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

Nadeema Appukutti,[a] Joseph R. Jones,[b] and Christopher J. Serpell*[a]

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

**We have used solid phase phosphoramidite synthesis commonly used to make DNA, to produce two sequence-isomeric polymers which display sequence-programmed folding and self-assembly, going beyond structures which would be trivially anticipated. These findings open up possibilities for more sophisticated sequence/structure relationships using the same synthetic platform.**

The functions of proteins and DNA are determined by a common blueprint: a specific sequence of monomers is converted into a 3D structure through supramolecular interactions. Accordingly, sequence-defined non-natural polymers are of great current interest.[1–3] However, these studies are typically synthetic, and exploration of the supramolecular chemistry of uniform polymers is rare, with examples focussing more often on length effects[4–9] rather than sequence architecture.[10,11] Conversely, disperse polymer chains can be folded to give single chain nanoparticles (SCNPs),[12] but these lack the monomer-level control of biopolymers.[13] Foldamers recapture protein conformational control but are rarely as large as proteins.[14,15] The solid phase phosphoramidite method[16] used to create DNA strands has emerged as a strategy for creation of macromolecular and sequence-defined non-nucleosidic polymers for controlling the properties of DNA,[17,18] and as a data storage medium.[19] Oligomers of chromophores produced this way combine self-assembly with photonic properties.[20] However, the supramolecular behaviour of non-biological, macromolecular sequence-defined polyphosphoesters has yet to be explored.[21]

We herein report the contrasting supramolecular behaviour of two sequence-isomeric polyphosphoesters. This represents an example of sequence-determined self-assembly in entirely non-biological macromolecules. We used two monomers, dodecane diol (**C12**) and hexa(ethylene glycol) (**HEG**) which are

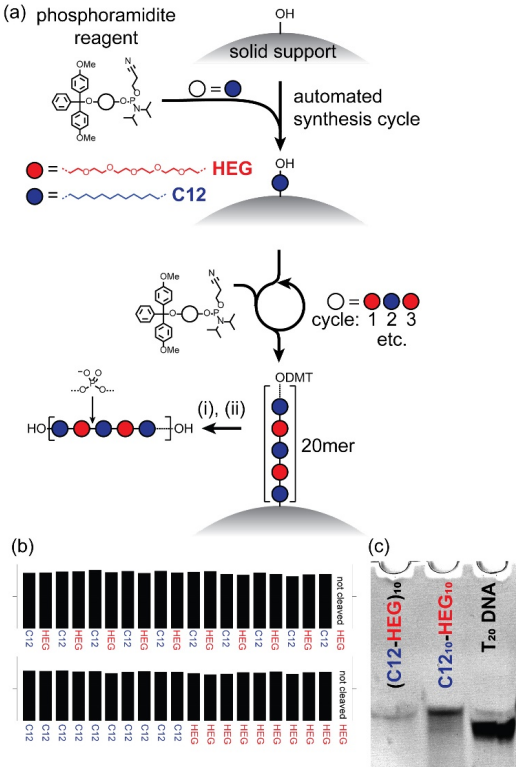


**Figure 1.** (a) Synthesis of sequence-defined polyphosphodiesters. Chemical structure of full products provided in Supporting Information Fig. S1. (b) Trityl histograms showing successful synthesis. (c) Denaturing polyacrylamide gel electrophoresis (20 %) showing single products ($M_w$ = 6021 g mol$^{-1}$). $T_{20}$ DNA ($M_w$ = 6019 g mol$^{-1}$) is provided as a molar mass comparison.

[a] *School of Physical Sciences, Ingram Building, University of Kent, Canterbury, Kent, CT2 7NH, UK*
[b] *Department of Chemistry, University of Warwick, Gibbet Hill, Coventry, CV4 7AL, UK*

Electronic Supplementary Information (ESI) available: synthetic protocols, MS, light scattering, spectroscopy, and microscopy. See DOI: 10.1039/x0xx00000x

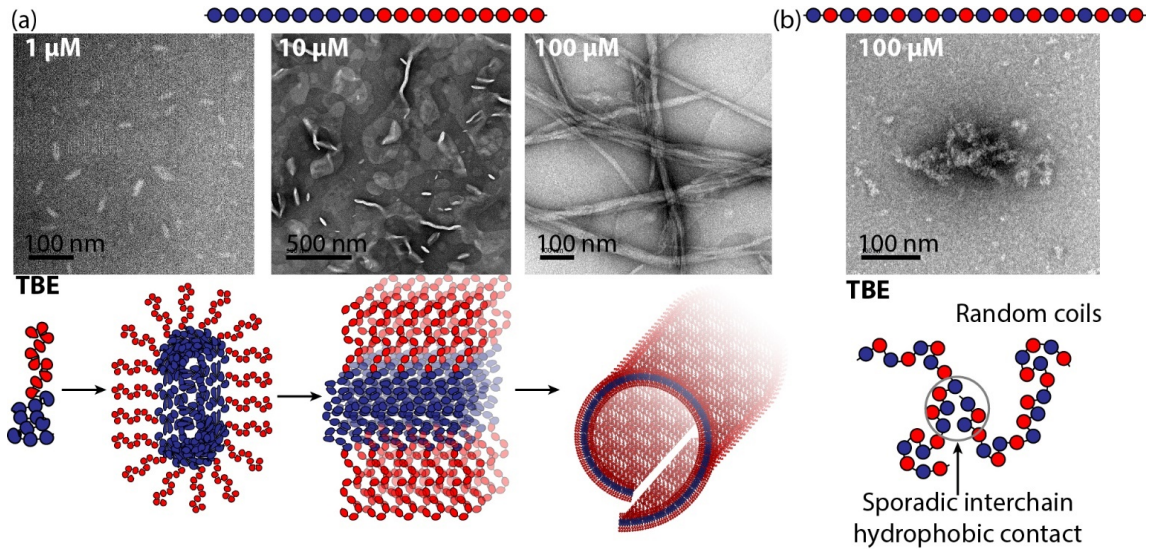


**Figure 2.** Self-assembly in TBE buffer of (a) **$C12_{10}$-$HEG_{10}$** and (b) **$(C12\text{-}HEG)_{10}$** as analysed by TEM.

commercially available as dimethoxytrityl (DMT)-protected phosphoramidite reagents (as 'spacers' in oligonucleotide strands). We used them to provide a hydrophobic (**C12**) or hydrophilic (**HEG**) region within the polymers, and designed two sequenced 20mers – a diblock arrangement **$C12_{10}$-$HEG_{10}$** and an alternating polymer **$(C12\text{-}HEG)_{10}$** (Fig. S1, Supporting Information for chemical structures). Because solid-phase synthesis has a stepwise yield of up to 99.5%, the polymers should have near-perfect sequence control and minimal dispersity, while the phosphates provide water solubility.

The polymers were produced using an automated synthesiser from controlled pore glass (CPG) supports via the UnyLinker,[22] eliminating the first nucleobase commonly installed on CPG supports (Fig. 1a). In-line monitoring of the outgoing DMT cation concentration by UV absorption at the deblock steps confirmed coupling success all the way to the 19mer (Fig. 1b). Due to the lack of UV-active groups in the final polymer which would enable UV quantification, the final DMT was left on the strand. After cleavage from the support, deprotection of the phosphate groups, and clean-up *via* size exclusion chromatography, the DMT groups were cleaved using 4:1 acetic acid:water, and the UV absorbance at 500 nm measured and compared against a standard curve to obtain yields. These yields (116% for **$C12_{10}$-$HEG_{10}$** and 59% for **$(C12\text{-}HEG)_{10}$** from nominally 1 μmol of starting functionality, subject to weighing, pipetting, and calibration-related errors) are similar to those obtained when synthesising DNA. We previously found stepwise coupling tended to decrease when the incoming monomer was different from the preceding one.[17] This pattern is seen here too, and we attribute it to sub-optimal non-covalent interactions between the two different monomers. Gel electrophoresis was used to confirm that the synthesis had been successful (Fig. 1c). Small differences in gel mobility relating to sequence of DNA strands of this length are common – the lack imperfect overlap seen here is to be expected with more exotic variations. To visualise the bands on the gel it was necessary to load a much greater quantity of the polymer (~1 nmol) onto polyacrylamide gels than would be needed for DNA (~10 pmol), using stains available for DNA gels. Successful synthesis was reaffirmed by mass spectrometric identification of tri- and tetraanionic peaks at m/z 2007 and 1505 respectively corresponding to the expected mono-isotopic mass of 6021 Da (Figs S2-S3). Uniformity was excellent, but not perfect - the [n-1] peaks could also be discerned as minor components.

We then assessed the effect sequence on self-assembly/folding. Previous results on similar polymers grown from DNA strands[17] suggested that self-assembly required continuous **C12** stretches and cations such as $Mg^{2+}$ to screen phosphate charges. We therefore hypothesised that in tris-acetate magnesium acetate buffer (TAMg, containing 12.5 mM $Mg^{2+}$) **$C12_{10}$-$HEG_{10}$** would give conventional spherical star micelles, whereas the alternating **$(C12\text{-}HEG)_{10}$** system would show no self-assembly. In $Mg^{2+}$-free conditions (tris-borate EDTA buffer, TBE) we expected no self-assembly from either polymer.

Self-assembly of both polymers was first screened in TBE buffer. We used dynamic light scattering (DLS) at timed intervals to give apparent hydrodynamic diameters (i.e. assuming a sphere; numbers should be considered only semi-quantitively). At 100 μM, **$C12_{10}$-$HEG_{10}$** increased in apparent hydrodynamic diameter from 72 nm to 615 nm over 4 hrs, and was then stable (Fig. S4); measured at 10 μM (Fig. S5) an initial peak at 274 nm evolved gradually to 714 nm, but had returned to 255 nm after 24 hours, while at 1 μM (Fig. S6), the apparent size fluctuated between 122 and 413 nm (although data quality was poor at this concentration). This suggests the formation of different superstructures at different concentrations. Conversely, the isomer **$(C12\text{-}HEG)_{10}$** at 100 μM produced assemblies with an apparent diameter of 615 nm from the outset which were stable over 24 hr (Fig. S4). At 10 μM (Fig. S5), the peak was again fairly stable, while at 1 μM (Fig. S6) the scattering was too weak to analyse with confidence.

The self-assembly was examined in parallel using the fluorescent membrane probe Nile Red (NR). NR fluoresces in hydrophobic media but is negligibly emissive in water – it can therefore be used to examine the extent of the hydrophobic microenvironment formed as a result of self-assembly.[23,24] For

**C12$_{10}$-HEG$_{10}$** at 100 μM, the integrated fluorescence intensity increased over the first four hours and then remained stable (Fig. S4). At 10 and 1 μM (Fig. S5, S6), the intensity decreased initially before stabilising. Again, this is consistent with a superstructure-concentration relationship. **(C12-HEG)$_{10}$** also gained intensity over time at 100 μM (Fig. S4), but lost intensity progressively at 10 μM (Fig. S5). At 1 μM (Fig. S6) there was no change in the already weak signal. The integrated fluorescence intensity arising from the block sequence sample was generally higher than that of the alternating polymer, indicating that the stretches of contiguous **C12** units provide a more significant hydrophobic microenvironment.

To examine the structure of the self-assembled nanomaterials, transmission electron microscopy (TEM) and atomic force microscopy (AFM) studies were conducted after a 24 hr incubation. For **C12$_{10}$-HEG$_{10}$** in TBE (Fig. 2a, S10-S12), at 100 μM, TEM revealed structures with widths of 10-40 nm, and multiple micrometers in length. Features consistent with twisting, fraying, and unrolling were seen. The same structures could be seen using AFM (Fig. S16). The model was tested by examining diluted samples. At 10 μM, **C12$_{10}$-HEG$_{10}$** produced primarily lamellae which exhibited 'pinching,' which would be the first stage in rolling of the lamellae to produce fibres with morphologies as seen at 100 μM. In a few areas, fibres were found, as well as smaller, prolate particles. At 1 μM, the prolate particles dominated (41 ± 8 nm long, 19 ± 3 nm wide). It appears that the prolate seeds merge to form lamellae, which then roll to give fibres, consistent with DLS and NR data indicating a structure/concentration relationship. Since these samples were made by dilution, the assemblies must be dynamic, allowing the fibres to break down into lamellae and seeds. Contrastingly, in the same buffer at 100 μM the sequence-isomer **(C12-HEG)$_{10}$** produced only very few irregular aggregates. We attribute this to the sample consisting of random coils which can aggregate through occasional patches of increased hydrophobicity (Fig 2b, S13).

Static light scattering (SLS) and small angle X-ray scattering (SAXS) experiments were conducted with a view to assessing whether the particles and their sizes as recorded by TEM and AFM were representative of the bulk solution and to investigate the apparent concentration dependence for structure formation (see Supporting Information). When **C12$_{10}$-HEG$_{10}$** was prepared in TBE buffer at $10\ \mu\mathrm{M}$, a population of particles of average size $\langle R_G \rangle_Z = 103 \pm 22\ \mathrm{nm}$ was identified via a Guinier plot of the SLS data, however the low-q SLS data suggested a second, less numerous, population of somewhat larger particles $\langle R_G \rangle_Z \approx 500\ \mathrm{nm}$. Samples then prepared at concentrations $20 \leq c \leq 100\ \mu\mathrm{M}$ exhibited incremental change in the scattering profile consistent with a gradual transition to large, higher order structures or else a network. This aligns with the observed transition from lamellae to scrolls. Unfortunately, SAXS experiments were unsuccessful, with the scattering profile for these samples being indistinguishable from the baseline (buffer) measurement.

The effects of sequence upon self-assembly were also manifested in TAMg buffer (Fig. S7-S9). As measured by DLS, the block copolymer **C12$_{10}$-HEG$_{10}$** at 100 μM immediately produced

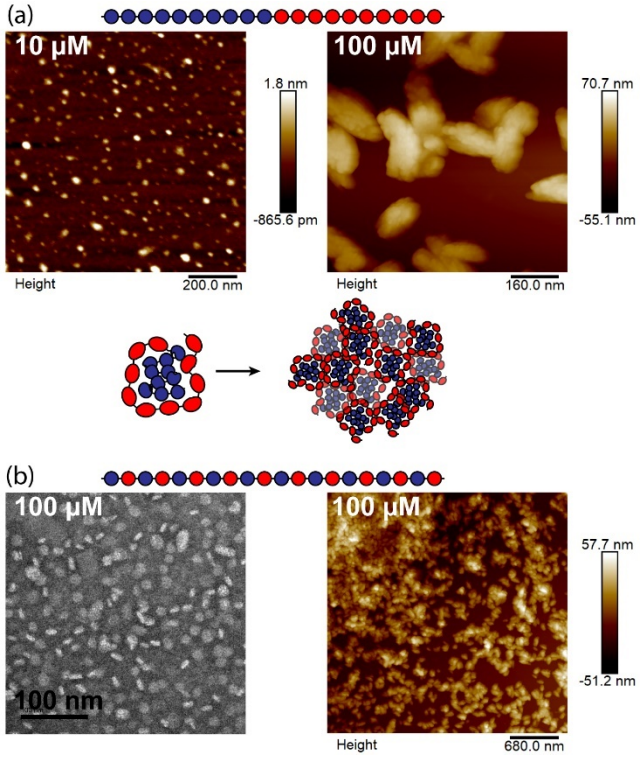


**Figure 3.** Self-assembly in TAMg buffer of (a) **C12$_{10}$-HEG$_{10}$** and (b) **(C12-HEG)$_{10}$**, analysed by TEM and AFM.

assemblies with an apparent diameter of 712 nm. The overall size distribution then decreased non-linearly over 24 hrs, to 295 nm. At 10 μM a stable population around 600 nm existed, and at 1 μM, over 24 hours particle size switched from around 115 nm to 16 nm. The latter would be consistent with SCNPs, but again data quality was low at this dilution. NR emission decreased with time in this buffer at all concentrations, being most pronounced at 100 μM, suggesting reorganisation of chains concomitant with the change in size. DLS showed that the alternating polymer **(C12-HEG)$_{10}$** at 100 μM gave a stable population with an apparent diameter of 615 nm, while the initially weak NR fluorescence decreased to barely-measurable levels. At lower concentrations the DLS data was not of sufficient quality to be used, and the NR emission was not distinguishable from the background.

**C12$_{10}$-HEG$_{10}$** at 100 μM was observed by TEM to consist of prolate particles smaller (18 ± 4 nm length by 7 ± 1 nm diameter) than those seen in TBE at 1 μM (Fig. S14), consistent with $Mg^{2+}$ promoting collapse of the polymers.[25] AFM (Fig. 3a, S17) showed these structures to be 'lumpy.' Upon dilution, AFM revealed spherical particles (1.3 ± 0.6 nm high, 22.5 ± 8.5 nm diameter, Fig. 3a, S18). Taking tip convolution into account, these measurements suggest formation of SCNPs. The lumpiness of the structures formed at 100 μM points to their composition as aggregates of SCNPs. With **(C12-HEG)$_{10}$**, at 100 μM discoid structures were seen by TEM (6.1 ± 1.0 nm by 14.6 ± 3.2 nm, Fig. 3b, S15) and confirmed by AFM (Fig. 3b, S19). It is not clear at this stage what kind of substructure could give this product since the smallest dimension of the particles is much larger than the maximum distance of a concertina conformation

of the polymer (~3.8 nm), as is seen in other alternating polymers.[26] To our knowledge, this is the first time that self-assembly of discs has been experimentally observed in alternating polymers.

Light scattering experiments were conducted similarly. **$C12_{10}$-$HEG_{10}$** samples prepared in TAMg buffer at 4 μM were observed to consist of non-isotropic particles of size $\langle R_G \rangle_Z = 107 \pm 5$ nm. Samples prepared at greater concentrations, in the range $7 \leq c \leq 100$ μM, exhibited incremental change in the scattering profile consistent with initial cluster formation and then formation of large, highly anisotropic, higher order structures from the particulate units (see Supporting Information, Figs S20-S24). **$(C12\text{-}HEG)_{10}$** did not scatter light with sufficient intensity for analysis, and again SAXS experiments did not yield useful data.

In summary, we have synthesised molecularly uniform, non-biological, sequence-defined, multi-kDa polymers which display self-assembly directly attributable to sequence. These findings are summarised in Table S2 which can be regarded as an approximate phase diagram for sequence, concentration, and buffer. The level of patterning explored here is at its simplest but the combination of macromolecular length and level of synthetic precision here is unique. Even these simple patterns have resulted not just in the production of unusual nanostructures, going beyond the classical star micelle/worm/vesicle schema, but also show interesting effects of dynamism not usually found in self-assembled polymer systems; indeed the outcomes have been far more interesting than our initial hypotheses anticipated. This is a strength of the charged phosphate group, which has allowed us to produce these nanostructures by direct dissolution, rather than a solvent switch, and also led to salt-driven structural switching. These findings raise interesting prospects for interface with biology, in which the ability to alter nanostructures based on the specific environment could be used to develop new diagnostics, drug carriers, or therapeutic systems themselves. We are working to diversify our monomer pool and develop methods to analyse self-assembly in a high-throughput manner to fully exploit the potential of the phosphoramidite method in development of next-generation materials.

## Conflicts of interest

There are no conflicts to declare.

## Notes and references


1 J.-F. Lutz, M. Ouchi, D. R. Liu and M. Sawamoto, *Science*, 2013, **341**, 1238149–1238149.

2 J.-F. Lutz, J.-M. Lehn, E. W. Meijer and K. Matyjaszewski, *Nat. Rev. Mater.*, 2016, **1**, 16024.

3 M. A. R. Meier and C. Barner-Kowollik, *Adv. Mater.*, 2019, 1806027.

4 D. R. Greer, M. A. Stolberg, S. Xuan, X. Jiang, N. P. Balsara and R. N. Zuckermann, *Macromolecules*, 2018, **51**, 9519–9525.

5 R. V. Mannige, T. K. Haxton, C. Proulx, E. J. Robertson, A. Battigelli, G. L. Butterfoss, R. N. Zuckermann and S. Whitelam, *Nature*, 2015, **526**, 415–420.

6 N. Gangloff, M. Höferth, V. Stepanenko, B. Sochor, B. Schummer, J. Nickel, H. Walles, R. Hanke, F. Würthner, R. N. Zuckermann and R. Luxenhofer, *Biopolymers*, 2019, **110**, e23259.

7 A. Das, K. Petkau-Milroy, G. Klerks, B. Van Genabeek, R. P. M. Lafleur, A. R. A. Palmans and E. W. Meijer, *ACS Macro Lett.*, 2018, **7**, 546–550.

8 B. Van Genabeek, B. F. M. De Waal, B. Ligt, A. R. A. Palmans and E. W. Meijer, *ACS Macro Lett.*, 2017, **6**, 674–678.

9 B. Van Genabeek, B. A. G. Lamers, B. F. M. De Waal, M. H. C. Van Son, A. R. A. Palmans and E. W. Meijer, *J. Am. Chem. Soc.*, 2017, **139**, 14869–14872.

10 M. R. Golder, Y. Jiang, P. E. Teichen, H. V.-T. Nguyen, W. Wang, N. Milos, S. A. Freedman, A. P. Willard and J. A. Johnson, *J. Am. Chem. Soc.*, 2018, **140**, 1596–1599.

11 W. Zhang, W. Shan, S. Zhang, Y. Liu, H. Su, J. Luo, Y. Xia, T. Li, C. Wesdemiotis, T. Liu, H. Cui, Y. Li and S. Z. D. Cheng, *Chem. Commun*, 2019, **55**, 636.

12 C. K. Lyon, A. Prasher, A. M. Hanlon, B. T. Tuten, C. A. Tooley, P. G. Frank and E. B. Berda, *Polym. Chem.*, 2015, **6**, 181–197.

13 A. M. Hanlon, C. K. Lyon and E. B. Berda, *Macromolecules*, 2016, **49**, 2–14.

14 X. Li, T. Qi, K. Srinivas, S. Massip, V. Maurizot and I. Huc, *Org. Lett.*, 2016, **18**, 1044–1047.

15 H. K. Murnen, A. R. Khokhlov, P. G. Khalatur, R. A. Segalman and R. N. Zuckermann, *Macromolecules*, 2012, **45**, 5229–5236.

16 M. H. Caruthers, *Acc. Chem. Res.*, 1991, **24**, 278–284.

17 T. G. W. Edwardson, K. M. M. Carneiro, C. J. Serpell and H. F. Sleiman, *Angew. Chemie - Int. Ed.*, 2014, **53**, 4567–4571.

18 D. Bousmail, L. Amrein, J. J. Fakhoury, H. H. Fakih, J. C. C. Hsu, L. Panasci and H. F. Sleiman, *Chem. Sci.*, 2017, **8**, 6218–6229.

19 A. Al Ouahabi, M. Kotera, L. Charles and J.-F. Lutz, *ACS Macro Lett.*, 2015, **4**, 1077–1080.

20 M. Vybornyi, Y. Vyborna and R. Häner, *Chem. Soc. Rev.*, 2019, 48, 4347–4360.

21 N. Appukutti and C. J. Serpell, *Polym. Chem.*, 2018, **9**, 2210–2226.

22 V. T. Ravikumar, R. K. Kumar, P. Olsen, M. N. Moore, R. L. Carty, M. Andrade, D. Gorman, X. Zhu, I. Cedillo, Z. Wang, L. Mendez, A. N. Scozzari, G. Aguirre, R. Somanathan and S. Berneès, *Org. Process Res. Dev.*, 2008, **12**, 399–410.

23 P. Greenspan, E. P. Mayer and S. D. Fowler, *J. Cell Biol.*, 1985, **100**, 965–973.

24 P. Greenspan and S. D. Fowler, *J. Lipid Res.*, 1985, **26**, 781–789.

25 A. Estévez-Torres and D. Baigl, *Soft Matter*, 2011, **7**, 6746.

26 Q. Xu, S. Li, C. Yu and Y. Zhou, *Chem. - A Eur. J.*, 2019, **25**, 4255–4264.