVGI
PGVGVPGVGVPGVGVPGVGVPGVGIPGVGVPGVGVPGVGVPGVGVPGVGIPGVGVPG
VGVPGVGVPGVGVPGVGIPGVGVPGVGVPGVGETTS*

$(X^1PGVG)_{50}$: $[(VPAVG)_2(IPAVG)(VPAVG)_2]_{10}$

MGWGSASGLVGVPAVGVPAVGIPAVGVPAVGVPAVGVPAVGVPAVGIPAVGVPAVGV
PAVGVPAVGVPAVGIPAVGVPAVGVPAVGVPAVGVPAVGIPAVGVPAVGVPAVGVPA
VGVPAVGIPAVGVPAVGVPAVGVPAVGVPAVGIPAVGVPAVGVPAVGVPAVGVPAVGI
PAVGVPAVGVPAVGVPAVGVPAVGIPAVGVPAVGVPAVGVPAVGVPAVGIPAVGVPA
VGVPAVGVPAVGVPAVGIPAVGVPAVGVPAVGETTS*

$(X^1PGVG)_{70}$: $[(VPAVG)_2(IPAVG)(VPAVG)_2]_{14}$

MGWGSASGLVGVPAVGVPAVGIPAVGVPAVGVPAVGVPAVGVPAVGIPAVGVPAVGV
PAVGVPAVGVPAVGIPAVGVPAVGVPAVGVPAVGVPAVGIPAVGVPAVGVPAVGVPA
VGVPAVGIPAVGVPAVGVPAVGVPAVGVPAVGVPAVGIPAVGVPAVGVPAVGVPAVGI
PAVGVPAVGVPAVGVPAVGVPAVGIPAVGVPAVGVPAVGVPAVGVPAVGIPAVGVPA
VGVPAVGVPAVGVPAVGIPAVGVPAVGVPAVGVPAVGVPAVGIPAVGVPAVGVPAVG
VPAVGVPAVGIPAVGVPAVGVPAVGVPAVGVPAVGIPAVGVPAVGVPAVGVPAVGVP
AVGIPAVGVPAVGVPAVGETTS*

$(X^1PGVG)_{120}$: $[(VPAVG)_2(IPAVG)(VPAVG)_2]_{24}$

MGWGSASGLVGVPAVGVPAVGIPAVGVPAVGVPAVGVPAVGVPAVGIPAVGVPAVGV
PAVGVPAVGVPAVGIPAVGVPAVGVPAVGVPAVGVPAVGIPAVGVPAVGVPAVGVPA
VGVPAVGIPAVGVPAVGVPAVGVPAVGVPAVGVPAVGIPAVGVPAVGVPAVGVPAVGI
PAVGVPAVGVPAVGVPAVGVPAVGIPAVGVPAVGVPAVGVPAVGVPAVGIPAVGVPA
VGVPAVGVPAVGVPAVGIPAVGVPAVGVPAVGVPAVGVPAVGIPAVGVPAVGVPAVG
VPAVGVPAVGIPAVGVPAVGVPAVGVPAVGVPAVGIPAVGVPAVGVPAVGVPAVGVP
AVGIPAVGVPAVGVPAVGVPAVGVPAVGIPAVGVPAVGVPAVGVPAVGVPAVGIPAVG
VPAVGVPAVGVPAVGVPAVGIPAVGVPAVGVPAVGVPAVGVPAVGIPAVGVPAVGVP
AVGVPAVGVPAVGIPAVGVPAVGVPAVGVPAVGIPAVGVPAVGVPAVGVPAVGIPAV
GVPAVGIPAVGVPAVGVPAVGVPAVGIPAVGVPAVGVPAVGVPAVGVPAVGIPAVGIP
AVGVPAVGVPAVGVPAVGVPAVGIPAVGVPAVGVPAVGETTS*

$(X^2PGVG)_{50}$: $[(IPAVGVPAVG)_2(IPAVG)]_{10}$

MGWGSASGLVGIPAVGVPAVGIPAVGVPAVGIPAVGIPAVGVPAVGIPAVGVPAVGIPA
VGIPAVGVPAVGIPAVGVPAVGIPAVGIPAVGVPAVGIPAVGVPAVGIPAVGIPAVGVPA
VGIPAVGVPAVGIPAVGIPAVGVPAVGIPAVGVPAVGIPAVGIPAVGVPAVGIPAVGVPA
VGIPAVGIPAVGVPAVGIPAVGVPAVGIPAVGIPAVGVPAVGIPAVGVPAVGIPAVGIPAV
GVPAVGIPAVGVPAVGIPAVGETTS*

Table S1: MALDI-TOF of elastin-like polypeptides investigated in this study.

Identifier	MALDI-TOF (kDa)	Theoretical (kDa)
$(X^1PGVG)_{50}$	21.8	22.1
$(X^1PAVG)_{50}$	22.6	22.8
$(X^1PAVG)_{70}$	31.2	31.3
$(X^1PAVG)_{120}$	54.1	52.6
$(X^2PAVG)_{50}$	22.9	23.0

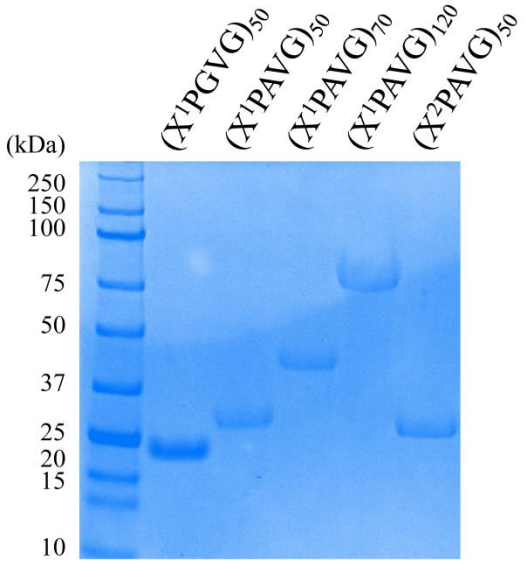

Figure S1: SDS-PAGE of elastin-like polypeptides investigated in this study.

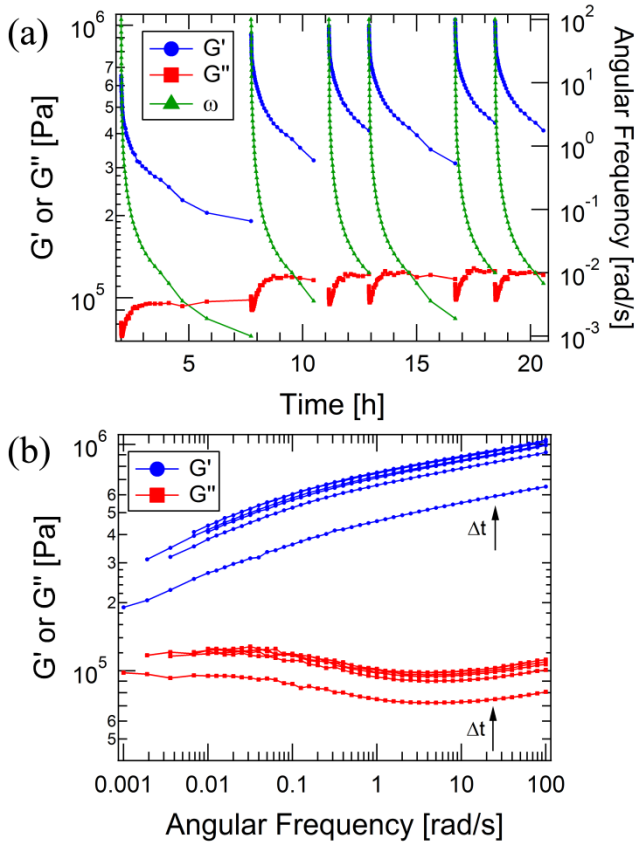

Figure S2: (a) Frequency sweep linear viscoelasticity data for $(X^1PAVG)_{120}$ at 20 wt% in water plotted as a function of (a) equilibration time at 37°C, and (b) frequency.

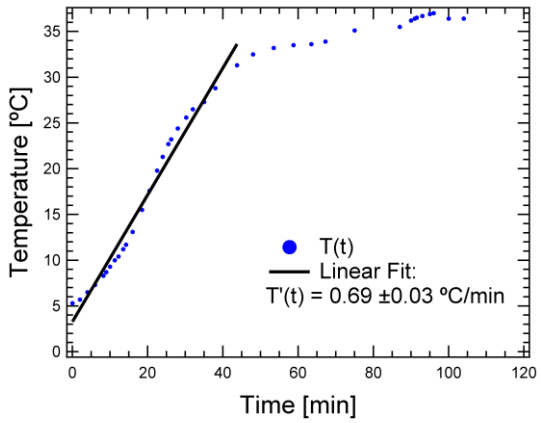

Figure S3: Temperature profile of the SANS sample holder during the heating step (heat source was a circulating water bath). The heating rate was approximately 0.7°C/min through 30°C. Measurements for the data analyzed in the body of the manuscript were taken at t = 150 min, an effective equilibration of 60 min at T = 37°C ± 0.5°C.

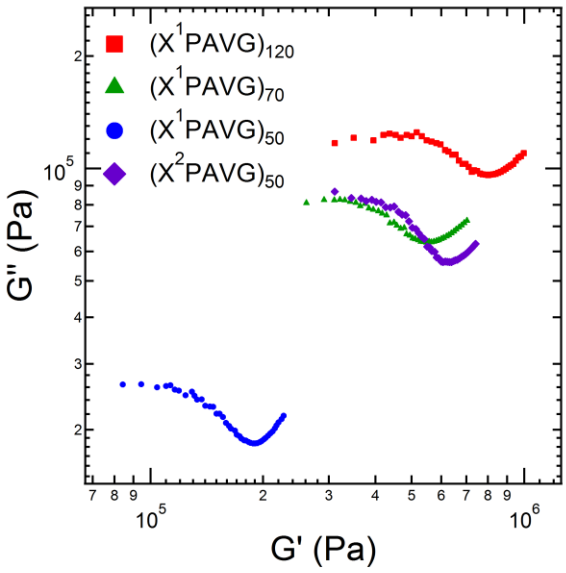

Figure S4: Cole-cole plots of linear viscoelasticity data for gels at 20 wt%. A simple Maxwell model, which would be represented by a semicircle on this plot, fails to describe the relaxation time distribution for these networks.

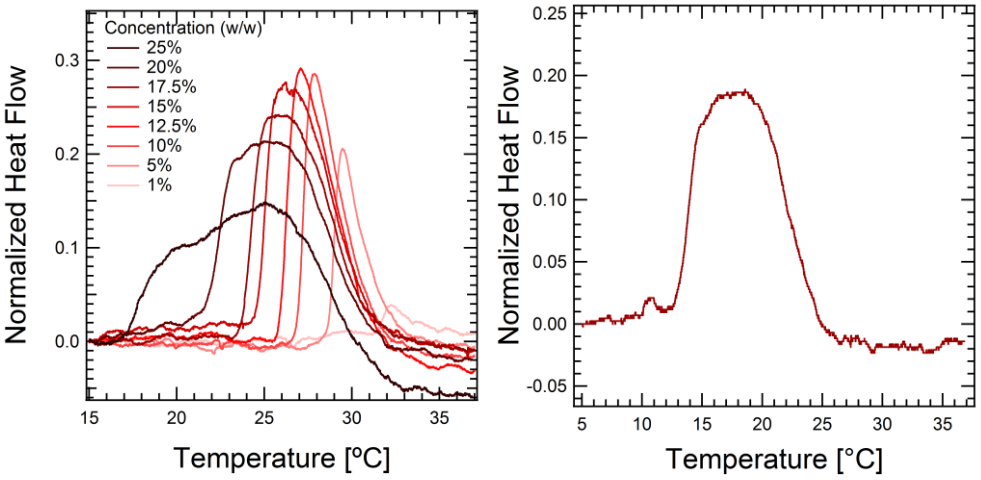

Figure S5: (a) DSC traces of solutions of $(X^1PAVG)_{50}$ on heating (dT/dt = 1°C/min) showing the effect of concentration on the onset and peak temperatures of the transition. (b) DSC trace of a 20 wt% solution of $(X^1PAVG)_{50}$ on heating (dT/dt = 1°C/min).

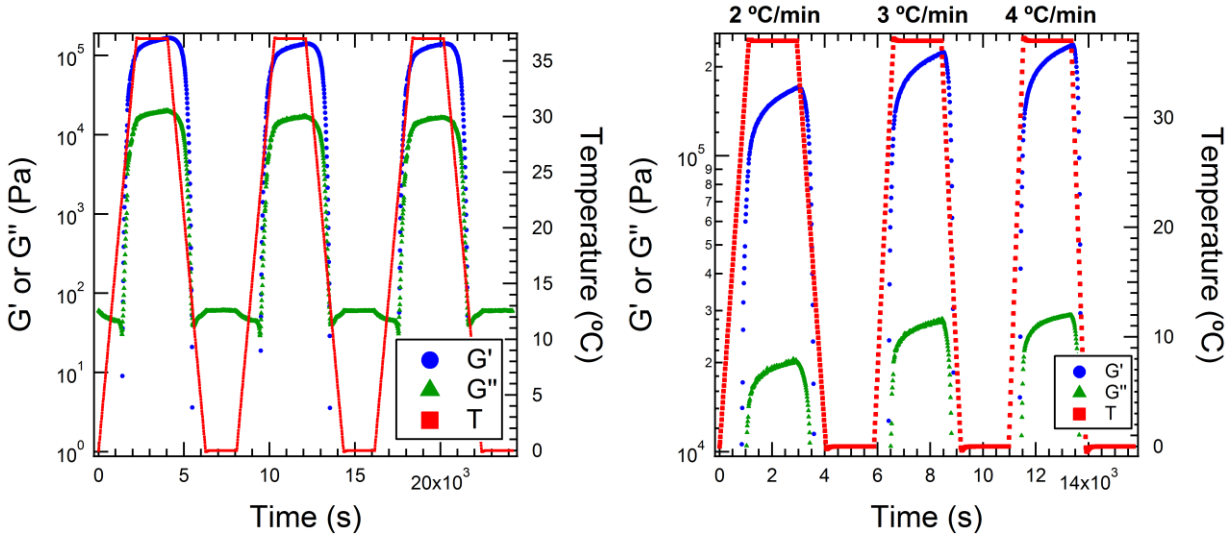

Figure S6: (a) Thermal cycling of $(X^1PAVG)_{50}$ at c = 20% (w/w) at a heating rate was 1°C/min. Measurements were taken at ω = 100 rad/s, γ = 0.01. Below the transition temperature in this sample (c.a. 20°C), $\tan(\delta) > 10^4$, and the G' data are out of range of the plot. (b) Effect of heating rate on the gelation of $(X^1PAVG)_{50}$ at 20 wt%. Moduli at the end of the heating step are 0.17, 0.22, 0.24 MPa for ramp rates of 2, 3, and 4°C/min, versus 0.15 ± 0.01 MPa at a heating rate of 1°C/min. Note that these data are generated on an independent sample from that shown in Figure 2c.

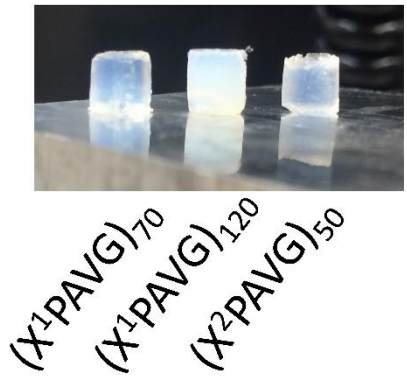

Figure S7: Additional images of 20 wt% gels.

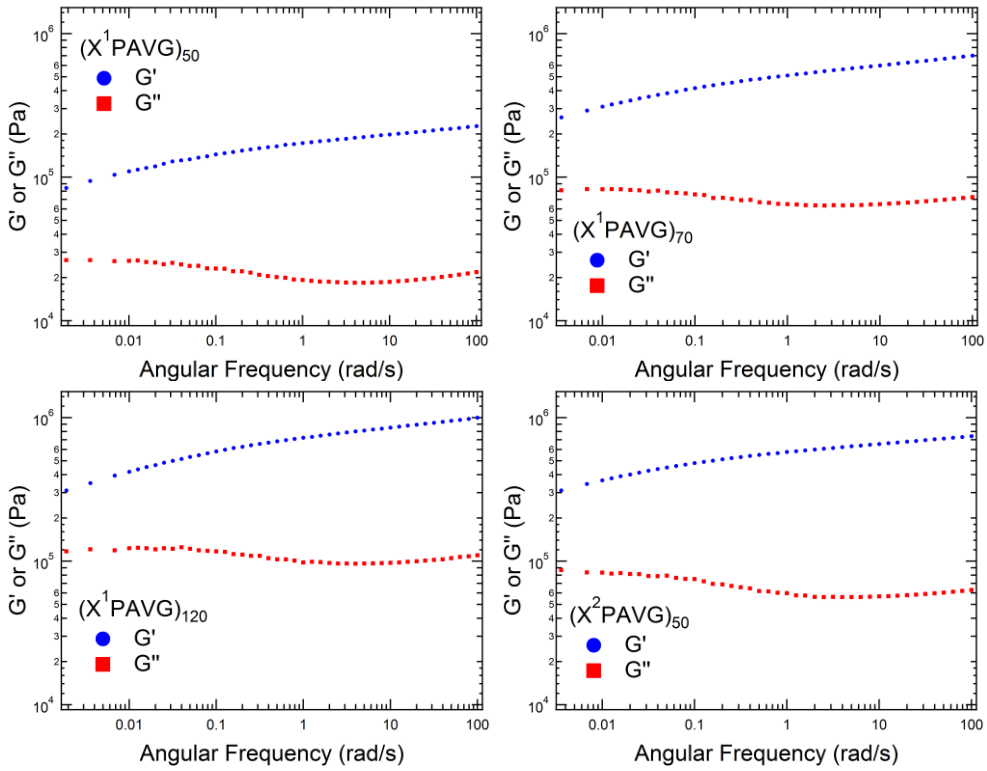

Figure S8: Representatitve frequency sweeps for 20 wt% gels at T = 37ºC, γ_o = 0.01.

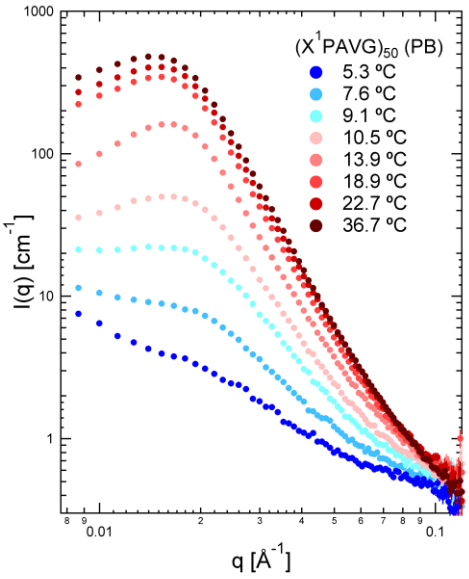

Figure S9: Small angle neutron scattering showing the thermal evolution of the scattering pattern on one $(X^1PAVG)_{50}$ in 100 mM sodium phosphate, pH= 7.6, at 20% (w/v). Data collected at the temperature extremes (T = 5.3°C and 36.7°C) were from samples equilibrated for at least 60 minutes. Data at intermediate temperatures were collected over 2 min during the

temperature ramp step without equilibration, with acquisitions starting at the temperature indicated in the legend.

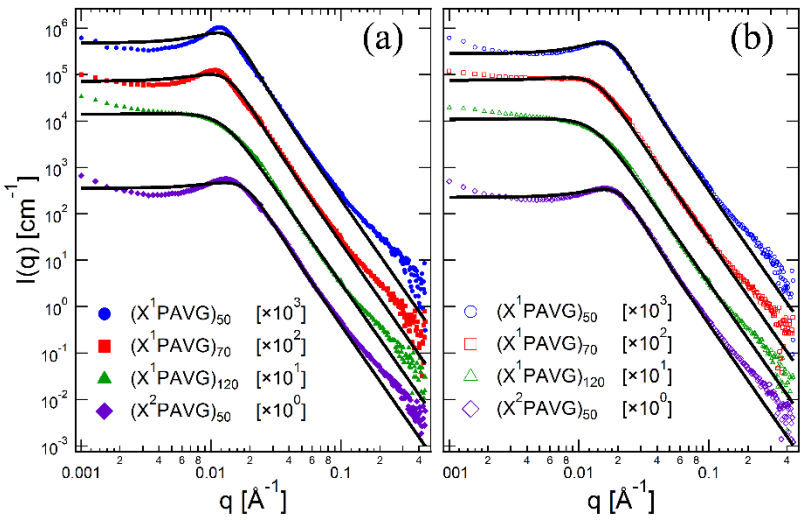

Figure S10: SANS intensity distributions and fits to the Teubner-Strey model (solid black lines; see below for model) for gels prepared in (a) D_2O or (b) 100 mM sodium phosphate buffer, pH = 7.6. For clarity, the data are shifted by the multiplicative factor indicated in the figure legend.

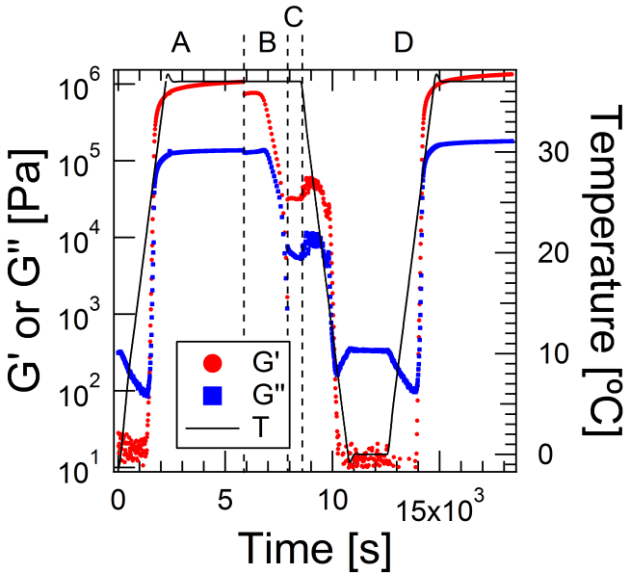

Figure S11: Recovery following large amplitude oscillatory shear rheology of $(X^1PAVG)_{120}$ at a concentration of 20 wt% in water. Region A: The sample is heated to 37°C and equilibrated, and SAOS is applied at ω = 100 rad/s and γ_o = 0.01. Region B: LAOS is applied at ω = 1 rad/s over a series of strain amplitudes from γ_o = 0.001 to γ_o = 2.5. Region C: SAOS is applied at ω = 100 rad/s and γ_o = 0.01 immediately, demonstrating the gel does not recover to the state prior to LAOS. Region D: After cooling, equilibrating at 0°C, then reheating and equilibrating at 37°C, stiff gels are reformed.

Large Amplitude Oscillatory Shear Rheology

The range of valid strain amplitudes for the LAOS measurements was assessed by examining the spectral purity of the input waveforms by Fourier transform analysis. While the third and fifth harmonic ratios of the strain waveform are less than 3%, a window considered valid for LAOStress,[2] apparent high frequency stress oscillations occur in elastic Lissajous-Bowditch plots at high strain amplitude and suggest poor control. Because these oscillations were observed to occur where the strain is small and shear rate is high, the spectral purity of the shear rate waveform ($\frac{d}{dt}\gamma(t) = \gamma_o\omega\cos(\omega t)$) was investigated for each sample.

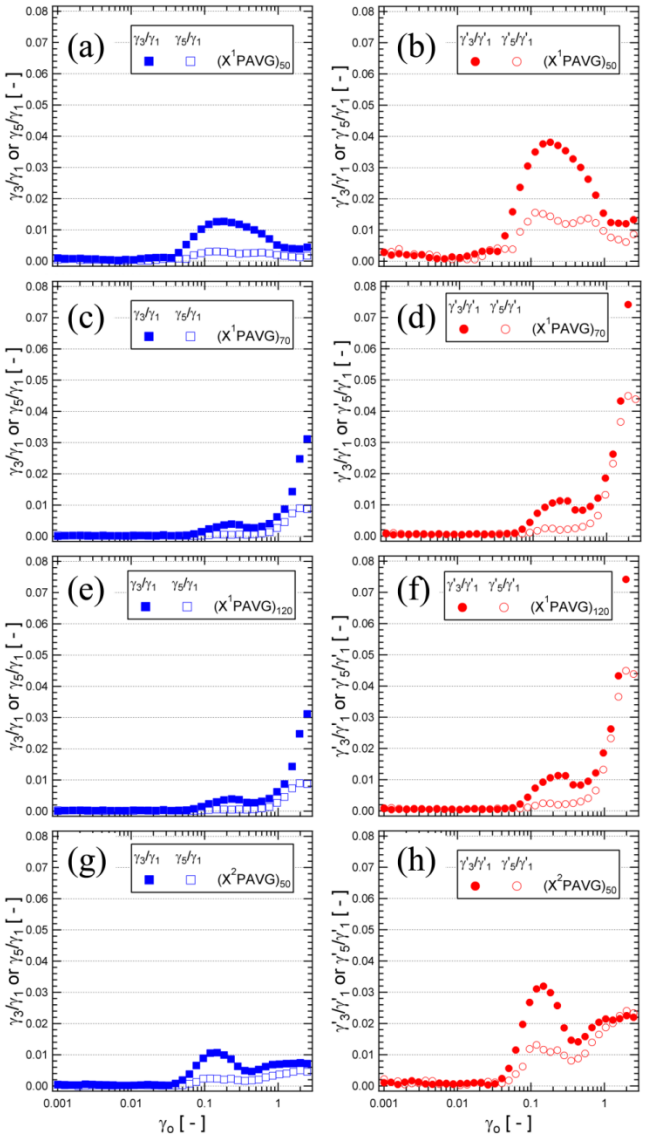

Figure S12: Spectral purity analysis of the strain and shear rate waveforms during LAOS.

Note that for each sample, there is typically an intermediate region of target strain amplitudes where the quality of the strain and shear rate waveforms deteriorate, as judged by the non-monotonic increase in the harmonic ratios. This non-ideality in the strain control does not increase significantly beyond 1% for any sample, but can peak sharply (i.e. in $(X^1PAVG)_{50}$ and $(X^2PAVG)_{50}$) above $\gamma_o = 0.1$. While in a strict sense the shear rate is not the control parameter in LAOS measurements, an ideal sinusoidal shear rate waveform is assumed when computing the first order measures of nonlinearity according to the Ewoldt/McKinley framework. In particular, the computation of G_M' and η_L' from the experimental data are assumed to be performed when the shear rate is maximized (and strain minimized), otherwise these measures do not represent elastic or viscous-like measurements, respectively. Based on the analysis of the spectral purity of the input stress waveforms for LAOStress provided in Dimitriou, *et al*, a conservative cutoff of 1% in the first harmonic ratio of the shear rate waveform was applied to determine data suitable for further analysis in this work.

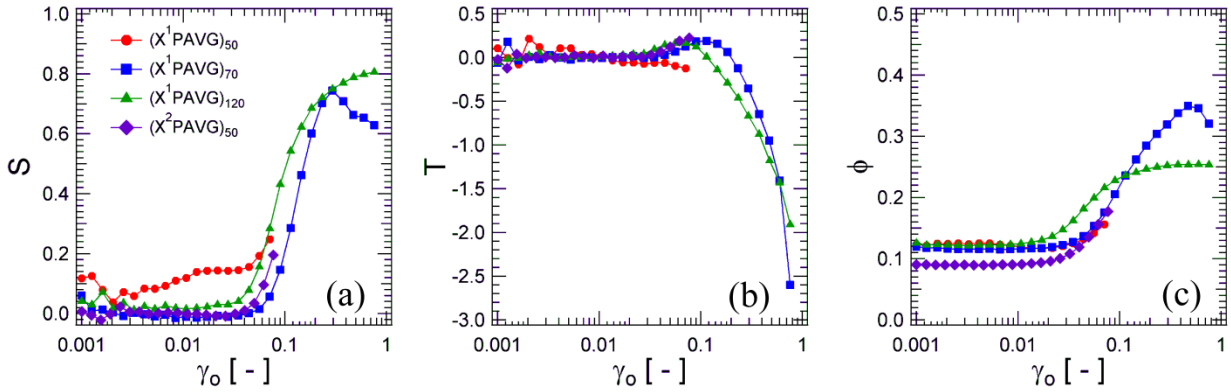

Figure S13: The (a) stiffening ratio, S, (b) thickening ratio, T, and (c) perfect plastic dissipation ratio, ϕ, for all gels during a single strain sweep out to failure.

Small Angle Neutron Scattering

Correction for the Incoherent Scattering Background

Data were corrected for the incoherent scattering background by fitting the high q region to a linear relation of the form:[1]

$$I(q)q^4 \rightarrow A + Bq^4 \tag{S1}$$

The value B was computed for each acquisition and subtracted off as shown below (Figure S14).

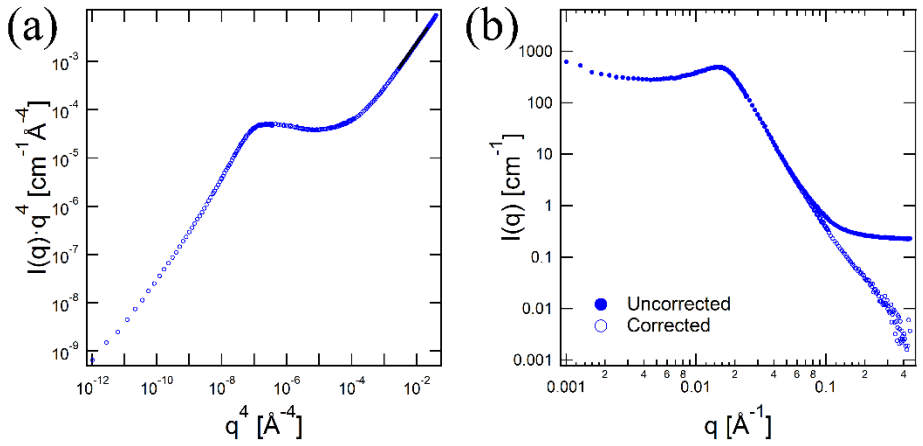

Figure S14: Example of procedure for correcting for the incoherent scattering background. Data are for $(X^1PAVG)_{50}$ in D_2O.

Supplemental Analysis of Scattering Data

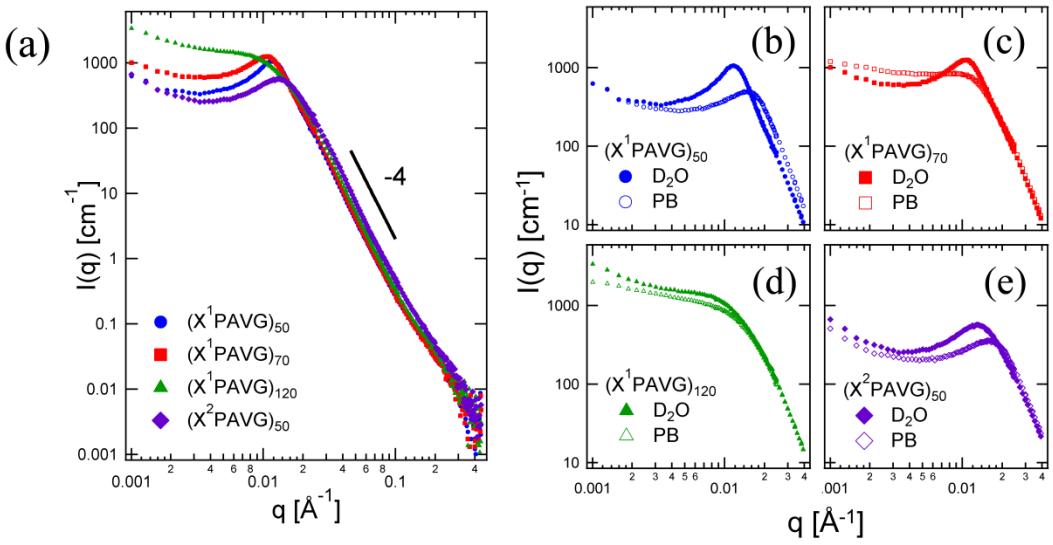

Figure S15: Small angle neutron scattering for 20% (w/v) gels heated at ca. 0.7°C/min. (a) Overlay of scattering patterns for gels prepared in D_2O to highlight the nearly identical scattering in the Porod region. (b-e) Effect of buffer on gel nanostructure (PB = 100 mM sodium phosphate buffer, pH = 7.6).

The Teubner-Strey model was also investigated here. This model was developed initially for the study of thermodynamically-stable biocontinuous microemulsions and has also been successfully applied to the study of phase separating polymer blends. Here, the model was utilized to extract the average domain spacing, d_{TS}, and correlation length, ξ_{TS}, from the scattering patterns for the ELP gels, where the scattering is fit to a function of the form:

$$I(q) - B = \frac{1}{a_2 + c_1 q^2 + c_2 q^4} \quad \text{(S2)}$$

$$\xi_{TS} = \left(\frac{1}{2}\left(\frac{a_2}{c_2}\right)^{1/2} + \frac{1}{4}\frac{c_1}{c_2}\right)^{-1/2} \quad \text{(S3)}$$

$$d_{TS} = 2\pi\left(\frac{1}{2}\left(\frac{a_2}{c_2}\right)^{1/2} - \frac{1}{4}\frac{c_1}{c_2}\right)^{-1/2} \quad \text{(S4)}$$

When $c_1 < 0$ and $a_2, c_2 > 0$, this function typically results in a peak in the scattered intensity, while retaining the q^{-4} behavior at high wave vectors. With this approach, the intradomain correlations and semiperiodic interdomain lengthscale can be estimated without assuming a form of the nanoscale morphology. The Teubner-Strey model assumes a correlation function of the form:

$$\gamma(r) = exp\left(\frac{-r}{\xi_{TS}}\right)\frac{d_{TS}}{2\pi r}\sin\left(\frac{2\pi r}{d_{TS}}\right) \quad \text{(S5)}$$

Here, ξ_{TS} captures the Porod-like inhomogeneity lengthscale, while d_{TS} is the interdomain correlation lengthscale.

While this model contains the minimum number of lengthscales that are evident in the scattered intensity distributions (one for the peak, and one for the Porod q^{-4} regime), the model inadequately fits the scattering data over the entire set of samples. The fits to the gels assembled in phosphate buffer are reasonable, but in D_2O, the model appears incapable of describing both the peak sharpness and the immediate transition to the Porod decay adequately. These poor fits motivate the use of a minimally more complex model, such as the CRW model described in the main text, to describe the assembly of all gels.

Table S2: Fit parameters to the Teubner-Strey model.

Identifier	Solvent*	a_2 [10^{-3}]	c_1	c_2 [10^3]	$B^{\dagger}$
$(X^1PGVG)_{50}$	D_2O	2.063 ± 0.005	-12.71 ± 0.04	49.60 ± 0.06	0.2248
	PB	3.509 ± 0.007	-13.83 ± 0.04	32.75 ± 0.04	0.2271
$(X^1PGVG)_{70}$	D_2O	1.381 ± 0.003	-8.19 ± 0.03	42.44 ± 0.05	0.1916
	PB	1.313 ± 0.003	-4.90 ± 0.03	36.99 ± 0.05	0.2068
$(X^1PGVG)_{120}$	D_2O	0.707 ± 0.002	-1.24 ± 0.02	30.93 ± 0.03	0.2178
	PB	0.898 ± 0.002	-1.14 ± 0.03	30.85 ± 0.04	0.2209
$(X^2PGVG)_{50}$	D_2O	2.759 ± 0.006	-8.28 ± 0.03	25.35 ± 0.03	0.2197
	PB	4.387 ± 0.009	-11.99 ± 0.04	24.70 ± 0.03	0.2414

* PB = 100 mM sodium phosphate in D_2O, pH = 7.6. †Background calculated directly from the scattering data as described above, and was held constant during the fitting procedure.

The Porod invariant, Q, provides a route to estimate the polypeptide concentration in the dense phase, and the results of this analysis are included in the table below (Table S3). First, Q is analyzed below (Figure S16c,d) to assess the suitability of the wavevector range investigated in the SANS measurements. The Porod invariant is related to the total scattering contrast $\langle \rho^2 \rangle$ in the sample volume, and is given by:

$$Q = \frac{1}{2\pi^2} \int_0^\infty dq\, q^2 I(q) = V \langle \rho^2 \rangle \tag{S6}$$

To determine Q precisely, the scattered intensity must be collected over a sufficiently broad range of wavevectors. The data processing steps are explicitly shown for the sample of $(X^1PAVG)_{50}$ in D_2O to evaluate the suitability of the analysis.

At low wavevectors, the invariant plateaus, although by the end of the experimentally-accessed q-range the invariant continues to increase slightly. As discussed above, this is potentially due to excess scattering on smaller lengthscales, but occurs over a range that too close to the incoherent background level to be clearly interpreted. Nevertheless, the value of the invariant with $q_{max} = 0.1$ or $q_{max} = 0.4$ varies by 13% for this characteristic example, setting a reasonable lower bound for the error in subsequent calculations.

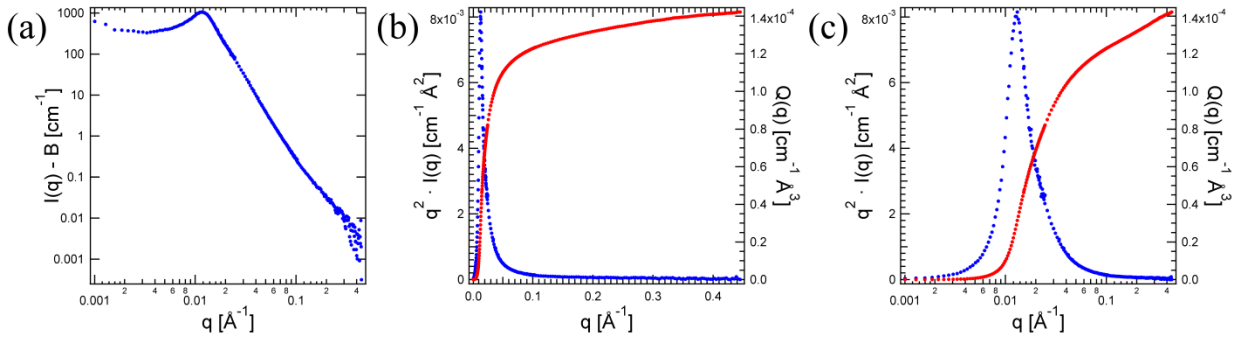

Figure S16: Analysis of Porod invariant in $(X^1PAVG)_{50}$ in D_2O. (a) Plot of experimental data determined in absolute units and corrected for the incoherent background. (b) Plot of the invariant, Q, (red) and its integrand, $q^2 I(q)$, (blue) as a function of q on a linear scale. (c) Identical plot as in (b), except with q on a log scale.

The dependence of Q on the protein volume fraction can be modeled as described in equations 7.1-7.3, and is shown for $(X^1PAVG)_{50}$ in D_2O below:

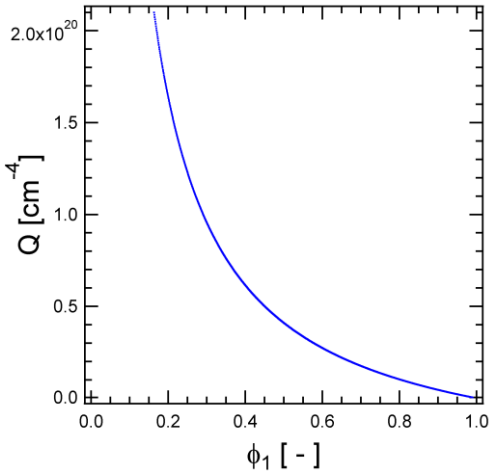

Figure S17: Example of the modeled dependence of the Porod invariant, Q, on the volume fraction of the protein-rich dense phase, ϕ_1, for $(X^1PAVG)_{50}$.

Table S3: Analysis of the volume fraction of the dense phase from the Porod invariant.

Identifier	Solvent*	Q [10^{-4} cm^{-1}Å^3]	$\rho_{Protein}$ † [10^{-6} Å^{-2}]	ϕ_1 [-]	$\phi_{protein}$ ‡ [-]
$(X^1PGVG)_{50}$	D_2O	1.42	2.483	0.22	0.70
	PB	1.67	2.483	0.20	0.79
$(X^1PGVG)_{70}$	D_2O	1.54	2.472	0.21	0.74
	PB	1.46	2.472	0.22	0.71
$(X^1PGVG)_{120}$	D_2O	1.72	2.460	0.19	0.81
	PB	1.57	2.460	0.21	0.75
$(X^2PGVG)_{50}$	D_2O	1.81	2.447	0.19	0.84
	PB	1.80	2.447	0.19	0.83

*PB = 100 mM sodium phosphate in D_2O, pH = 7.6. †Computed using NIST's online scattering length density calculator assuming complete deuteration at all exchangeable proton positions. ρ_{D2O} was estimated to be 6.402 × 10^{-6} Å^{-2}. ‡$\phi_{protein} = \phi_1/\phi_{Protein,gel}$.

Supplemental Equations for the CRW Model

The full version of the two-point correlation function derived by Choi and Chen,[3] which is represented in a minimal form in the text, is reproduced here:

$$g(r) = \frac{4bc(a^2+(b+c)^2)^2}{(b+c)r}\left[e^{-cr}\left(\frac{(a^2-b^2+c^2)}{(4a^2b^2+(a^2-b^2+c^2)^2)^2}+\frac{r}{4c((a^2+b^2)^2+2(a^2-b^2)c^2+c^4)}\right)+ e^{-br}\left(\frac{-8a^2b^2+\left(a^2+b^2\right)^2+2\left(a^2-b^2\right)c^2+c^4}{4ab(4a^2b^2+(a^2-b^2+c^2)^2)^2}\sin(ar)-\frac{\left(a^2-b^2+c^2\right)}{(4a^2b^2+(a^2-b^2+c^2)^2)^2}\cos(ar)\right)\right] \quad \text{(S7)}$$

To convert the fit parameters to the curvature measures, the following relations are used:

$$\langle H\rangle = \frac{\beta}{2}\sqrt{\frac{\pi}{6}\langle k^2\rangle} \quad \text{(S8)}$$

$$\langle K\rangle = \frac{1}{6}\langle k^2\rangle(\beta^2-1) \quad \text{(S9)}$$

$$\langle H^2\rangle = \frac{1}{6}\langle k^2\rangle\left(\beta^2+\frac{6\langle k^4\rangle}{5\langle k^2\rangle^2}-1\right) \quad \text{(S10)}$$

where $\langle H\rangle$ is the mean curvature, $\langle K\rangle$ is the Gaussian curvature, and $\langle H^2\rangle$ is the square-mean curvature, which are functions of the clipping parameter, β, (related to the volume fractions) in addition to second and fourth moments of the inverse eighth order spectral function proposed by Choi and Chen. The equations for those moments are:

$$\langle k^2\rangle = \frac{c(a^2+b^2+bc)}{(b+c)} \quad \text{(S11)}$$

$$\langle k^4\rangle = \frac{c(a^4+2a^2b^2+b^4+4a^2bc+4b^3c+4b^2c^2+bc^3)}{(b+c)} \quad \text{(S12)}$$

References

1. P. Perrin and R. E. Prud'Homme, *Macromolecules*, 1994, **27**, 1852-1860.
2. C. J. Dimitriou, R. H. Ewoldt and G. H. McKinley, *Journal of Rheology*, 2013, **57**, 27-70.
3. S. Choi and S. Chen, *Progr. Colloid Polym. Sci.*, 1997, **106**, 14-23.