

## Sequence Isomerism in Uniform Polyphosphoesters Programmes Self-Assembly and Folding

*Nadeema Appukutti*<sup>1</sup>, *Joseph R. Jones*<sup>2</sup> and *Christopher J. Serpell*<sup>1\*</sup>

<sup>1</sup> School of Physical Sciences, Ingram Building, University of Kent, Canterbury, Kent, CT2 7NH, UK

<sup>2</sup> Department of Chemistry, University of Warwick, Gibbet Hill, Coventry CV4 7AL

\*Email: C.J.Serpell@kent.ac.uk / Twitter: @CJSerpell

### Supplementary Information

#### Contents

Materials, instrumentation, and general considerations	2
Synthesis of sequence-defined polyphosphoesters	3
Mass spectrometry	3
Fluorescence and DLS data	6
Supplementary transmission electron microscopy images	12
Supplementary atomic force microscopy images	18
Light scattering data and analysis	22
Table of self-assembled structures	28
References	29

## Materials, instrumentation, and general considerations

Spacer-CE Phosphoramidite C12 (**C12**), Spacer-CE Phosphoramidite 18 (**HEG**) and all synthesiser reagents were purchased from LinkTech, and the universal 1000 Å CPG solid support was purchased from GeneCust. All other solvents, reagents, and buffer components were purchased from Fisher.

TAMg buffer is composed of 45 mM Tris and 12.5 mM Mg(OAc)<sub>2</sub>·6H<sub>2</sub>O with pH adjusted to 8.0 using glacial acetic acid. TBE buffer is 90 mM Tris, 90 mM boric acid and 1.1 mM EDTA with a pH of 8.0.

**PAGE:** For gel electrophoresis, the polymers were loaded (100 μM, 10 μL, to which 10 μL of 8 M urea was added) on to 20 % polyacrylamide gel (TBE buffer plus 2.4 M urea) and electrophoresis carried out at room temperature for 1 hour at constant current of 15 mA. The gels were stained to visualize the individual bands created by electrophoresis with methylene blue stain.

**Dynamic light scattering:** The self-assembled structures of sequenced polymers in liquid environments were investigated using DLS with the ZetaSizer Nano-ZS by Malvern Instruments Limited. Samples were prepared in TBE and TAMg buffer (100 μL of 100 μM). For the dilution studies, 100 μL of 10 μM was prepared by diluting 100 μM sample 10x with 1x TBE and 1x TAMg buffer, and 500 μL of 1 μM was prepared by diluting the 10 μM samples with buffer. All the measurements were carried out at 25 °C and measurements were repeated three times in order to check their reproducibility.

**Fluorescence:** The sequence-defined polymer samples at 100 μM, 10 μM and 1 μM (100 μL, 100 μL and 500 μL respectively) were prepared, from stocks in pure water, in TBE buffer and in TAMg buffer, and 4.5 μL of Nile Red working solution (0.001 M, acetone) was added for 100 μL samples, and 22.5 μL for 500 μL samples. The samples were vortexed briefly and fluorescence spectra were recorded using an Agilent Cary Eclipse fluorescence spectrophotometer at room temperature in low volume quartz cuvettes. All the spectra were taken using an excitation wavelength of 535 nm, and monitoring emission between 560 and 750 nm. The samples were sealed and incubated overnight at room temperature in the absence of light before recording the 24 hour timepoint.

**TEM:** Samples for TEM were prepared in TBE and TAMg buffer and left overnight for self-assembly to stabilise, and then dropped onto a carbon coated copper 200 mesh TEM grids and negatively stained using 2% uranyl acetate (5 μL, 30 seconds). The stain was wicked off. Samples were imaged on a Jeol 1230 TEM, operating at an accelerating voltage of 80 kV and the images were recorded with a Gatan Multiscan 790 digital camera.

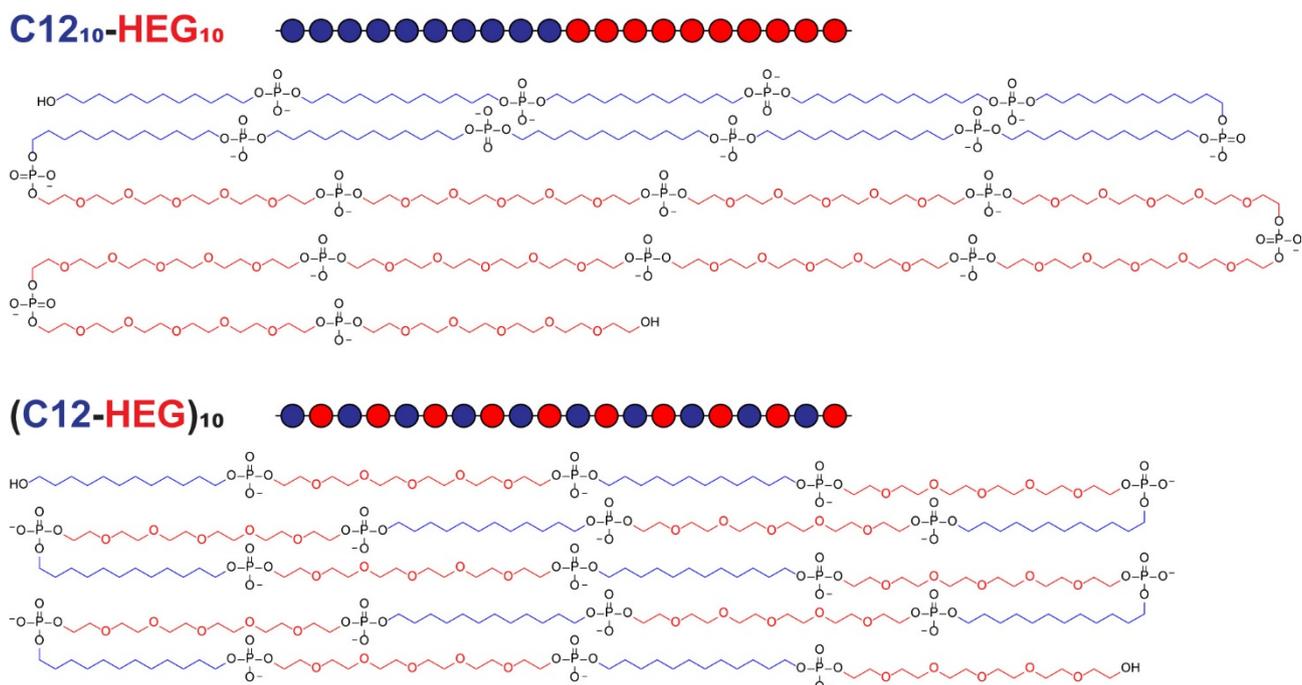
**AFM:** Samples for AFM were prepared in TBE and TAMg buffer (100 μM, 5 μL). Samples were deposited on freshly cleaved mica sheets for one minute, followed by 5 water washes with 5 μL of water. Excess liquid was blotted off with the edge of a filter paper. Prepared mica plates were dried under vacuum for one hour. Measurements were conducted on a Bruker Multi-Mode microscope with a Quadrex Nanoscope III controller using Bruker ScanAsyst-Air silicon tip on nitride lever, with frequency of 70 Hz. Data were analyzed using Nanoscope Analysis 1.5 software (Bruker, CA, US).

**Scattering experiments:** Scattering experiments were conducted for samples of C12<sub>10</sub>-HEG<sub>10</sub> only (using both TBE and TAMg buffers) since this polymer appeared to produce particles and assemblies of far greater mass than (C12-HEG)<sub>10</sub> and so would be expected to effect far greater scattering per sample concentration.

**Small angle X-ray scattering (SAXS):** Small-angle X-ray scattering (SAXS) measurements were made using a Xenocs Xeuss 2.0 equipped with a micro-focus Cu K<sub>α</sub> source collimated with Scatterless slits. The scattering was measured using a Pilatus 300k detector with a pixel size of 0.172 mm x 0.172 mm. A radial integration as a function of scattering length,  $q$ , was performed on the 2-dimensional scattering profile and the resulting data corrected for absorption and background from the sample holder.

## Synthesis of sequence defined polyphosphoesters

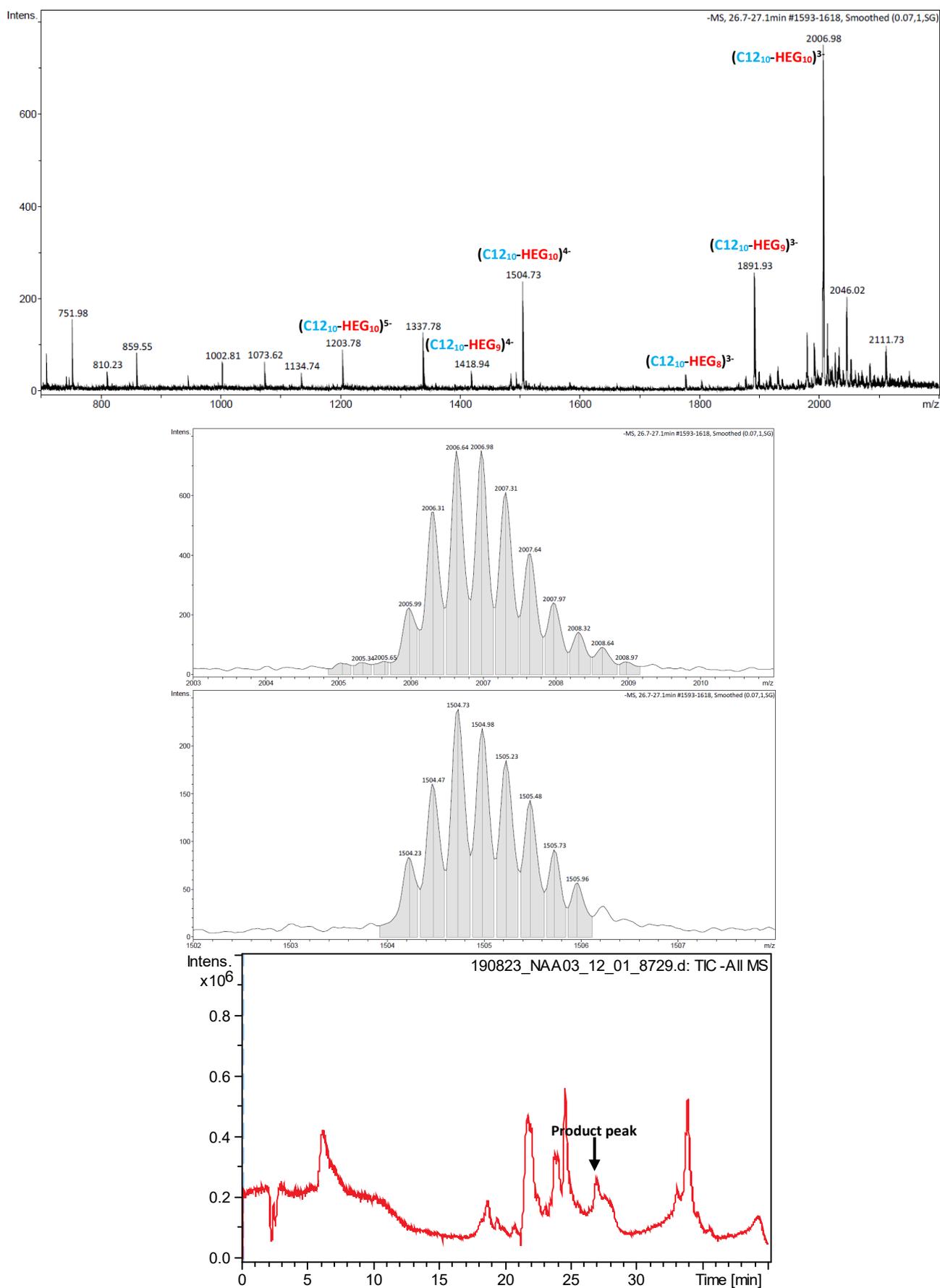
Using commercially purchased C12 and HEG phosphoramidite reagents, 20 monomer unit long polymer strands with different sequence patterns were synthesized, using automated solid phase synthesis. The automated solid-phase synthesis was performed on an ABI Expedite™ 8909 synthesiser on a 1  $\mu\text{mol}$  scale on a universal 1000 Å CPG solid support. The synthesis method was adapted from the supplier defaults only in the use of an extended (600 seconds) coupling time. Coupling efficiency was monitored automatically with the removal of dimethoxytrityl (DMT) 5'-OH protecting groups by 3% dichloroacetic acid in dichloromethane. The final DMT group was not removed from the completed strands. The polymers were cleaved and deprotected with ammonium hydroxide at 60 °C. The desalted, purified product was obtained by size exclusion chromatography with zetadex media. The final DMT group was then removed manually using 4:1 acetic acid:water. The collected DMT cation was used for quantification with Thermo Scientific™ NanoDrop™ One Spectrophotometer. The concentration was calculated using the calibration curve generated from DMT-Cl treated with the same acetic acid mix.



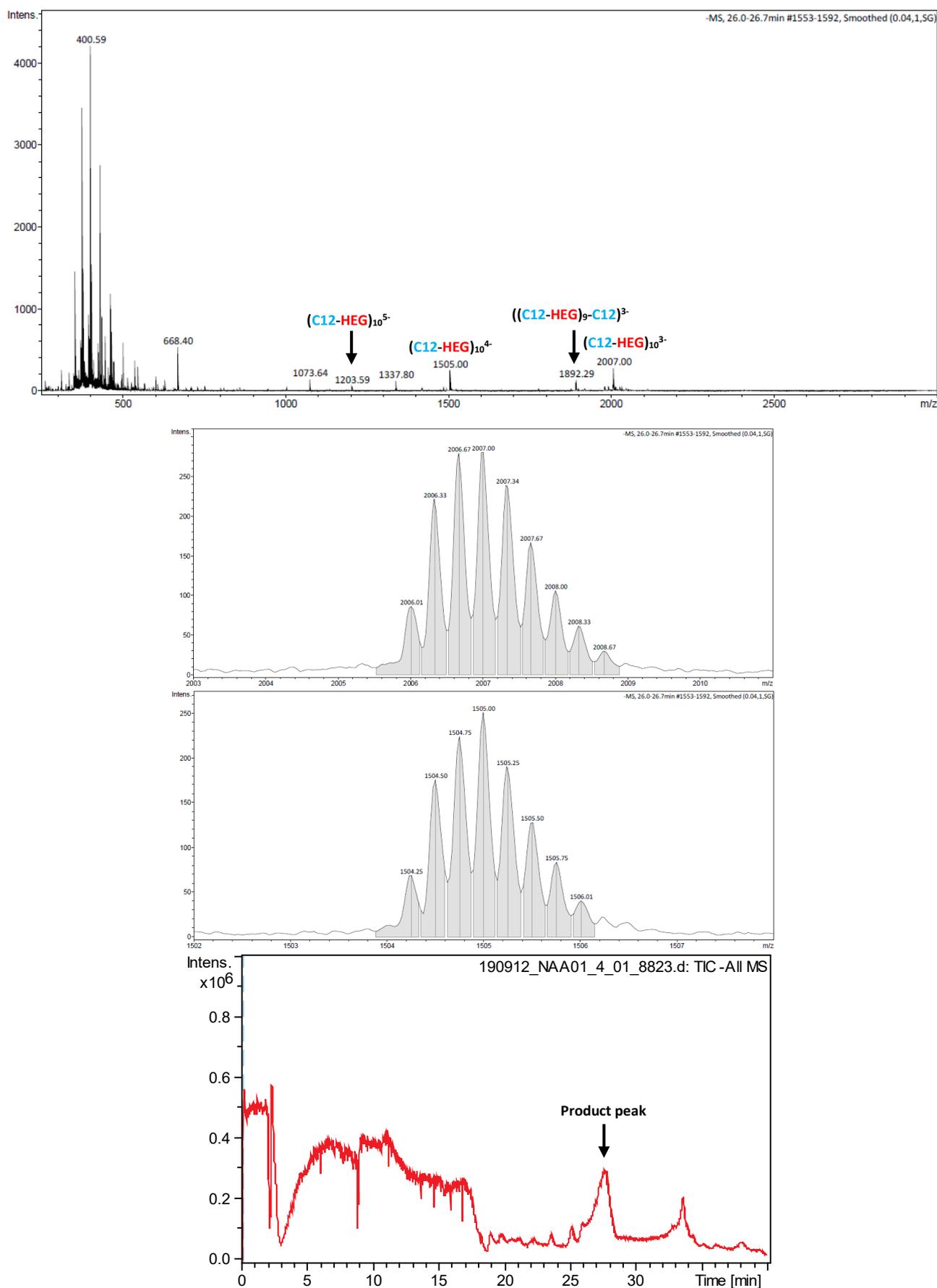
**Figure S1.** Alignment of abbreviations used against full chemical structures of the polymers synthesised

## Mass spectrometry

Electrospray mass spectra were recorded on a Bruker micrOTOF-Q II mass spectrometer. Samples were separated on-line by reverse-phase HPLC on a Phenomenex Nucleosil C18 column (3  $\mu\text{m}$ , 150 x 2.0 mm) running on an Agilent 1100 HPLC system at a flow rate of 0.2 mL/min using a water, 15 mM triethylamine, 400 mM hexafluoroisopropanol (HFIP), methanol gradient: mobile phase (A), 15 mM TEA, 400 mM HFIP in water; mobile phase (B), methanol; 10% (B) for 5 minutes, then a linear gradient 10 – 100% (B) in 25 minutes, then 100% (B) for 5 minutes, before returning to the initial conditions. The eluant was monitored at 260 nm and then directed into the electrospray source at -3.5 kV and mass spectra recorded from 250-3000 m/z. Data was analysed with Bruker's Compass Data Analysis software.

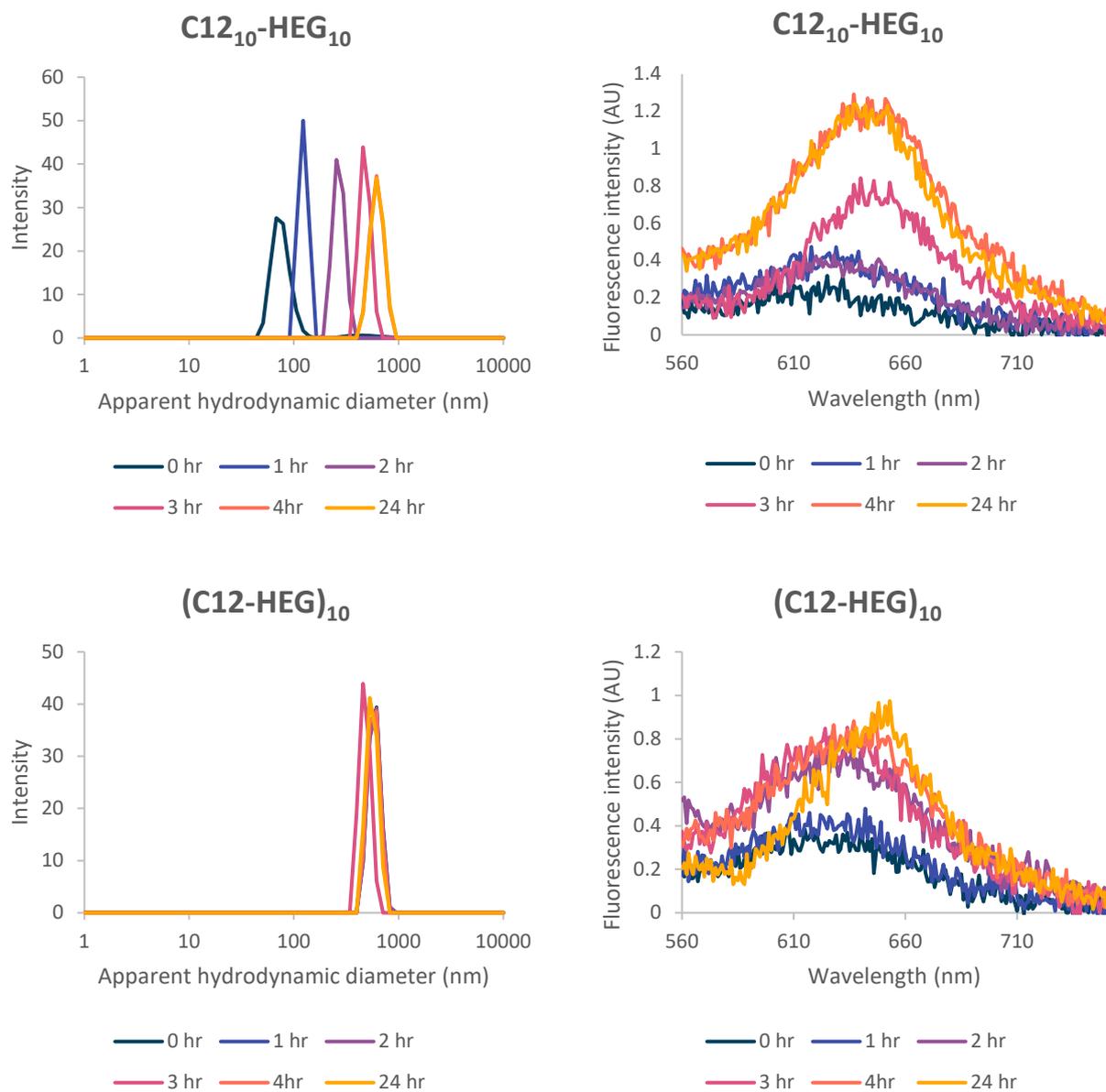


**Figure S2.** Mass spectrum of  $C_{12_{10}}-HEG_{10}$ , with expansions of the triply and quadruply charged peaks corresponding to a monoisotopic mass of 6020.93 Da (calculated exact mass is 6020.77 Da), and the total ion current chromatogram (TIC). No macromolecular products were observed at other retention times.

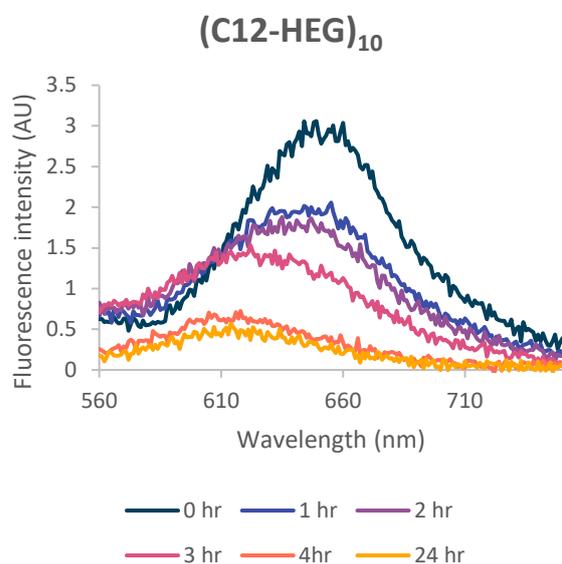
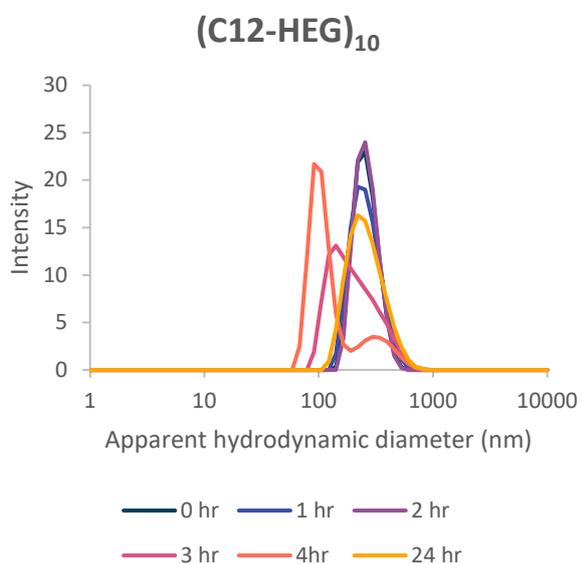
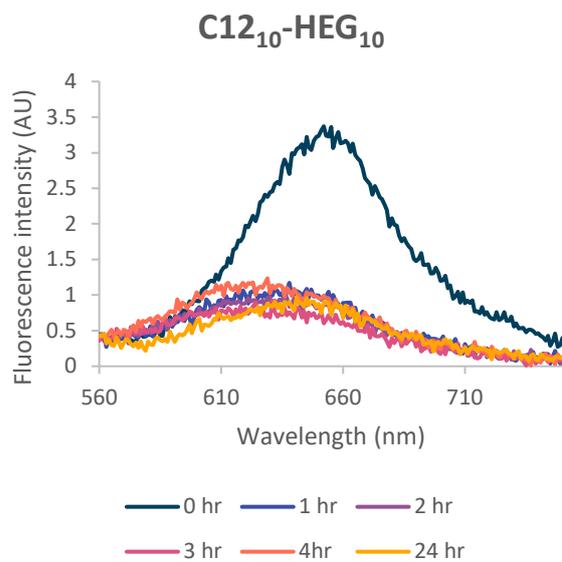
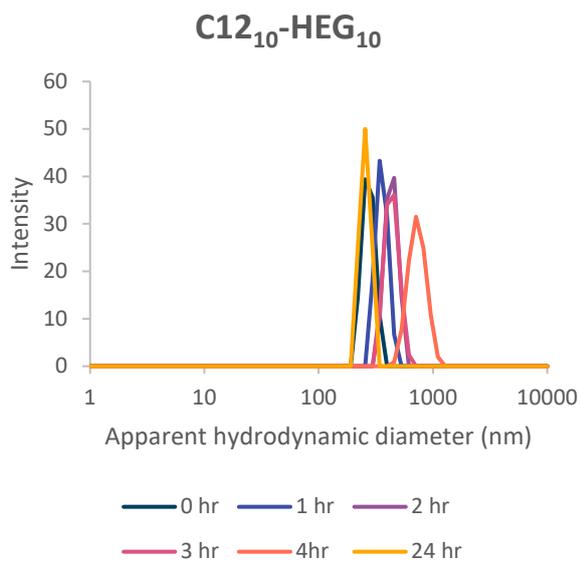


**Figure S3.** Mass spectrum of  $(\text{C12-HEG})_{10}$ , with expansions of the triply and quadruply charged peaks corresponding to a monoisotopic mass of 6021.00 Da (calculated exact mass is 6020.77 Da), and the total ion current chromatogram (TIC). No macromolecular products were observed at other retention times.

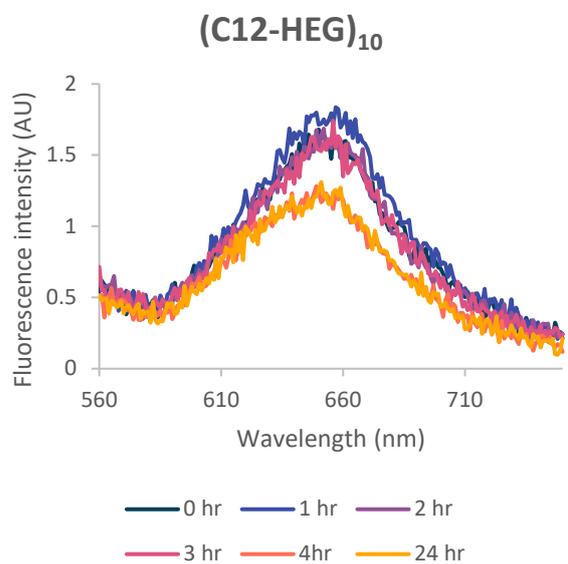
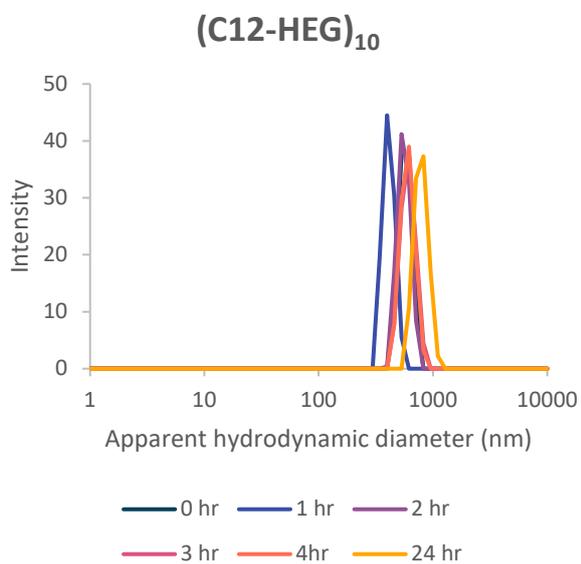
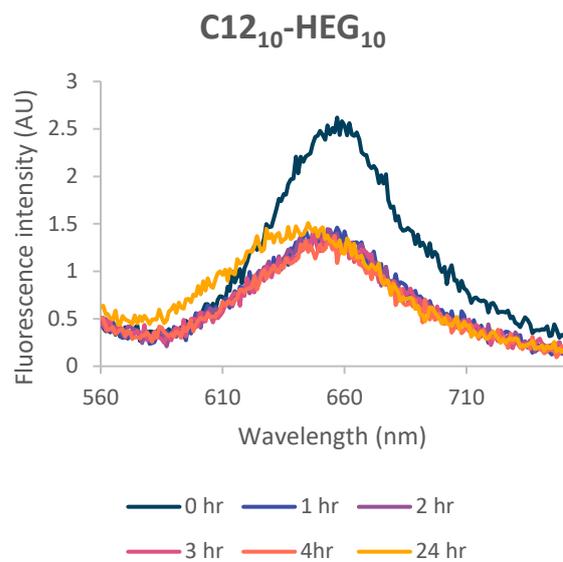
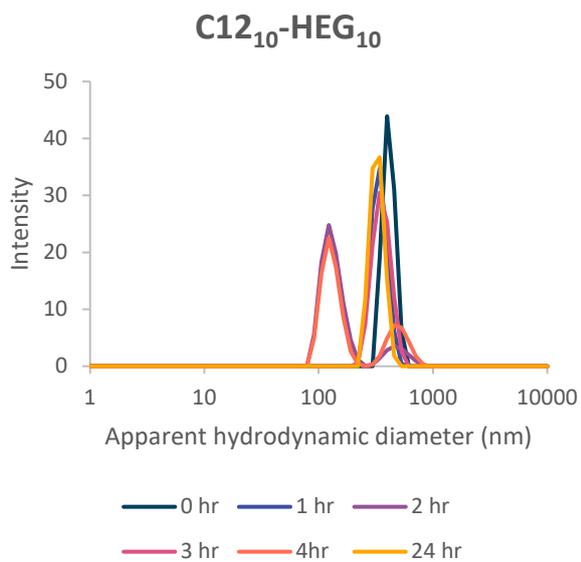
## Fluorescence and DLS data



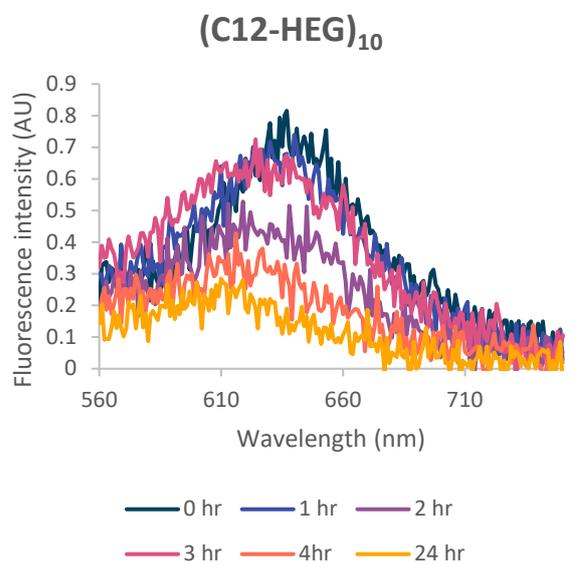
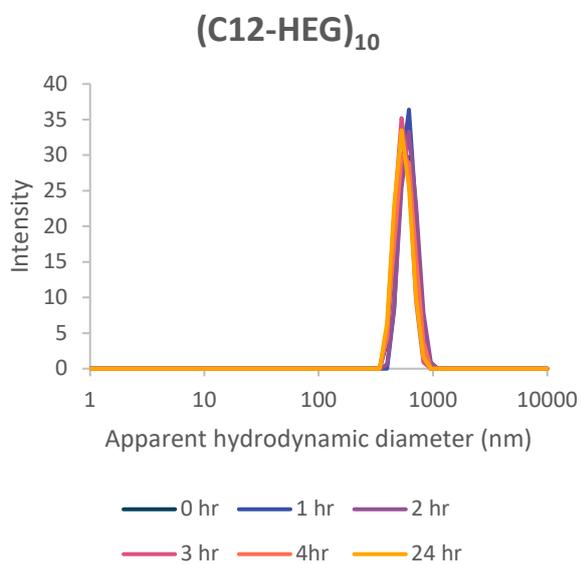
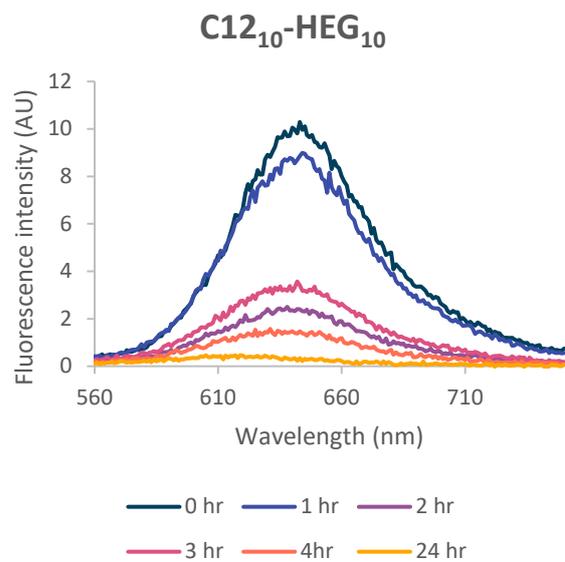
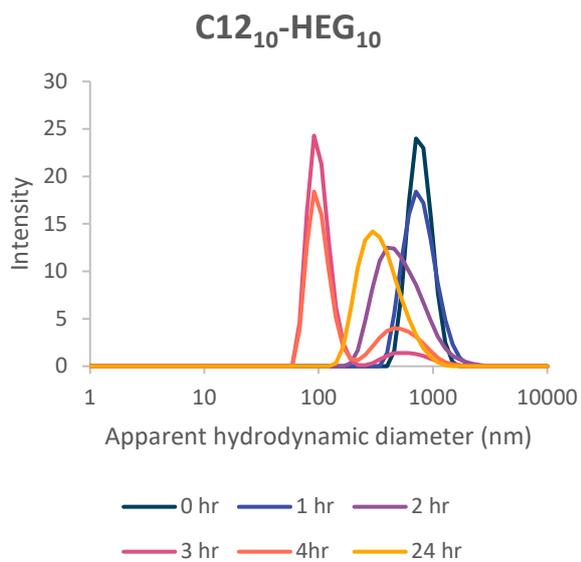
**Figure S4.** DLS (left) and NR emission measurements (right) on **C12<sub>10</sub>-HEG<sub>10</sub>** (top) and **(C12-HEG)<sub>10</sub>** (bottom) in TBE buffer at 100  $\mu$ M.



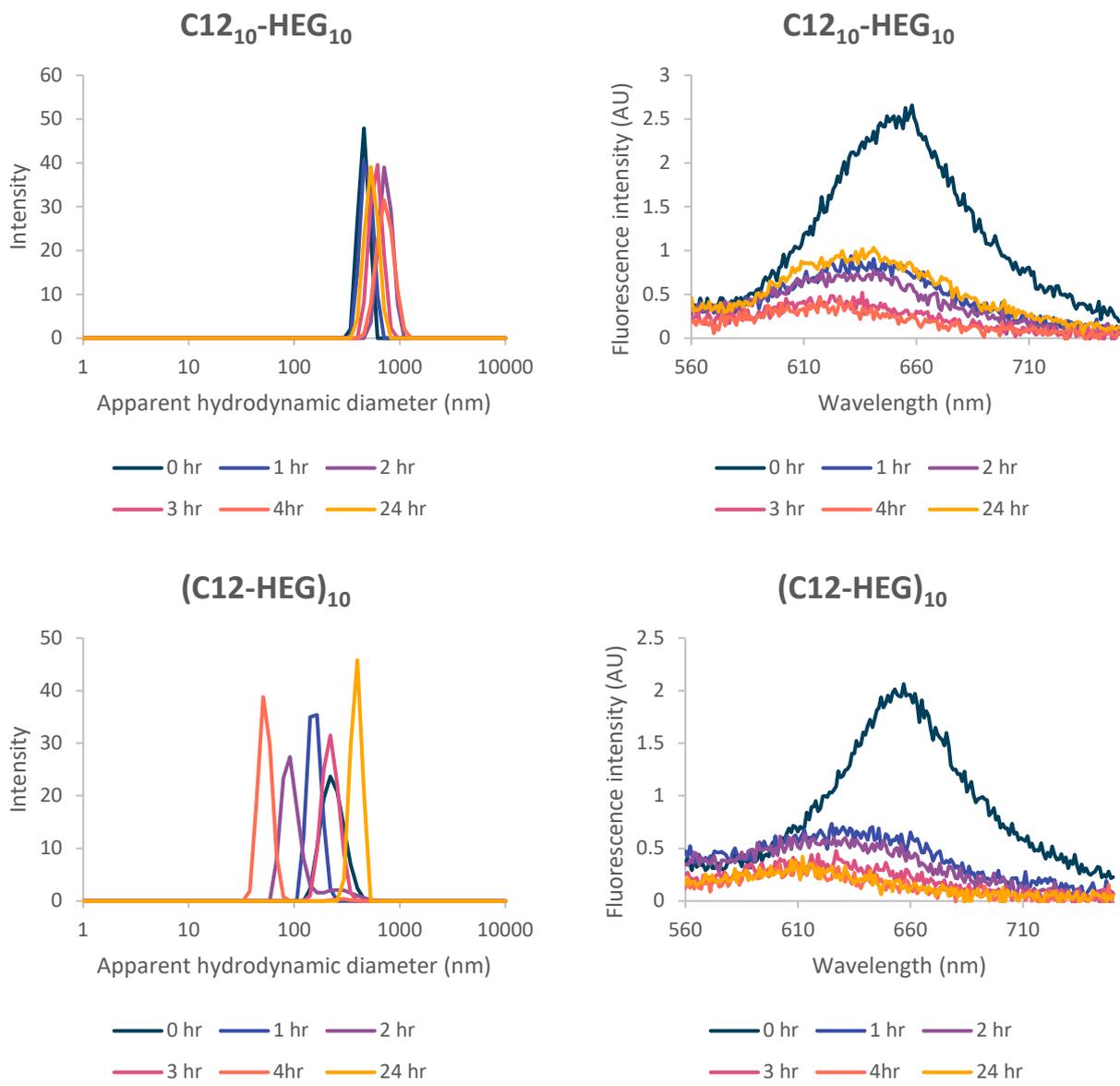
**Figure S5.** DLS (left) and NR emission measurements (right) on **C12<sub>10</sub>-HEG<sub>10</sub>** (top) and **(C12-HEG)<sub>10</sub>** (bottom) in TBE buffer at 10  $\mu$ M.



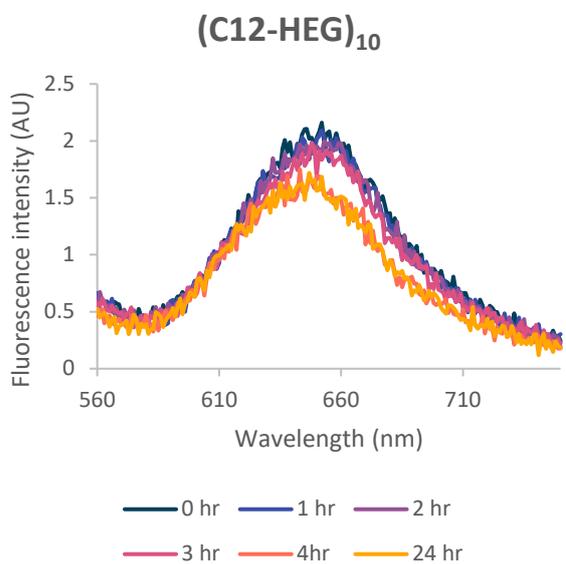
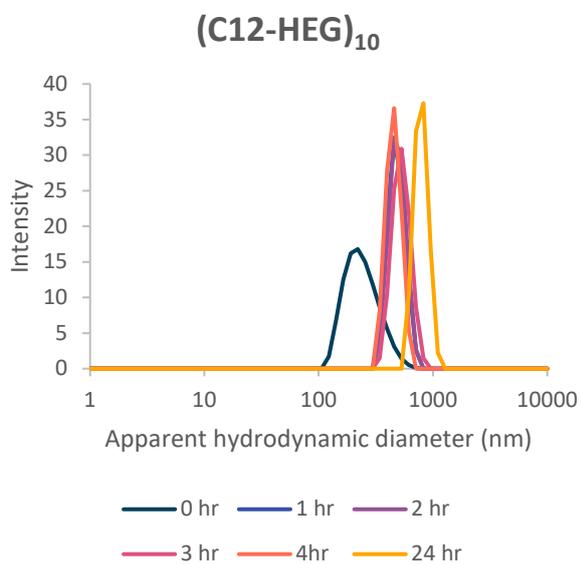
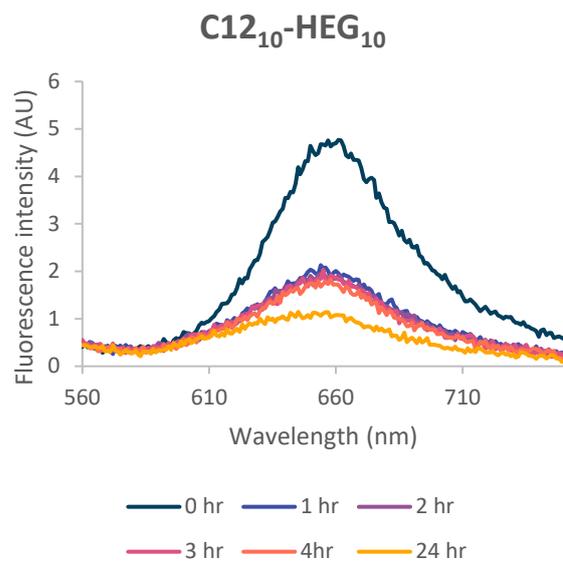
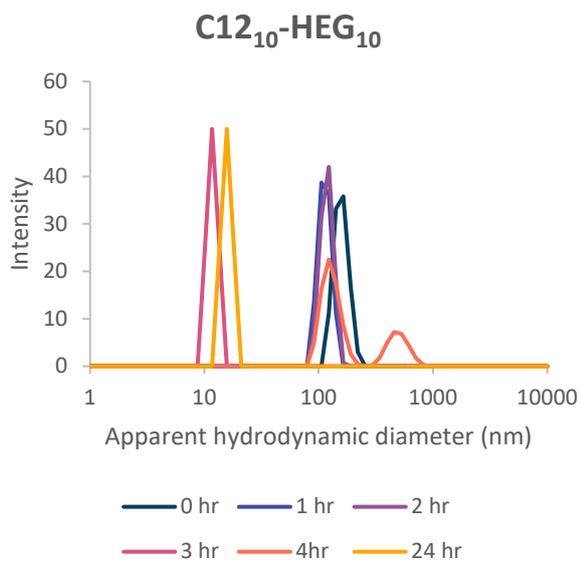
**Figure S6.** DLS (left) and NR emission measurements (right) on **C12<sub>10</sub>-HEG<sub>10</sub>** (top) and **(C12-HEG)<sub>10</sub>** (bottom) in TBE buffer at 1  $\mu$ M.



**Figure S7.** DLS (left) and NR emission measurements (right) on **C12<sub>10</sub>-HEG<sub>10</sub>** (top) and **(C12-HEG)<sub>10</sub>** (bottom) in TAMg buffer at 100  $\mu$ M.

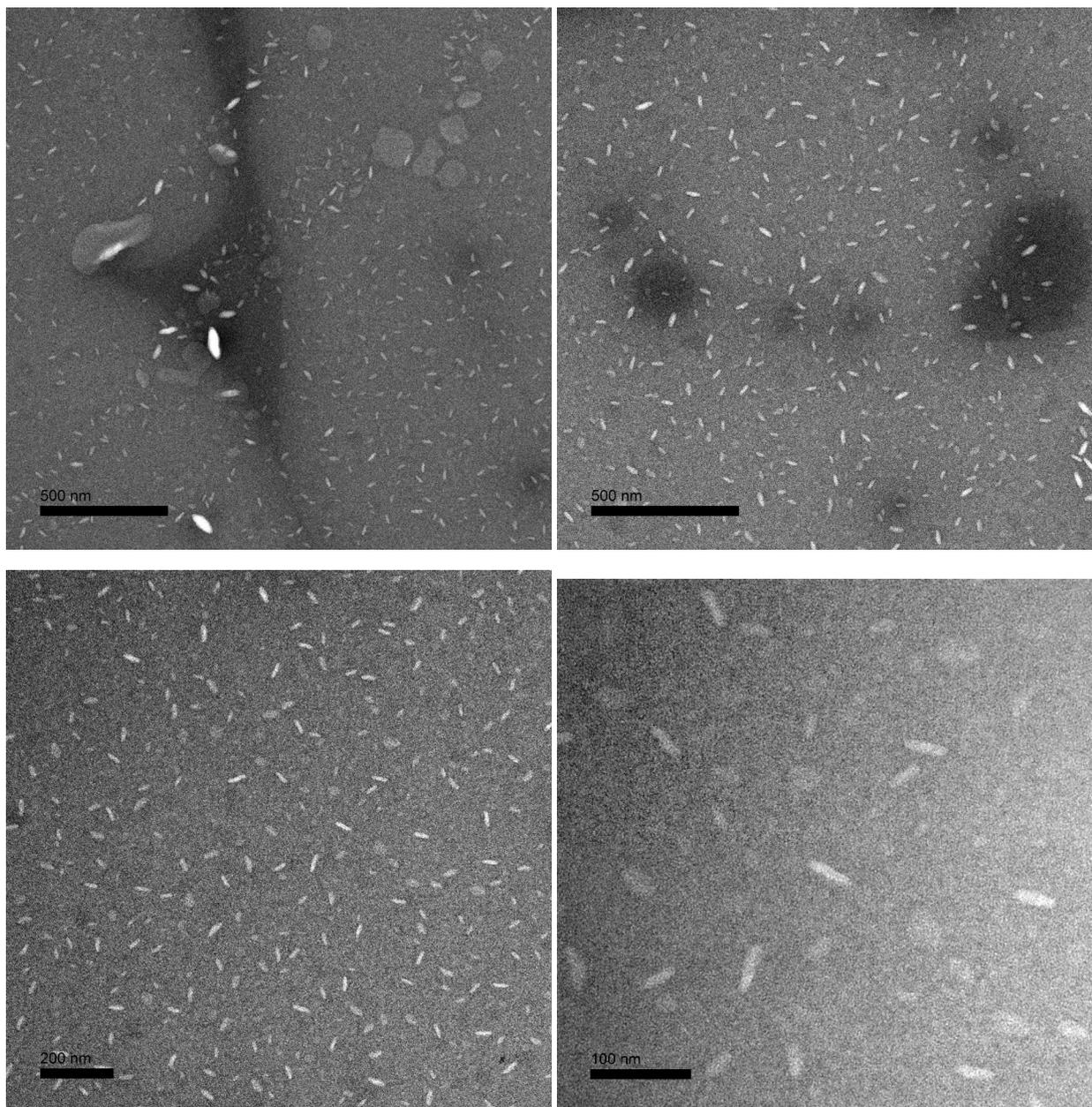


**Figure S8.** DLS (left) and NR emission measurements (right) on  $C12_{10}\text{-HEG}_{10}$  (top) and  $(C12\text{-HEG})_{10}$  (bottom) in TAMg buffer at  $10\ \mu\text{M}$ .

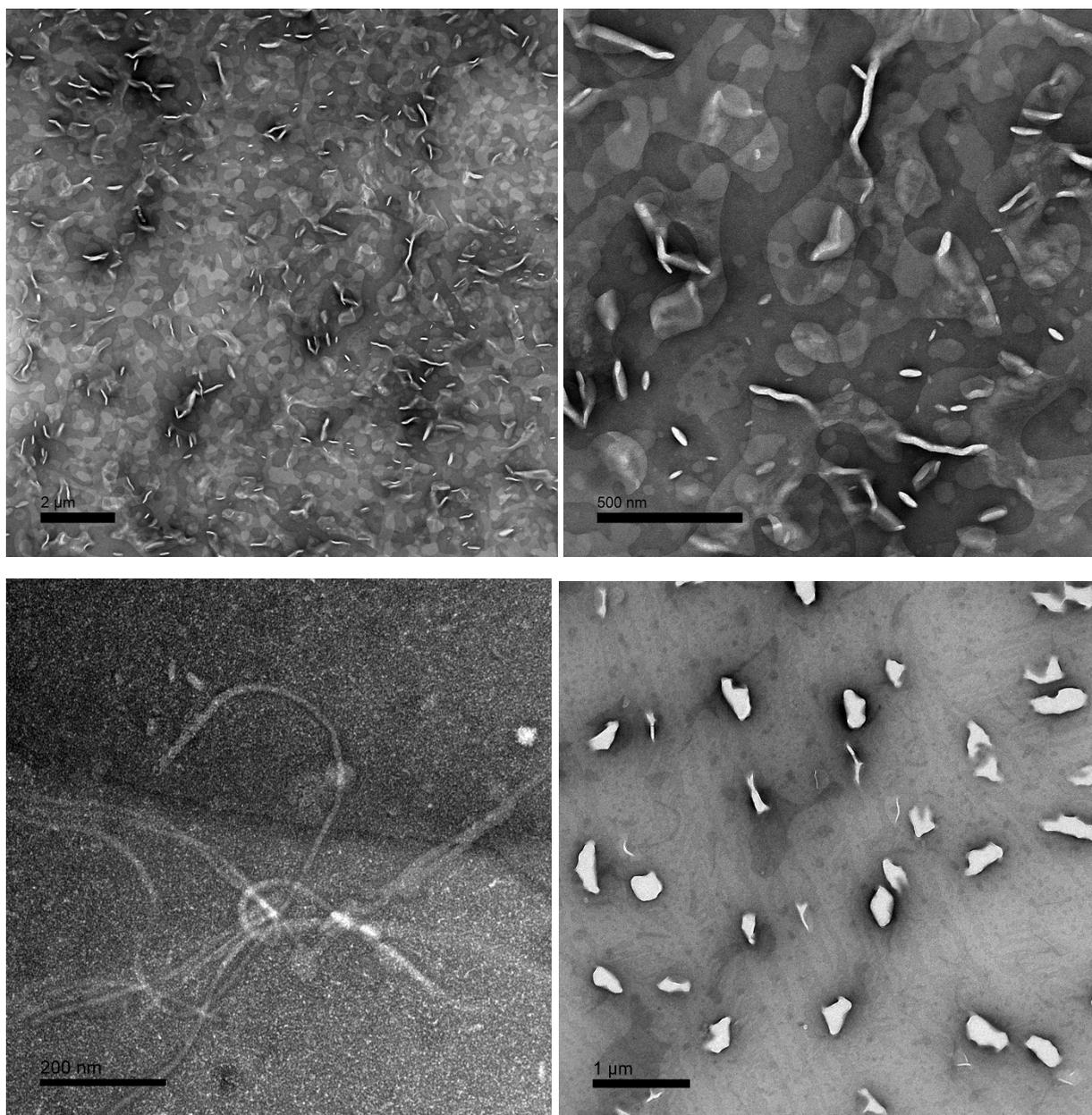


**Figure S9.** DLS (left) and NR emission measurements (right) on **C12<sub>10</sub>-HEG<sub>10</sub>** (top) and **(C12-HEG)<sub>10</sub>** (bottom) in TAMg buffer at 1  $\mu$ M.

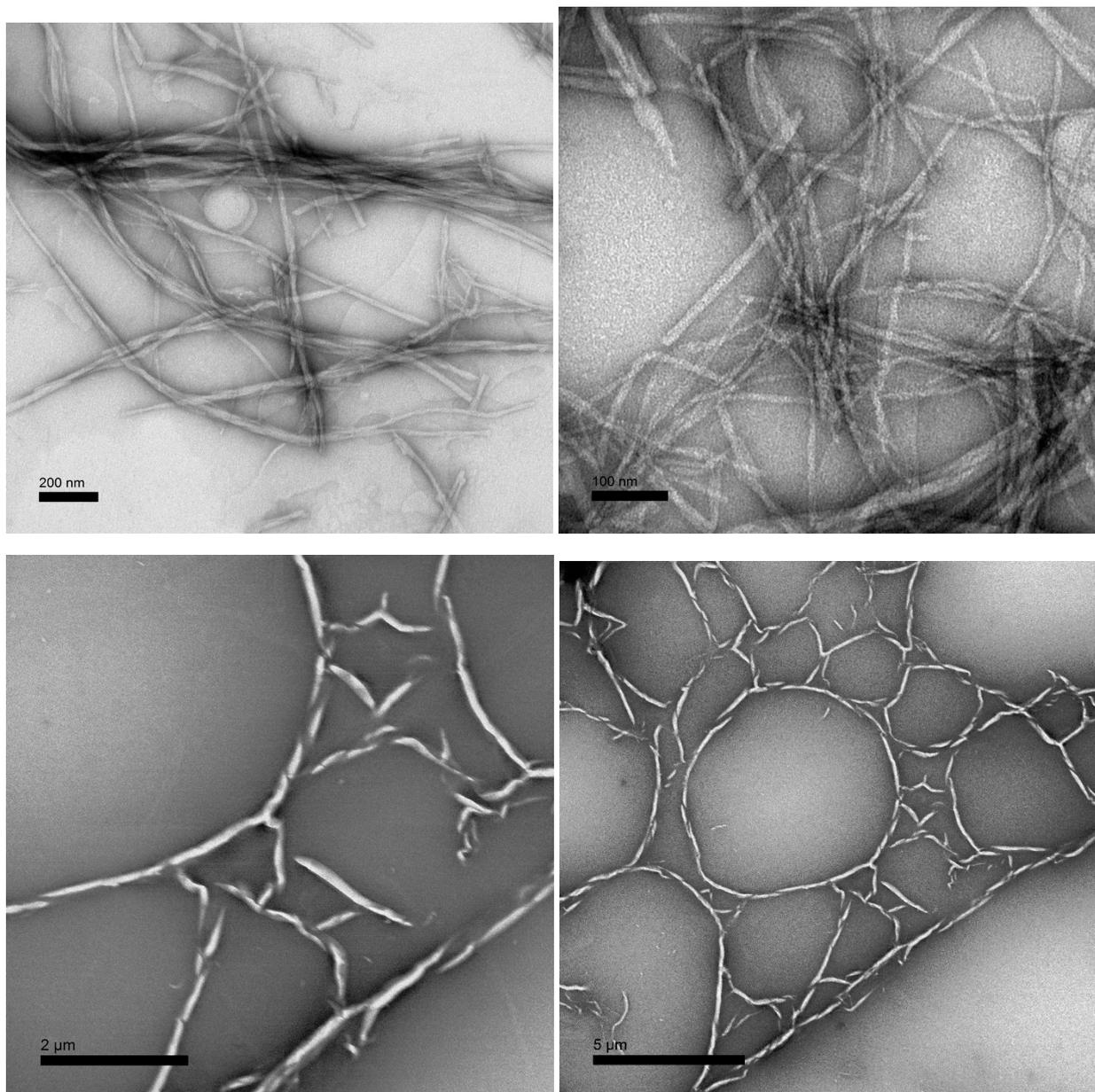
## Supplementary transmission electron microscopy images



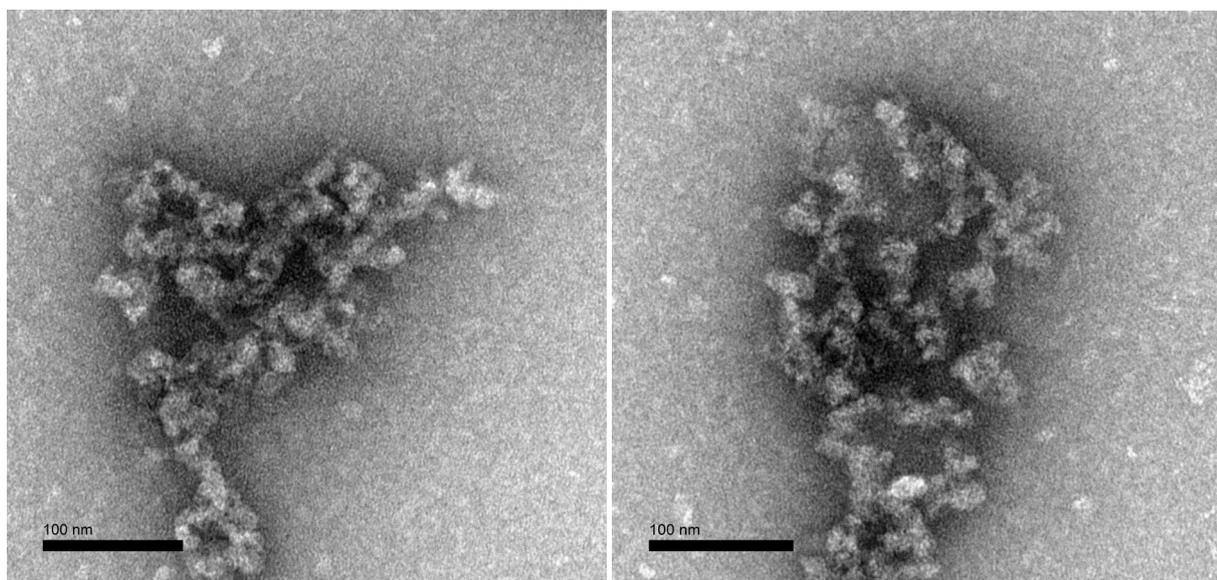
**Figure S10.** TEM images of C12<sub>10</sub>-HEG<sub>10</sub>, 1 μM, TBE buffer.



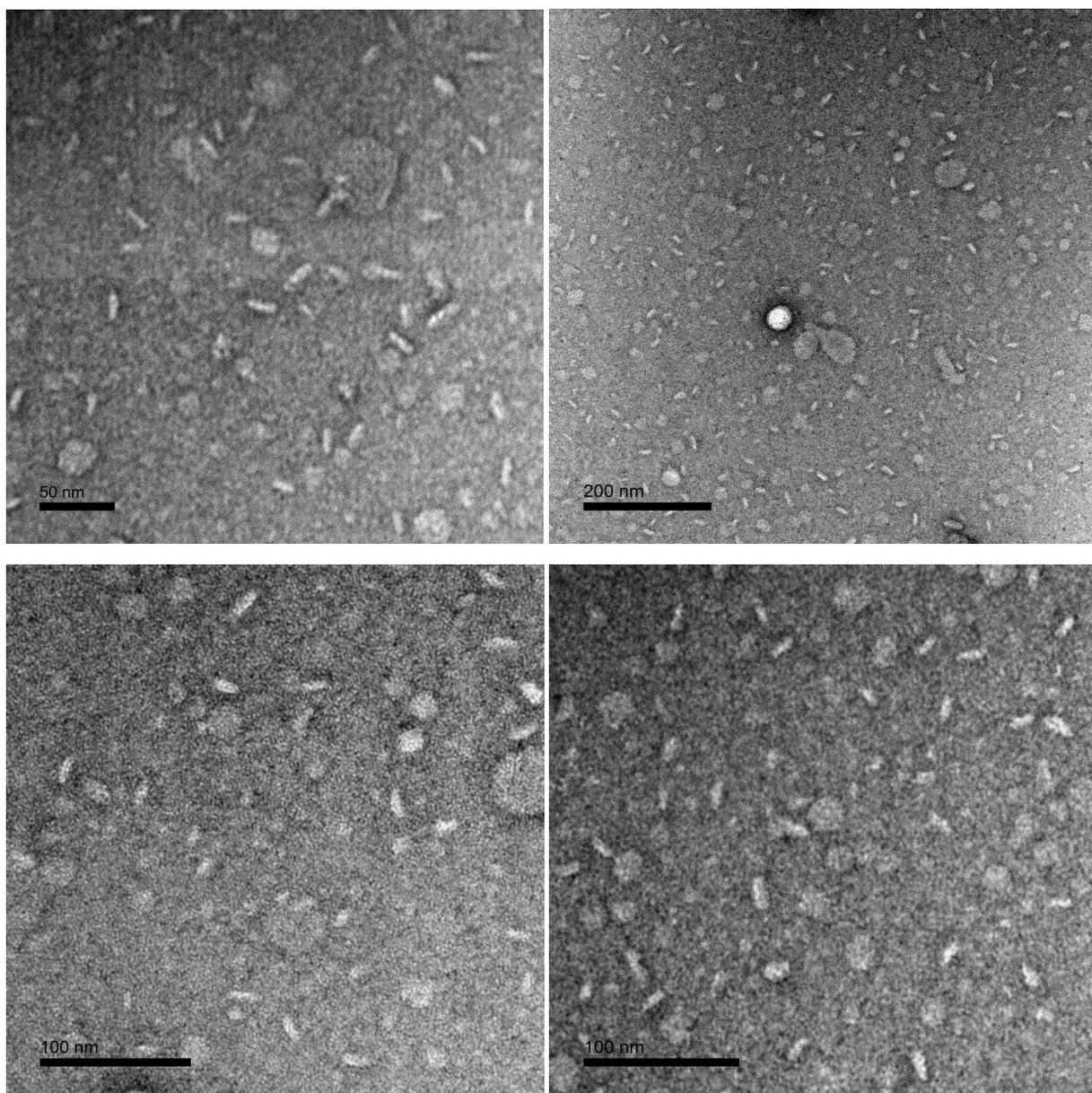
**Figure S11.** TEM images of C12<sub>10</sub>-HEG<sub>10</sub>, 10 μM, TBE buffer.



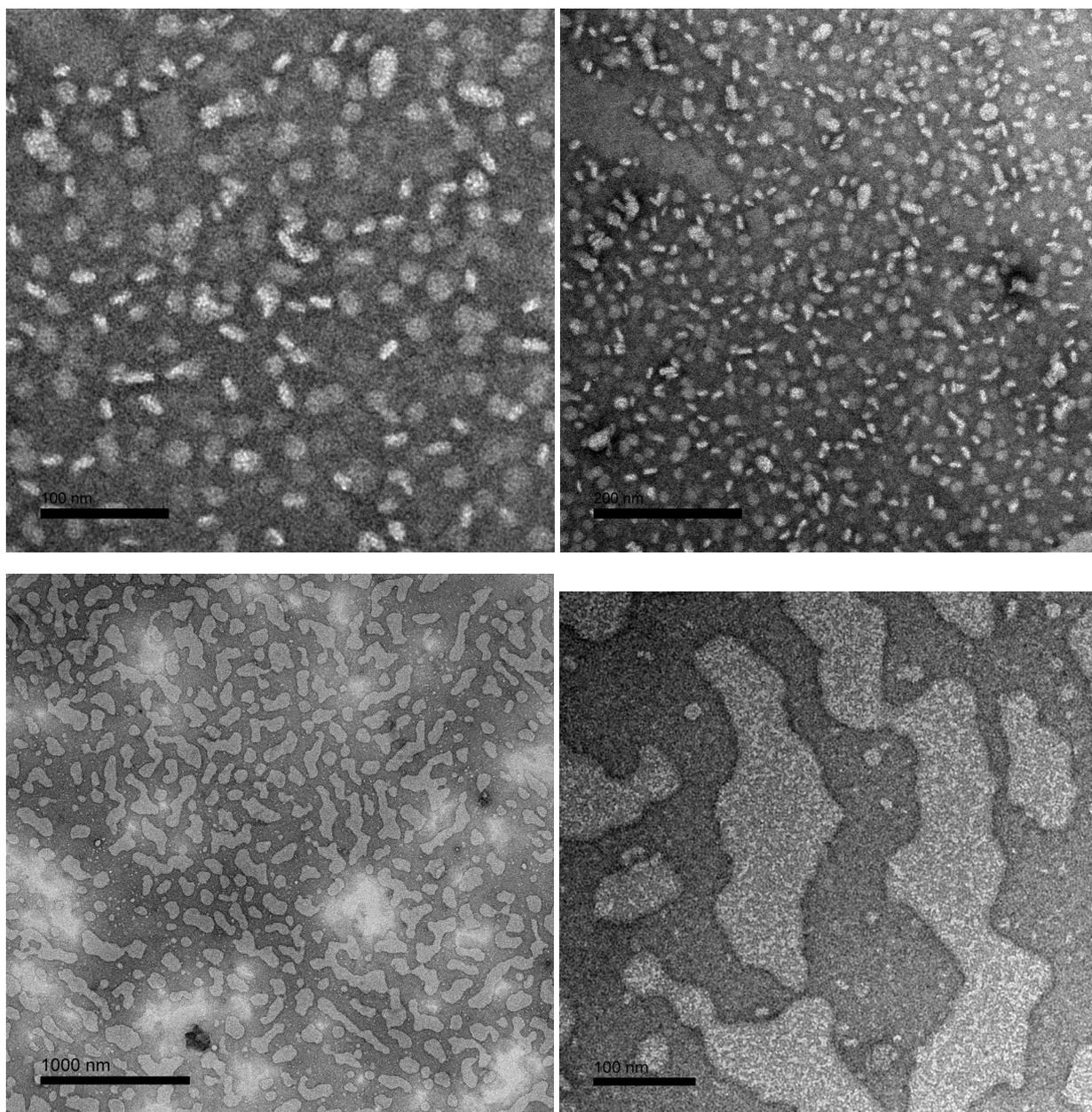
**Figure S12.** TEM images of **C12<sub>10</sub>-HEG<sub>10</sub>**, 100 μM, TBE buffer.



**Figure S13.** TEM images of **(C12-HEG)<sub>10</sub>**, 100  $\mu$ M, TBE buffer.

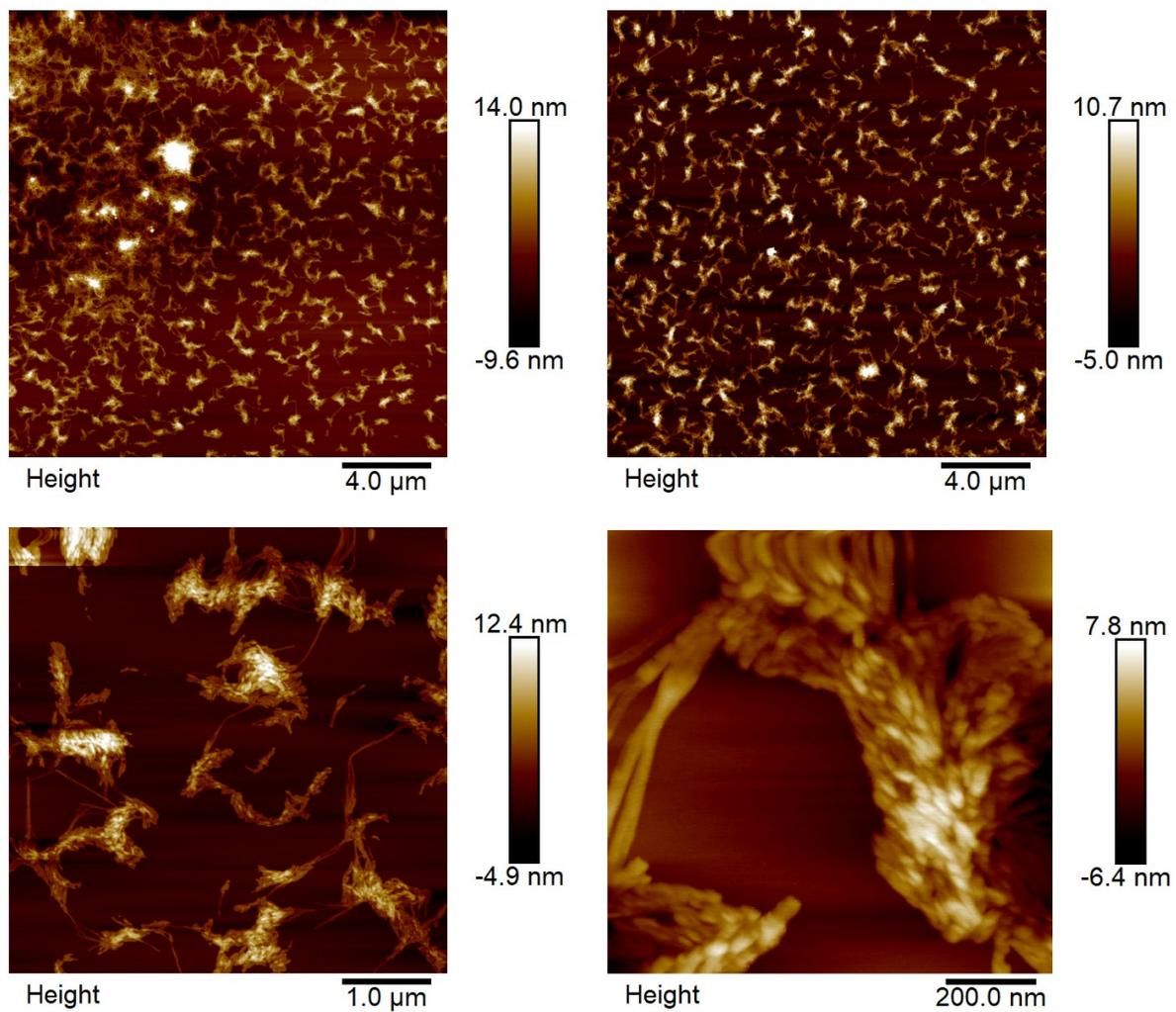


**Figure S14.** TEM images of **C12<sub>10</sub>-HEG<sub>10</sub>**, 100  $\mu$ M, TAMg buffer.

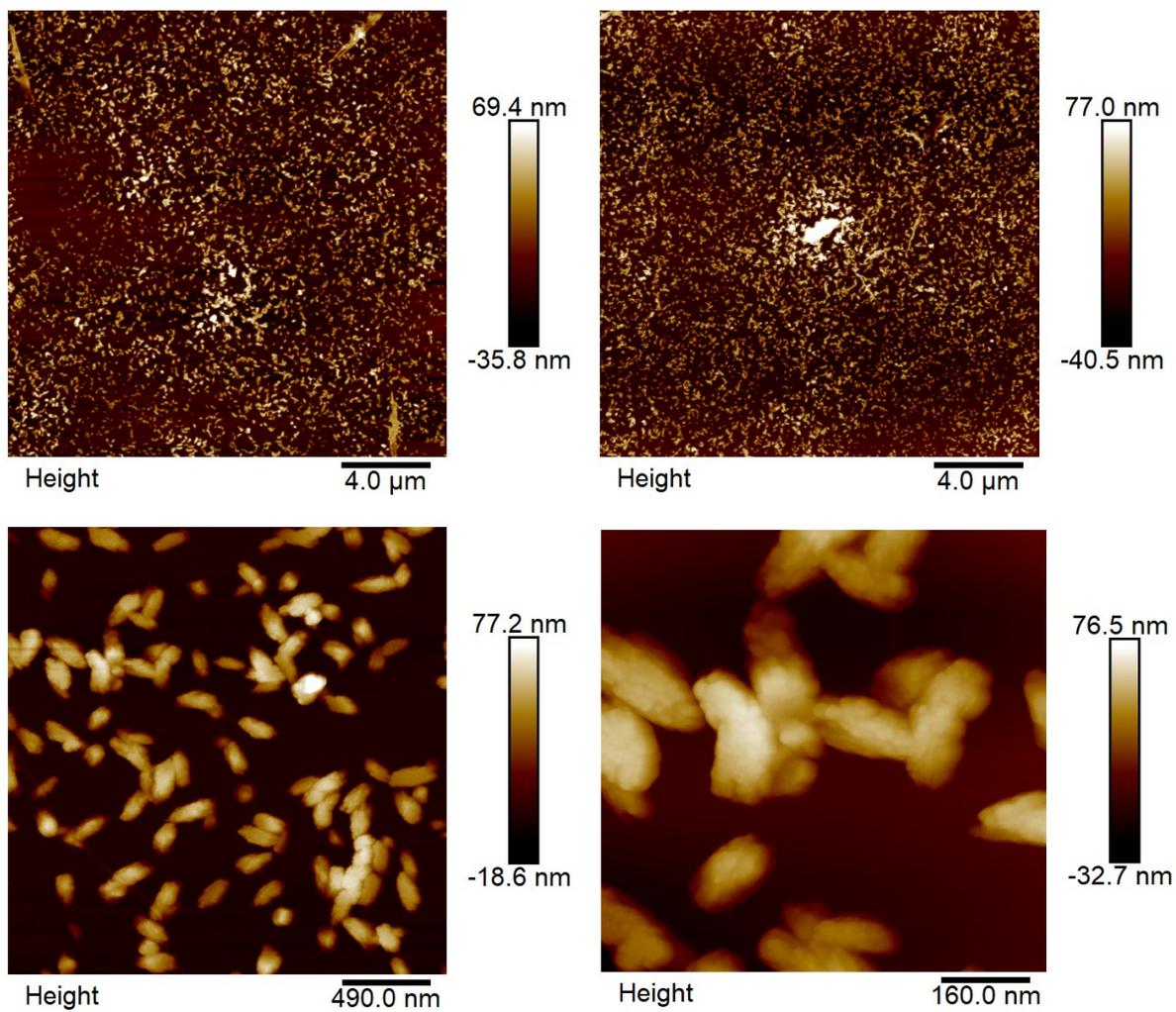


**Figure S15.** TEM images of (C12-HEG)<sub>10</sub>, 100 μM, TAMg buffer.

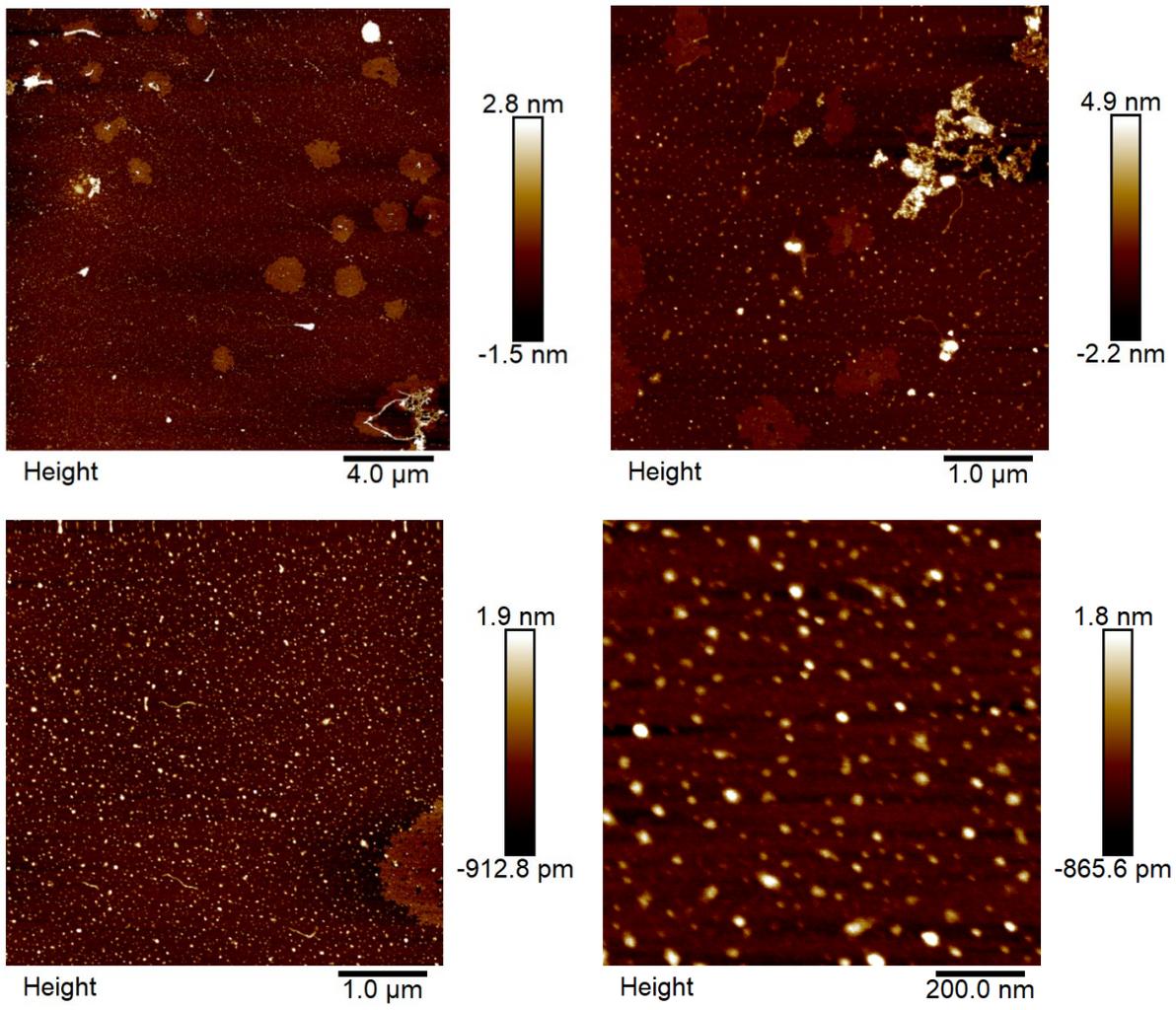
## Supplementary atomic force microscopy images



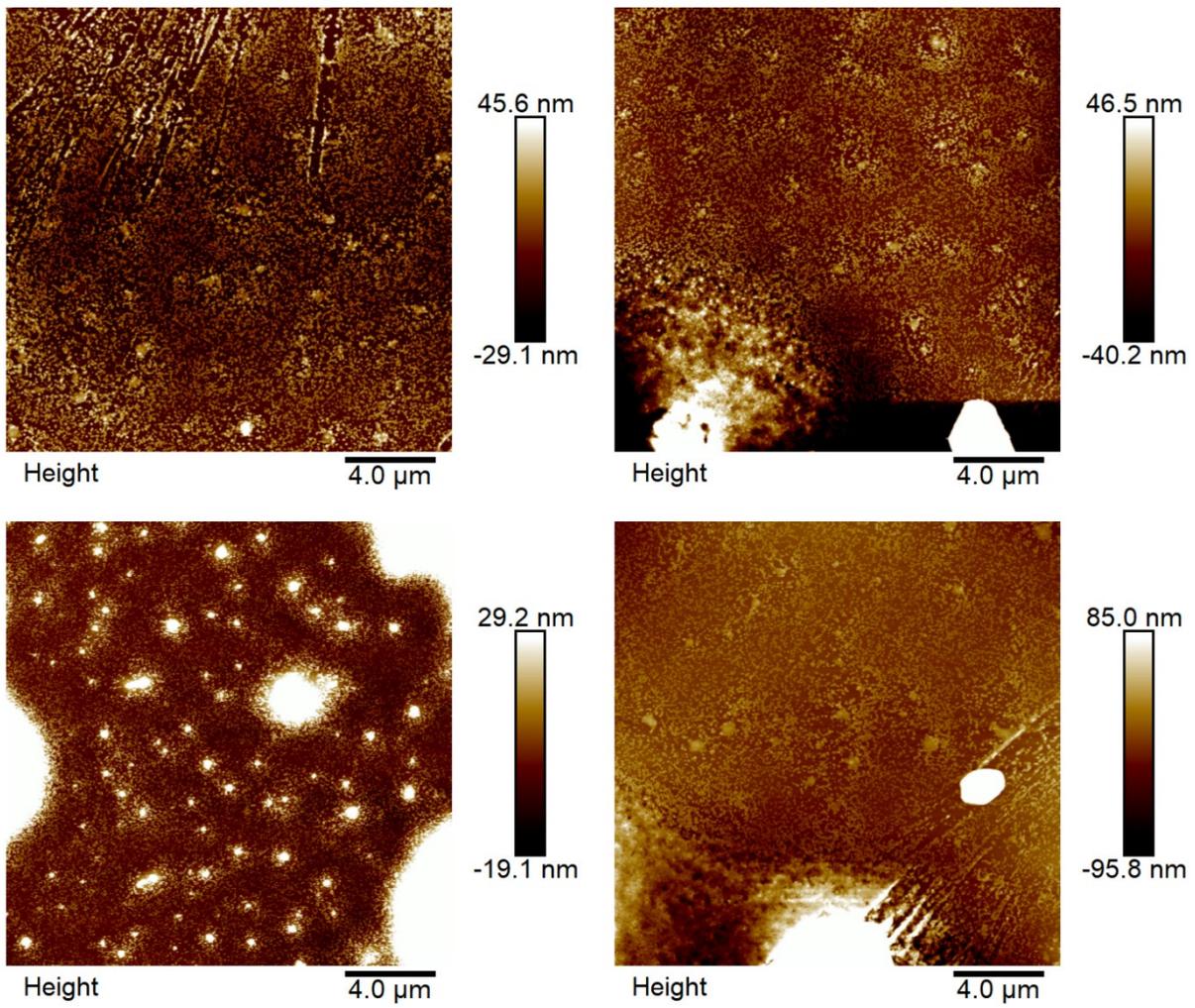
**Figure S16.** AFM images of C12<sub>10</sub>-HEG<sub>10</sub>, 100 μM, TBE buffer.



**Figure S17.** AFM images of C12<sub>10</sub>-HEG<sub>10</sub>, 100 μM, TAMg buffer.



**Figure S18.** AFM images of C12<sub>10</sub>-HEG<sub>10</sub>, 10 μM, TAMg buffer.



**Figure S19.** AFM images of (C12-HEG)<sub>10</sub>, 100 μM, TAMg buffer.

## Analysis of static light scattering (SLS)

Scattering experiments were conducted using an ALV-CGS3 goniometer-based system supplied by ALV GmbH, operating a laser at wavelength  $\lambda = 633$  nm. Time averaged scattering intensity data were recorded over the angular range  $12 < \theta < 150^\circ$  at  $1^\circ$  intervals.

Typically, scattering techniques use either standard linear plots (e.g. Zimm or Guinier plots) or employ least squares fitting to an appropriate model for estimating an intensity-weighted average particle size,  $\langle R_G \rangle_Z$ , or else in the case of highly anisotropic particles one uniform dimension---e.g. lamella thickness---might be identifiable at an appropriate length scale, even if the other dimension---e.g. lamella radius---were to be highly variable.<sup>1,2</sup> Alternatively, in the case of scattering from non-spherical objects that exhibit complex phase behaviour a generalised empirical model is available, which characterises different kinds of structure in terms of dimensionality parameters.<sup>3</sup>

Application of a generalised Guinier-Porod model to empirical data and interpretation of the fitted parameters are explained in detail by Hammouda.<sup>2</sup> Briefly, a minimal complexity empirical model may be formulated in terms of two dimensionality parameters,  $s$  and  $d$ , and a characteristic size,  $R$ .

The model takes the form:

$$I(q) = \left(\frac{G}{q^s}\right) \cdot \exp\left(-\frac{q^2 R^2}{3-s}\right) \quad , q \leq q^*$$

$$I(q) = \frac{D}{q^d} \quad , q \geq q^*$$

Where  $q^* \sim \frac{1}{R}$  and  $D$  and  $G$  are scaling factors calculated to make the function  $I(q)$  and its derivative continuous at  $q^*$ .

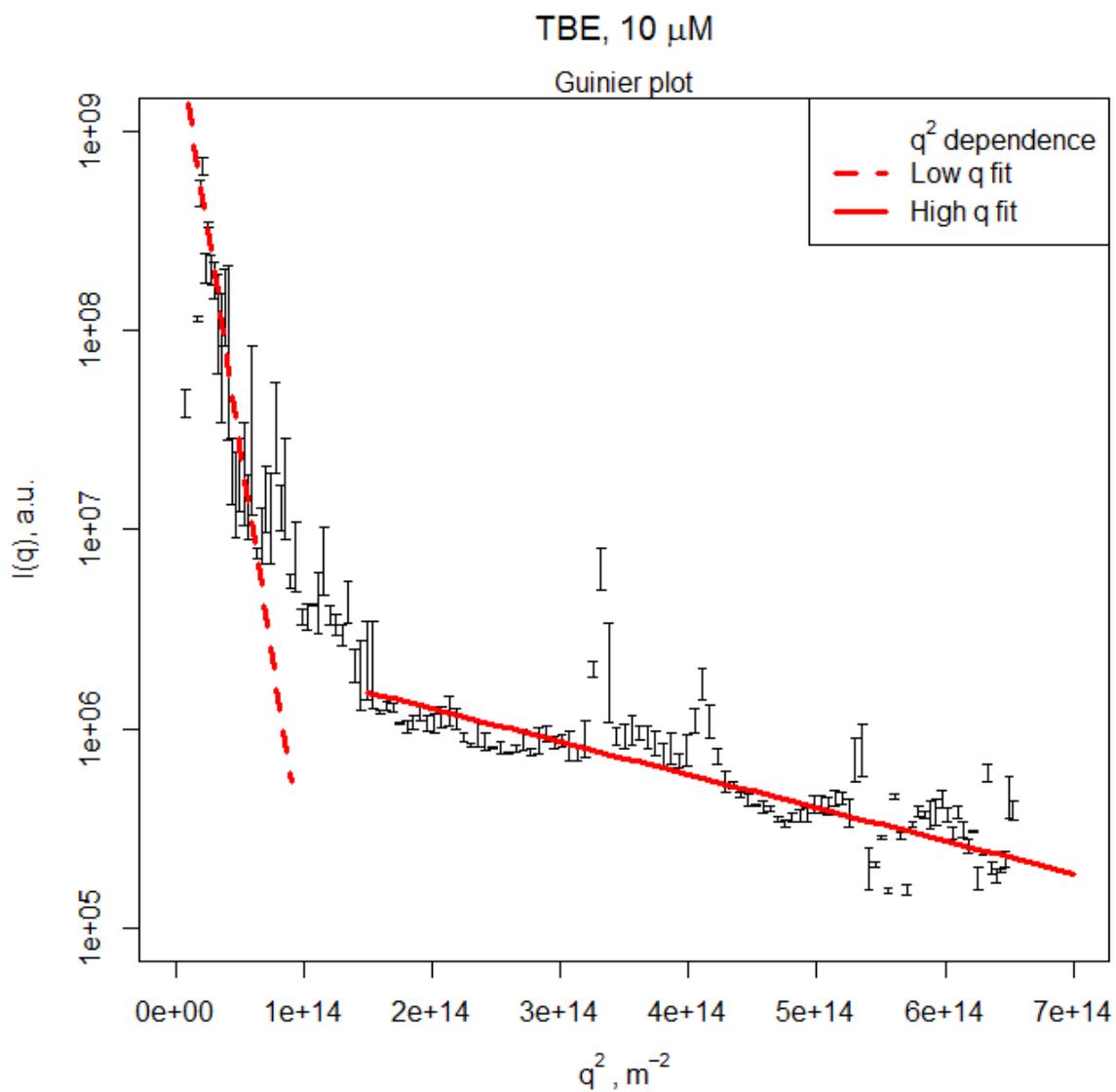
In this context, we wish to show that

- our interpretation of the tuple  $(s, d)$  is consistent with the structures observed by TEM and so these images may be considered representative of the bulk solution; and,
- the transition from one kind of structure to another is incremental, according to change in the sample concentration.

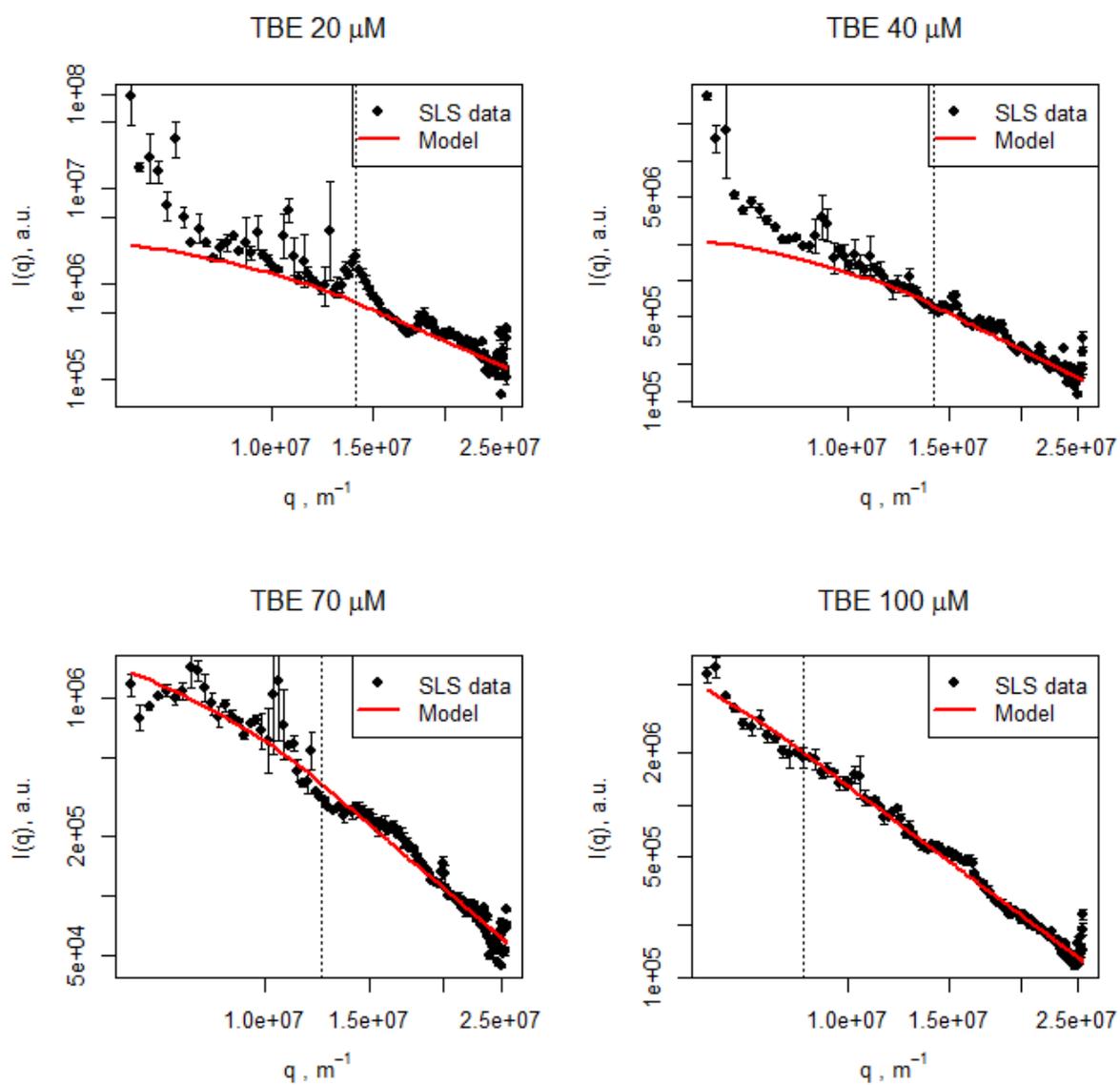
For **C12<sub>10</sub>-HEG<sub>10</sub>** in TBE buffer at  $10 \mu\text{M}$ , a Guinier plot was used to identify a characteristic size,  $\langle R \rangle_Z \approx 103 \text{ nm}$  (Figure S20) and the values  $(s, d) = (0, 2.65)$  would be consistent with an anisotropic particle, assumed to be a lamella or sheet. As the sample concentration increases the dimensionality parameters, again, change incrementally. At  $100 \mu\text{M}$  the values  $(s, d) = (1.42, 2.51)$  would be consistent with a network of cylindrical forms, as observed by TEM. Fits are shown in Figure S21.

For **C12<sub>10</sub>-HEG<sub>10</sub>** in TAMg buffer at  $4 \mu\text{M}$ , a Guinier plot of the SLS data suggests  $\langle R_G \rangle_Z \approx 107 \text{ nm}$  (Fig. S22) and the combination of dimensionality parameters  $(s, d) = (0.14, 3.2)$  are consistent with an ellipsoidal particle. As shown in Table S1, as the sample concentration increases there is a corresponding increase in the value of  $s$  and decrease in the value of  $d$ . The combination  $(s, d) = (1.47, 2.67)$  at  $100 \mu\text{M}$  is consistent with higher order arrays made up of the ellipsoidal unit, making for a coarse surface. Figure S23 shows the corresponding fits.

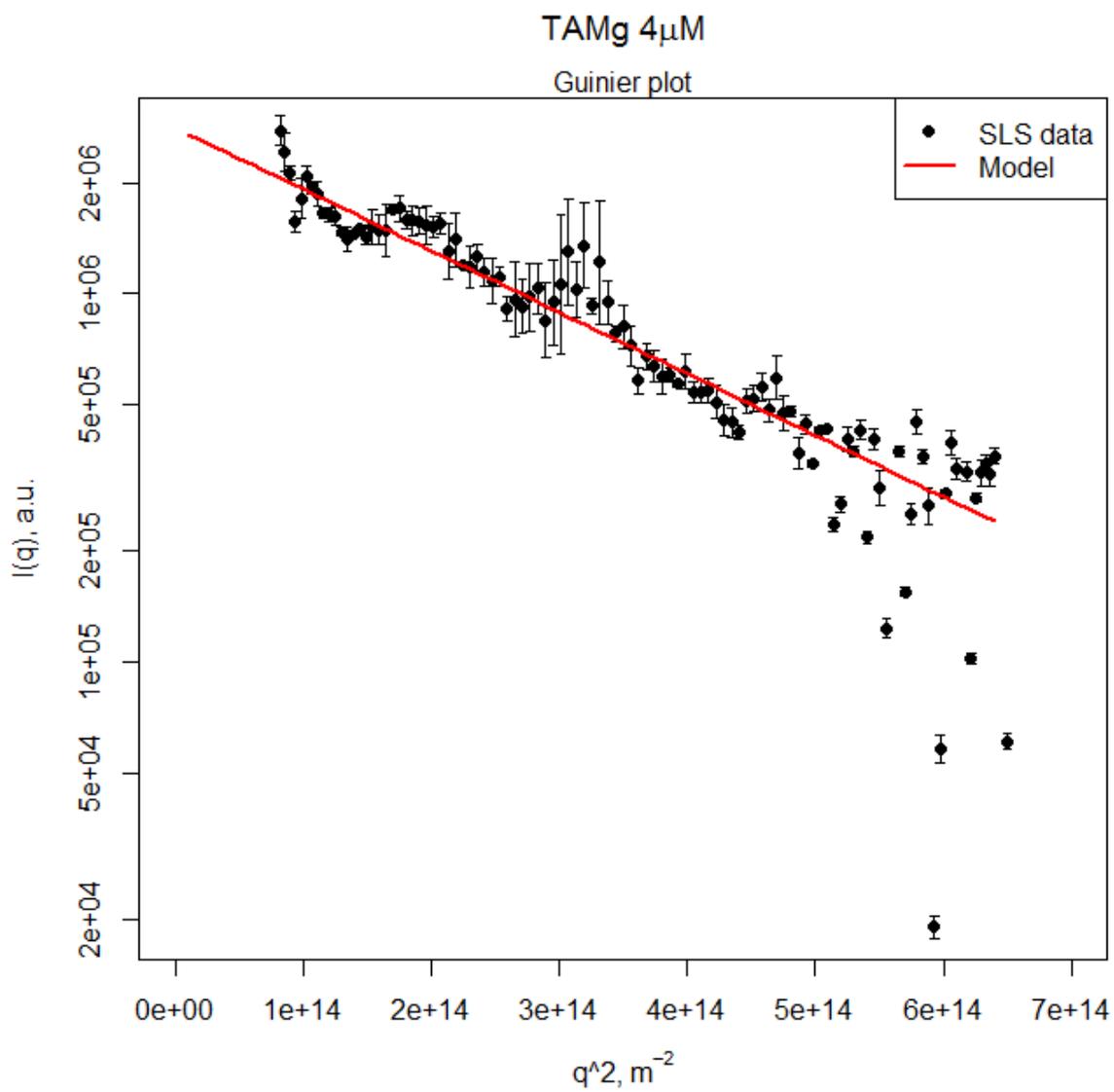
By inspection, when comparing the empirical scattering profiles at lower concentrations to the fitted model (Fig. S21, S23) the residuals exhibit a more or less pronounced oscillatory feature. Tentatively, we would suggest this might be explained with reference to a structure factor<sup>2</sup>,  $S(q)$ , to incorporate the effect of an excluded volume due to particle size (Fig. S24). We note this effect appears to diminish over the accessible  $q$ -range as the average size of structure increases according to sample concentration, which would be consistent with an increase in the average excluded volume and corresponding decrease in the number density of particles – both of which we may associate with an aggregation/assembly process.



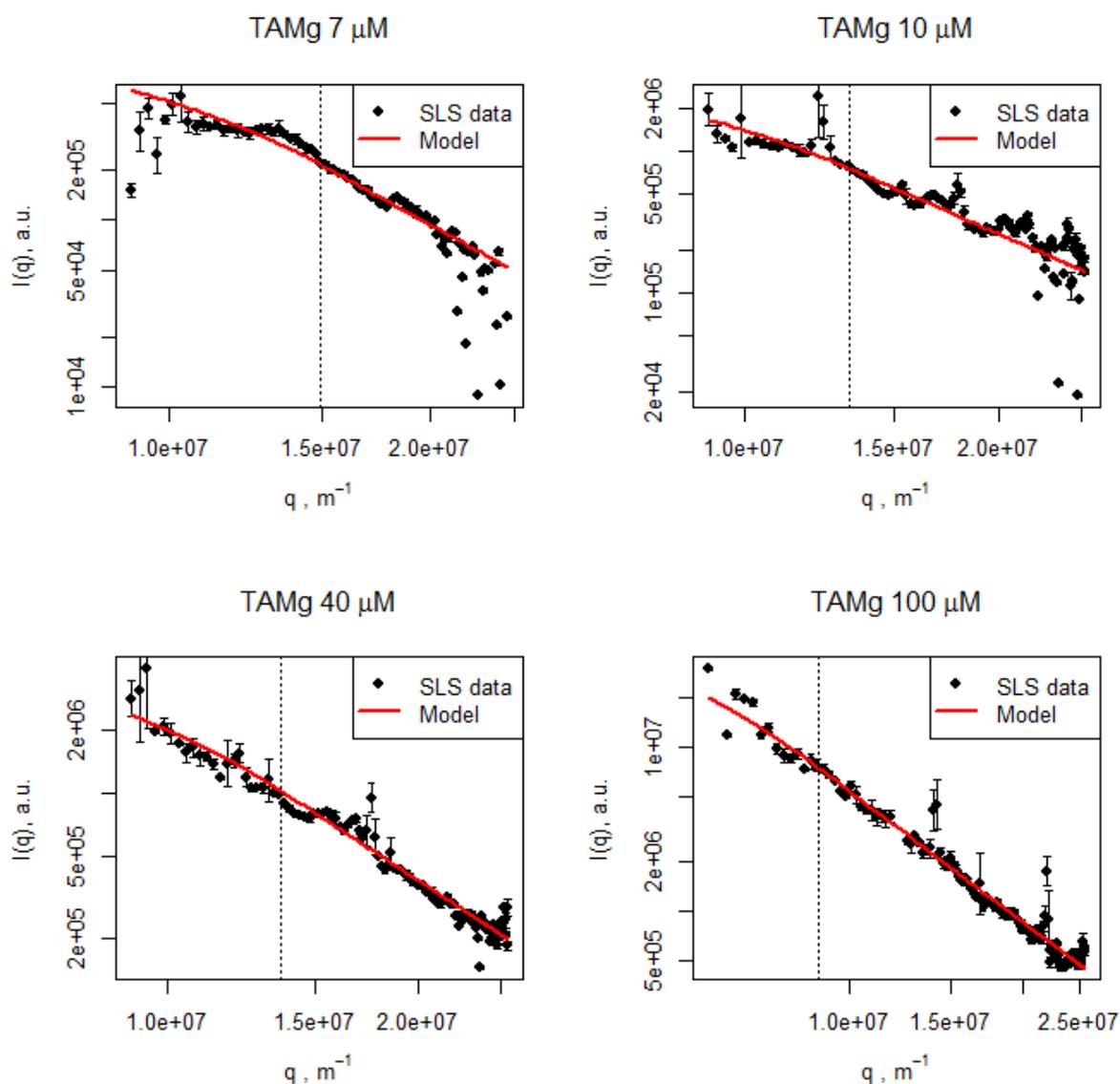
**Figure S20.** Guinier plot to estimate the mean particle size,  $\langle R_G \rangle_Z$ , for **C12<sub>10</sub>-HEG<sub>10</sub>** in TBE buffer at 10  $\mu$ M. The high-q fit identifies the most numerous particle population. The low-q fit indicates the presence of some larger structures.



**Figure S21.** Fits of the generalised Guinier-Porod model to data recorded for **C12<sub>10</sub>-HEG<sub>10</sub>** in TBE buffer, showing change in the scattering profile according to sample concentration.



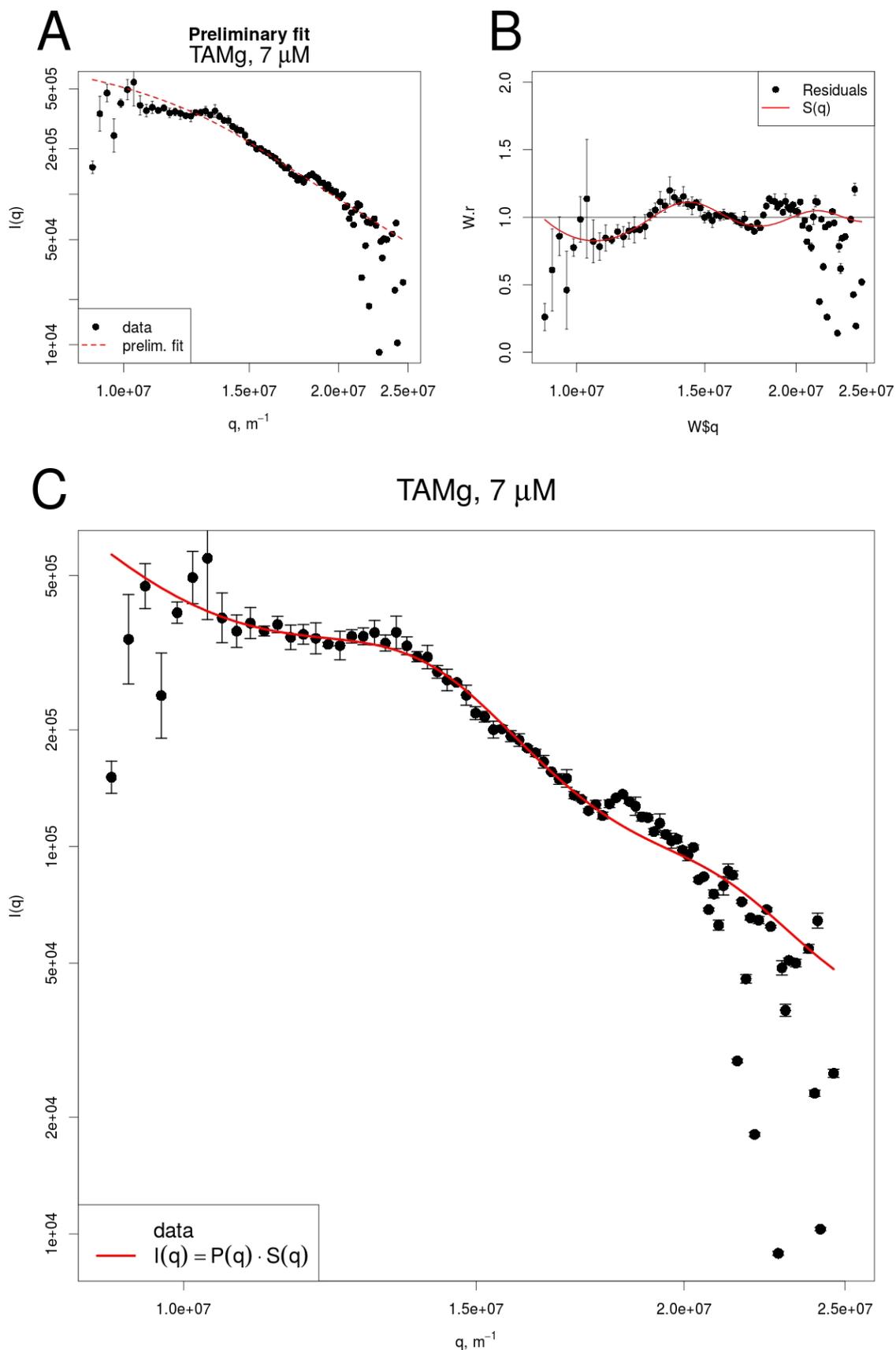
**Figure S22.** Guinier plot to estimate the mean particle size,  $\langle R_G \rangle_Z$ , for **C12<sub>10</sub>-HEG<sub>10</sub>** in TAMg buffer at 4  $\mu$ M.



**Figure S23.** Fits of the generalised Guinier-Porod model to data recorded for **C12<sub>10</sub>-HEG<sub>10</sub>** in TAMg buffer, showing change in the scattering profile according to sample concentration.

**Table S1**

Buffer	Sample concentration,	Fitted parameters	
	$\mu\text{M}$	s	d
TAMg	4	0.14	3.2
	7	0.65	2.86
	10	0.78	2.62
	40	0.73	2.64
	100	1.47	2.67
TBE	10	0	2.65
	20	0	2.65
	40	0.5	2.42
	70	0.85	2.58
	100	1.42	2.51



**Figure S24.** Fit to scattering data for **C12<sub>10</sub>-HEG<sub>10</sub>** in TAMg buffer at 7  $\mu\text{M}$ . [A] Upon fitting the generalised Guinier-Porod model,  $P(q)$ , an oscillatory pattern to the residuals may be more or less apparent, according to concentration. [B] We suggest this pattern might be explained by a structure factor,  $S(q)$ , thereby incorporating the effect of an excluded volume due to particle size over the measured  $q$ -range. [C] shows the resulting composite model for the empirical scattering profile,  $I(q)=P(q)S(q)$ .

**Table S2.** Summary of self-assembly by sequence, buffer, and concentration. <sup>[a]</sup> No observable self-assembly. Critical aggregation concentrations can be inferred from the degree of dilution required to achieve this.

Buffer	C12 <sub>10</sub> -HEG <sub>10</sub>		(C12-HEG) <sub>10</sub>	
	TBE	TAMg	TBE	TAMg
100 μM	Scrolls	Lumpy particles/ clusters	Random aggregates	Discs, film
10 μM	Lamellae	SCNPs	Random coil	Film
1 μM	Prolate seeds	[a]	[a]	[a]
0.1 μM	[a]	[a]	[a]	[a]

## References

1. M. Andersson, B. Wittgren and K-G. Wahlund, *Anal. Chem.*, 2003, **75**, 4279-4291
2. B. Hammouda, A new Guinier-Porod model, *J. Appl. Cryst.*, 2010, **43**, 716-719
3. J. S. Pedersen, Modelling of Small-Angle Scattering data from Colloids and Polymer Systems, in *Neutrons, X-rays and Light*, Ed. P. Lindner and Th. Zemb, 2002, Elsevier.