

Supporting Information for

Refining the design of diblock elastin-like polypeptides for self-assembly into nanoparticles

Michèle Dai,[§] Evangelos Georgilis,[§] Guillaume Goudounet, Bertrand Garbay, Jan Pille, Jan C. M. van Hest, Xavier Schultze, Elisabeth Garanger* and Sébastien Lecommandoux*

[§] Authors contributed equally to the work.

Gene and protein sequences of ELP-[M₁V₃-*i*]-[I-20] (*i* = 20, 40, 60) diblocks

ELP-[M₁V₃-20]-[I-20]

atgtggggttccaggcgttggagtgccaggcatgggcgtaggaggtgtgggagttccaggt
M W V P G V G V P G M G V P G V G V P G
gttggggtaccgggcgtcggagttcctgggatgggagttccgggagttggtgtgccgggt
V G V P G V G V P G M G V P G V G V P G
gtcgggtgtgcctggggtgggtgttccaggtatgggggttccgggtgtcggcgttcccggc
V G V P G V G V P G M G V P G V G V P G
gttgggtgttccaggcgtaggtgtgccgggaatgggggttccgggagttggtgtacctggc
V G V P G V G V P G M G V P G V G V P G
gtgggagtagctggagtcggcgtgcctggtagtggcgtagctggcgtacctggc
V G V P G V G V P G M G V P G V G V P G
gtaggtgttccaggcattggagtgccaggcattggcgtaggagttccaggt
V G V P G I G V P G I G V P G I G V P G
attggggtaccgggcgtcggagttcctgggatcggagttccgggaattggtgtgccgggt
I G V P G I G V P G I G V P G I G V P G
atcgggtgtgcctgggatcgggtgttccaggtatcgggggttccgggtatcggcgttcccggc
I G V P G I G V P G I G V P G I G V P G
attggtgttccaggcatcgggtgtgccgggaatgggggttccgggagttggtgtacctggc
I G V P G I G V P G I G V P G I G V P G
attggggtacctggaatcggcgtgcctggtagtggcgtagcctggcatcggcgttccctggc
I G V P G I G V P G I G V P G I G V P G
attggttgctaa
I G C -

ELP-[M₁V₃-40]-[I-20]

atgtggggttccaggcgttggagtgccaggcatgggcgtaggaggtgtgggagttccaggt
M W V P G V G V P G M G V P G V G V P G
Gttggggtaccgggcgtcggagttcctgggatgggagttccgggagttggtgtgccgggt
V G V P G V G V P G M G V P G V G V P G
gtcgggtgtgcctggggtgggtgttccaggtatgggggttccgggtgtcggcgttcccggc
V G V P G V G V P G M G V P G V G V P G
gttgggtgttccaggcgtaggtgtgccgggaatgggggttccgggagttggtgtacctggc
V G V P G V G V P G M G V P G V G V P G
gtgggagtagctggagtcggcgtgcctggtagtggcgtagcctggcgtacctggc
V G V P G V G V P G M G V P G V G V P G
gtaggtgttccaggcgttggagtgccaggcatgggtaggaggtgtgggagttccaggt
V G V P G V G V P G M G V P G V G V P G
gttggggtaccgggcgtcggagttcctgggatgggagttccgggagttggtgtgccgggt
V G V P G V G V P G M G V P G V G V P G
gtcgggtgtgcctggggtgggtgttccaggtatgggggttccgggtgtcggcgttcccggc
V G V P G V G V P G M G V P G V G V P G
gttgggtgttccaggcgtaggtgtgccgggaatgggggttccgggagttggtgtacctggc
V G V P G V G V P G M G V P G V G V P G
gtgggagtagctggagtcggcgtgcctggtagtggcgtagcctggcgtacctggc
V G V P G V G V P G M G V P G V G V P G
gtaggtgttccaggcattggagtgccaggcattggcgtaggagttccaggt
V G V P G I G V P G I G V P G I G V P G
attggggtaccgggcgtcggagttcctgggatcggagttccgggaattggtgtgccgggt

I G V P G I G V P G I G V P G I G V P G
 atcgggtgtgcctgggatcgggtgttccaggtatcgggggttccgggtatcggcgttcccggc
 I G V P G I G V P G I G V P G I G V P G
 attggtgttccaggcatcgggtgtgccgggaattgggggttccggggattggtgtacctggc
 I G V P G I G V P G I G V P G I G V P G
 attgggggtacctggaatcggcggtgcctgggtattggcggtgcctggcatcggcgttcctggc
 I G V P G I G V P G I G V P G I G V P G
 attggttgctaa
 I G C -

ELP-[M₁V₃-60]-[I-20]

atgtgggggttccaggcgttggagtgccaggcatgggcgtaccagggtgtgggagttccaggt
 M W V P G V G V P G M G V P G V G V P G
 gttgggggtaccgggcgtcggagttcctgggatgggagttccgggagttggtgtgccgggt
 V G V P G V G V P G M G V P G V G V P G
 gtcgggtgtgcctgggggtgggtgttccaggtatgggggttccgggtgtcggcgttcccggc
 V G V P G V G V P G M G V P G V G V P G
 gttggtgttccaggcgtaggtgtaccgggaatgggggttccgggagttggtgtacctggc
 V G V P G V G V P G M G V P G V G V P G
 gtgggagtacctggagtcggcggtgcctgggtatgggcgtgcctggcggtcggcgtagctggc
 V G V P G V G V P G M G V P G V G V P G
 gtaggtgttccaggcgttggagtgccaggcatgggcgtaccagggtgtgggagttccaggt
 V G V P G V G V P G M G V P G V G V P G
 gttgggggtaccgggcgtcggagttcctgggatgggagttccgggagttggtgtgccgggt
 V G V P G V G V P G M G V P G V G V P G
 gtcgggtgtgcctgggggtgggtgttccaggtatgggggttccgggtgtcggcgttcccggc
 V G V P G V G V P G M G V P G V G V P G
 gttggtgttccaggcgtaggtgtgccgggaatgggggttccgggagttggtgtacctggc
 V G V P G V G V P G M G V P G V G V P G
 gtgggagtacctggagtcggcggtgcctgggtatgggcgtgcctggcggtcggcgtagctggc
 V G V P G V G V P G M G V P G V G V P G
 gtaggtgttccaggcattggagtgccaggcattggcgtagctaccagggtattggagttccaggt
 V G V P G I G V P G I G V P G I G V P G
 attgggggtaccgggcgtcggagttcctgggatcggagttccgggaattggtgtgccgggt
 I G V P G I G V P G I G V P G I G V P G
 atcgggtgtgcctgggatcgggtgttccaggtatcgggggttccgggtatcggcgttcccggc
 I G V P G I G V P G I G V P G I G V P G
 attggtgttccaggcatcgggtgtgccgggaattgggggttccggggattggtgtacctggc
 I G V P G I G V P G I G V P G I G V P G
 attgggggtacctggaatcggcggtgcctgggtattggcggtgcctggcatcggcggttccctggc

I G V P G I G V P G I G V P G I G V P G
attgggttgctaa
I G C -

Figure S1. Nucleotide sequences of the ELP-[M₁V₃-*i*]-[I-20] (*i* = 20, 40, 60) genes and corresponding amino acids sequences.

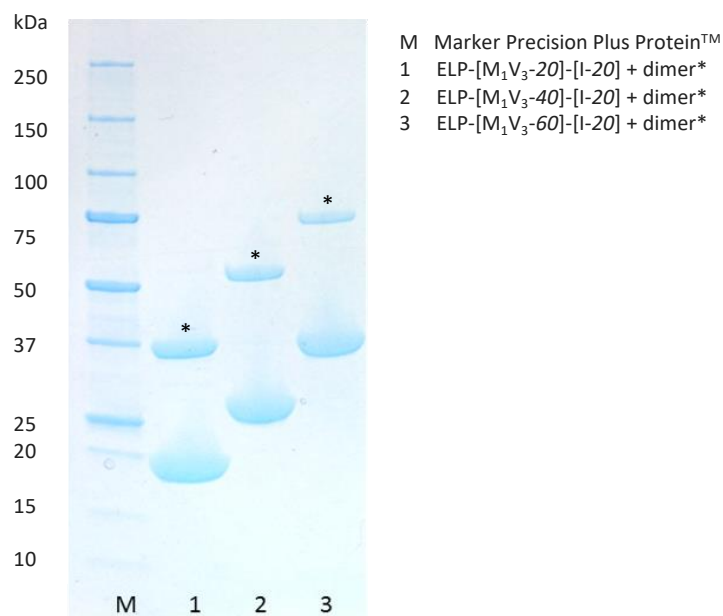
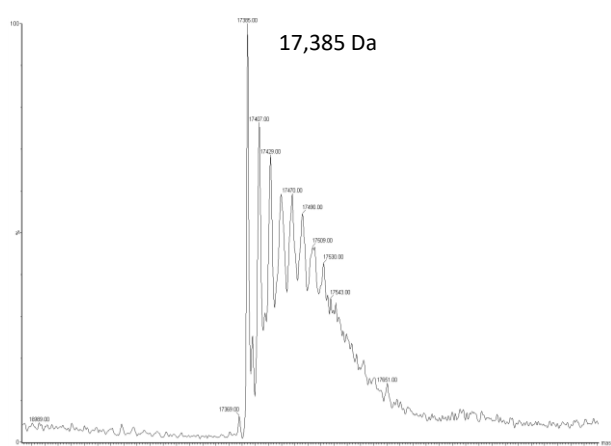
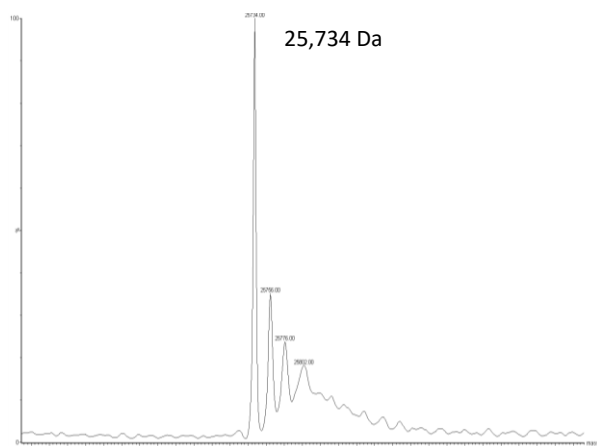


Figure S2. Coomassie-stained SDS-PAGE gel of purified ELP-[M₁V₃-*i*]-[I-20] (*i* = 20, 40, 60) before capping the cysteines located at the C-terminal of the different ELP diblocks (with their respective dimers ELP-[M₁V₃-*i*]-[I-20]-[I-20]-[M₁V₃-*i*]).

(A)



(B)



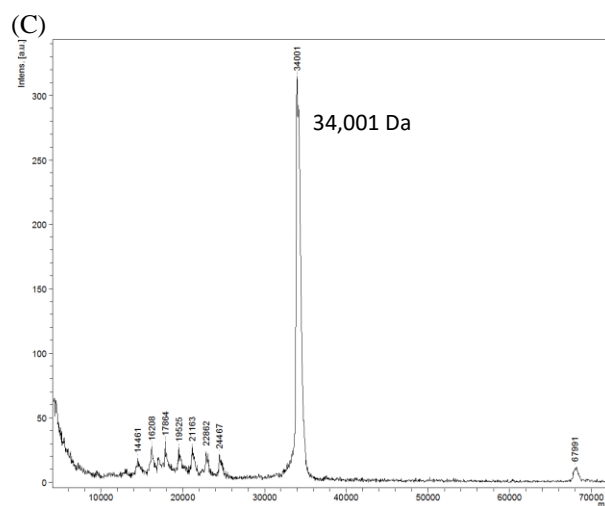
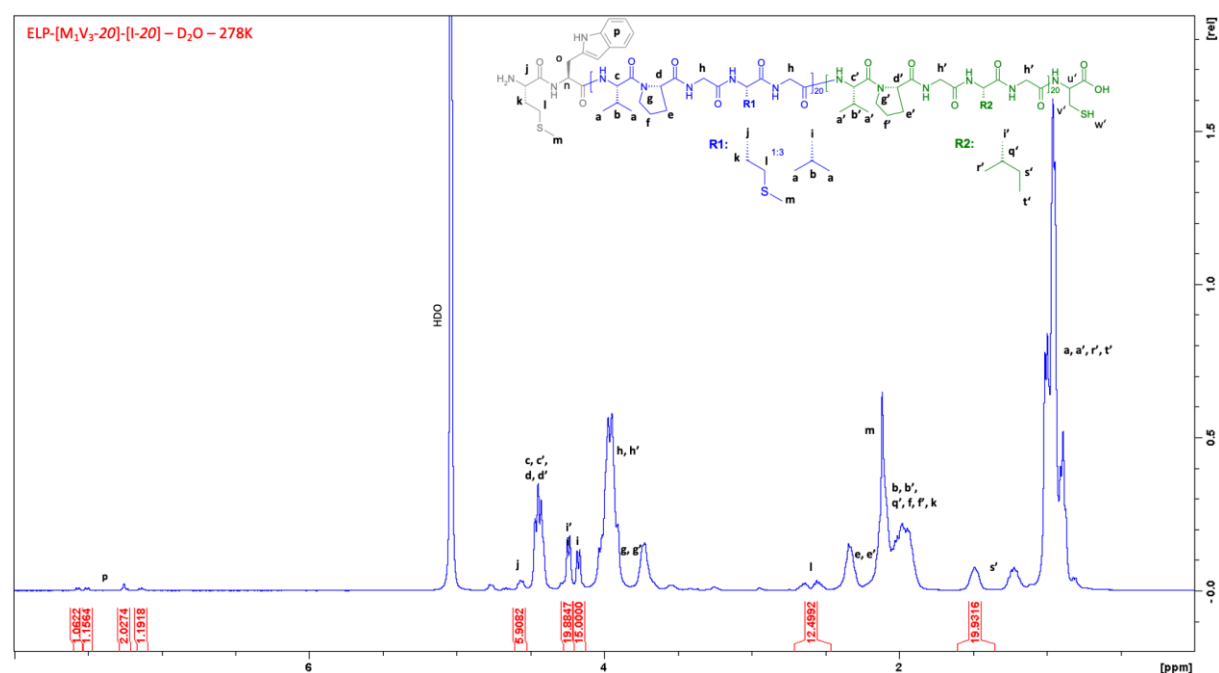
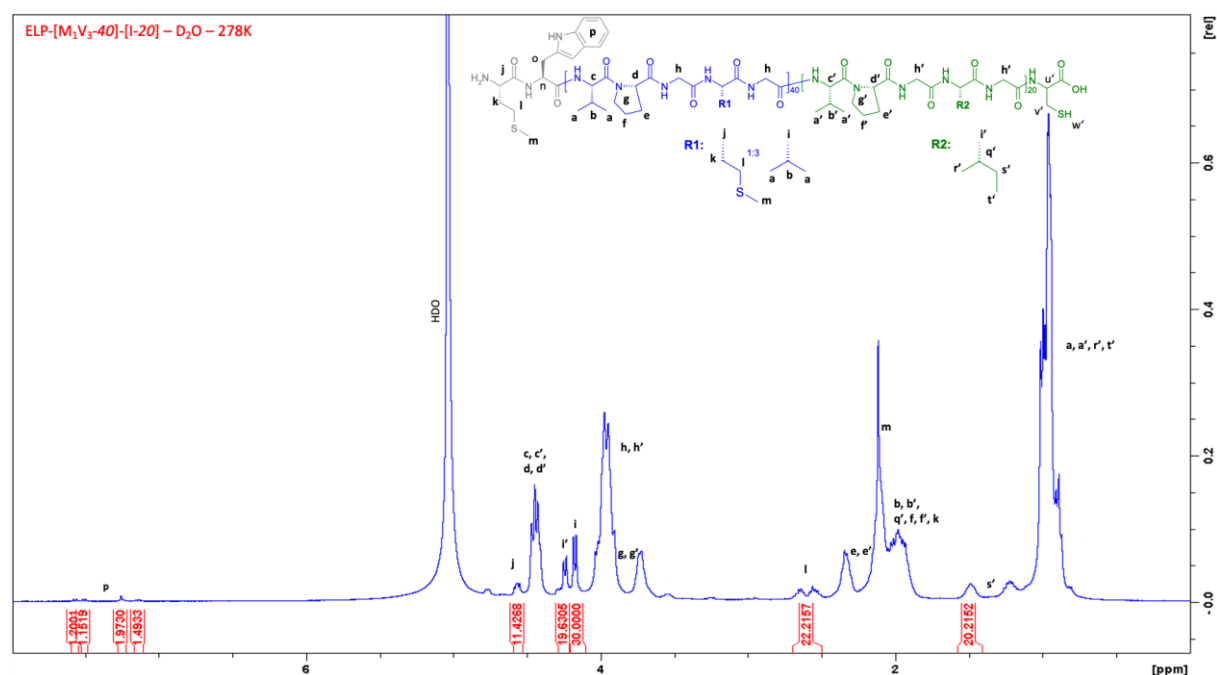


Figure S3. Mass spectra of (A) *ESI-MS of non-capped ELP-[M₁V₃-20]-[I-20], (B) ESI of non-capped ELP-[M₁V₃-40]-[I-20], (C) **MALDI-TOF spectrum of non-capped ELP-[M₁V₃-60]-[I-20]. *ESI-MS = electrospray ionization mass spectrum. **MALDI-TOF = Matrix Assisted Laser Desorption Ionization-time-of-flight.

(A)



(B)



(C)

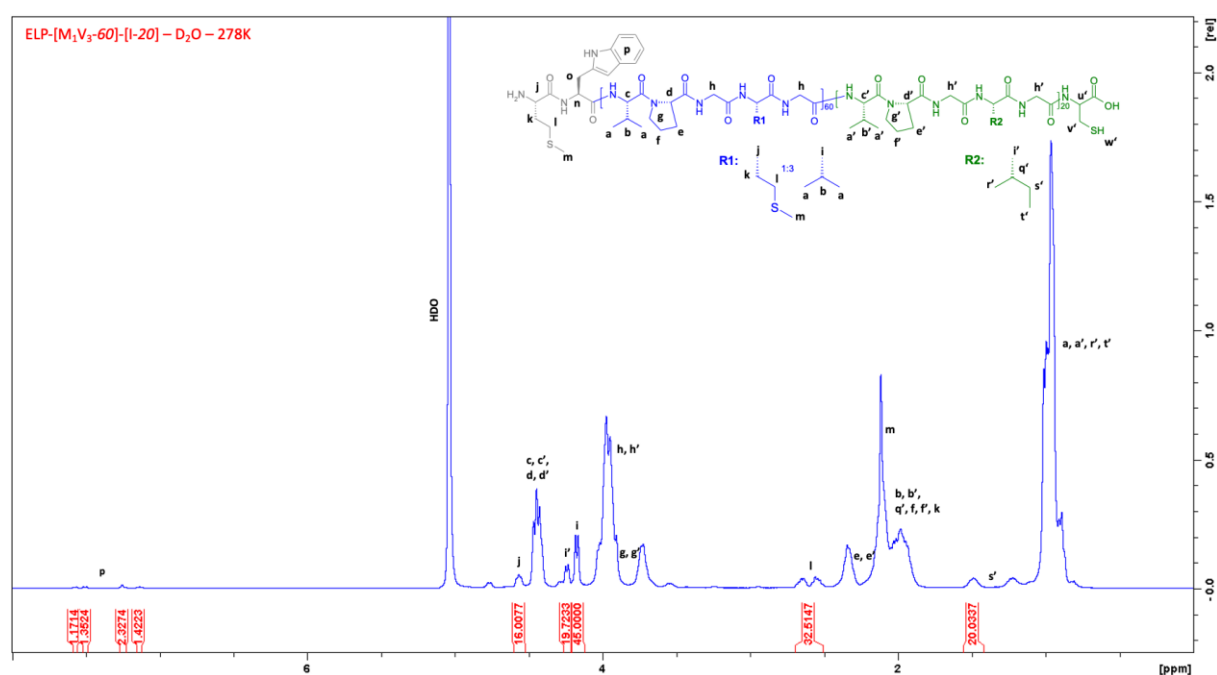


Figure S4. NMR spectra in D₂O of non-capped (A) ELP-[M₁V₃-20]-[I-20], (B) ELP-[M₁V₃-40]-[I-20] and (C) ELP-[M₁V₃-60]-[I-20].

Cysteines from the diblocks ELP-[M₁V₃-*i*]-[I-20] (*i* = 20, 40, 60) were modified to avoid disulfide bridge formation. The reducing agent tris(2-carboxyethyl)phosphine hydrochloride (TCEP·HCl) was added to break every S-S bond between cysteines and *N*-ethyl maleimide (NEM) was introduced by thiol-Michael addition in basic conditions with *N*-ethyldiisopropylamine (DIPEA) (Figure S5A). After isolation by extensive dialysis against ultrapure water and freeze-drying, the products were characterized as ELP-[M₁V₃-*i*]-[I-20] (*i* = 20, 40, 60) by SDS-PAGE (Figure S5B)

and size-exclusion chromatography (SEC) in DMF (Figure S5C) showing both disappearance of dimer traces. To be consistent, the monoblocks ELP-[I-*j*] (*j* = 20, 40, 60) were also synthesized with the same procedure than for the diblocks. As ELP-[M₁V₃-*i*] do not have cysteine residues, no modification was performed.

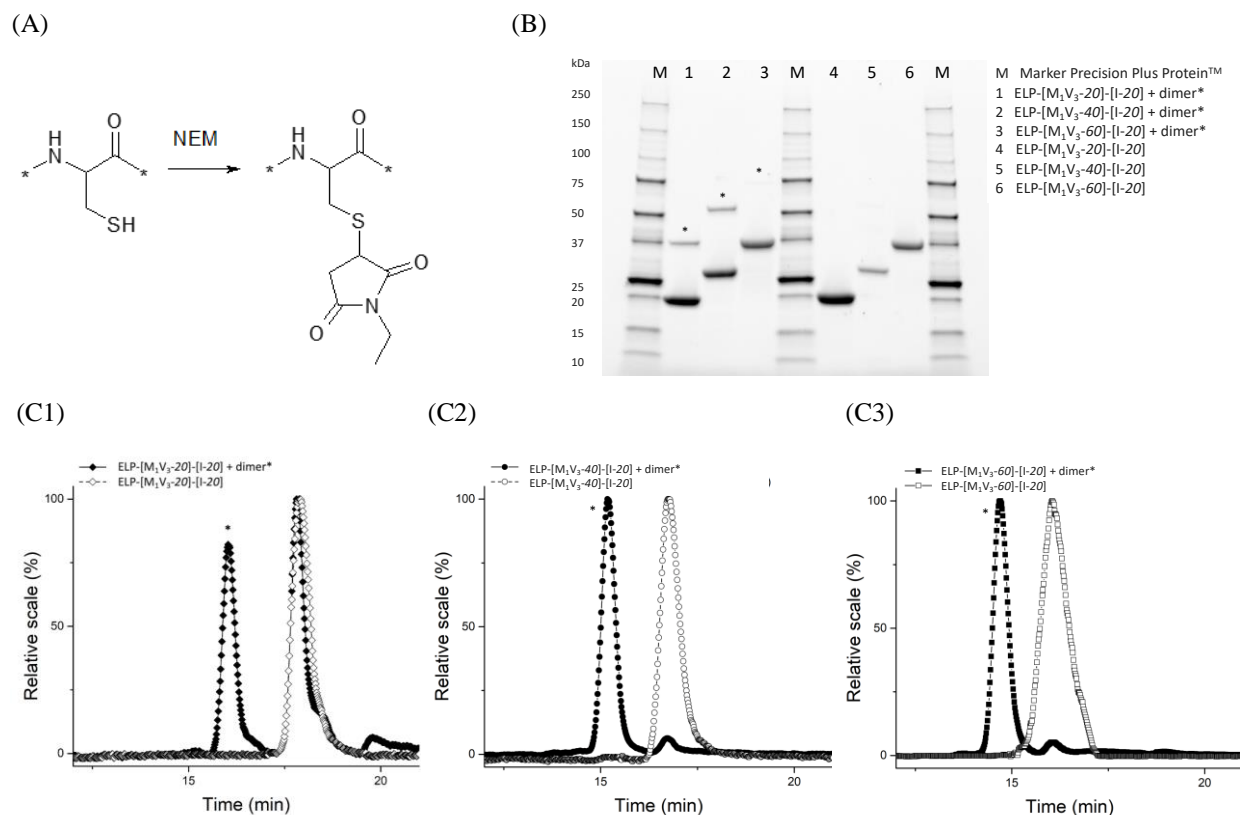


Figure S5. (A) Scheme of the introduction of a *N*-ethyl maleimide (NEM) into cysteine thiol *via* thiol-Michael addition. (B) Stain-free gel SDS-PAGE of produced diblock elastin-like polypeptides (1-2-3), and after introducing *N*-ethyl maleimide (4-5-6). (C) ELP-[M₁V₃-*i*]-[I-20] (*i* = 20 (C1), 40 (C2), 60 (C3)) size-exclusion chromatography spectra in DMF.

(A) Gene and protein sequences of ELP-[M₁V₃-*i*]-[I-90] (*i* = 40, 60) diblocks

ELP-[M₁V₃-40]-[I-90]

```

atgtggggtccaggcgttgagtgccaggcatgggcgtaccagggtgtgggagttccaggt
M W V P G V G V P G M G V P G V G V P G
gttggggtaccgggcgtcggagttcctgggatgggagttccgggagttggtgtgccgggt
V G V P G V G V P G M G V P G V G V P G
gtcgggtgtgcctggggtgggtgttcagggtatgggggttccgggtgtcggcgttcccggc
V G V P G V G V P G M G V P G V G V P G
gttgggtgttcaggcgttaggtgtgccgggaatgggggttccgggagttggtgtacctggc
V G V P G V G V P G M G V P G V G V P G
gtgggagtagctggagtcggcgtgcctgggtatgggcgtgcctggcgtcggcgtacctggc
V G V P G V G V P G M G V P G V G V P G
gtaggtgttcaggcgttgagtgccaggcatgggcgtaccagggtgtgggagttccaggt
V G V P G V G V P G M G V P G V G V P G
gttggggtaccgggcgtcggagttcctgggatgggagttccgggagttggtgtgccgggt
V G V P G V G V P G M G V P G V G V P G

```

[illegible]

P G I G V P G I G V P G Y - - -

ELP-[M₁V₃-60]-[I-90]

atgtggggttccaggcgttggagtgccaggcatgggcgtaccagggtgtgggagttccagggt
M W V P G V G V P G M G V P G V G V P G
gttggggtaccgggcgtcggagttcctgggatgggagttccgggagttggtgtgccgggt
V G V P G V G V P G M G V P G V G V P G
gtcgggtgtgcctggggtgggtgttccagggtatgggggttccgggtgtcggcgttcccggc
V G V P G V G V P G M G V P G V G V P G
gttgggtgttccaggcgtaggtgtgccgggaatgggggttccgggagttggtgtacctggc
V G V P G V G V P G M G V P G V G V P G
gtgggagtacctggagtcggcgtgcctgggtatgggcgtgcctggcgtcggcgtacctggc
V G V P G V G V P G M G V P G V G V P G
gtaggtgttccaggcgttggagtgccaggcatgggcgtaccagggtgtgggagttccagggt
V G V P G V G V P G M G V P G V G V P G
gttggggtaccgggcgtcggagttcctgggatgggagttccgggagttggtgtgccgggt
V G V P G V G V P G M G V P G V G V P G
gtcgggtgtgcctggggtgggtgttccagggtatgggggttccgggtgtcggcgttcccggc
V G V P G V G V P G M G V P G V G V P G
gttgggtgttccaggcgtaggtgtgccgggaatgggggttccgggagttggtgtacctggc
V G V P G V G V P G M G V P G V G V P G
gtgggagtacctggagtcggcgtgcctgggtatgggcgtgcctggcgtcggcgtacctggc
V G V P G V G V P G M G V P G V G V P G
gtagggcgtacctggagtcggcgtgcctgggtatgggcgtgcctggcgtcggcgtacctggc
V G V P G V G V P G M G V P G V G V P G
gtagggcgtacctggagtcggcgtgcctgggtatgggcgtgcctggcgtcggcgtacctggc
V G V P G V G V P G M G V P G V G V P G
gtagggcgtacctggagtcggcgtgcctgggtatgggcgtgcctggcgtcggcgtacctggc
V G V P G V G V P G M G V P G V G V P G
gtagggcgtacctggagtcggcgtgcctgggtatgggcgtgcctggcgtcggcgtacctggc
V G V P G V G V P G M G V P G V G V P G
gtagggcatcgggtgtaccgggtatcgggtgtgccgggcatcgggtgtgccgggtatcgggggtt
V G I G V P G I G V P G I G V P G I G V
ccaggcatcgggtgtaccaggcatcgggggtaccgggtatcgggtgtgccgggcatcgggtgtg
P G I G V P G I G V P G I G V P G I G V
ccgggtatcgggggttccaggcatcgggtgtaccaggcatcgggtgtaccgggtatcgggtgtg
P G I G V P G I G V P G I G V P G I G V
ccgggcatcgggtgtgccgggtatcgggggttccaggcatcgggtgtaccaggcatcgggtgtg
P G I G V P G I G V P G I G V P G I G V
ccaggcatcgggggtaccgggtatcgggtgtgccgggcatcgggtgtgccgggtatcgggggtt
P G I G V P G I G V P G I G V P G I G V
ccaggcatcgggtgtaccaggcatcgggtgtaccgggtatcgggtgtgccgggcatcgggtgtg
P G I G V P G I G V P G I G V P G I G V
ccgggtatcgggggttccaggcatcgggtgtaccaggcatcgggtgtaccgggtatcgggtgtg
P G I G V P G I G V P G I G V P G I G V
ccgggcatcgggtgtgccgggtatcgggggttccaggcatcgggtgtaccaggcatcgggggtg

P G I G V P G I G V P G I G V P G I G V
 ccgggtatcgggtgtgccgggcatcgggtgtgccgggtatcgggggttccaggcatcgggtgta
 P G I G V P G I G V P G I G V P G I G V
 ccaggcatcgggtgtaccgggtatcgggtgtgccgggcatcgggtgtgccgggtatcgggggtt
 P G I G V P G I G V P G I G V P G I G V
 ccaggcatcgggtgtaccaggcatcgggtgtaccgggtatcgggtgtgccgggcatcgggtgtg
 P G I G V P G I G V P G I G V P G I G V
 ccgggtatcgggggttccaggcatcgggtgtaccaggcatcgggggtaccgggtatcgggtgtg
 P G I G V P G I G V P G I G V P G I G V
 ccgggcatcgggtgtgccgggtatcgggggttccaggcatcgggtgtaccaggcatcgggtgta
 P G I G V P G I G V P G I G V P G I G V
 ccgggtatcgggtgtgccgggcatcgggtgtgccgggtatcgggggttccaggcatcgggtgta
 P G I G V P G I G V P G I G V P G I G V
 ccaggcatcgggtgtaccgggtatcgggtgtgccgggcatcgggtgtgccgggtatcgggggtt
 P G I G V P G I G V P G I G V P G I G V
 ccaggcatcgggtgtaccaggcatcgggggtaccgggtatcgggtgtgccgggcatcgggtgtg
 P G I G V P G I G V P G I G V P G I G V
 ccgggtatcgggggttccaggcatcgggtgtaccaggcatcgggtgtaccgggtatcgggtgtg
 P G I G V P G I G V P G I G V P G I G V
 ccgggcatcgggtgtgccgggtatcgggggttccaggcatcgggtgtaccaggcatcgggtgta
 P G I G V P G I G V P G I G V P G I G V
 ccgggtatcgggtgtgccgggcatcgggtgtgccgggtatcgggggttccaggcatcgggtgta
 P G I G V P G I G V P G I G V P G I G V
 ccaggcatcgggggtaccgggtatcgggtgtgccgggcatcgggtgtgccgggtatcgggggtt
 P G I G V P G I G V P G I G V P G I G V
 ccaggcatcgggtgtaccaggcatcgggtgtaccgggtatcgggtgtgccgggcatcgggtgtg
 P G I G V P G I G V P G I G V P G I G V
 ccgggtatcgggggttccaggcatcgggtgtaccaggctactgataatga
 P G I G V P G I G V P G Y - - -

(B)

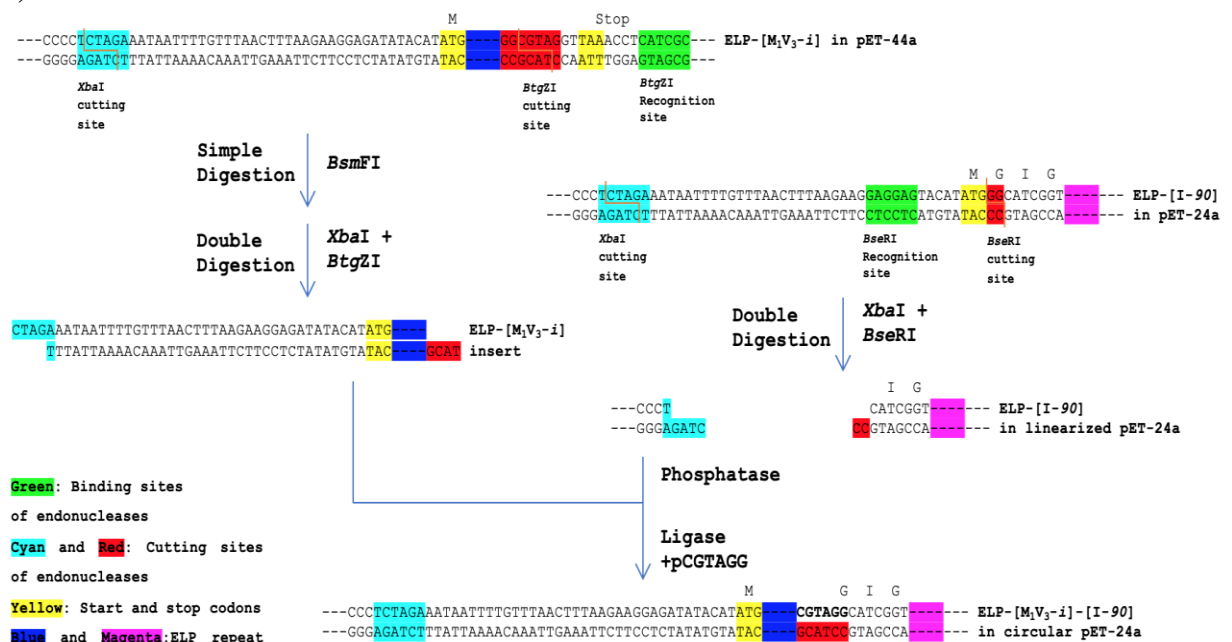


Figure S6. (A) Sequences of the ELP-[M₁V₃-i]-[I-20] ($i = 20, 40, 60$) genes and of their corresponding proteins. (B) Schematic representation of the cloning strategy followed for the construction of the ELP-[M₁V₃-i]-[I-90] library.

(A) MS/MS spectrum of protein 55,314 Da. The x-axis represents m/z from 52,000 to 58,000, and the y-axis represents relative intensity from 0 to 100. The base peak is at m/z 55,314.30 (27960). Other labeled peaks include 52403.30 (1191), 53193.40 (1060), 53407.90 (1412), 53919.90 (1017), 54345.10 (1529), 55292.30 (1840), 55334.80 (6694), 55373.60 (3115), 55395.60 (1678), 56443.70 (1368), 57060.40 (1208), 57621.40 (1121), and 58000.

(B) MS/MS spectrum of protein 63,664 Da. The x-axis represents m/z from 62,000 to 66,000, and the y-axis represents relative intensity from 0 to 100. The base peak is at m/z 63663.80 (4930). Other labeled peaks include 62028.70 (916), 62054.00 (458), 62735.40 (379), 62811.60 (653), 62964.00 (446), 63637.20 (625), 63724.70 (923), 63776.70 (603), 64365.50 (585), 64390.50 (968), 64528.20 (1008), 64551.10 (632), 64578.50 (421), 65264.40 (605), 65429.70 (438), 65735.90 (399), and 66000.

(C) MS/MS spectrum of protein 55,523 Da. The x-axis represents m/z from 30,000 to 60,000, and the y-axis represents relative intensity from 0 to 5000. The base peak is at m/z 55522.800. Other labeled peaks include 30719.750 and 57048.000.

(D) MS/MS spectrum of protein 63,949 Da. The x-axis represents m/z from 30,000 to 60,000, and the y-axis represents relative intensity from 0 to 6000. The base peak is at m/z 63948.000. Other labeled peaks include 30719.750 and 57048.000.

Figure S9. Mass spectra of (A) *ESI-MS of ELP-[M₁V₃-40]-[I-90], (B) *ESI-MS of ELP-[M₁V₃-60]-[I-90], (C)

**MALDI-TOF of ELP-[M_1V_3-40]-[I-90] and (D) **MALDI-TOF of ELP-[M_1V_3-60]-[I-90]. *ESI-MS = electrospray ionization mass *spectrum*. **MALDI-TOF = Matrix Assisted *Laser* Desorption Ionization-time-of-flight.

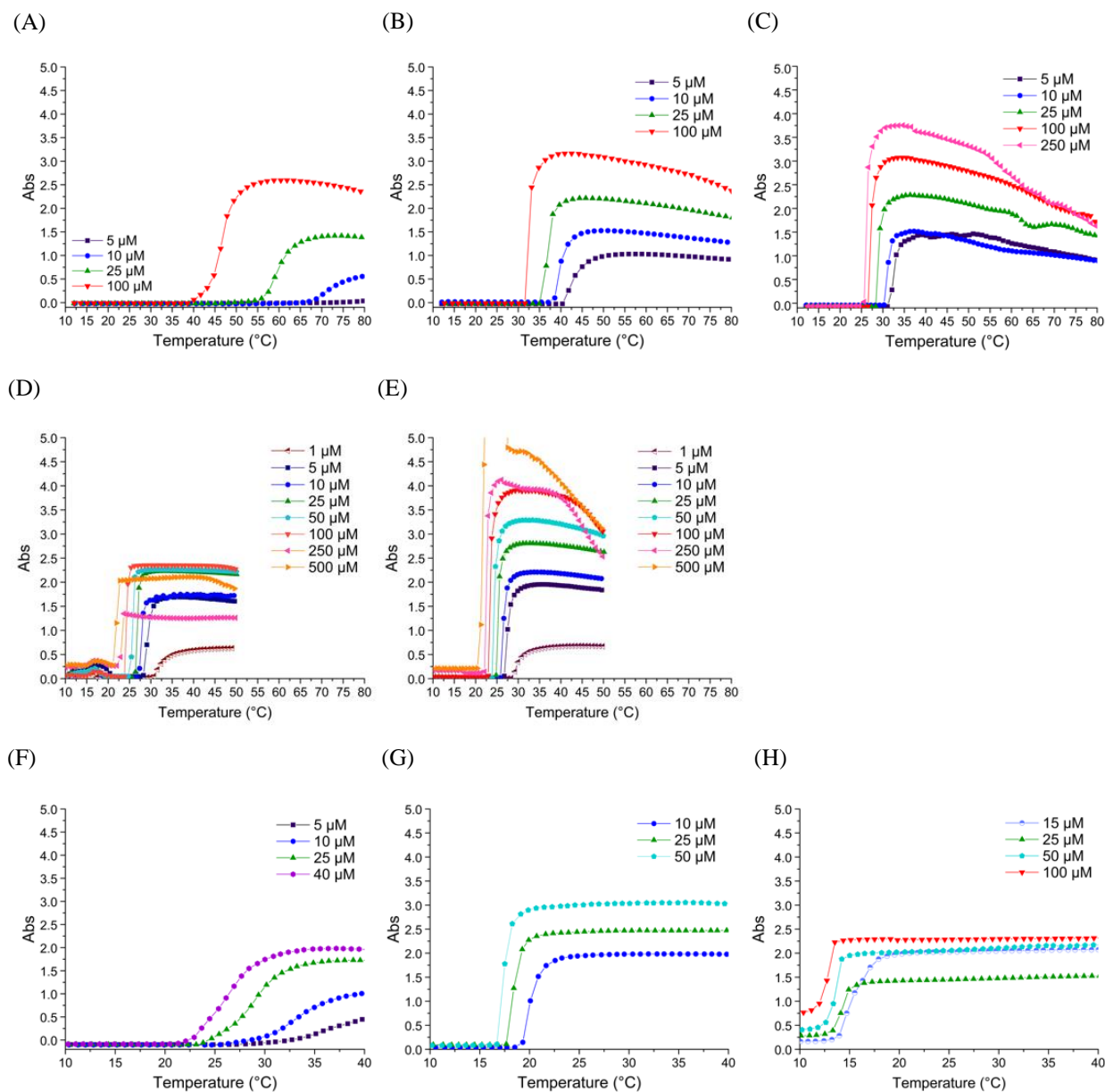
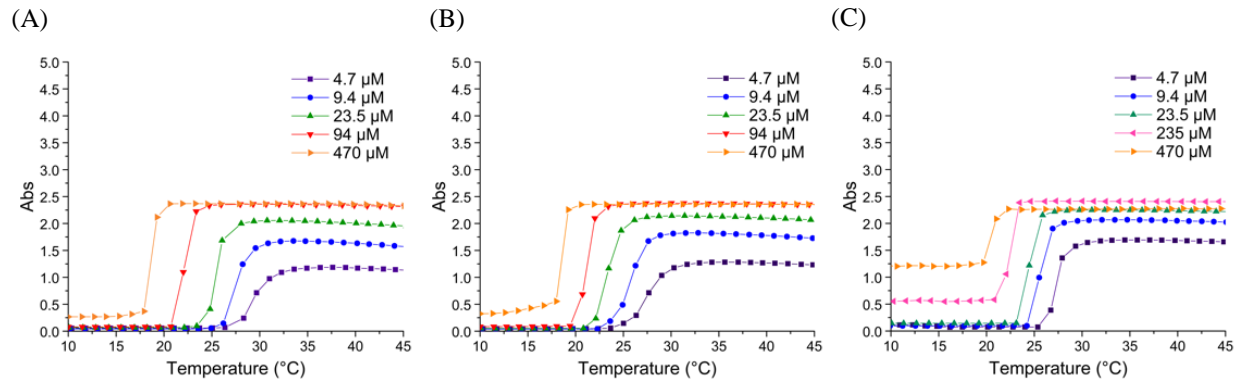


Figure S10. Turbidity assays at 350 nm with a rate of 1°C/min of (A, B, C, D, E) ELP-[M_1V_3-i] ($i = 20, 40, 60, 80, 100$ respectively) and (F, G, H) ELP-[I- j] ($j = 20, 40, 60$ respectively) in PBS.

Table S1. The linear dependence of T_t versus $\log(C)$ correlated to eq. 1, for ELP monoblocks experimental data.

Monoblock name	T_t versus $\log(C)$	R^2
ELP-[M_1V_3-20]	$y = -10.28 \cdot \ln(x) + 93.20$	0.99
ELP-[M_1V_3-40]	$y = -3.28 \cdot \ln(x) + 46.79$	0.99
ELP-[M_1V_3-60]	$y = -2.26 \cdot \ln(x) + 36.48$	0.99
ELP-[M_1V_3-80]	$y = -1.77 \cdot \ln(x) + 31.32$	0.99
ELP-[M_1V_3-100]	$y = -1.46 \cdot \ln(x) + 28.79$	0.99
ELP-[I-20]	$y = -4.33 \cdot \ln(x) + 42.71$	0.99
ELP-[I-40]	$y = -1.62 \cdot \ln(x) + 23.00$	0.99
ELP-[I-60]	$y = -1.32 \cdot \ln(x) + 18.58$	0.99



For temperatures below T_t when the chain length of the hydrophilic block was long and the solution was concentrated, absorbance was increased. This was possibly caused by the existence of some aggregates favored by the increased concentration. Similarly, the turbidity of solution above their T_t , were higher when the concentration of the solution was raised. The same reason can explain this phenomenon: the higher the concentration of the solution, the more objects were formed and the more turbid the solution will be.

Figure S11. Turbidity assays at 350 nm with a rate of $1^\circ\text{C}/\text{min}$ of (A, B, C) ELP-[M_1V_3-i]-[I-20] ($i = 20, 40, 60$ respectively) in PBS.

Table S2. The linear dependence of T_t versus $\log(C)$ correlated to eq. 1, for ELP diblocks experimental data.

Monoblock name	T_t versus $\log(C)$	R^2
ELP-[M_1V_3-20]-[I-20]	$y = -2.45 \cdot \ln(x) + 32.29$	0.99
ELP-[M_1V_3-40]-[I-20]	$y = -1.94 \cdot \ln(x) + 30.05$	0.98
ELP-[M_1V_3-60]-[I-20]	$y = -1.49 \cdot \ln(x) + 27.93$	0.99
ELP-[$*M_1V_3-20$]-[I-20]	$y = -4.34 \cdot \ln(x) + 59.37$	0.99
ELP-[$*M_1V_3-40$]-[I-20]	$y = -3.07 \cdot \ln(x) + 53.29$	0.99
ELP-[$*M_1V_3-60$]-[I-20]	$y = -1.87 \cdot \ln(x) + 47.06$ ($T_{t\text{bulk}}$)	0.99
ELP-[$*M_1V_3-60$]-[I-20]	$y = -1.99 \cdot \ln(x) + 41.35$ (CMT)	0.95
ELP-[M_1V_3-40]-[I-90]	$y = -0.61 \cdot \ln(x) + 17.22$	0.90
ELP-[M_1V_3-60]-[I-90]	$y = -0.54 \cdot \ln(x) + 16.82$	0.91
ELP-[$*M_1V_3-40$]-[I-90]	$y = -0.38 \cdot \ln(x) + 26.22$ ($T_{t\text{bulk}}$)	0.65
ELP-[$*M_1V_3-40$]-[I-90]	$y = -0.76 \cdot \ln(x) + 19.99$ (CMT)	0.89
ELP-[$*M_1V_3-60$]-[I-90]	$y = -0.34 \cdot \ln(x) + 31.68$ ($T_{t\text{bulk}}$)	0.94
ELP-[$*M_1V_3-60$]-[I-90]	$y = -0.39 \cdot \ln(x) + 19.78$ (CMT)	0.93

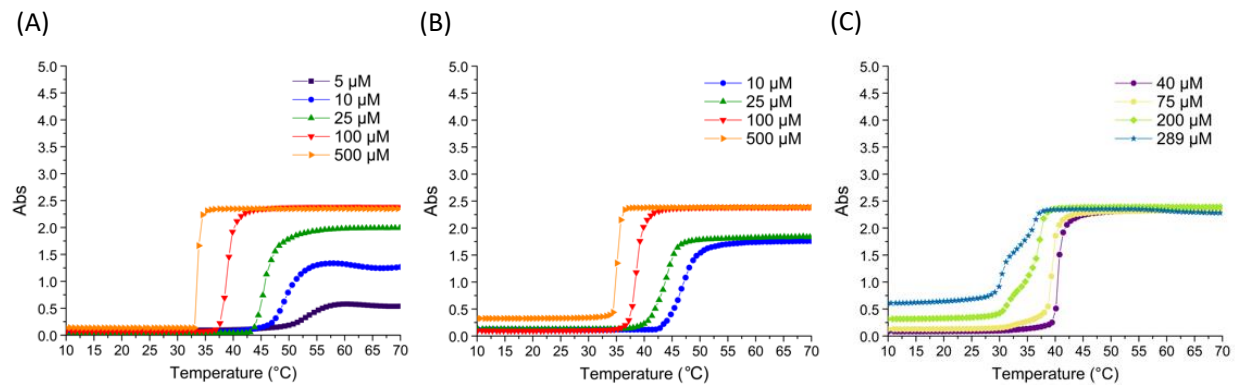


Figure S12. Turbidity assays at 350 nm with a rate of $1^\circ\text{C}/\text{min}$ of (A, B, C) ELP-[$*M_1V_3-i$]-[I-20] ($i = 20, 40, 60$ respectively) in ultrapure water.

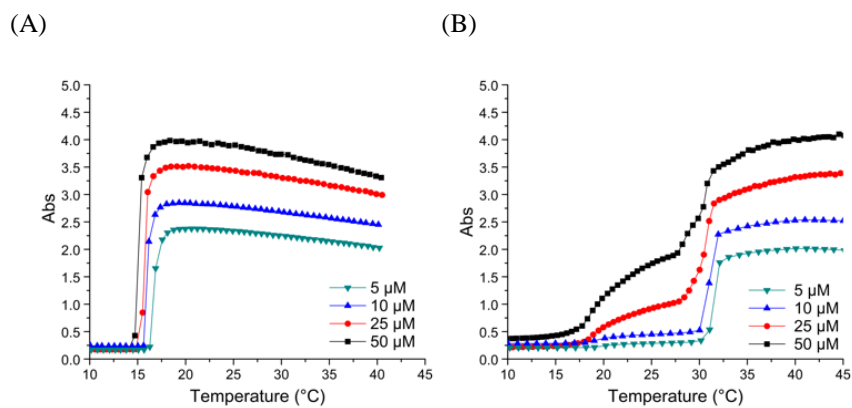


Figure S13. Turbidity assays at 350 nm with a rate of 1°C/min of (A) ELP-[M₁V₃-60]-[I-90] and (B) its oxidized counterpart, ELP-[*M₁V₃-60]-[I-90], in PBS.

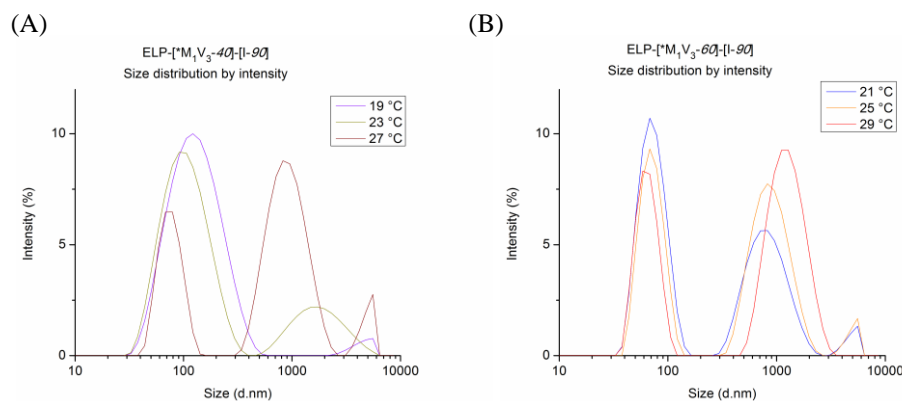


Figure S14. Size distribution by intensity of (A, B) ELP-[M₁V₃-*i*]-[I-90] (*i* = 40, 60, respectively) at 10 μM in PBS determined *via* DLS in the particle formation regime.