Supporting Information

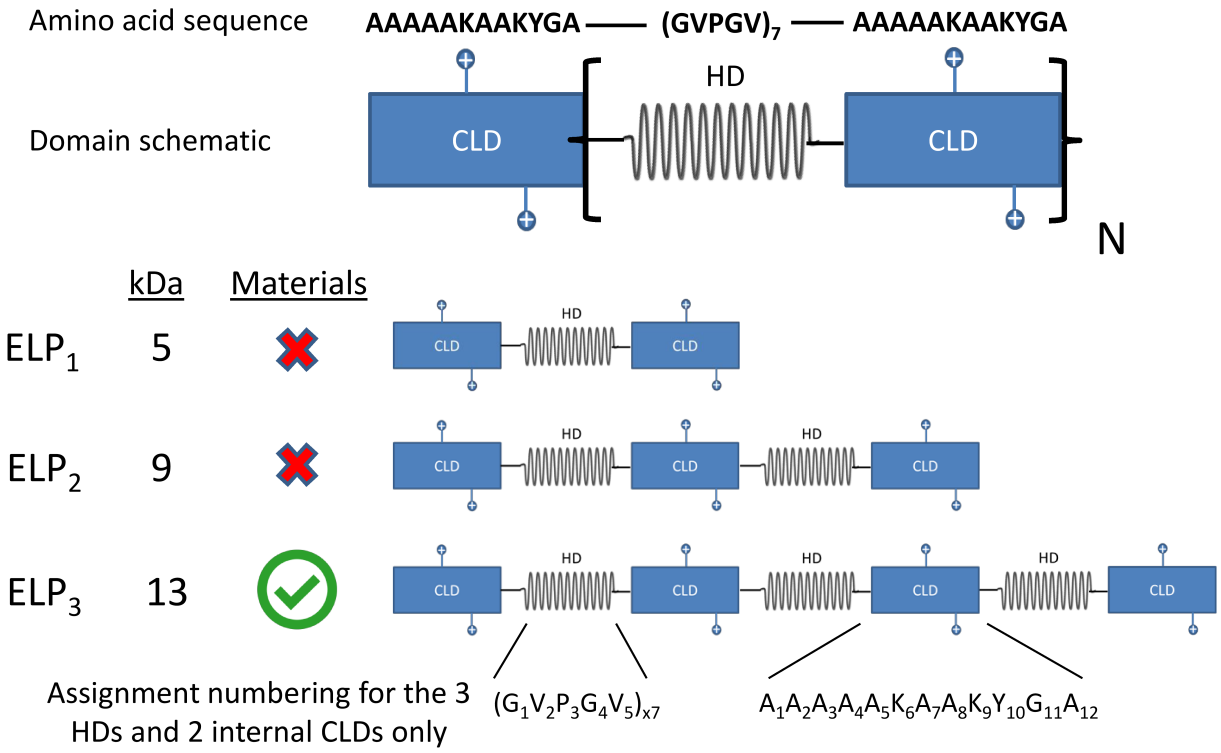

Figure S1. **The alternating domain architecture of the ELP constructs.** The repeating amino acid sequence of the ELPs is shown above the general schematic of the repeating domain architecture. Crosslinking domains (CLD) and hydrophobic domains (HD) are represented as blue rectangles and springs respectively. ELP_1, ELP_2 and ELP_3 have the same basic amino acid sequence motifs but vary by the total number of C-terminal HD-CLD repeats (N). At least four CLDs were needed to make crosslinked materials from an ELP coacervate, as ELP_3 was the only ELP that made resilient biomaterials when its coacervate was crosslinked with genipin. The numbering of amino acids within the HD and CLD repeating units, as used in discussion of the NMR data, is shown below the ELP_3 schematic.

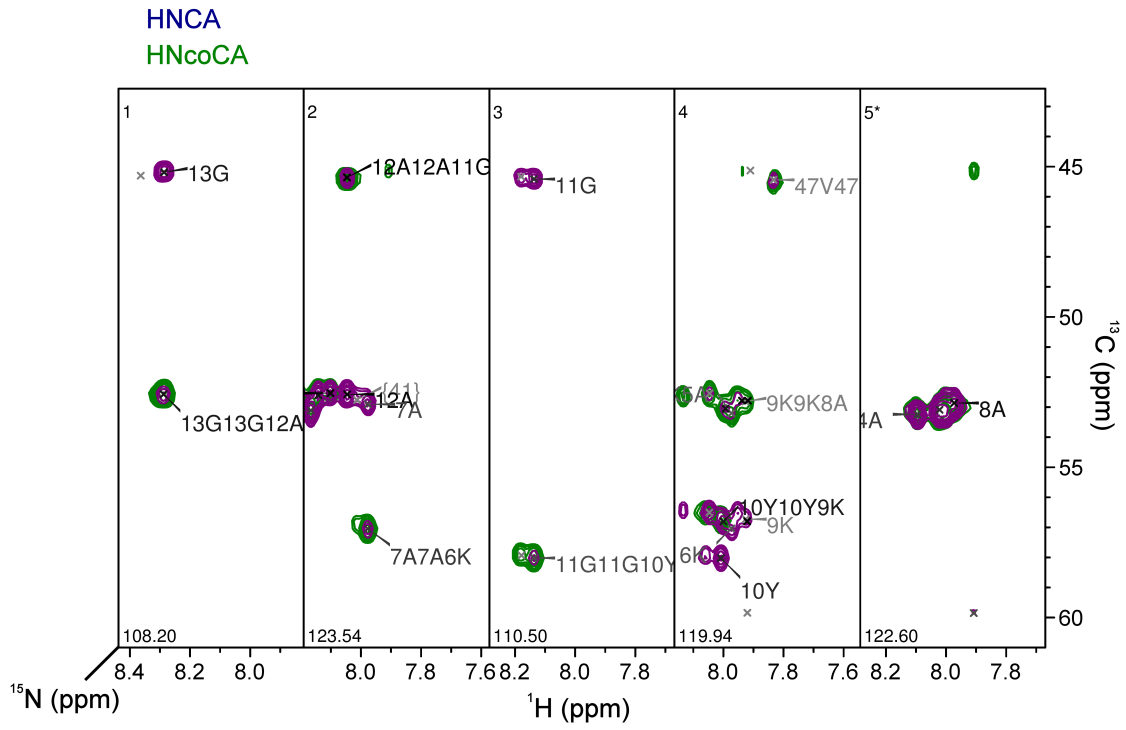

Figure S2. **Assignment of NMR resonances in the C-terminus of the CLDs of ELP$_3$.** Overlay of HNCA (navy blue) and HN(CO)CA (green) triple resonance spectra showing the assignment of ^{1}HN, ^{15}N and ^{13}Cα chemical shifts to the C-terminus of the internal CLDs. This example assignment starts in strip one with the glycine immediately following the CLD (13G), progressing CO to N through the peptide bonds, concluding in strip five at position 8A.

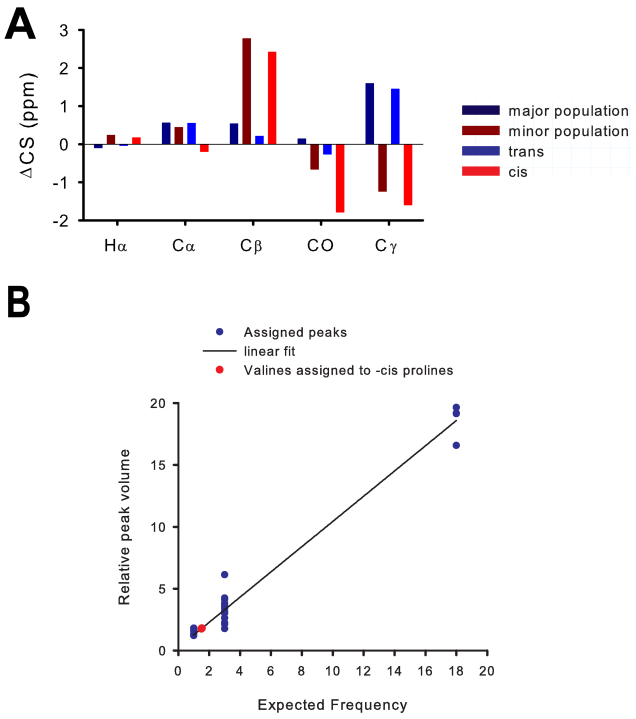

Figure S3. Identification of cis-prolines and calculation of their abundance. (**A**) Secondary chemical shifts (deviation from reported random coil values) for the major and minor populations of proline identified during resonance assignment of the ELP_3 HDs. Also shown are the reported secondary shifts expected for proline in the trans- or cis- conformation (1), the latter of which closely matches the minor population of prolines. This has allowed specific assignment of resonances arising from HD repeats containing trans- versus cis-proline. (**B**) Due to the repeating domain architecture and the repetitive HD sequences, the chemical shifts of most amino acids at a given repeat position were identical, with differences only occurring at the interfaces between domains. To confirm this analysis, the relative peak volume of assigned peaks from the ELP_3 ^{1}H-^{15}N HSQC was plotted against the expected frequency of each site (amino acid and position). The first 5 residues at the N- and C-terminus, which could be differentiated from internal CLD residues, were assigned an expected frequency of 1. The internal CLD residues and amino acids at the border between the HDs and CLDs were assigned an expected frequency of 3. Since each amino acid in the internal HD GVPGV repeats, which did not border the CLDs, had the same chemical shifts, their expected frequency was 18 (21 minus 3), assuming all prolines had a trans backbone conformation. The linear fit by least-squares regression displayed a strong correlation between peak volume and expected amino acid frequency (slope of 1.019, a y-intercept of 0.2348, and an r^2 of 0.974). This supports our assignments and suggests the HSQC is reporting on all the average chemical environments of every amino acid in ELP_3. The minor population of valines assigned as being adjacent to cis- prolines is indicated in red. Using the correlation described above results in a 1:9 ratio of cis-trans isomers, which is high relative to that observed for folded globular proteins, but closely matches the ratios reported for glycine-rich hydrophobic polypeptides.

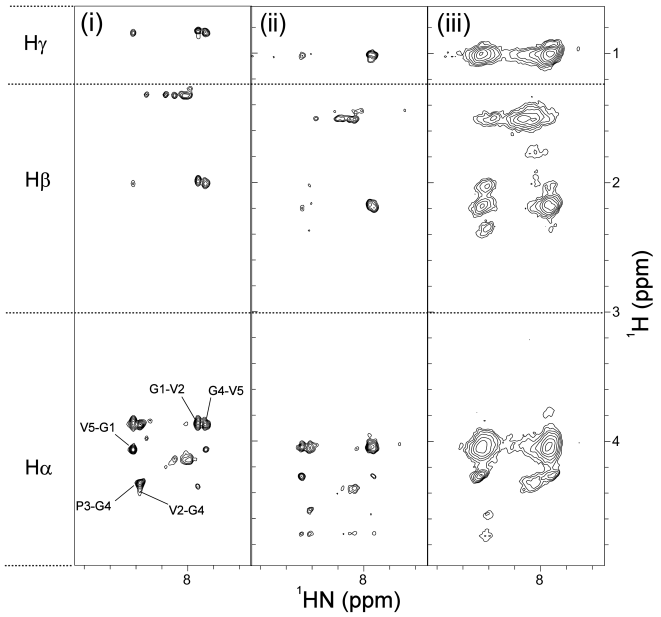

Figure S4. Transient β–turns are observed in NOE spectra of both the monomer and the coacervate. Regions of ^{1}H-^{1}H NOESY spectra showing NOEs between sidechain protons (Hα, Hβ and Hγ) and backbone amide protons (HN) are shown for monomeric ELP_3 at 20 °C (*i*), and 37 °C (*ii*) and for coacervated ELP_3 at 37 °C (*iii*). Labels in (*i*) indicate key inter-residue NOEs observed for the HD domains of the monomer.

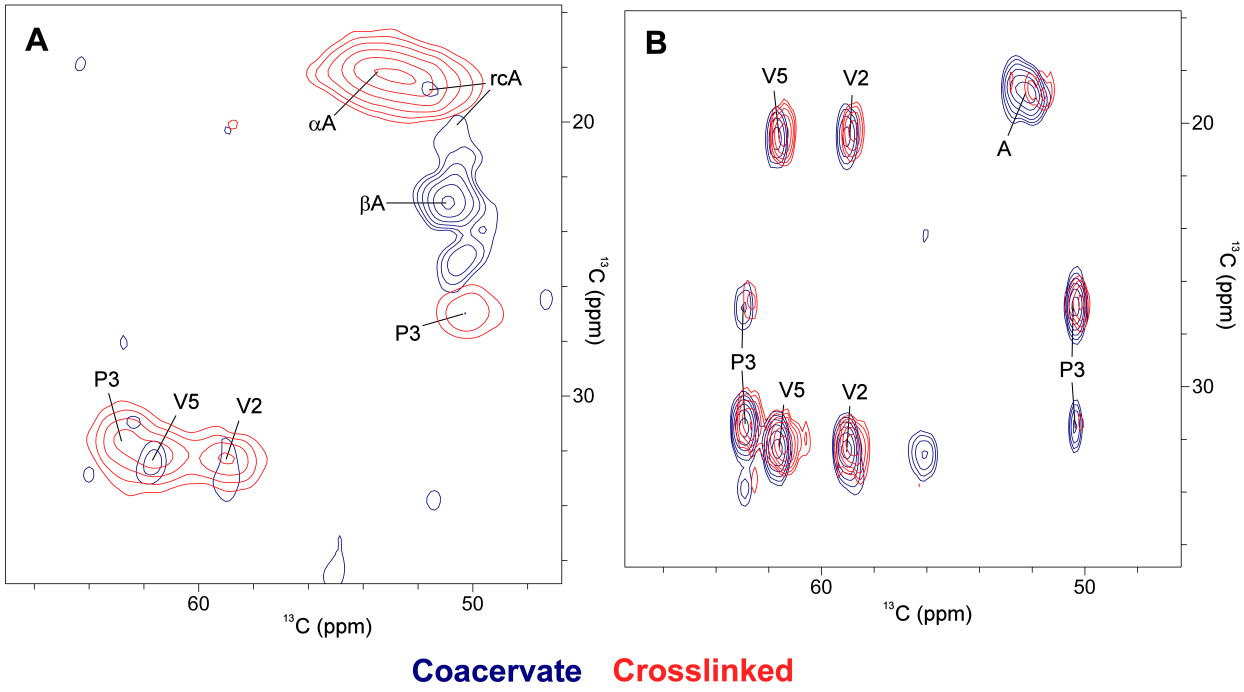

Figure S5. The secondary structure of ELP_3 CLDs is altered by chemical crosslinking, while the HD remain highly disordered. The Cβ-Cα/Cγ-Cα region of ^{13}C-^{13}C correlation spectra of phase-separated (blue) and crosslinked (red) ELP_3, recorded under magic angle spinning conditions, are shown. Peak labels describe amino acid assignments established from solution state NMR experiments and supported by spin system assignments in the solid state NMR experiments. Alanine peak labels (**a**) are preceded by their assigned secondary structure type based on Cα and Cβ chemical shift values: α-helix (α), random coil (rc) and β-strand (β). Spectra were obtained using cross-polarization followed by a proton-driven spin diffusion (PDSD) mixing time of 25 ms (**A**), or INEPT polarization of ^{13}C followed by a 9.8 ms TOBSY mixing period (**B**).

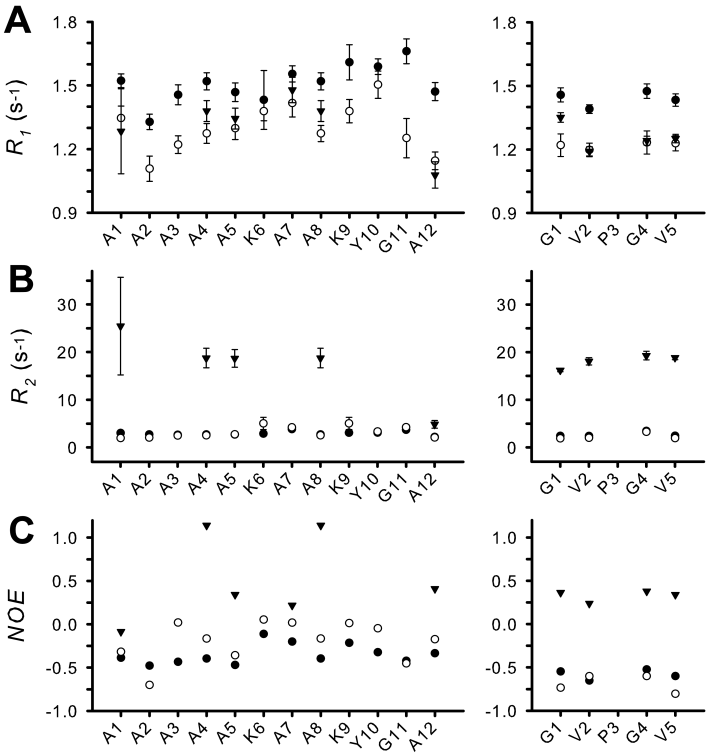

Figure S6. **Sequence specific backbone amide NMR relaxation rates of ELP_3 monomer and coacervate.** The ^{15}N R_1 (***A***), ^{15}N R_2 (***B***) and heteronuclear ^{1}H-^{15}N NOE values (**C**) of the monomer (●), the residual monomer post-phase transition (**O**), and the slow diffusing coacervated ELP_3 species (▼) are shown. Error bars describe the SEM. All data were obtained at 37 °C. A4 and A8 had overlapping ^{1}H-^{15}N resonances so that the values reported represent the average relaxation rates for both sites.

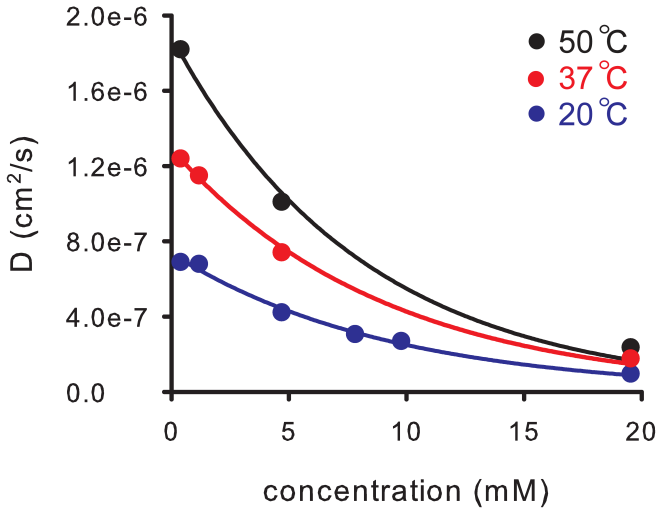

Figure S7. ELP_3 diffusion measured by NMR provides an estimate of protein concentration in the coacervate droplets. The rate of diffusion was measured for monomeric ELP_3 as a function of both temperature and protein concentration, accounting for the effects of varying thermal energy and solution viscosity, respectively. Following a previously reported approach (2), these data can be fit to obtain a protein dependent constant b, and the rate of diffusion at infinite dilution (D_0, Y-intercept of the fit line) at each temperature, using the following equation: $D = D_0^{(-b*[protein])}$. These fits were then used to convert PFG diffusion rates obtained for ELP_3 within the coacervate to an estimated protein concentration, as reported in Table S3. This assumes that increased protein density is the primary contributor to reduced rates of diffusion in the protein-rich phase, which is supported by our data demonstrating that any intermolecular interactions present are transient, non-specific and short-lived on the NMR time-scale. Using a partial specific volume of 0.73 cm^3/g for ELP_3, the water content of the protein rich phase can be estimated at 62.5% by weight.

Table S1. ^{1}H, ^{15}N and ^{13}C chemical shifts (ppm) for monomeric ELP_3 at 20 ^{o}C

Sequence position	^{1}HN	$H^{15}N$	$^{1}H\alpha$	$^{1}H\beta a$	$^{1}H\beta b$	$^{1}H\gamma a$	$^{1}H\gamma b$	$^{1}H\delta a$	$^{1}H\delta b$	^{13}CO	$^{13}C\alpha$	$^{13}C\beta$	$^{13}C\gamma$	$^{13}C\delta$
CLD														
A1	8.32	127.3	4.22	1.33						178.2	53.0	19.0		
A2	8.17	123.4	4.17	1.33						178.3	53.2	19.0		
A3	8.10	122.5	4.17	1.33						178.4	53.2	19.0		
A4	8.01	122.7	4.17	1.33						178.4	53.1	19.0		
A5	8.01	122.7	4.17	1.34						178.6	53.2	19.0		
K6	7.97	119.9	4.15							176.9	57.0	32.9		
A7	7.98	123.7	4.17	1.34						177.9	52.9	19.1		
A8	8.01	122.7	4.17	1.34						177.9	52.9	19.1		
K9	7.94	119.7	4.12							176.5	56.7	33.0		
Y10	8.01	120.0	4.51	2.86	3.03					176.5	58.0	38.9		
G11	8.15	110.6	3.81							173.7	45.4			
A12	8.05	123.6	4.26	1.34						178.2	52.6	19.4		
HD - trans prolines														
G1	8.42	112.6	3.86							173.6	45.2			
V2	7.92	121.1	4.36	1.98		0.83	0.88			174.6	59.9	32.8		
P3		139.0	4.33	1.85	2.21	1.89	1.98	3.61	3.81	177.4	63.7	32.2	27.6	51.1
G4	8.37	109.4	3.87							174.1	45.3			
V5	7.86	119.1	4.07	2.01		0.84	0.85			176.6	62.5	32.8		
HD - cis prolines														
G1	8.42	112.6	3.89							172.8	45.3			
V2	7.64	119.2	4.20	1.92						174.8	59.3			
P3		138.0	4.66	2.10	2.28	1.79	1.87	3.44	3.54	176.6	63.5	34.5	24.8	50.1
G4										173.8	45.4			
V5	8.01	118.5	4.08							176.7	62.2			

Table S2. ^{1}H, ^{15}N and ^{13}C chemical shifts (ppm) for ELP_3 coacervate at 37 ^{o}C

Sequence position	^{1}HN	$H^{15}N$	$^{1}H\alpha$	$^{1}H\beta a$	$^{1}H\beta b$	$^{1}H\gamma a$	$^{1}H\gamma b$	$^{1}H\delta a$	$^{1}H\delta b$	^{13}CO	$^{13}C\alpha$	$^{13}C\beta$	$^{13}C\gamma$	$^{13}C\delta$
CLD														
A1	8.33	126.1								178.3	53.2			
A2	8.18	122.4								178.3	52.5			
A3	8.12	121.9								177.7	52.9	18.9		
A4	8.02	121.9								177.9	53.1			
A5	8.02	121.9								178.0	52.9	18.8		
K6	8.01	119.0								177.5	57.1			
A7	8.03	123.0								177.8	52.8			
A8	8.02	121.9								177.8	53.1			
K9	8.01	119.0								176.3	56.9			
Y10	8.11	123.3								176.4	57.9			
G11	8.23	110.1								173.6	45.6			
A12	8.09	123.3								177.2	52.6			
HD - trans prolines														
G1	8.44	115.6	4.03							173.6	45.4			
V2	7.93	120.2	4.55	2.18		1.01	1.07			174.5	59.9	33.0		
P3			4.51	2.05	2.39	2.09	2.17	3.83	3.97	178.1	63.7	32.3	27.5	51.1
G4	8.38	108.7	4.05							174.0	45.5			
V5	7.88	118.3	4.28	2.21		1.04	1.04			176.5	62.5	32.9		

Table S3. Protein concentrations inside ELP_3 coacervates formed under varied solution conditions, calculated using PFG diffusion data as described in Fig. S8.

$[ELP_3]$ (mM)	$[ELP_3]$ (mg/mL)	[NaCl] (mM)	Temperature (^{o}C)	[coacervate] (mM)	[coacervate] (mg/mL)
2.7	34	300	50	38.3	490
4.7	60	300	50	37.5	480
5.9	75	300	50	36.7	470
9.8	125	300	50	39.1	500
7.8	100	600	37	42.3	550

References

1. Shen Y, Bax A (2010) Prediction of Xaa-Pro peptide bond conformation from sequence and chemical shifts. *J Biomol NMR* 46(3):199–204.

2. O'Leary TJ (1987) Concentration dependence of protein diffusion. *Biophys J* 52(1):137–139.