PAPER

Check for updates

Secondary structure drives self-assembly in weakly segregated globular protein–rod block copolymers†

Helen Yao, [ID] ‡[a] Kai Sheng,‡[b] Jialing Sun,[b] Shupeng Yan,[b] Yingqin Hou,[b] Hua Lu [ID] [b] and Bradley D. Olsen [ID] *[a]

Protein–polymer bioconjugates represent a class of materials that integrate protein functionality with polymer material properties and block copolymer self-assembly. To investigate the effect of polymer block secondary structure and chirality on self-assembly of globular protein–helix diblock copolymers, four types of bioconjugates consisting of a poly(amino acid) and enhanced green fluorescent protein (eGFP) were synthesized and compared: two homochiral, α-helix-forming bioconjugates incorporating either L- or D-type poly(amino acids), a 1 : 1 blend of the L- and D-type bioconjugates, and a bioconjugate incorporating structureless, achiral poly(amino acids). Poly(amino acids) (PAAs) were synthesized *via* N-carboxy anhydride (NCA) polymerization, and PAAs were conjugated to eGFP *via* native chemical ligation. All bioconjugates with a helical block self-assembled into lamellae at all concentrations measured (20 to 60 wt%). In contrast, the random copolymer of L- and D-type monomers did not self-assemble at any concentration or temperature. This was shown to be an effect of a non-repulsive interaction between the flexible PAA and the eGFP blocks, which is strong enough to affect the protonation state of the eGFP chromophore in water. Therefore, secondary structure of the polymer block can modulate the effective segregation strength between blocks and drive self-assembly even in systems with non-repulsive blocks.

Introduction

Proteins such as enzymes and antibodies have seen wide use in industries such as food,[1] medicine,[2] textiles,[3] defense,[4] and diagnostics.[5,6] Unfortunately, since most proteins evolved to work in biological systems, they are often unstable in non-native operating conditions that may be more relevant for a given application.[7] There are a variety of techniques to stabilize enzymes in foreign environments, including directed evolution,[8] rational protein engineering,[9] immobilization onto support matrices,[10] chemical crosslinking,[10] addition of osmolytes,[10] and bioconjugation to a polymer block.[11–14] Bioconjugation can promote the retention of enzyme functionality, enhance stability, impart useful material properties, and direct self-assembly[15] due to the presence of a polymer block.

Notably, self-assembly of globular protein–coil block copolymers (GCBCs) has been demonstrated to preserve protein function in myoglobin,[16] antibodies,[17] and nanobodies,[18] and even to enhance the performance of the overall material.

Much work has been done to characterize the self-assembly of such bioconjugate materials in dilute aqueous solution. Solution-state self-assembly is largely governed by solvent–bioconjugate interactions. Generally, the protein block is hydrophilic, while the polymer block is hydrophobic, which causes these "giant amphiphiles" to spontaneously self-assemble into micelles or vesicles, with the hydrophobic block segregated away from solvent.[19] Depending on protein and polymer chemistry, these nanostructures can take on a wide variety of shapes, including spherical micelles, rod-like micelles, vesicles, and toroids.[20] Conjugation to polymers such as poly(N-isopropylacrylamide) (PNIPAM) or elastin-like polypeptide (ELP) also imparts thermoresponsive properties to these materials.[21] Biocompatibility and the ability to tune the type of nanostructure formed based on the temperature or pH of the environment has led to great interest in using GCBCs as drug carriers.[22]

In concentrated solution and in the solid-state, GCBCs have been shown to self-assemble into many of the same phases that are observed in traditional coil–coil block copolymers,

[a]Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. E-mail: bdolsen@mit.edu;
Tel: +(617) 715-4548
[b]College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, P. R. China
†Electronic supplementary information (ESI) available. See DOI: 10.1039/c9py01680e
‡These authors contributed equally to this work.

such as lamellae and cylinders. However, the phase behavior is substantially different from that of traditional coil–coil block copolymers.[23] GCBCs in solution have asymmetric phase diagrams, lack inverse phases at high coil fractions, and exhibit a reentrant order–disorder transition (ODT) at high concentrations.[24–27] Unlike many coil–coil block copolymers, changing the polymer chemistry of the coil block can change the type of phases observed, as well as shift the critical concentration at which order first appears (C_{ODT}).[25] The phase behavior of GCBCs and discrepancies from the phase behavior of coil–coil block copolymers can be attributed to many factors, including the presence of a non-uniform charged surface on the protein,[28] interactions with hydration networks,[25] protein size, and protein shape.[29] GCBC phase behavior can be tuned by temperature, pH, and ionic strength, since proteins are charged species, and the polymer chemistry can be chosen to be charged[27,30] and/or thermoresponsive.[28,31] Electrostatic repulsion between highly charged proteins has been shown to decrease the propensity for order; this effect can be modulated by ionic strength.[32]

In addition to the folded structure of the protein block, there is potential for the chain structure of the polymer block to impact self-assembly. Incorporation of helical domains into block copolymers is common in rod–coil and rod–rod systems.[33,34] The rod block can be made from synthetic polymers,[33] poly(amino acids) (PAAs),[35] polypeptides,[36] or polypeptoids.[37] These structures have been investigated in both dilute and concentrated solution. In dilute solution, rod–coil block copolymers with a helical block largely behave like GCBCs, with solvent selectivity dictating the location of each block relative to the solvent.[33] However, the rigid rod imposes an additional packing restriction which favors phases such as pucks or vesicles rather than spherical micelles for certain rod-to-coil length ratios.[33,38] The presence of the rod also leads to liquid crystalline ordering into smectic or nematic phases.[38,39] In concentrated solution and the solid-state, many rod–coil block copolymers favor lamellar nanostructures to balance interfacial energy between blocks and rod alignment constraints.[33,39] In some cases, depending on the chemistry and processing conditions, it is also possible to observe more complex nanostructures such as zigzag, arrowhead, and pucks.[38] While the rod block is more rigid than the coil block, this does not prevent it from folding, bending, or deforming in other ways to pack optimally within self-assembled domains, as is seen in helical PAAs.[40–44] The use of PAAs also allows for responsive materials that undergo helix-to-sheet,[45,46] helix-to-coil,[45,46] or *trans*-to-*cis*[47] transitions upon an environmental change, such as in temperature, pH, or ionic strength. For example, ABA triblock copolypeptides consisting of poly(alanine) and poly(glutamic acid) have been shown to shift from tape-like structures into spherical clusters as pH increases. Increasing the ionic strength had the effect of driving a transition of the poly(alanine) to β-sheets, allowing the formation of fibrils.[48]

Besides chain structure, chirality of the polymer block can also change phase behavior in various polypeptide systems.

For example, chirality can impact the formation of complex coacervates. Homochiral polypeptides have been shown both experimentally[49] and in simulation[50] to precipitate rather than coacervate, as opposed to achiral polypeptides. In the block copolymer literature, there is also evidence that chirality in helical blocks can lead to order in the helical homopolymer *via* homochiral evolution, which, in turn, can influence the global chirality of a block copolymer.[51] For example, in a system of polystyrene-*b*-poly-L-lactide (PS-*b*-PLLA), chirality of the PLLA block drives the formation of a stable hexagonally packed "helical" phase of PLLA in a matrix of PS.[52]

Recently, a versatile and efficient conjugation method was developed to bioconjugate proteins and poly(amino acids) (PAAs) synthesized from *N*-carboxyanhydride (NCA) polymerization with topological control over the resulting structure of the bioconjugate.[53] This allows for the design of bioconjugates with an α-helical PAA block with tunable chirality (L or D) and a globular protein block. Incorporation of a PAA block into a GCBC has two important advantages: first, it imparts control over the secondary structure and chirality of the polymer block, allowing for the simultaneous analysis of the influence of these parameters on phase behavior. Second, having a PAA instead of a synthetic polymer makes the overall bioconjugate more biocompatible and biodegradable, allowing these materials to be seamlessly integrated into biomedical applications.[54]

In this work, self-assembly of globular protein–helix diblock copolymers in concentrated solution is reported. Structurally, these bioconjugates are globular protein–rod block copolymers, with shape rigidity in both blocks. Bioconjugates were synthesized *via* a native chemical ligation of enhanced green fluorescent protein (eGFP) to homochiral L- or D- and achiral L/D-poly(γ-2-(2-(2-methoxyethoxy)ethoxy)ethyl glutamate) (P(EG)₃Glu) PAAs. The monomers were functionalized with a small PEG trimer to increase solubility of the polymer block and to eliminate charge. Self-assembled nanostructures in dilute and concentrated solution were characterized by small-angle X-ray scattering and dynamic light scattering to illustrate the effect of PAA chirality and secondary structure on the self-assembly process.

Methods

Instrumentation

Proton nuclear magnetic resonance (NMR) spectra were recorded on a 400 MHz Bruker ARX400 FT-NMR spectrometer using D₂O (Cambridge Isotope Laboratories, Inc.) as the solvent. Tandem gel permeation chromatography (GPC) experiments were performed on a system equipped with an isocratic pump (Model 1100, Agilent Technology, Santa Clara, CA), a DAWN HELEOS 9-angle laser light scattering detector with a fixed wavelength of 658 nm (Wyatt Technology, Santa Barbara, CA) and an Optilab rEX refractive index detector (Wyatt Technology, Santa Barbara, CA). Separations were performed on serially connected size exclusion columns (103 Å, 104 Å,

105 Å Phenogel columns, 5 μm, 7.8 × 300 mm, Phenomenex, Torrance, CA) at 50 °C using 0.1 M LiBr in *N*,*N*-dimethylformamide (DMF) as the mobile phase. Ultraperformance liquid chromatography-electrospray ionization mass spectrometry (UPLC-ESIMS) was performed on a tandem system equipped with an ACQUITY H-Class UPLC (Waters Corp.) and a quadrupole rods SQ Detector 2 mass spectrometer (Waters Corp.). Separation was performed on a protein BEH C4 column (Waters 300, 1.7 2.150 mm) with ultrapure water (with 0.1% formic acid) and acetonitrile as the mobile phase. Denaturing protein gel electrophoresis (SDS-PAGE) images were recorded on a Typhoon FLA 9500 laser scanner (GE Healthcare Corp.). Fast protein liquid chromatography (FPLC) separation was performed on an ÄKTA FPLC system (GE Healthcare, Inc.) using a Superdex 200 Increase 10/300 GL, a Superdex75 HiLoad 16/600 pg, or a MonoQ 5/50 GL column. The choice of column depended on the identity of the polymer, as detailed in the purification section below. The Superdex 200 Increase 10/300 GL was used to test size-exclusion purification conditions on a small scale, with subsequent large-scale purification being done on the Superdex75 HiLoad 16/600 pg column. Protein yield was assessed by Nanodrop (NanoPhotometer P-class, Germany). Secondary structure was confirmed using circular dichroism (CD) spectroscopy, which was recorded for a wavelength range of 200–250 nm with a scan rate of 100 nm min^{-1} and a bandwidth of 1 nm (Bio-Logic Science Instruments, France).

Materials

All chemicals were purchased from commercial sources and used as received unless otherwise specified. Phenyl trimethylsilyl sulfide (PhS-TMS) was purchased from Sigma-Aldrich (St Louis, USA). Anhydrous *N*,*N*-dimethylformamide (DMF) was purchased from Sigma-Aldrich and treated with methyl isocyanate bound to polystyrene beads to remove free amines (Sigma-Aldrich, St Louis, USA) prior to polymerization. Scheme 1 shows the overall synthetic scheme for the preparation of the (EG)$_3$-L-Glu-NCA monomer (2$_L$), NCA ring-opening polymerization (ROP), and the subsequent protein–polymer conjugation. For D and achiral L/D polymers, a similar scheme was used but with (EG)$_3$-D-Glu-NCA monomer (2$_D$) and a 1 : 1 mixture of L and D monomers, respectively. Table 1 summarizes the polymers and bioconjugates used in this study, along with the shorthand notation used to describe each sample.

Preparation of poly(amino acids) with different chirality

The synthesis of L- and D-(EG)$_3$Glu (1) and (EG)$_3$Glu-NCA (2) was carried out according to the protocol described by Chen *et al.*; the characterization data (^{1}H and ^{13}C NMR spectra) of compounds (1) and (2) synthesized here were identical to those reported by Chen *et al.*[55]

The ROP of NCAs was initiated by phenyl trimethylsilyl sulfide (PhSTMS) as reported previously.[54,56] All polymers were synthesized for a target degree of polymerization of 100, corresponding to a molecular weight of about 27.5 kDa. For the synthesis of L-P(EG)$_3$Glu-SPh (3$_L$), 2.0 g (EG)$_3$-L-Glu-NCA

(2$_L$) was weighed out and dissolved in 10 mL DMF in a glove box, to which 0.126 mL of 0.500 mM PhSTMS was added while stirring at room temperature. After 36 hours, 0.130 mL acetic anhydride was added to quench the polymerization. The reaction mixture was precipitated into 225 mL of chilled diethyl ether. The resultant white pellet was dissolved in 10 mL of a 0.1% aqueous solution of acetic acid and dialyzed three times against 1 L 0.1% aqueous acetic acid at 4 °C for 4 hours. The solution was lyophilized to give 1.2 g polymer (3) as a white solid. The synthesis of D- and L/D-P(EG)$_3$Glu-SPh (3$_D$) was performed the same way as above but by using (EG)$_3$-D-Glu-NCA (2$_D$) and an equimolar amount of (EG)$_3$-L-Glu-NCA (2$_L$) and (EG)$_3$-D-Glu-NCA (2$_L$) as co-monomers, respectively.

All polymers were characterized using NMR spectroscopy and GPC. GPC traces (Fig. S1†) and ^{1}H NMR spectra (Fig. S2–S4†) are provided in the ESI† of this article. Polymer molar mass, dispersity, and degree of polymerization are reported in Table 1.

Expression of CG-eGFP and TEV protease digestion of TEV-CG-eGFP

The expression and proteolytic cleavage of enhanced green fluorescent protein (eGFP) with a tobacco etch virus (TEV) cleavage site and a chemical linker (TEV-CG-eGFP) were both reported previously.[57] The amino acid sequence of TEV-CG-eGFP is provided in the ESI.† Briefly, the plasmid pET28a-eGFP encoding TEV-CG-eGFP was constructed by PolePolar Biotechnology Co., Ltd (Beijing). The recombinant eGFP has a cleavable site (shown in red in the TEV-CG-eGFP sequence) recognizable by the TEV protease. After cleavage, the product CG-eGFP (4) consists of eGFP with an exposed N-terminal cysteine tethered to a Gly$_4$Ser linker (shown in blue in the sequence). The C-terminus of the protein is fused with a 6xHis tag for Ni-NTA purification.

The pET28a-eGFP plasmid was transformed into *Escherichia coli* (*E. coli*) BL21(DE3) cells. The cells were cultured in lysogeny broth (LB) medium supplemented with kanamycin (50 μg mL^{-1}) and induced at OD$_{600}$ = 1.0 with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). The cells were cultured at 30 °C and 180 rpm for another 6–8 hours before being harvested by centrifugation (6500 rpm for 30 min) at 4 °C. The cell pellet was resuspended in 40 mL of lysis buffer (20 mM Tris-HCl, 150 mM NaCl, pH 8.0) and lysed by sonication. The insoluble cell debris was removed by centrifugation (12 000*g* for 60 min) at 4 °C. The supernatant was clarified with a 0.22 μm filter, and the protein was purified by Ni-NTA affinity chromatography. The yield of TEV-CG-eGFP was 40 mg L^{-1}.

50 mg TEV-CG-eGFP was diluted into 50 mL of cleavage buffer (0.5 mM EDTA, 1 mM DTT, 25 mM Tris-HCl, pH = 8.0). 0.5 mg TEV protease (0.01 equiv.) was added. The reaction mixture was incubated in a dark environment at room temperature for 2 hours. UPLC-ESIMS was used to monitor reaction progress, which was allowed to proceed until completion. The buffer was exchanged into 50 mM Tris-HCl pH = 7.4, and the product CG-eGFP (4) was concentrated to 15 mg mL^{-1} using AMICON Ultra ultracentrifugation filters with a molecular
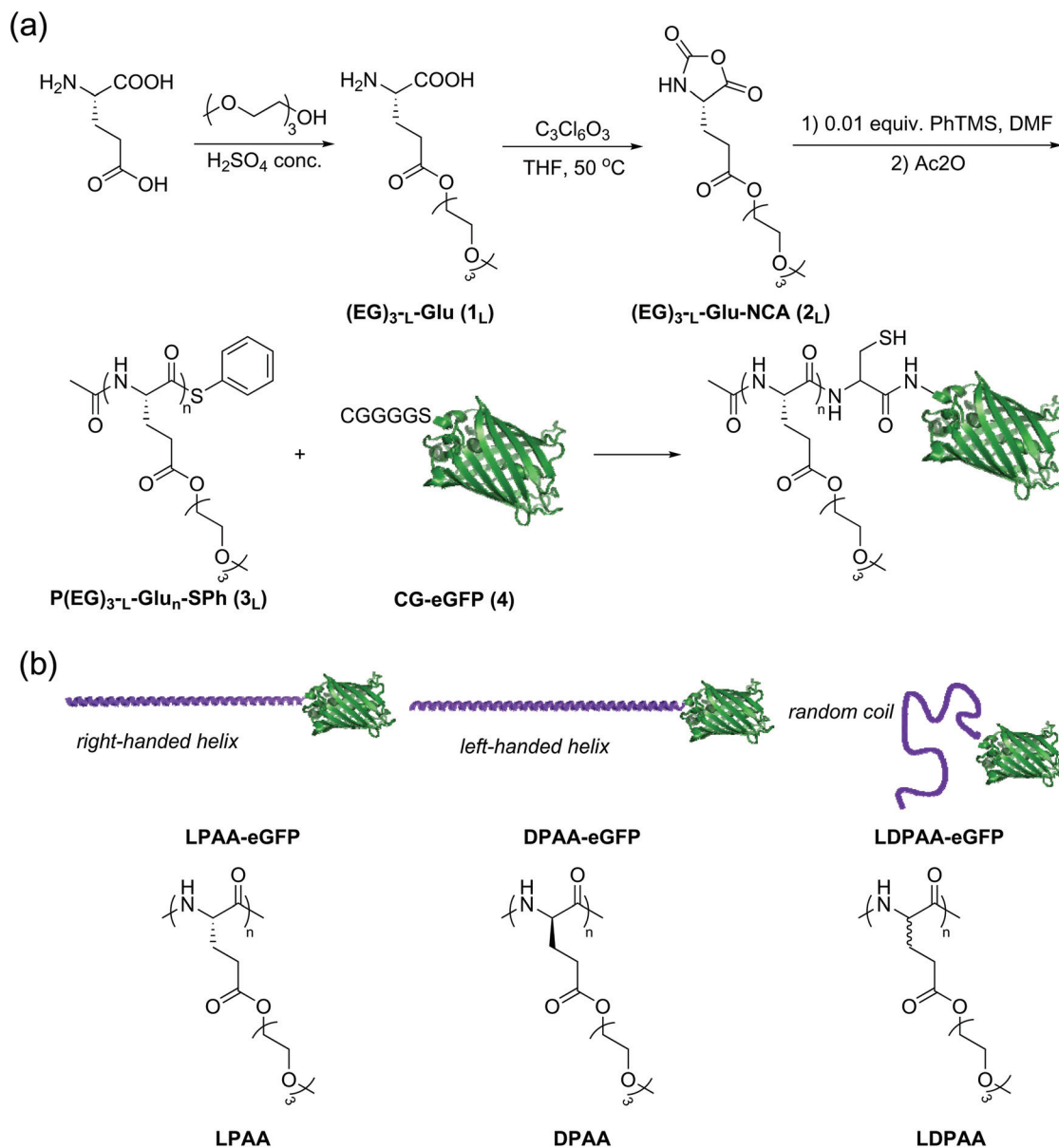
Scheme 1 (a) Synthesis of P(EG)$_3$-L-Glu-SPh and the subsequent site-specific conjugation with CG-eGFP to afford the LPAA-eGFP conjugate. The other variants of conjugates were synthesized by a similar route. The D homochiral variant, P(EG)$_3$-D-Glu-SPh, was synthesized by using the D chirality monomer. The achiral random copolymer P(EG)$_3$-L/D-Glu-SPh was synthesized using a 1 : 1 mixture of L and D monomers. (EG)$_3$ signifies an ethylene glycol trimer. The target degree of polymerization is 100; actual degrees of polymerization can be found in Table 1. (b) LPAA-eGFP, DPAA-eGFP, and LDPAA-eGFP conjugates along with the chemical structure of the corresponding polymer block. Helices were rendered in PyMOL with length approximately to scale with eGFP.

weight cutoff of 10 kDa. Protein molar mass before and after cleavage was confirmed by mass spectrometry (Fig. S5†).

Synthesis and purification of bioconjugates

The protein–polymer conjugation was carried out by following a protocol similar to what has been described previously.[53,54,58] 3.3 mL of 15 mg mL^{-1} CG-eGFP (4) (1.8 μmol) was added into 150 mg L-, D-, or L/D-P(EG)$_3$Glu-SPh (3) (5.4 μmol). The mixture was vortexed until clear. The reaction mixture was incubated in darkness at room temperature for 12 hours. After 12 hours,

36 μL of 50 mM TCEP was added to prevent dimerization, and the reaction mixture was incubated at room temperature for 1 hour. Characterization data for the bioconjugate samples is summarized in Table 1 along with shorthand notation for naming each sample.

The conjugates were first purified by fast protein liquid chromatography (FPLC) on an ÄKTA FPLC system (GE Healthcare, Inc.) using a Superdex75 HiLoad 16/600 pg size-exclusion column. The purification was done in 1× PBS buffer with a flow rate of 0.8 mL min^{-1}. Eluted fractions were col-

Table 1 Properties of bioconjugate materials

Sample	Polymer block	Protein molar mass (kDa)	Polymer molar mass (kDa)	Polymer dispersity index ($Đ$)	Weight fraction of polymer block
LPAA-eGFP	P(EG)$_3$-L-Glu$_{88}$	28.1	24.3	1.02	0.464
DPAA-eGFP	P(EG)$_3$-D-Glu$_{97}$	28.1	26.7	1.02	0.487
LDPAA-eGFP	P(EG)$_3$-L/D-Glu$_{100}$	28.1	27.5	1.02	0.495
L : D 1 : 1-eGFP	This is a 1 : 1 mixture (by mass of bioconjugate) of the LPAA-eGFP and DPAA-eGFP bioconjugates				

lected at a rate of 0.8 mL per tube. Samples with the retention volume ranging from 43 mL to 61 mL were collected, and unreacted CG-eGFP was discarded. For further purification of LPAA-eGFP or DPAA-eGFP, unreacted L- or D-P(EG)$_3$Glu-SPh (3) was removed by Ni-NTA chromatography. 10 mg of LPAA-eGFP or DPAA-eGFP purified from the Superdex75 HiLoad 16/600 pg column were bound to 5 mL Ni-NTA resin. The resin was washed with 50 mL NTA0 buffer (50 mM Tris, 150 mM NaCl, pH 8.0) to remove unreacted L- or D-P(EG)$_3$Glu-SPh (3). Then LPAA-eGFP or DPAA-eGFP was eluted with 15 mL NTA300 buffer (50 mM Tris, 150 mM NaCl, 300 mM imidazole, pH 8.0). The elution fraction was exchanged into 1× PBS buffer by dialysis.

LDPAA-eGFP was further purified by a Mono Q 5/50 GL column to remove unreacted P(EG)$_3$-L/D-Glu-SPh (3_{LD}). 10 mg of LDPAA-eGFP were dialyzed against 20 mM Tris (pH 8.0) in preparation for FPLC and loaded onto a Mono Q 5/50 GL column using the same loading rate as described above. The column was washed with 20 mL 20 mM Tris (pH 8.0) to remove unreacted P(EG)$_3$-L/D-Glu-SPh (3_{LD}). Then LDPAA-eGFP was eluted with 20 mM Tris (pH 8.0) containing NaCl (linear elution, 0 to 0.5 M NaCl, 10 mL). Pure LDPAA-eGFP was dialyzed against 1× PBS.

In preparation for lyophilization, the conjugates were dialyzed against 0.1 M NH$_4$HCO$_3$ solution to remove extra nonvolatile salts. The bioconjugates were then lyophilized to give the final products. The purity of the conjugates was assessed using SDS-PAGE (Fig. 1).

The L : D 1 : 1-eGFP blend was prepared by dissolving LPAA-eGFP and DPAA-eGFP separately into Milli-Q water at 50 mg mL^{-1} each and mixing the two solutions after complete sol-

vation. The solutions were mixed and equilibrated at 4 °C overnight before lyophilization to produce the 1 : 1 blend of LPAA-eGFP and DPAA-eGFP. Protein secondary structure after lyophilization was confirmed using CD spectroscopy (Fig. S17†). The four types of bioconjugates used for the self-assembly study are summarized in Table 1.

Materials structure characterization

Structure and phase behavior in both dilute and concentrated solution were characterized for the four types of bioconjugates listed in Table 1 using circular dichroism (CD) spectroscopy, small-angle X-ray scattering (SAXS), depolarized light scattering (DPLS), dynamic light scattering (DLS), and UV-Vis spectroscopy. The protein fold for both the PAA and the eGFP was confirmed using CD spectroscopy. All measurements were taken at 37 °C using a 1 mm path length quartz cuvette. Conjugates were dissolved in PBS at a concentration of 13 μM eGFP (overall conjugate concentration depended on degree of polymerization of the polymer), and polymers were dissolved in PBS at a concentration of 0.3 mg mL^{-1}. CD spectra were converted into molar ellipticity and analyzed for secondary structure content using the BeStSel CD analysis[59] and PDB2CD tools.[60]

Small-angle X-ray scattering (SAXS) measurements on PAA-eGFP conjugates were taken at the 11-BM beamline at the National Synchrotron Light Source II (NSLS-II) at Brookhaven National Lab. Samples were dissolved in Milli-Q purified water at concentrations between 20 and 60 wt% in 5 wt% intervals. All samples were equilibrated for at least 3 hours at 4 °C prior to measurement. After equilibration, samples were loaded into 1 mm path length aluminum washers and sealed with Kapton tape. Samples were measured at temperatures from 10–50 °C in 5 °C intervals; prior to measurement, all samples were equilibrated at the desired temperature for 10 minutes. Acquisition times (1 s or 10 s) were selected such that there was no detectable beam damage to the sample's nanostructure. To normalize between the two acquisition times, data taken at 10 s were scaled by a factor of 10. 2D SAXS data were background corrected using eqn (1) for empty cell and dark field scattering and transformed into 1D data *via* an azimuthal average of the 2D SAXS pattern.

$$I(q) = \frac{I_{\text{sample}}}{T_{\text{sample}}} - \frac{t_{\text{sample}}}{t_{\text{empty}}} \frac{I_{\text{empty}}}{T_{\text{empty}}} \tag{1}$$

Here, I is intensity, t is thickness, and T is transmission.

DPLS was used to measure turbidimetry and birefringence for all samples. Samples were dissolved in Milli-Q water using
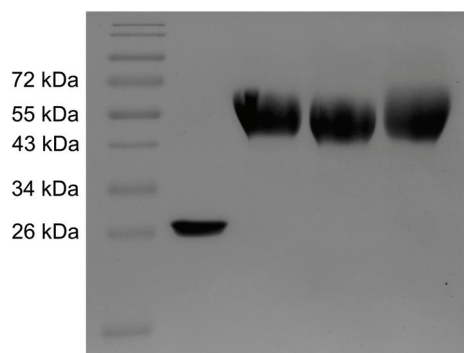

Fig. 1 SDS-PAGE gels for TEV-CG-eGFP and conjugates. The lanes are as follows: (1) ladder (2) TEV-CG-eGFP (3) LPAA-eGFP (4) DPAA-eGFP (5) LDPAA-eGFP.

the same procedure as for SAXS. After equilibration at 4 °C, the samples were loaded into a 1 mm thick Teflon washer, which was sandwiched between two quartz windows and held in place by a brass outer cell. All measurements were taken using a Coherent OBIS LX660 laser (λ = 662 nm). A rear polarizer was used for birefringence but was removed for transmission. Data were collected during both a heating and a cooling ramp from 10–50 °C with a ramp rate of 1 °C min^{-1}. Birefringence measurements were corrected using a dark field and transmission correction, and transmission measurements were normalized to the maximum laser transmission.

DLS measurements were made using a Wyatt DynaPro Plate Reader with an 850 nm laser at the MIT Koch Institute Swanson Biotechnology Center. Bioconjugates were dissolved in Milli-Q water at a concentration of 1 mg mL^{-1} and filtered using a 0.2 μm cellulose acetate syringe filter (VWR). 50 μL of each type of bioconjugate was loaded in replicates of three into a 384-well glass-bottom plate. DLS measurements were taken using a 10 s acquisition time, with 10 acquisitions per well, leading to a total of 30 measurements per bioconjugate type. The Dynamics 7.5 software was used to convert raw intensity data into autocorrelation curves and size distributions for each bioconjugate.

UV-Vis spectroscopy was measured using a Varian Cary 50 spectrometer with a wavelength range of 200–800 nm and a scan rate of 600 nm min^{-1}, with data recorded at 1 nm intervals. Samples were prepared at 0.5 mg mL^{-1} in Milli-Q water and filtered using a 0.2 μm cellulose acetate syringe filter (VWR) prior to loading into a quartz cuvette with a path length of 1 cm. Three wavelength scans were taken for each sample and averaged. The baseline was corrected by subtracting a Milli-Q water background measured in the same quartz cuvette.

Results and discussion

Secondary structure of bioconjugates

All PAAs and bioconjugates used in this study have the expected chirality-dependent helicity. Fig. 2 shows the CD spectra for the poly(amino acids) prior to bioconjugation (Fig. 2a) and the CD spectra for the bioconjugates with each type of PAA (Fig. 2b), along with eGFP. LPAA and DPAA both exhibit α-helical secondary structure, with the DPAA signal opposite in sign to that of LPAA, indicating that it is of the opposite chirality. LDPAA has no CD signal, suggesting that it does not have significant secondary structure and is most likely a random coil. However, it is possible that it may have small stretches of secondary structure that can form if there are long enough segments of consecutive homochiral monomer units. Achiral polypeptides have been shown to sample such structured transient states in simulation.[50]

The Beta Structure Selection (BeStSel) tool was used to analyze the composition of different elements of secondary structure present in each bioconjugate based on the CD spectra presented in Fig. 2.[59] This tool was chosen due to its enhanced performance for β-type structures compared to other CD analysis tools, given that the protein block is a β-barrel. However, BeStSel (as well as other CD analysis tools) is not designed for amino acids with D chirality, since natural amino acids have L chirality. Thus, for the PAAs alone, the DPAA signal was negated to transform it into an equivalent LPAA signal before inputting into BeStSel. Table 2 summarizes the results from BeStSel. For eGFP and LPAA-eGFP, the molar ellipticity for each was directly input into BeStSel. To examine just the structure of LPAA after conjugation, the eGFP signal was subtracted from the LPAA-eGFP CD signal. For DPAA-eGFP, the eGFP signal was first subtracted from the bioconjugate signal. Next, the difference was negated to transform DPAA into LPAA, and the eGFP signal was added again to obtain DPAA-eGFP after modifying with L chirality. To understand the structure of only the DPAA block after conjugation, the eGFP signal was subtracted, and the resulting difference was negated to account for chirality. Secondary structure analysis was not performed for LDPAA materials because they contain random copolymers of L and D chirality with no method of correcting for the D signal quantitatively. Fig. 3 shows the resulting CD spectra for bioconjugates, PAAs only, eGFP only, PAAs after
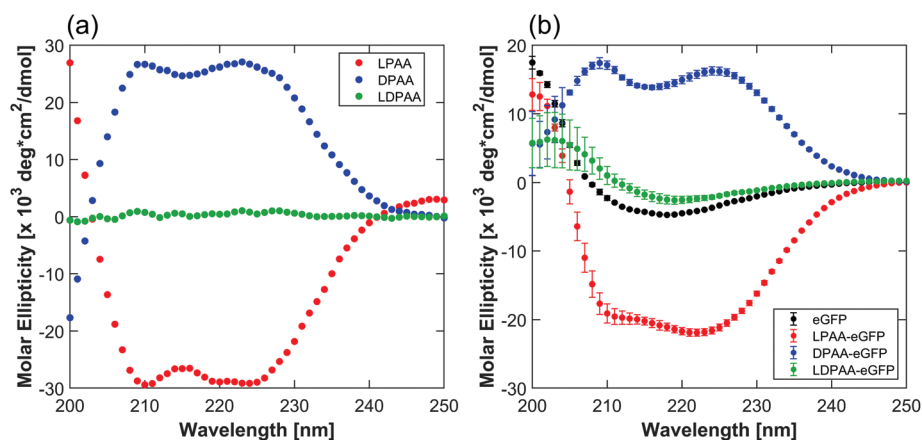

Fig. 2 CD spectra for (a) poly(amino acids) alone and (b) bioconjugates and eGFP. All raw CD signals (measured in mdeg) were converted into molar ellipticity. Error bars are ± 1σ.

Table 2 Secondary structure analysis of CD spectra from BeStSel

	% Helix	% Antiparallel	% Parallel	% Turn	% Others
eGFP only	9.6	45.1	0.0	14.6	30.7
LPAA only	84.5	2.3	0.0	13.1	0.0
LPAA, after conjugation	No fit possible				
LPAA-eGFP	35.6	27.9	15.4	14.1	7.1
DPAA only	69.7	0.4	2.9	10.5	16.4
DPAA, after conjugation	65.8	2.0	26.1	6.1	0.0
DPAA-eGFP	27.1	0.0	72.9	0.0	0.0
eGFP, after LPAA conjugation	0.4	36.5	0.6	16.3	46.3
eGFP, after DPAA conjugation	2.8	41.1	0.0	11.0	45.1
eGFP, after LDPAA conjugation	1.4	43.9	0.0	13.3	41.4
eGFP (2Y0G crystal structure)[60]	4.5	47.1	0.0	8.1	40.4
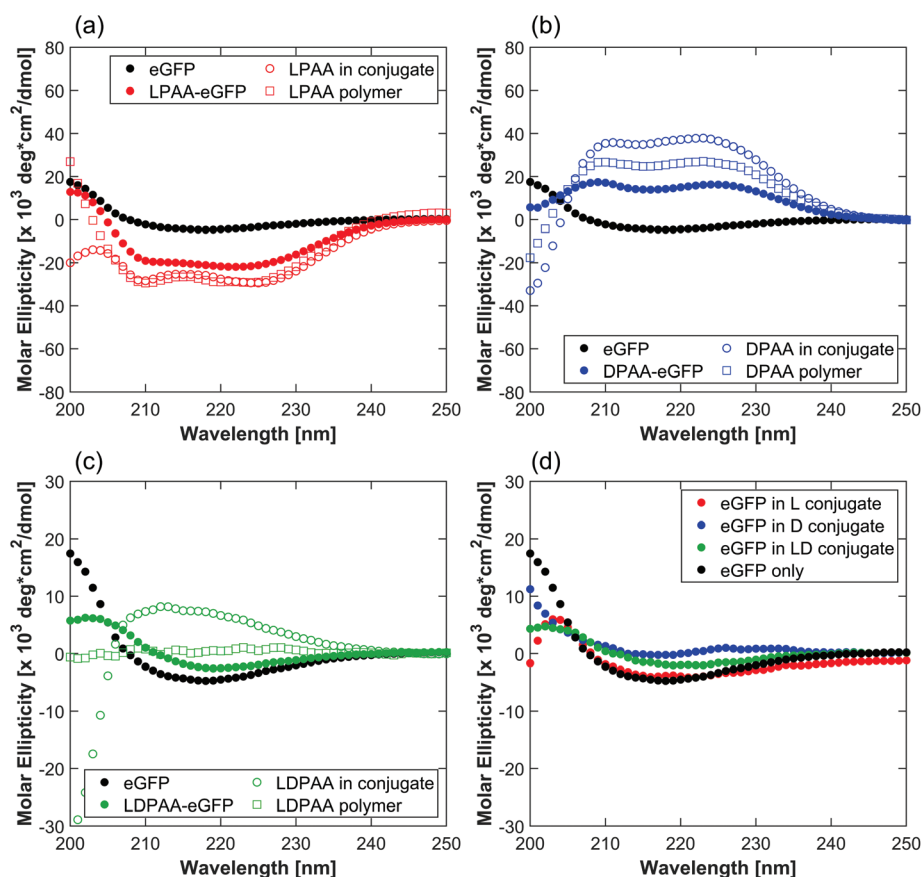
Fig. 3 CD spectra for variants of (a) LPAA, (b) DPAA, (c) LDPAA, and (d) eGFP. For all four figures, the eGFP spectrum measured from a solution of only eGFP is shown in closed black circles. For figures (a)–(c), the CD spectra for the bioconjugates (XPAA-eGFP) are shown in closed colored circles (red for LPAA, blue for DPAA, and green in LDPAA). The CD spectra for the PAA while in the bioconjugate are shown in open circles; this was calculated by subtracting out the eGFP signal. The CD spectra for the PAA only (not in a bioconjugate) are shown in open squares. For figure (d), the CD spectra for eGFP in each type of conjugate after subtracting out the PAA signal is shown in colored circles (red for LPAA, blue for DPAA, and green for LDPAA).

eGFP is subtracted, and eGFP after PAAs are subtracted. For LPAA and DPAA, Fig. 3a and b show that the structure of the PAA after conjugation is very similar to that of the PAA alone. However, for LDPAA, the signal differs between LDPAA alone and LDPAA from the bioconjugate in the low wavelength regime, as shown in Fig. 3c. This suggests that in the bioconjugate, the LDPAA may take on some secondary structure.

Bioconjugation does not change the secondary structure of eGFP significantly. The secondary structure of the eGFP alone is similar to the predicted eGFP secondary structure based on its crystal structure (PDB: 2Y0G) using projected CD data from PDB2CD,[60] with a mainly antiparallel β-sheet contribution (~40–50%), a small β-turn contribution, a very minor α-helix contribution, and the rest in "other", which captures imperfect

or distorted formations and random coils (Table 2). In all cases, eGFP appears to retain its basic secondary structure even after conjugation, which shows that the presence of the poly(amino acid) does not significantly affect its folding.

Bioconjugation does not reduce α-helical character in either homochiral polymer. Fractional helicity for DPAA was calculated in two ways: using BeStSel and using the characteristic CD signal at 222 nm for an α-helix (eqn (2)).[61]

$$f_{\text{helicity, 222 nm}} = \frac{\theta^{\text{exp}} - \theta^{0\%}}{\theta^{100\%} - \theta^{0\%}}. \qquad (2)$$

In eqn (2), fraction helicity ($f_{\text{helicity, 222 nm}}$) is defined as the ratio of the experimental molar ellipticity at 222 nm (θ^{exp}) to the theoretical molar ellipticity of a perfect α-helix at 222 nm ($\theta^{100\%} = -39\,500$ deg cm^2 dmol^{-1}), normalized by the molar ellipticity of a 0% helicity protein ($\theta^{0\%} = -3000$ deg cm^2 dmol^{-1}). From BeStSel, the secondary structure of the DPAA only is predominantly α-helix (69.7%), but less so than LPAA (84.5%), with contributions from β-turns and other structures. After conjugation, the DPAA appears to be 65.8% α-helical, with the rest of the contributions from parallel β-sheets and turns. This change in the other elements may be a result of secondary structure change upon conjugation. From eqn (2), DPAA after conjugation has a calculated helicity of 95.2%, compared to 65% for DPAA only. This is because the signal at 222 nm for conjugated DPAA is higher than that for DPAA only. For LPAA after conjugation, the helical content could only be estimated using eqn (2) due to the small dip in the CD signal after subtraction at ~205 nm, and CD fitting cannot be performed reliably without including the entire range of 200–240 nm.[59,62,63] Outside of this distortion, the CD spectra of the LPAA in the conjugate aligns very well to the LPAA polymer only (Fig. 3a). From eqn (2), LPAA before and after conjugation has a helical content of 71%.

Self-assembly of bioconjugates

All solution-state eGFP conjugates incorporating a helical polymer block (LPAA-eGFP, DPAA-eGFP, and the 1 : 1 blend of LPAA- and DPAA-eGFP) form a lamellar phase for the entire range of concentrations investigated, as shown in the phase diagrams in Fig. 4. Ordered (i.e. below the order–disorder transition temperature, T_{ODT}), low concentration helical conjugates exhibit SAXS curves with three to four broad peaks indexed at q^*, $2q^*$, $3q^*$, and $4q^*$. This is illustrated in Fig. 5 for LPAA-eGFP; the other bioconjugates exhibit similar SAXS patterns (Fig. S7–S9†). At concentrations below 40 wt%, the lamellar phase transitions into a disordered phase when heated, with T_{ODT} increasing with increasing concentration. For all biocon-
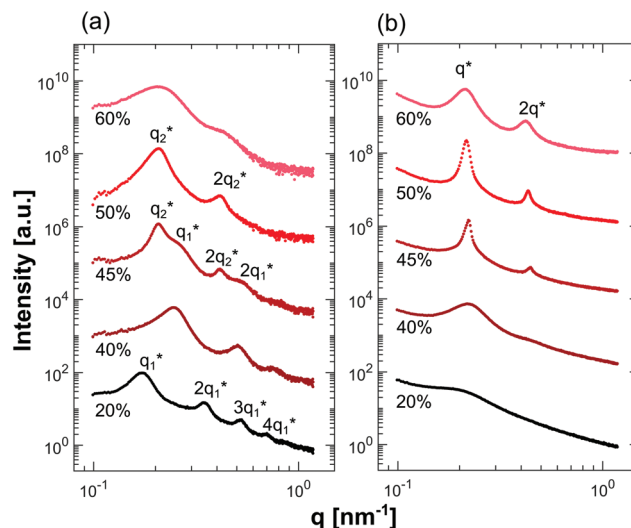

Fig. 5 SAXS intensity curves for LPAA-eGFP at (a) 10 °C and (b) 50 °C for selected concentrations (wt%). The lamellar peaks (q^*, $2q^*$, $3q^*$, $4q^*$) are labeled for both lamellar phases (low concentration 1 and high concentration 2). SAXS patterns are offset for clarity. 45 wt% forms both lamellar phases.
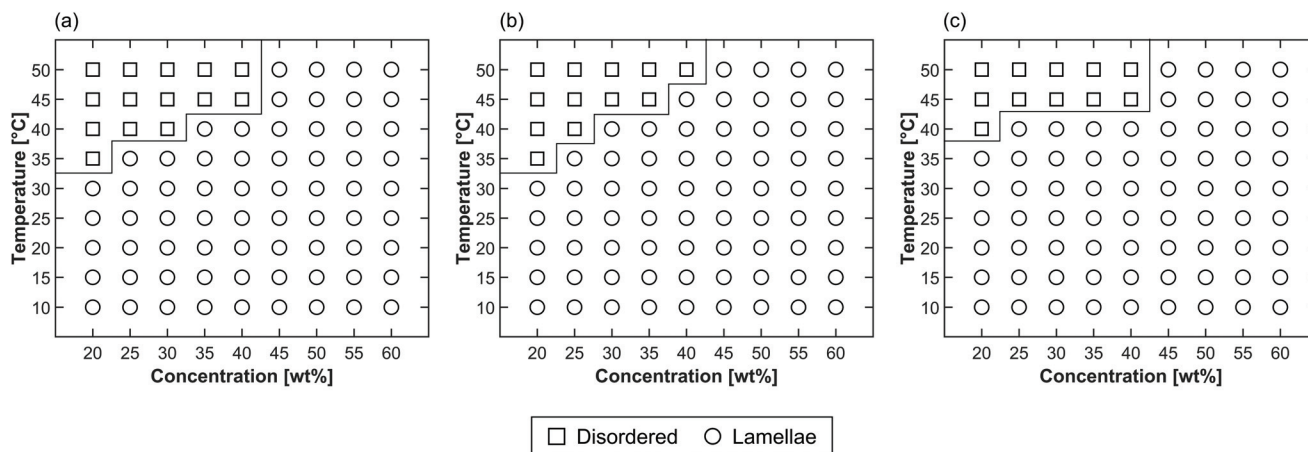

Fig. 4 Phase diagrams for (a) LPAA-eGFP, (b) DPAA-eGFP, and (c) 1 : 1 blend of LPAA-eGFP and DPAA-eGFP (L : D 1 : 1-eGFP). None of the samples showed significant birefringence, with all power fractions below 0.15. Birefringence curves for each sample can be found in the ESI.†
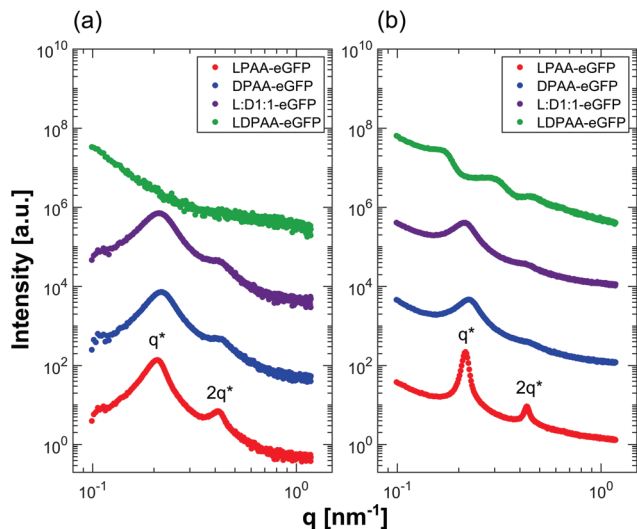
Fig. 6 SAXS intensity curves for all bioconjugates at 50 wt% for (a) 10 °C and (b) 50 °C. The lamellar peaks ($q*$, $2q*$) are labeled for the LPAA-eGFP curve. SAXS patterns are offset for clarity. All peaks observed in LPAA-eGFP, DPAA-eGFP, and L : D 1 : 1-eGFP correspond to lamellae. The peaks in (b) correspond to aggregation at high temperature in LDPAA-eGFP samples.

jugates, the disordered phase remains a homogeneous single phase. At concentrations above 40 wt%, the lamellar phase persists from 10 to 50 °C. However, as shown in Fig. 5, the $3q*$ and $4q*$ higher order peaks disappear as concentration increases from 40 wt% to 50 wt%. This is accompanied by a simultaneous discontinuous shift in the domain spacing (Fig. 7) and sharpening of the $q*$ and $2q*$ lamellar peaks, which is most pronounced at higher temperatures (Fig. 5 and 6). The changes in the SAXS patterns and in the domain spacing at high concentration suggest that there is a phase transition at around 45 wt% between two distinct lamellar phases. While turbidimetry did not reveal macrophase separation consistent with a coexistence region, phase coexistence is permitted by the phase rule and consistent with the simultaneous observation of two series of peaks by SAXS which both index to different lamellae. At concentrations higher than 50 wt%, the $q*$ and $2q*$ peaks broaden once more, indicating a decrease in ordering quality. It is unclear whether the bioconjugates may be going through a reentrant ODT, as this feature is not as apparent as observed for globular protein–coil block copolymers.[15,24–27] Above 40 wt%, for bioconjugates with the same concentration, the $q*$ and $2q*$ peaks sharpen as tempera-
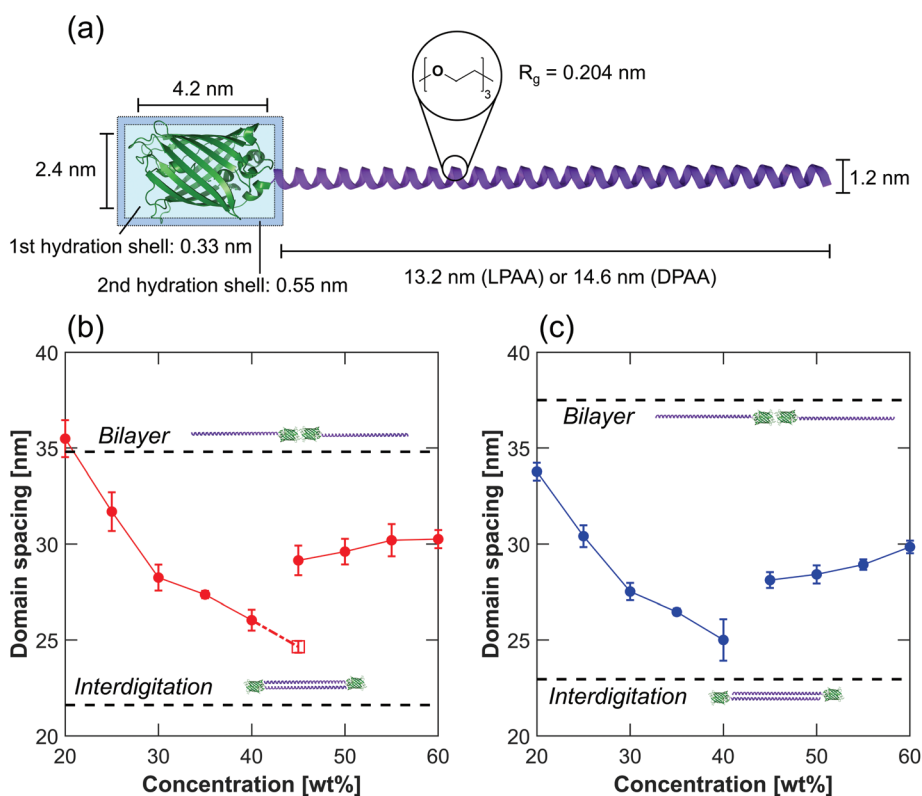

Fig. 7 (a) Dimensions of a bioconjugate. The eGFP block is a β-barrel modeled as a cylinder. It has two hydration shells. The PAA block is an α-helix with (EG)$_3$ side chains (shown in inset). Average domain spacing calculated from the $q*$ peak for (b) LPAA-eGFP and (c) DPAA-eGFP. The domain spacing was averaged over all the temperatures over which a specific sample formed the same phase, since it is independent of temperature. Error bars represent $\pm 1\sigma$. For all three types of bioconjugates, the domain spacing has a non-monotonic trend with the concentration. The closed circles represent domain spacing calculated from the main SAXS peak. The open square is the domain spacing calculated for the second set of peaks for LPAA-eGFP at 45 wt%, which exhibited two lamellar phases. The dashed lines (- -) represent the theoretical domain spacing if the bioconjugates arranged into bilayers or interdigitated layers, as shown in the inset schematics. The L : D 1 : 1-eGFP blend (Fig. S16†) was similar to both the LPAA and DPAA, with two lamellar phases at 45 wt%.

ture increases. Based on these features, the optimal ordering concentration appears to be around 50 wt% (plotted in Fig. 6 for different bioconjugates), which has also been observed to be the best regime for ordering in globular protein–coil block copolymers.[25,26]

For all three bioconjugates, no macrophase separation was detected for any concentration. Birefringence was not observed for any helical bioconjugate except in the most concentrated samples of the 1 : 1 blend (Fig. S13†), and even in such cases, it was very weak (Fig. S11–S14†). This is unusual for block copolymers with a rod block, as rod alignment tends to lead to liquid crystalline phases.[33] Here, the lack of liquid crystalline phases may be due to small grain sizes or steric hindrance from the globular protein block, preventing the rod-like helices from aligning easily.

Domain spacing calculated from the position of the q^* peak ($d = 2\pi/q^*$) provides further evidence for a phase transition between two distinct lamellar phases at ~45 wt%. The lamellar domain sizes change non-monotonically and discontinuously with concentration for LPAA-eGFP, DPAA-eGFP, and L : D 1 : 1-eGFP (see Fig. 7). The domain spacing is relatively insensitive to temperature (Fig. S15†). In the low concentration lamellar phase (<45 wt%), the domain spacing indicates a transition from bilayers to interdigitated lamellae as concentration increases. Assuming that eGFP is 4.2 nm in length[64] and that the α-helix is 13.2 nm or 14.6 nm in length (LPAA or DPAA, respectively, with perfect helices),[65] a bilayer morphology would have a domain spacing of approximately 34.8 and 37.5 nm for LPAA-eGFP and DPAA-eGFP, respectively (top dashed line in Fig. 7). As the concentration increases, the conjugates are forced to pack more tightly, leading to interdigitating helices. Interdigitation of helices has been observed in many polypeptide rod–coil and rod–rod block copolymers.[66–69] A fully interdigitated helix layer with head-to-head eGFP blocks (as illustrated in Fig. 7) is estimated to have a width of 21.6 nm (LPAA-eGFP) or 23.0 nm (DPAA-eGFP). Here, the domain spacing reaches a minimum between 24–26 nm at 45 wt% in LPAA-eGFP and at 40 wt% in DPAA-eGFP. An analysis of interfacial area indicates that interdigitation is possible (dimensions shown in Fig. 7a). GFP has a diameter of approximately 2.4 nm.[64] In addition, it is surrounded by two hydration shells that are 0.33 and 0.55 nm,[70] extending the GFP diameter to about 4.16 nm. A typical α-helix has a diameter of 1.2 nm.[71] Since the PAAs have an additional small PEG trimer side group, the rod block actually has a diameter of 1.61 nm, assuming that a PEG trimer has a radius of gyration of 0.204 nm in good solvent.[72] Interdigitated PAA rods would thus have a collective diameter of 3.22 nm, which is less than the total hydrated GFP diameter.

When bioconjugate concentration further increases (>45 wt%), there is a phase transition into a new lamellar phase with domain spacing between bilayer and fully interdigitated helices. Several factors may be at play in driving this phase transition. At higher concentration, there are fewer solvent molecules in the system. For coil–coil block copolymers, solvent selectivity has been shown to affect domain

spacing.[73] Non-selective solvents increase the domain spacing by partitioning to the space between blocks to reduce unfavorable interactions.[73] At high concentrations of bioconjugates, the helix block may also distort, leading to changes in the domain spacing. The observation of helical distortions is not unusual; for many rod–coil systems that incorporate a helical block, the helix often shows evidence of bending, folding, or stretching, depending on a number of factors, including backbone interactions, confinement conditions, chemical identity of the other block, and solvent environment.[33,37,42,74] Helical distortion is consistent with the absence of birefringence in the materials. Due to the presence of multiple competing effects, it is difficult to definitively determine the specific morphology of the high concentration lamellar phase.

The strong similarities in the phase behavior of LPAA-eGFP, DPAA-eGFP, and L : D 1 : 1-eGFP (Fig. 4 and 6) suggest that chain structure (i.e. rod rigidity) plays a more dominant role in driving self-assembly than chirality. All three bioconjugates have nearly identical phase diagrams (Fig. 4), with the same types of nanostructures and phase transitions being observed. The main difference between the SAXS data for LPAA-eGFP and DPAA-eGFP is the width of the lamellar peaks, indicating differences in grain size, quality of ordering, and/or interfacial sharpness. LPAA-eGFP exhibited the best order of the three helical bioconjugates, with the narrowest peaks across all lamellar phases. This indicates that the chirality of the polymer backbone may impact molecular packing. DPAA is also slightly longer than LPAA as synthesized, which may lead to a higher frequency of distorted helices, which can reduce order. The LPAA-eGFP also exhibited a transitional SAXS pattern at 45 wt%, where both the 4-peak lamellar and 2-peak lamellar phases were apparent (see Fig. 5a). DPAA-eGFP also exhibited both lamellar phases, but coexistence was never observed by SAXS. The L : D 1 : 1-eGFP blend exhibits self-assembly behavior that is more similar to the DPAA-eGFP. The only features that point to the presence of LPAA-eGFP in the blend are the observation of a transitional SAXS pattern at 45 wt% and slightly sharper peaks than a pure DPAA-eGFP bioconjugate. The strong similarities between L : D 1 : 1-eGFP and DPAA-eGFP are due to the higher volume fraction of DPAA in this bioconjugate, since it is longer than LPAA. These minor differences in phase behavior among the helical bioconjugates suggest that chirality is not a driving force for self-assembly but can impact the packing of the nanostructures.

In contrast to the LPAA-eGFP and the DPAA-eGFP, when eGFP is conjugated to an achiral random copolymer of $(EG)_3$-L-Glu and $(EG)_3$-D-Glu (LDPAA-eGFP), there is no self-assembly observed at any temperature or concentration. The only SAXS features that appear are wide correlation peaks that are indicative of aggregation at intermediate concentrations (35–50 wt%) and high temperature (>40 °C) (Fig. 6). As temperature and concentration increases, solvent selectivity and solubility change, leading to aggregation of bioconjugates. This can be observed as a slight decrease in light transmission with temperature as measured by DPLS for 35–50 wt% bioconjugates (Fig. S14†). These peaks can be fit to a hard sphere structure
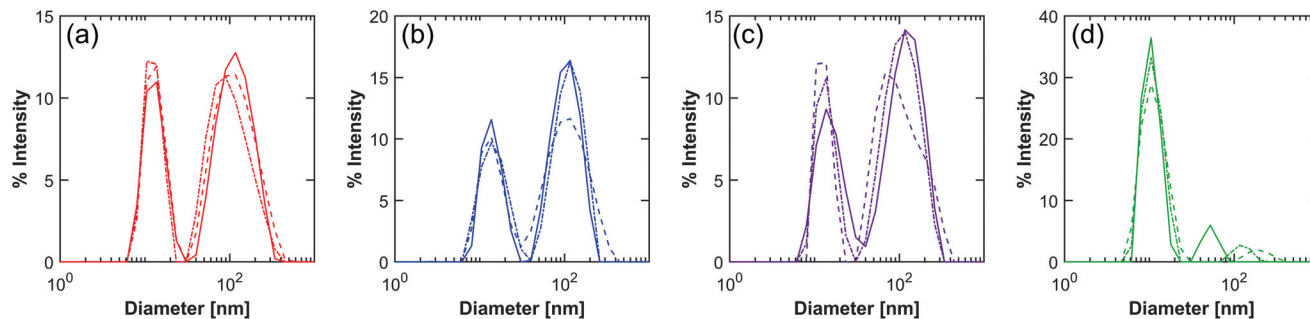
Fig. 8 Size distribution histograms measured from DLS (represented as %intensity as a function of diameter) for (a) LPAA-eGFP, (b) DPAA-eGFP, (c) L : D 1 : 1-eGFP, and (d) LDPAA-eGFP at 1 mg mL⁻¹ in water. Each curve represents a replicate.

Table 3 Peak diameters for bioconjugates from DLS

Sample	Peak 1 Diameter [nm]	Peak 2 Diameter [nm]
LPAA-eGFP	13.2 ± 0.4	123 ± 10
DPAA-eGFP	14.8 ± 0.5	118 ± 7
L : D 1 : 1-eGFP	14 ± 2	123 ± 7
LDPAA-eGFP	11.3 ± 0.5	Not significant

factor with Percus–Yevick closure[75] and the form factor for spheres (Fig. S18†). While this structure factor captures the main features of the SAXS peaks, it is an imperfect fit, indicating that aggregation occurs but not necessarily as spherical structures. At very high concentrations above 50 wt%, the bioconjugates are not soluble in water at any temperature in the range of 10–50 °C. At these very high concentrations, the samples macrophase separate into a bioconjugate-rich and a bioconjugate-poor phase, both of which are disordered, as reflected in the corresponding SAXS curves (Fig. S10†).

The lack of self-assembly behavior in the LDPAA-eGFP suggests that the PAA in a flexible form may not be repulsive enough to the eGFP block to drive self-assembly. In the case of the helical PAAs, the rigidity prevents association between protein and polymer, leading to self-assembly behavior. These interactions are reflected in the self-assembly behavior in dilute solution, where helical bioconjugates formed different structures from LDPAA bioconjugates. DLS shows that at 1 mg mL⁻¹, bioconjugates with a helical block form two populations: a population in the tens of nm corresponding to a single molecule of bioconjugate, and a population at about 100 nm corresponding to a soluble aggregate (see Fig. 8 and Table 3). As shown in Table 3, the single molecule diameters are similar across all bioconjugates, with the LDPAA-eGFP being slightly smaller. The diameters of the LPAA-eGFP and DPAA-eGFP are very close to the predicted dimensions of a protein–rod dumbbell. The aggregated population is not present in the LDPAA-eGFP, suggesting that even in dilute solution, there is self-assembly in the bioconjugates with a helical block but not in those with a random copolymer block.

Further evidence of interaction between LDPAA and eGFP can be seen in the UV-Vis absorbance spectra in Fig. 9. It is

known that the GFP chromophore can take on two configurations that absorb at two different wavelengths.[76,77] eGFP was designed to absorb maximally at around 485 nm, which means the chromophore is deprotonated, whereas the wild-type GFP (wtGFP) absorbs maximally at around 395 nm, which results from a protonated chromophore. Here, when dissolved in non-buffered Milli-Q water, the eGFP chromophore shows characteristics of both the eGFP and the wtGFP (Fig. 9) when eGFP is conjugated to a helical block (LPAA or DPAA). However, when eGFP is conjugated to LDPAA, there is only one absorption peak at 485 nm, which is the expected excitation wavelength for eGFP. The eGFP chromophore can become protonated when it is in a low pH environment.[76,77] Since the water is not buffered, eGFP can become protonated when conjugated to a helical block that does not interact with it. This is not unprecedented, as the same characteristic two-peak absorption spectra have been observed for eGFP conjugated to poly(N-isopropylacrylamide) (PNIPAM), which is another self-assembling GCBC system with a non-interacting polymer.[28,78] However, since LDPAA-eGFP seems to keep its deprotonated chromophore even though it is in the same solvent, this is evidence that the LDPAA interacts with the eGFP in a way that
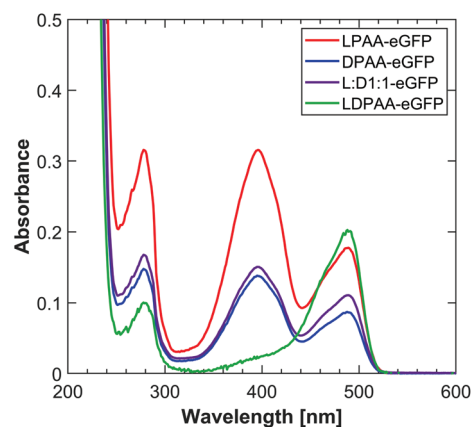

Fig. 9 UV-Vis absorbance curves for LPAA-eGFP (red), DPAA-eGFP (blue), L : D 1 : 1-eGFP (purple), and LDPAA-eGFP (green) at 0.5 mg mL⁻¹ in water.

prevents the eGFP chromophore from interacting with additional water molecules. These observations suggest that conjugation to an interacting polymer may also be able to tune the chromophores of other fluorescent proteins, which are also sensitive to pH.[79]

The non-repulsive interactions between LDPAA and eGFP may be mediated by the short poly(ethylene glycol) (PEG)-like side chain $(EG)_3$, which was added to the PAA to increase solubility in water and remove the possibility of charge effects when using poly(glutamic acid). PEG and PEG-like polymers have been shown in the literature to associate with and stabilize proteins.[80,81] Huang *et al.* showed that poly(oligoethylene glycol acrylate) (POEGA), which has PEG-like functionalities, is able to form an adsorption layer around mCherry, another fluorescent β-barrel protein.[78] The Xu group synthesized and stabilized α-helical bundle peptides by attaching PEG side chains.[82] Simulations of these conjugates showed that the PEG chain can interact with residues on the peptide surface to promote intramolecular hydrogen bonding[83] and reduce solvent-accessible surface area for hydrophobic residues.[84] These previous works indicate that certain globular proteins can have non-repulsive and even attractive interactions with PEG and PEG-like functional groups.

Compared to globular protein–coil block copolymers, globular protein–rod bioconjugates with a helical block exhibited a wider window of self-assembly with respect to concentration in solution. In addition, since the PAA block is uncharged, it is expected that factors such as pH and ionic strength would affect these globular protein–rod bioconjugates in a similar manner as for GCBCs that have been previously studied.[28,32] Like many other globular protein–coil block copolymers with a symmetric coil fraction, the LPAA-eGFP, DPAA-eGFP, and their blend self-assemble into a lamellar phase, which is the most commonly observed nanostructure for these systems.[25,27] However, here, a lamellar phase is observed even at the lowest concentration studied at SAXS (20 wt%) for all three helical block systems. Other globular protein–coil block copolymers that have previously been studied have a critical order–disorder concentration (C_{ODT}) below which the system is disordered that lies above 20 wt%.[25,27] The prevalence of lamellae and the lack of any other nanostructure (*e.g.* hexagonally packed cylinders, spheres, or gyroid) is due to the rod-like nature of the helical PAA block. In rod–coil block copolymers, curved interfaces are unfavorable due to rod alignment effects.[33] Block structure of the protein has also been shown to be a strong indicator of self-assembly in other GCBCs. In a panel of 15 GCBCs, each with a different protein block, it was found that self-assembly and ordering quality is largely predicted by β-sheet content and molar mass of the protein.[29]

The LDPAA-eGFP behaves very differently from other GCBCs because it appears to lack sufficient segregation strength between blocks to drive self-assembly. The loss of chirality prevents the formation of secondary structure in the polymer block. Without the rigidity of the α-helix, non-repulsive interactions between the LDPAA random copolymer and eGFP dominate, leading to disorder. However, the types of

chemical functionalities in LDPAA have not prevented other polymer coil blocks from self-assembling. For example, bioconjugates incorporating mCherry and POEGA can self-assemble into lamellae, hexagonally packed cylinders, and gyroid phases at a similar coil fraction.[25] In addition, fusion proteins of mCherry and elastin-like polypeptides (ELPs), which do have a peptide backbone but a less uniform peptide sequence, also exhibit self-assembly behavior.[31,85] Here, the combination of the peptide backbone, the glutamic acid, and the PEG-based side chain appears to decrease the repulsive potential between the LDPAA polymer and the eGFP protein. Thus, the LDPAA-eGFP bioconjugate does not have a strong enough driving force for the formation of self-assembled nanostructures.

Conclusions

Self-assembly was demonstrated in globular protein–helix block copolymers, with similar lamellar phases being formed for both L and D chirality α-helices. Using NCA polymerization and native chemical ligation, tightly controlled, tunable poly(amino acid) blocks with different chirality can be conjugated easily to proteins. Secondary structure of the polymer block played a more significant role in driving self-assembly than type of chirality. In contrast, no self-assembly was found in the LDPAA random copolymer-based block copolymer, suggesting that rigidity and chemical structure play a significant role in tuning segregation strength between the poly(amino acid) and protein blocks. The weak segregation between LDPAA and eGFP may be due to the $(EG)_3$ side chain of the PAA, preventing the formation of nanostructures. These results have important implications for designing devices that integrate protein functionality with block copolymer self-assembly. Secondary structure-induced rigidity in the polymer block can increase the effective segregation strength between blocks and drive self-assembly even in systems with weak to no repulsive interactions between blocks.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

References

1 B. Hussain, M. Yüce, N. Ullah and H. Budak, in *Nanobiosensors*, ed. A. M. Grumezescu, Academic Press, 2017, ch. 3, pp. 93–127, DOI: 10.1016/B978-0-12-804301-1.00003-5.

2 J. F. Liang, Y. T. Li and V. C. Yang, *J. Pharm. Sci.*, 2000, **89**, 979–990.

3 M. Coradi, M. Zanetti, A. Valério, D. de Oliveira, A. da Silva, S. M. A. G. U. de Souza and A. A. Ulson de Souza, *Mater. Chem. Phys.*, 2018, **208**, 28–34.

4 A. J. Russell, J. A. Berberich, G. F. Drevon and R. R. Koepsel, *Annu. Rev. Biomed. Eng.*, 2003, **5**, 1–27.

5 A. R. Freedman, G. Galfre, E. Gal, H. J. Ellis and P. J. Ciclitira, *J. Immunol. Methods*, 1987, **98**, 123–127.

6 Y. Lee, T. Ishii, H. J. Kim, N. Nishiyama, Y. Hayakawa, K. Itaka and K. Kataoka, *Angew. Chem.*, 2010, **122**, 2606–2609.

7 S. t. Janeček, *Process Biochem.*, 1993, **28**, 435–445.

8 V. G. H. Eijsink, S. Gåseidnes, T. V. Borchert and B. van den Burg, *Biomol. Eng.*, 2005, **22**, 21–30.

9 V. G. H. Eijsink, A. Bjork, S. Gaseidnes, R. Sirevag, B. Synstad, B. van den Burg and G. Vriend, *J. Biotechnol.*, 2004, **113**, 105–120.

10 V. M. Balcão and M. M. D. C. Vila, *Adv. Drug Delivery Rev.*, 2015, **93**, 25–41.

11 U. T. Bornscheuer, *Angew. Chem., Int. Ed.*, 2003, **42**, 3336–3337.

12 U. Hanefeld, L. Gardossi and E. Magner, *Chem. Soc. Rev.*, 2009, **38**, 453–468.

13 S. Jevsevar, M. Kunstelj and V. G. Porekar, *Biotechnol. J.*, 2010, **5**, 113–128.

14 E. M. Pelegri-O'Day, E. W. Lin and H. D. Maynard, *J. Am. Chem. Soc.*, 2014, **136**, 14323–14332.

15 C. S. Thomas, M. J. Glassman and B. D. Olsen, *ACS Nano*, 2011, **5**, 5697–5707.

16 A. Huang, G. Qin and B. D. Olsen, *ACS Appl. Mater. Interfaces*, 2015, **7**, 14660–14669.

17 X. H. Dong, A. C. Obermeyer and B. D. Olsen, *Angew. Chem., Int. Ed.*, 2017, **56**, 1273.

18 J. M. Paloni, E. A. Miller, H. D. Sikes and B. D. Olsen, *Biomacromolecules*, 2018, **19**, 3814–3824.

19 K. Velonia, A. E. Rowan and R. J. M. Nolte, *J. Am. Chem. Soc.*, 2002, **124**, 4224–4225.

20 C. Lavigueur, J. G. Garcia, L. Hendriks, R. Hoogenboom, J. J. L. M. Cornelissen and R. J. M. Nolte, *Polym. Chem.*, 2011, **2**, 333–340.

21 B. Jeong and A. Gutowska, *Trends Biotechnol.*, 2002, **20**, 305–311.

22 A. S. Hoffman, P. S. Stayton, V. Bulmus, G. Chen, J. Chen, C. Cheung, A. Chilkoti, Z. Ding, L. Dong, R. Fong, C. A. Lackey, C. J. Long, M. Miura, J. E. Morris, N. Murthy, Y. Nabeshima, T. G. Park, O. W. Press, T. Shimoboji, S. Shoemaker, H. J. Yang, N. Monji, R. C. Nowinski, C. A. Cole, J. H. Priest, J. M. Harris, K. Nakamae, T. Nishino and T. Miyata, *J. Biomed. Mater. Res.*, 2000, **52**, 577–586.

23 M. W. Matsen and F. S. Bates, *Macromolecules*, 1996, **29**, 1091–1098.

24 C. S. Thomas and B. D. Olsen, *Soft Matter*, 2014, **10**, 3093–3102.

25 D. Chang, C. N. Lam, S. Tang and B. D. Olsen, *Polym. Chem.*, 2014, **5**, 4884–4895.

26 C. N. Lam and B. D. Olsen, *Soft Matter*, 2013, **9**, 2393–2402.

27 D. Chang and B. D. Olsen, *Polym. Chem.*, 2016, **7**, 2410–2418.

28 C. N. Lam, M. Kim, C. S. Thomas, D. Chang, G. E. Sanoja, C. U. Okwara and B. D. Olsen, *Biomacromolecules*, 2014, **15**, 1248–1258.

29 A. Huang, J. M. Paloni, A. Wang, A. C. Obermeyer, H. V. Sureka, H. Yao and B. D. Olsen, *Biomacromolecules*, 2019, **20**, 3713–3723.

30 C. E. Mills, Z. Michaud and B. D. Olsen, *Biomacromolecules*, 2018, **19**, 2517–2525.

31 G. Qin, M. J. Glassman, C. N. Lam, D. Chang, E. Schaible, A. Hexemer and B. D. Olsen, *Adv. Funct. Mater.*, 2015, **25**, 729–738.

32 C. N. Lam, H. Yao and B. D. Olsen, *Biomacromolecules*, 2016, **17**, 2820–2829.

33 B. D. Olsen and R. A. Segalman, *Mater. Sci. Eng., R*, 2008, **62**, 37–66.

34 A. Kros, W. Jesse, G. A. Metselaar and J. J. L. M. Cornelissen, *Angew. Chem., Int. Ed.*, 2005, **44**, 4349–4352.

35 R. J. I. Knoop, G. J. M. Habraken, N. Gogibus, S. Steig, H. Menzel, C. E. Koning and A. Heise, *J. Polym. Sci., Part A: Polym. Chem.*, 2008, **46**, 3068–3077.

36 A. Carlsen and S. Lecommandoux, *Curr. Opin. Colloid Interface Sci.*, 2009, **14**, 329–339.

37 E. C. Davidson, A. M. Rosales, A. L. Patterson, B. Russ, B. Yu, R. N. Zuckermann and R. A. Segalman, *Macromolecules*, 2018, **51**, 2089–2098.

38 V. Pryamitsyn and V. Ganesan, *J. Chem. Phys.*, 2004, **120**, 5824–5838.

39 B. D. Olsen, M. Shah, V. Ganesan and R. A. Segalman, *Macromolecules*, 2008, **41**, 6809–6817.

40 A. Douy and B. Gallot, *Polymer*, 1987, **28**, 147–154.

41 B. Perly, A. Douy and B. Gallot, *Macromol. Chem. Phys.*, 1976, **177**, 2569–2589.

42 A. Douy and B. Gallot, *Polymer*, 1982, **23**, 1039–1044.

43 J.-P. Billot, A. Douy and B. Gallot, *Macromol. Chem. Phys.*, 1977, **178**, 1641–1650.

44 J.-P. Billot, A. Douy and B. Gallot, *Macromol. Chem. Phys.*, 1976, **177**, 1889–1893.

45 A. I. Triftaridou, F. Chécot and I. Iliopoulos, *Macromol. Chem. Phys.*, 2010, **211**, 768–777.

46 K. E. Gebhardt, S. Ahn, G. Venkatachalam and D. A. Savin, *J. Colloid Interface Sci.*, 2008, **317**, 70–76.

47 R. Graf, H. W. Spiess, G. Floudas, H. J. Butt, M. Gkikas and H. Iatrou, *Macromolecules*, 2012, **45**, 9326–9332.

48 M.-T. Popescu, G. Liontos, A. Avgeropoulos and C. Tsitsilianis, *Soft Matter*, 2015, **11**, 331–342.

49 S. L. Perry, L. Leon, K. Q. Hoffmann, M. J. Kade, D. Priftis, K. A. Black, D. Wong, R. A. Klein, C. F. Pierce Iii, K. O. Margossian, J. K. Whitmer, J. Qin, J. J. de Pablo and M. Tirrell, *Nat. Commun.*, 2015, **6**, 6052.

50 K. Q. Hoffmann, S. L. Perry, L. Leon, D. Priftis, M. Tirrell and J. J. de Pablo, *Soft Matter*, 2015, **11**, 1525–1538.

51 T. Wen, H.-F. Wang, M.-C. Li and R.-M. Ho, *Acc. Chem. Res.*, 2017, **50**, 1011–1021.

52 R.-M. Ho, Y.-W. Chiang, C.-K. Chen, H.-W. Wang, H. Hasegawa, S. Akasaka, E. L. Thomas, C. Burger and B. S. Hsiao, *J. Am. Chem. Soc.*, 2009, **131**, 18533–18542.

53 Y. Hou, J. Yuan, Y. Zhou, J. Yu and H. Lu, *J. Am. Chem. Soc.*, 2016, **138**, 10995–11000.

54 Y. Hou, Y. Zhou, H. Wang, J. Sun, R. Wang, K. Sheng, J. Yuan, Y. Hu, Y. Chao, Z. Liu and H. Lu, *ACS Cent. Sci.*, 2019, **5**, 229–236.

55 C. Chen, Z. Wang and Z. Li, *Biomacromolecules*, 2011, **12**, 2859–2863.

56 J. Yuan, Y. Sun, J. Wang and H. Lu, *Biomacromolecules*, 2016, **17**, 891–896.

57 G. Wu and H. Lu, *ChemNanoMat*, 2016, **2**, 354–363.

58 Y. Hou, Y. Zhou, H. Wang, R. Wang, J. Yuan, Y. Hu, K. Sheng, J. Feng, S. Yang and H. Lu, *J. Am. Chem. Soc.*, 2018, **140**, 1170–1178.

59 A. Micsonai, F. Wien, L. Kernya, Y.-H. Lee, Y. Goto, M. Réfrégiers and J. Kardos, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, E3095–E3103.

60 L. Mavridis and R. W. Janes, *Bioinformatics*, 2017, **33**, 56–63.

61 Y. Wei, A. A. Thyparambil and R. A. Latour, *Biochim. Biophys. Acta*, 2014, **1844**, 2331–2337.

62 L. Whitmore and B. A. Wallace, *Nucleic Acids Res.*, 2004, **32**, W668–W673.

63 L. Whitmore and B. A. Wallace, *Biopolymers*, 2008, **89**, 392–400.

64 M. Ormo, A. B. Cubitt, K. Kallio, L. A. Gross, R. Y. Tsien and S. J. Remington, *Science*, 1996, **273**, 1392–1395.

65 J. E. Murray, N. Laurieri and R. Delgoda, in *Pharmacognosy*, ed. S. Badal and R. Delgoda, Academic Press, Boston, 2017, ch. 24, pp. 477–494, DOI: 10.1016/B978-0-12-802104-0.00024-X.

66 F. Zhou, T. Ye, L. Shi, C. Xie, S. Chang, X. Fan and Z. Shen, *Macromolecules*, 2013, **46**, 8253–8263.

67 M. Lee, B.-K. Cho, H. Kim, J.-Y. Yoon and W.-C. Zin, *J. Am. Chem. Soc.*, 1998, **120**, 9168–9179.

68 H. Schlaad, H. Kukula, B. Smarsly, M. Antonietti and T. Pakula, *Polymer*, 2002, **43**, 5321–5328.

69 M. Gkikas, J. S. Haataja, J. Seitsonen, J. Ruokolainen, O. Ikkala, H. Iatrou and N. Houbenov, *Biomacromolecules*, 2014, **15**, 3923–3930.

70 S. Perticaroli, G. Ehlers, C. B. Stanley, E. Mamontov, H. O'Neill, Q. Zhang, X. Cheng, D. A. A. Myles, J. Katsaras and J. D. Nickels, *J. Am. Chem. Soc.*, 2017, **139**, 1098–1105.

71 C. Branden and J. Tooze, *Introduction to Protein Structure*, Garland Publishing, New York, 2nd edn, 1999.

72 Z. Ma, D. N. LeBard, S. M. Loverde, K. A. Sharp, M. L. Klein, D. E. Discher and T. H. Finkel, *PLoS One*, 2014, **9**, e112292.

73 C. Lai, W. B. Russel and R. A. Register, *Macromolecules*, 2002, **35**, 4044–4049.

74 G. Floudas, P. Papadopoulos, H. A. Klok, G. W. M. Vandermeulen and J. Rodriguez-Hernandez, *Macromolecules*, 2003, **36**, 3673–3683.

75 J. K. Percus and G. J. Yevick, *Phys. Rev.*, 1958, **110**, 1–13.

76 L. M. Oltrogge, Q. Wang and S. G. Boxer, *Biochemistry*, 2014, **53**, 5947–5957.

77 U. Haupts, S. Maiti, P. Schwille and W. W. Webb, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 13573–13578.

78 A. Huang, H. Yao and B. D. Olsen, *Soft Matter*, 2019, **15**, 7350–7359.

79 H. Shinoda, M. Shannon and T. Nagai, *Int. J. Mol. Sci.*, 2018, **19**, 1548.

80 J. Wu, C. Zhao, W. Lin, R. Hu, Q. Wang, H. Chen, L. Li, S. Chen and J. Zheng, *J. Mater. Chem. B*, 2014, **2**, 2983–2992.

81 J. K. Dozier and M. D. Distefano, *Int. J. Mol. Sci.*, 2015, **16**, 25831–25864.

82 J. Y. Shu, C. Tan, W. F. DeGrado and T. Xu, *Biomacromolecules*, 2008, **9**, 2111–2117.

83 A. Jain and H. S. Ashbaugh, *Biomacromolecules*, 2011, **12**, 2729–2734.

84 E. Hamed, T. Xu and S. Keten, *Biomacromolecules*, 2013, **14**, 4053–4060.

85 G. Qin, P. M. Perez, C. E. Mills and B. D. Olsen, *Biomacromolecules*, 2016, **17**, 928–934.