



Cite this: DOI: 10.1039/d0me00051e

Placement of tyrosine residues as a design element for tuning the phase transition of elastin-peptide-containing conjugates: experiments and simulations†

Phillip A. Taylor,^{‡a} Haofu Huang,^{‡b} Kristi L. Kiick^{id}*^b and Arthi Jayaraman^{id}*^{ab}

Elastin-like polypeptides (ELP) have been widely used in the biomaterials community due to their controllable, thermoresponsive properties and biocompatibility. Motivated by our previous work on the effect of tryptophan (W) substitutions on the LCST-like transitions of short ELPs, we studied a series of short ELPs containing tyrosine (Y) and/or phenylalanine (F) guest residues with only 5 or 6 pentapeptide repeat units. A combination of experiments and molecular dynamics (MD) simulations illustrated that the substitution of F with Y guest residues impacted the transition temperature (T_t) of short ELPs when conjugated to collagen-like-peptides (CLP), with a reduction in the transition temperature observed only after substitution of at least two residues. Placement of the Y residues near the N-terminal end of the ELP, away from the tethering point to the CLP, resulted in a lower T_t than that observed for peptides with the Y residues near the tethering point. Atomistic and coarse-grained MD simulations indicated an increase in intra- and inter-peptide hydrogen bonds in systems containing Y guest residues that are suggested to enhance the ability of the peptides to coacervate, with a concomitantly lower T_t . Simulations also revealed that the placement of Y-containing pentads near the N-terminus (*i.e.*, away from the CLP tethering point) *versus* C-terminus of the ELP led to more π - π stacking interactions at low temperatures, in agreement with our experimental observations of a lower T_t . Overall, this study provides mechanistic insights into the driving forces for the LCST-like transitions of ELPs and offers additional means for tuning the T_t of short ELPs for biomedical applications such as on-demand drug delivery and tissue engineering.

Received 30th April 2020,
Accepted 3rd July 2020

DOI: 10.1039/d0me00051e

rsc.li/molecular-engineering

Design, System, Application

Modern biomaterials with controllable chemical and mechanical properties are needed to tackle complex problems in science and engineering, including but not limited to, on-demand drug delivery and nanomedicine. This study investigates how optimization of amino acid sequence tunes the structure and phase transitions of biomaterials composed of elastin-like peptides (ELPs) and collagen-like peptides (CLPs). This study highlights a modular approach for synthesizing short ELPs for which atomistic and coarse-grained molecular dynamics simulations can be used to obtain mechanistic insights into the driving forces for self-assembly, such as hydrogen bonding and π - π stacking. In this study, we show how the substitution and placement of tyrosine (Y) residues in sequences of short ELPs can be used to control the lower critical solution temperature (LCST)-like transitions of ELP-CLP bioconjugates. We observe a reduction in the transition temperature (T_t) with an increasing number of Y *vs.* phenylalanine (F) substitutions, and a reduced T_t for ELPs with Y located at the N-terminus *vs.* C-terminus. Overall, our combined experimental and computational approaches can be used to streamline the development of thermoresponsive nanocarriers for on-demand drug delivery under physiologically relevant conditions.

Introduction

Responsive polypeptides have garnered significant interest over the past few decades since they display tunable chemical and mechanical properties in response to external stimuli such as heat, light, pH, and salt.^{1–4} Such chemical and physical properties include robust thermal stability⁵ (*e.g.*, high melting temperatures) and controllable morphologies^{6,7} (*e.g.*, vesicles *vs.* micelles), which are directly linked to the molecular features of the peptide chains such as their self-

^a Department of Chemical and Biomolecular Engineering, University of Delaware, Newark, DE 19716, USA. E-mail: arthij@udel.edu

^b Department of Materials Science and Engineering, University of Delaware, Newark, DE 19716, USA. E-mail: kiick@udel.edu

† Electronic supplementary information (ESI) available. See DOI: 10.1039/d0me00051e

‡ Equal contributions.

assembly and phase transitions. As a result, peptides have been used in biomedical applications to create on-demand drug delivery vehicles which can release their therapeutic cargo in response to external stimuli such as light, heat, and pH,^{8–11} and hydrogels which can either release cargo or change their mechanical properties in response to external stimuli such as heat.^{12,13} One such class of polypeptides, the elastin-like polypeptides (ELP), are composed of multiple repeats of the amino acid sequence, (V-P-G-X-G)_n, where V, P, and G are valine, proline, and glycine, respectively. The fourth residue, X, in the pentad can be any amino acid except proline.¹⁴ ELPs are known to undergo a lower critical solution temperature (LCST)-like phase transition in aqueous solutions in which the ELP is soluble below the transition temperature, T_t , and insoluble above the T_t .^{15,16} Previous studies of high molecular-weight ELPs have shown that the T_t of ELPs can be controlled by adjusting ELP molecular weight, concentration, guest residue composition, pressure, pH, buffer composition, salt concentration or by conjugation of ELPs with other peptides and polymers.^{1,17–23} Recent studies, however, have been mainly focused on long ELPs containing tens and even hundreds of repeats of (V-P-G-X-G) pentapeptides while shorter ELPs containing fewer than ten repeat units have not been widely studied due to their high T_t .^{24–30} Shorter ELPs could be preferable for applications in drug delivery, tissue engineering, and device manufacture owing to their scalable solid-phase manufacturing, and thus shorter ELPs which have low molecular weights but also display T_t values near-physiological temperatures are of considerable interest.

In order to tune the T_t of ELPs and understand the driving forces for the LCST-like transition, there have been several fundamental studies focused on elucidating the molecular driving forces governing the assembly of ELPs at the single chain level.^{31–37} Studies have reported that hydrogen bonding is particularly important for the aggregation of ELP chains as they undergo the LCST-like transition.^{33,34,38} Previous reports have shown that at low temperatures below the T_t , the ELP chains behave as hydrophilic species in which they form a large number of hydrogen bonds between the peptide chains and surrounding water molecules. As the temperature is increased above the T_t , the ELP chains form more inter- and intra-peptide hydrogen bonds than peptide–water hydrogen bonds, and secondary structures, (β , α , π) turns, have also been observed above the T_t both in molecular dynamics simulations and circular dichroism (CD) measurements.^{39,40} For example, Zhao *et al.* performed a hydration analysis in which the number of hydrogen bonds between peptide and water and the number of water molecules in the first hydration shell were used to determine the T_t in all-atom molecular dynamics simulations.³⁶ They observed that the dependence of the T_t of ELPs on their chain lengths agreed well with experiments in that both had the same power-law exponents.

Past computational work from the Jayaraman group and experimental work from the Kiick group has also shown that the incorporation of bulky, aromatic guest residues such as tryptophan (W) impacts the LCST-like transition *via* induced

chain stiffness, hydrophobic interactions, and also turn structures.³⁸ The correlation between an increased propensity for turn structures and lower T_t in ELPs has also been suggested by Nuhn and Klok based on circular dichroism (CD) measurements; their studies also suggested that hydrophobicity is not the sole determinant of the LCST-like behavior of short ELPs.⁴⁰ In addition to tuning the T_t , other work from Kiick and co-workers⁴¹ has shown that the substitution of F with W can also tune the morphology of the assembled ELP bioconjugates, supporting the formation of platelets at temperatures above the T_t rather than vesicles as reported by Luo and Kiick.⁴² Therefore, the incorporation of aromatic guest residues offers a means for adjusting the T_t and assemblies of short ELPs to tune their physicochemical and morphological properties for specific applications.

ELPs can be mixed with polymer solutions and conjugated to other classes of thermoresponsive peptides such as collagen-like peptide (CLP) triple helices in order to tune the T_t of the ELP block.^{12,13,17,42} CLPs are composed of three peptide chains where each CLP chain is made up of repeat units of amino acid triplets, (X-Y-G), where X and Y are usually proline (P) and hydroxyproline (O), respectively. Three CLP chains then self-assemble *via* hydrogen bonds between adjacent chains to form the CLP triple helix which, at temperatures below the melting temperature (T_m) of the triple helix, mimics the triple helical structure of native collagen.⁴³ Short CLPs have been used in thermoresponsive fabrics,⁴⁴ and also as therapeutic matrices and molecules⁴⁵ owing to their heat-triggered folding-unfolding behavior. The conjugation of short ELP to CLP can be used to control the T_t of the ELP.⁴² ELP–CLP conjugates are also dually thermoresponsive such that vesicles are formed above the T_t of the ELP–CLP triple helix conjugate, while vesicles dissociate at higher temperatures above the T_m of the CLP triple helix. Also, studies have illustrated that CLPs can be used to stain collagen derived from human tissues, especially those with high ECM turnover (*e.g.* joints and articular cartilage⁴⁶), which makes this bilayer vesicle structure amenable to the delivery of molecules to collagen-rich regions in human tissues. Molecular dynamics simulations by Jayaraman and co-workers showed that the decrease in T_t of the ELP–CLP triple helix conjugates as compared to the free ELP is a result of the local crowding of the ELP chains when conjugated to CLP triple helix.¹⁷ Past work by Dunshee *et al.* has also shown that the choice of CLP sequence impacts the T_t of ELP–CLP conjugates through the thermal stability of the CLP domain (indicated by the T_m) and overall hydrophilicity of the ELP–CLP conjugate.⁴⁷ Similarly, ELPs have also been combined (but not conjugated) with hydrophilic polymers such as PEG to create thermoresponsive hydrogels which display controllable optical and mechanical properties.^{12,13}

In this paper, we investigate the impact of tyrosine (Y) substitutions on the LCST-like transitions of ELP–CLP conjugates using a combination of molecular dynamics simulations (atomistic and coarse-grained) and experiments.

Inspired by our previous work on W-containing ELPs, Y substitutions are used to tune the T_t of the ELP block such that it undergoes the LCST-like transition near-physiological temperatures, making it more suitable for drug delivery applications. The impact of substituting F with Y is not intuitive as hydrophobicity scales disagree on the relative hydrophobicities of F and Y.^{48,49} To the best of our knowledge, previous studies have also not investigated the impact of π - π stacking interactions on the LCST-like transitions of ELPs containing aromatic guest residues. Therefore, we hypothesize that the stacking interactions between nearby aromatic side-chain groups on Y and F would facilitate the aggregation of the ELP chains above the T_t . The impact of sequence order on the T_t is investigated by keeping the guest residue composition of the ELP chain constant and changing the relative positions of pentads containing the Y guest residues. Although there have been studies which have shown that the choice of sequence order can lead to hysteresis during the LCST-like transition,⁵⁰ there have not been any studies focused on the impact of sequence order on the value of the T_t .

Experimental methods

Materials

All Fmoc-protected amino acids (glycine, proline, hydroxyproline, valine, phenylalanine, tyrosine and Fmoc-propargyl glycine), 4-azidobutanoic acid, hexafluorophosphate (HBTU), oxyma Pure, N,N' -diisopropylcarbodiimide (DIC) for solid-phase peptide synthesis (SPPS) were purchased from Chempep (Wellington, FL). HPLC-grade acetonitrile, dimethylformamide (DMF) used as solvents in SPPS and HPLC purification were purchased from Fisher Scientific (Fairlawn, NJ). triisopropylsilane (TIS), diisopropylethylamine (DIPEA), trifluoroacetic acid (TFA), copper(I) acetate and dehydrated DMF and DMSO used in cleavage were purchased from Sigma-Aldrich (St. Louis, MO).

Synthesis of collagen-like and elastin-like-peptides

The collagen-like-peptide with the sequence, (GPO)₈GG, and elastin-like-peptide with sequences, (VPGFG)_x(VPGYG)_yG' and (VPGYG)_y(VPGFG)_xG' (G': propargyl glycine) were synthesized *via* traditional solid-phase peptide synthesis using the Liberty Blue™ automated microwave peptide synthesizer (CEM Corporation, Charlotte, NC). The additional GG group of the CLP sequence was added for ease of experimental synthesis as it inhibits potential side reactions at the C-terminus of the CLP triple helix. The general synthesis scale was 0.25 mmol using a 0.19 mmol g⁻¹ low loading Rink amide ProTide resin. Double coupling and double deprotection were conducted on both CLP and ELP syntheses and DIC and Oxyma were used to do the carboxylic acid activation. The deprotection of the Fmoc group was conducted using 20% v/v piperidine. The single coupling cycle was 125 seconds (70 °C for 15 seconds and 90 °C for 110 seconds). 4-Azidobutanoic acid was manually coupled to the N-terminus of the CLP when on the

resin. Double coupling with a 4:1 amino acid/resin ratio was used for the conjugation. Cleavage of the peptides from the resin was conducted in 93:4.5:2.5 (v:v:v) TFA/TIS/H₂O for 3 hours and precipitated in cold ether. Crude peptides were purified *via* reverse-phase HPLC (C₁₈ column). The mobile phase comprised gradients of degassed deionized water with 0.1% TFA and acetonitrile with 0.1% TFA, at a flow rate of 30 mL min⁻¹. Then, the peptide was detected by a UV detector at 214 nm. Fractions with product were collected and lyophilized. The molecular weight of the peptides was confirmed *via* electrospray ionization mass spectrometry (ESI-MS Waters™ Xevo TQ-GC mass spectrometry system, Milford, MA) (Fig. S1–S13†) and the purity of the peptide was confirmed *via* reverse-phase HPLC (Waters™, Milford, MA) (Fig. S14 and S15†).

Preparation of ELP-b-CLP conjugates

The CLP sequence, (GPO)₈GG, was conjugated to various ELPs *via* the copper(I)-catalyzed azide-alkyne cycloaddition (CuAAC) “click” reaction to generate ELP-b-CLP. As the total length of each conjugate is approximately 60 amino acids, traditional SPPS methodologies to synthesize the full-length ELP-CLP would provide low yields of peptides. Therefore, we performed a ‘click’ reaction for the conjugation, which provides desired ease of synthesis and versatility for conjugating any type of molecule/peptide to either the ELP or the CLP in future studies. Solutions of CLP (4.5 μmol), ELP (3 μmol) and Cu(I) acetate (0.25 equiv. to alkyne) in 1.5 mL 7:3 (v:v) anhydrous DMF & DMSO were added to a nitrogen-purged vial along with a clean Cu wire to prevent Cu(I) from oxidation. The alkyne group on the ELP conjugates with the azide group on the CLP forms a triazole group (illustrated in Fig. 1). The mixture was stirred at 80 °C overnight and precipitated in 5-fold cold ether and re-dissolved in 3:7 (v:v) ACN/H₂O for further HPLC purification. The crude copolymer was purified *via* reverse-phase HPLC (C₁₈ column). The mobile phase comprised gradients of degassed deionized water with 0.1% TFA and acetonitrile with 0.1% TFA, at a flow rate of 30 mL min⁻¹. The column was heated to 65 °C to unfold the CLP triple helix, and the eluted peptide was detected by a UV detector at 214 nm. Fractions with product were collected and lyophilized. The molecular weight of the peptides was confirmed *via* electrospray ionization mass

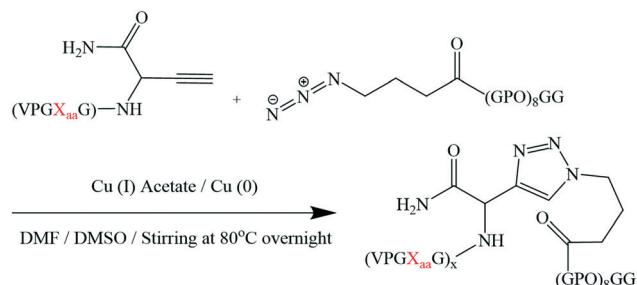


Fig. 1 Reaction scheme for ELP-CLP conjugation.

spectrometry (ESI-MS Waters™ Xevo TQ-GC mass spectrometry system, Milford, MA) (Fig. S1–S13†). The purity of the conjugates was confirmed by RP-analytical HPLC (C₁₈ column) and the results are shown in Fig. S14 and S15.†

Self-assembly of ELP–CLP conjugates

Samples (1 mg) of purified ELP–CLP conjugates were weighed and added into glass vials containing 1 mL of HPLC-grade H₂O. The solutions were heated to 80 °C and incubated at 80 °C for 2 hours to unfold the CLP triple helix. Solutions were transferred into 10 mm glass dynamic light scattering (DLS) cuvettes after being filtered at 80 °C. The annealing process from 80 °C to 25 °C was conducted using DLS for 3 hours on a Zetasizer Nano series (Nano ZS, Malvern Instruments, U.K.) at a scattering angle of 173° and data fitting using the cumulant method. Solutions were incubated at 25 °C in DLS overnight before other measurements. The T_t of a given ELP–CLP conjugate was obtained by measurement of the average hydrodynamic size of nanoparticles at temperatures from 4 °C to 80 °C with an interval of 3 °C. Samples were incubated at each temperature for 5 min before measurement. Each data point in our DLS results was calculated from the average of 3 different measurements and the error bar was calculated from their standard deviation. The transition temperature of the conjugate was determined using the first derivative^{51–54} and the corresponding temperature was labelled as the T_t .

Circular dichroism spectroscopy (CD) characterization

Circular dichroic spectroscopy (Jasco 810 circular dichroism spectropolarimeter, Jasco Inc., Easton, MD) was conducted for the characterization of the secondary structure of the CLP domain. CLP and ELP–CLP conjugates were dissolved at a concentration of 100 µM in PBS (10 mM, pH 7.4, 137 mM NaCl and 2.7 mM KCl) and incubated at 4 °C overnight before measurement. The CD spectra were recorded using quartz cells with a 0.2 cm optical path length. Full wavelength scans were collected to study the conformation of the peptide domain at 4 °C. The scanning rate was 50 nm min^{−1}, with a response time of 4 s. The wavelength scans were obtained from 200 to 250 nm and were recorded every 1 nm. To measure the melting temperature of the CLP domain, variable temperature experiments were conducted at the maximum wavelength in each ELP–CLP conjugate (e.g., 225 nm) with a 10 °C h^{−1} heating rate from 4 to 80 °C.

Transmission electron microscopy (TEM) characterization

ELP–CLP conjugates were prepared on carbon-coated 213 copper grids. Grids, pipet, pipet tips, and samples were prepared in an isothermal oven at desired temperatures for at least one hour before sample preparation. 5 µL of sample solution was dropped on the grid and blotted after 60 seconds. For staining, 1% phosphotungstic acid (PTA; pH adjusted to 7.0 using 1 M NaOH) as a negative stain was used. A total of 3 µL of the PTA solution was dropped on the grid and blotted after 10 s. The grid was left inside the oven

to dry for at least 1 hour before imaging. TEM images were taken on a Tecnai G2 at a 120 keV accelerating voltage.

Simulations

Atomistic simulations

Atomistic molecular dynamics simulations were performed using the GROMACS 4.6.7 simulation package⁵⁵ to investigate the effect of tyrosine substitutions on the LCST-like transitions of tethered ELPs which mimic the conjugation of ELP to CLP (Fig. 2). ELP sequences containing either five or six repeat units and a variable number of tyrosine (Y) guest residues ranging from zero to two tyrosine residues per ELP sequence were studied. The nomenclature ‘F₅’ series is used to refer to ELP sequences containing five repeat units and ‘F₆’ series used for ELP sequences containing six repeat units. ELP sequences in general are denoted only by the guest residues. For example, the ELP sequence (VPGFG)₄(VPGYG)₂ is referred to as F₄Y₂.

Tethered ELPs were simulated by first generating an ELP chain with the appropriate sequence in a β-spiral initial configuration using the PyMOL peptide builder tool.⁵⁶ The N-terminus of each ELP chain was capped with a charged, NH₃⁺ group while the C-terminus was capped with an uncharged, carboxyl (COOH) group. In order to reduce the computational resources needed to atomistically simulate an entire ELP–CLP conjugate, the impact of ELP conjugation to CLP was reproduced in atomistic simulations by tethering 3 ELP chains in space with the tethering points positioned in a

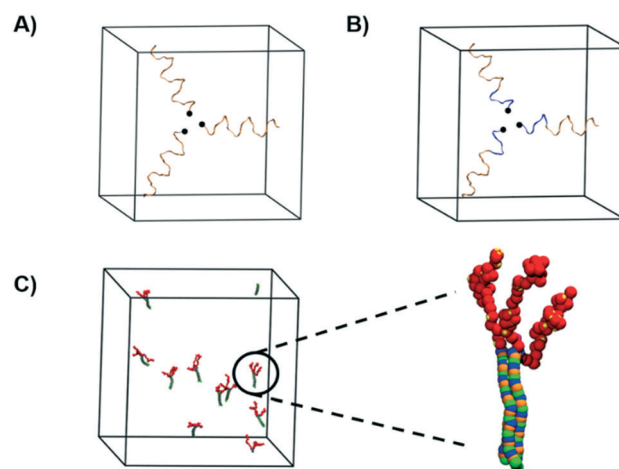


Fig. 2 Atomistic models of tethered ELPs and coarse-grained ELP–CLP models used in this study. (A and B) Ribbon rendering of the atomistic ELPs tethered in space at their C-termini for F₆ (left, (VPGFG)₆) and F₄Y₂ (right, (VPGFG)₄(VPGYG)₂). Explicit water molecules and salt are used in the simulations but not shown here for clarity. ELP repeat units containing phenylalanine (F) guest residues are shown in orange while repeat units containing tyrosine (Y) guest residues are shown in blue. (C) Coarse-grained (CG) model of F₄Y₂–(POG)₈ with an expanded highlight of a single ELP–CLP conjugate. Briefly, ELP backbone CG beads are shown in red with their yellow attractive CG beads.

triangular pattern, mimicking their native configuration when attached to the CLP triple helix (Fig. 2A and B). We set the distance between tethered C-termini of the 3 ELP chains as ~ 1.03 nm so that the diameter of a circle circumscribing the tethering points is approximately equal to the reported diameter of a CLP triple helix, ~ 1.2 nm.⁴³ Moreover, the tethering of the C-terminus was achieved by defining a position restraint on the carbon atom of the C-terminal carbonyl group as reported by Prhashanna *et al.*³⁸ Using GROMACS, position restraints were modelled using harmonic bonds between the positions of peptide atoms at a specific point in time and their initial positions as.

$$U_{\text{res}}(x, y, z) = \frac{1}{2} k_{\text{res}} \left[(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 \right] \quad (1)$$

In eqn (1) U_{res} is the positional restraint potential, k_{res} is the harmonic bond force constant set at 10^4 kJ (mol nm²)⁻¹, (x, y, z) is the current position of the atom in Cartesian coordinates, and (x_0, y_0, z_0) is the initial position of the atom. In our simulations, the exact chemistries of the end groups for the experimentally produced peptides were not captured to keep the simulation results general, irrespective of the peptide synthesis technique.

We chose the OPLS-AA/L force field⁵⁷ and TIP4P water model⁵⁸ for our atomistic simulations. The initial configuration of the tethered ELPs was solvated in TIP4P water in cubic box sizes of 9 nm for the F₅ series and 10 nm for the F₆ series. Furthermore, since the N-terminus of the ELP chain contains a charged, NH₃⁺ group, we included 3 chloride (Cl⁻) ions to maintain charge neutrality. The peptides used in the experiments, on the other hand, have an acetylated N-terminus and the charge neutrality of the experimental molecules was captured with a synthesis-agnostic, neutral end group in the simulation. Each system was energy minimized using a steepest descent method with the maximum force tolerance level set as 900 kJ mol⁻¹ nm⁻¹. Molecular dynamics simulations were performed in the NPT ensemble at a desired temperature and a pressure of 1 bar for 200 ns. We chose temperatures ranging from 5 °C to 65 °C to mimic the experimental temperature range. A time step of 2 fs was used for the numerical integration. Bonds involving hydrogen atoms were constrained using the LINCS method.⁵⁹ Temperature and pressure were maintained constant using the stochastic velocity rescaling method⁶⁰ and Berendsen barostat,⁶¹ respectively, with time constants of 0.1 and 2.0 ps, respectively. All electrostatic interactions were modelled using the particle-mesh Ewald (PME) method using fourth-order cubic interpolation.⁶²

Several studies have shown that hydrogen bonding is a major driving force for the LCST-like transitions of elastin-like peptides.^{32,36,38} Therefore, we quantified peptide-peptide (inter-peptide, intra-peptide, and total) and peptide-water hydrogen bonds as a measure of the extent of assembly and aggregation before and after the LCST-like transition. We do not claim to predict the exact T_t using our atomistic simulations but rather use the hydrogen bonds to obtain

molecular insights into the effect of tyrosine substitutions on the chain conformations (*i.e.* structure) and thermodynamics of the ELPs chain as they undergo the LCST-like transition. Hydrogen bonds were defined based on distance and geometric criteria, in which a hydrogen-bonding donor and acceptor are separated by less than 0.35 nm and the angle between the acceptor atom, donor atom, and hydrogen must be less than 30°. We also quantified turn structures in our simulations where the β , α , or π turn structures were defined as an intra-peptide hydrogen bond formed between residues i and $i + n$, where ($n = 3, 4, 5$, respectively).

In this paper, we hypothesize that π - π stacking also plays an important role as a driving force for the LCST-like transition due to the inclusion of aromatic guest residues such as phenylalanine (F) and tyrosine (Y). Inspired by the works of Headen *et al.*,⁶³ to investigate the impact of tyrosine substitutions on the extent of stacking interactions between aromatic guest residues, the radial and angular distribution functions are quantified as

$$g(r, \theta) = \frac{\Delta n(r, \theta)}{4\pi r^2 \sin\theta \cdot \Delta r \cdot \Delta\theta \cdot \rho} \quad (2)$$

where $\Delta n(r, \theta)$ is the number of aromatic residues in the distance range Δr and angle $\Delta\theta$, and ρ is the bulk number density. The factor of $1/\sin\theta$ accounts for the dependence of the solid angle on θ and degeneracies in the angular distribution.

For each system, the last 100 ns of the simulation was employed for data analysis to ensure that each system is properly equilibrated. The mean and standard deviation of these last 10 000 configurations were calculated and reported.

Coarse-grained (CG) simulations

In this work, we used a modified version of the coarse-grained ELP model of Prhashanna *et al.*³⁸ based on previous work by Jayaraman and co-workers that capture directional attractive interactions like H-bonds and π - π stacking in macromolecules.⁶⁴⁻⁶⁷ Each ELP residue was modelled using two coarse-grained beads – one bead for the peptide backbone (shown in red in Fig. 2C) and another smaller attractive “H-bond” bead for the hydrogen bonding donor or acceptor on the amino acid (shown in yellow in Fig. 2C). This modified CG model was motivated by atomistically observed increased propensity of tyrosine guest residues to form hydrogen bonds relative to other amino acids in ELP. With this CG model, every amino acid in the ELP can partake in hydrogen bonding, while tyrosine has the stronger propensity to form hydrogen bonds.

The characteristic length, σ , and characteristic energy, ϵ , for the CG model were 0.5 nm and 0.1 kcal mol⁻¹, respectively, such that the reduced temperature, $T^* = 5.92$ is related to room temperature, $T = 25$ °C. Furthermore, the characteristic mass, m , was chosen arbitrarily since the goal of the CG model is not to capture the dynamics of these systems but rather the correct thermodynamics. Therefore,

the bead masses do not need to represent the true mass of the corresponding amino acid. In our CG model, each ELP backbone bead had a mass of $1.0m$ and a diameter of 1.0σ , while each H-bond bead had a mass of $1.0m$ and a diameter of 0.3σ .

Adjacent backbone (BB) beads were connected using harmonic bond potentials, where the bond length was 0.84σ and force constant was $1000\epsilon/\sigma^2$. Each H-bond bead was connected to its parent backbone bead *via* a harmonic bond potential with a bond length of 0.37σ and a force constant of $1000\epsilon/\sigma^2$. In this work, we did not explore the impact of the guest residue on the stiffness of the ELP chain as described in our previous work.³⁸ Therefore, we assumed that both tyrosine and phenylalanine guest residues impart the same degree of local stiffness to the ELP chain and included the harmonic angle potentials between G–X–G residues (beads) as described by Prhashanna *et al.*³⁸ with the force constant and equilibrium angle of π implemented for both Y and F. Next, to ensure that hydrogen bond formation occurs perpendicular to the peptide backbone, an H-bond bead–BB bead–adjacent BB bead angle potential was included with a force constant of $300\epsilon/\text{rad}^2$ and equilibrium angle of 90° .

As in our previous work,³⁸ we captured the increasing tendency of an ELP chain to collapse with increasing temperature using increasing attraction strength between non-bonded ELP CG backbone beads, ϵ_{EB} , *via* the Lennard-Jones (LJ) potential⁶⁸ described as

$$U^{\text{EB}}(r) = \begin{cases} \epsilon_{\text{EB}} \left[\left(\frac{\sigma^{\text{EB}}}{r} \right)^{12} - \left(\frac{\sigma^{\text{EB}}}{r} \right)^6 \right]; & r < 2.5\sigma^{\text{EB}} \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

where σ^{EB} is the diameter of an ELP backbone bead and ϵ_{EB} is the interaction strength between non-bonded EB beads. We also captured all non-bonded interactions between H-bond beads in the ELP chain *via* another Lennard-Jones potential as shown below in eqn (4),

$$U_{ij}^{\text{HB}} = \epsilon^{\text{HB}} \left[\left(\frac{\sigma^{\text{HB}}}{r} \right)^{12} - \left(\frac{\sigma^{\text{HB}}}{r} \right)^6 \right] \quad (4)$$

with the bead diameter, σ^{HB} , set at 0.3σ and the interaction strength, ϵ^{HB} , set at 50.4ϵ for Y–Y pairs, 31.9ϵ for Y–X pairs, and 20.2ϵ for X–X pairs where Y is tyrosine and X can be any amino acid besides tyrosine, Y. These interaction strengths were chosen as the minimum interaction strength that were required to reproduce the LCST trends observed in experiments for two extreme cases: 1) an ELP sequence without Y substitutions and 2) another ELP sequence including two Y substitutions. Details of our choice of interaction strengths and parameterization are available in the ESI† (Fig. S16).

The (POG)₈ sequence was used for CLP in all simulations to match the experimental (GPO)₈GG sequence. The additional GG group used in experiments to facilitate synthesis was removed from our computational model as it does not affect the stability of the CLP triple helix. In this

work, each CLP strand was a chain of POG triplets in which each (POG) triplet was represented using a proline backbone (PB) bead, a proline H-bond (PH) acceptor bead, a hydroxyproline backbone bead (OB), a glycine backbone (GB) bead, and a glycine H-bond (GH) donor bead. All H-bond beads in the CLP model had a diameter of 0.3σ and a mass of $1.0m$ while all backbone beads had a diameter of 1.0σ and a mass of $3.0m$. This CLP model, as discussed in the work of Condon and Jayaraman,⁶⁴ captures the directionality and specificity of the inter-CLP strand hydrogen bonds using a combination of isotropic, bonded and non-bonded, interactions involving H-bond beads and the adjacent backbone beads, and bead sizes. It is also important to note that the CG CLP model does not capture the experimentally observed helicity of the CLP triple helix but reflects the correct trends in how the melting temperature of the CLP triple helix varies with CLP design as described previously by Condon and Jayaraman. The details of all bonded and non-bonded interactions are available in the original paper by Condon and Jayaraman.⁶⁴

Finally, using the above ELP–CLP CG models, Langevin dynamics was performed in the NVT ensemble on 10 ELP–CLP conjugates in a cubic simulation box of size 140σ using the LAMMPS simulation package.⁶⁹ In these simulations, the bonded interactions were integrated with a time step of 0.0005τ and non-bonded interactions with a time step of 0.001τ . Every system was first equilibrated for 10^8 -time steps followed by a 10^7 -time step production run in which snapshots were stored every 100 000-time steps. All ELP–CLP simulations were performed at $T^* = 3.0$, below the melting temperature of the (POG)₈ triple helix. Furthermore, to differentiate between conformation (ELP) states before and after the LCST-like transition, we determined the ensemble-average number of ELP bead pairwise contacts, $\langle N_{\text{contacts}} \text{ per bead} \rangle$, as a function of ELP solvophobicity, ϵ_{EB} . The number of contacts was computed for each frame in a simulation by counting the total number of unique ELP backbone bead pairs which are separated by 2.5σ or less, then dividing the total by the number of ELP backbone beads in the simulation. Next, the ensemble-average number of ELP bead contacts per ELP bead was calculated by averaging over all configurations and across three trials with different initial configurations and velocities. The error bars shown in the results are the standard deviations across three trials. For each system, the onset of aggregation (transition point) was defined as the value of ϵ_{EB} at which the plot of $\langle N_{\text{contacts}} \text{ per bead} \rangle$ vs. ϵ_{EB} reaches an inflection point.

Results and discussion

In this work, we seek to provide insight into the role of tyrosine as a guest residue on the T_t of short ELPs and self-assembly of ELP–CLP conjugates, while comparing our current findings to our previous studies on (VPGFG)₆–(GPO)₄GFOGER(GPO)₄GG and our W-containing ELP–CLP conjugates.^{38,42} First, we substitute different numbers of

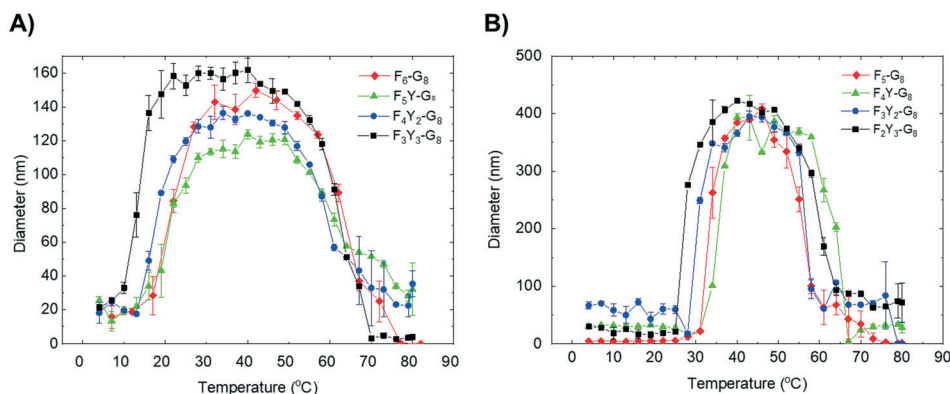


Fig. 3 Average hydrodynamic diameter as a function of temperature obtained from dynamic light scattering (DLS) characterization for A) Y-containing F₆ series ELP-CLPs; B) Y-containing F₅ series ELP-CLPs. The error bars represent the standard deviation of three different measurements from a representative synthetic batch.

phenylalanine (F) guest residues with tyrosine (Y) in both the F₆ and F₅ series to investigate the impact of the number of tyrosine residues on the T_t . As noted above our nomenclature uses G₈ to represent the CLP domain (GPO)₈GG, and only the guest residues in the ELP domain are used to represent the whole pentapeptide of the ELP domain. For example, (VPGFG)₄(VPGYG)₂ is denoted as F₄Y₂ and (VPGFG)₃(VPGYG)₂ as F₃Y₂. We do not expect that the triazole linkage should have any significant impact on the assembly properties of these molecules, given its relative size and position in the conjugate, as the assembly relies on the properties of the CLP domain and ELP domain separately.

In Fig. 3, we show the average hydrodynamic diameter obtained from dynamic light scattering (DLS) for each ELP-CLP conjugate. In Fig. 3A, for the F₆ series, the diameter rises steeply as the temperature is increased from 10 °C to 30 °C and each curve has a different temperature at which the diameter abruptly increases, which we denote as the ‘onset temperature’. For temperatures below the onset temperature, the diameter is approximately 20 nm and transmission electron microscopy (TEM) measurements indicate that the conjugate is not self-assembled (Fig. 4). As the temperature is increased, the diameter increases until it reaches ~150 nm for the self-assembled vesicles (confirmed by TEM measurements, shown in Fig. 4 and Table S1†). Then, as the temperature is increased further above 55 °C, the diameter decreases due to the unfolding of the CLP triple helix (confirmed by circular dichroism measurements, shown in Fig. S17†). We have confirmed, in DLS experiments, that the “free” ELPs containing tyrosine substitutions lack an LCST-like transition across the full range of temperatures (Fig. S18†). Our observations are in agreement with similar studies on other short ELPs,^{25–30} which do not display T_t below the boiling point of water. In previous work reported by Luo *et al.*,⁴² as the temperature was increased above the T_t , the diameter slightly increased presumably because the unfolding of the CLP domain causes the structure to become less compact. In this work, we did not observe this phenomenon likely because the T_t and T_m are closer in value

for these conjugates than for those previously reported.⁴² As the number of F-to-Y substitutions increases, (*i.e.*, from F₆-G₈ to F₄Y₂-G₈, to F₃Y₃-G₈) the onset temperature decreases by ~3 °C for each Y substitution, but the changes in diameter with temperature for all conjugates show similar behavior. We note that for F₅Y-G₈ and F₆-G₈ the onset temperature and curves overlap, suggesting that one Y substitution is not sufficient to induce a shift.

In Fig. 3B, for the F₅ series, at temperatures below the onset point, the small diameters observed likely represent an unassembled state. As the temperature is increased from 25 °C to 40 °C, the diameter rises steeply to approximately 400 nm. Although our studies of F₆-based conjugates⁴⁷ suggests that such conjugates yield nanostructures with diameters of approximately 100 nm, previous studies by Qin *et al.*⁴¹ showed that when G₈ was conjugated with shorter ELPs, the dimensions of self-assembled species were of similar dimensions as those reported here. Consistent with previous reports of an increasing vesicle diameter with decreasing

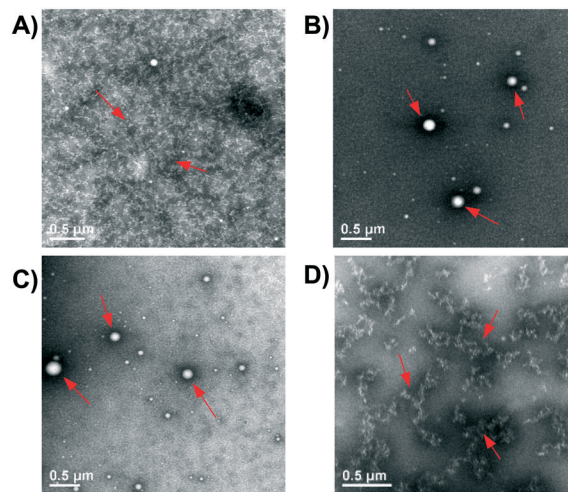


Fig. 4 Transmission electron microscopy (TEM) images showing representative assembled structures for F₄Y₂-G₈ at A) 4 °C; B) 25 °C; C) 50 °C; D) 80 °C.

Table 1 Transition temperatures (T_t) obtained from DLS measurement of aqueous solutions of conjugates with an increasing number of Y guest residues at the C-termini in the F_6 and F_5 series ELP-CLPs. The error represents the standard error in the T_t values obtained from three different synthetic repeats along with three different measurements

ELP-CLP sequences	T_t
F_6 -(GPO) $_8$ GG	21.2 ± 2.0 °C
F_5 Y-(GPO) $_8$ GG	20.2 ± 0.9 °C
F_4 Y $_2$ -(GPO) $_8$ GG	16.4 ± 0.4 °C
F_3 Y $_3$ -(GPO) $_8$ GG	12.9 ± 0.2 °C
F_5 -(GPO) $_8$ GG	33.2 ± 0.5 °C
F_4 Y-(GPO) $_8$ GG	34.0 ± 0.4 °C
F_3 Y $_2$ -(GPO) $_8$ GG	31.3 ± 0.3 °C
F_2 Y $_3$ -(GPO) $_8$ GG	28.0 ± 0.5 °C

layer thickness,⁷⁰ we observe an increase in the vesicle diameter for the vesicles comprising shorter ELP domains. The temperature-dependent behavior of the F_5 series is similar to that of the F_6 series; upon substituting two Ys in place of F, the onset temperature decreases by ~ 3 °C, but for a single Y substitution, the onset temperature remains unchanged as compared to F_5 . The T_t values for each conjugate are determined from the steepest slope for each conjugate's diameter vs. temperature curve (see the ESI†), and the calculated T_t values for all conjugates are shown in Table 1.

In Fig. 4, we show the representative TEM data for F_4 Y $_2$ -G $_8$ to confirm the morphology and dimensions of the self-assembled vesicles. As the T_t of F_4 Y $_2$ -G $_8$ is 16.4 °C, there are no large vesicles formed at 4 °C (indicated in red arrows in Fig. 4A). When the temperature is increased to 25 °C and 50 °C (Fig. 4B and C), there are vesicles present with a diameter of approximately 130 nm (indicated in red arrows), which is consistent with our DLS data, and the observed vesicles are significantly larger than the expected diameter of micelles that might be formed from these ELP-CLP conjugates (*ca.* 20 nm). The smaller spheres present in the background in Fig. 4B and C may result from staining artifacts or small aggregated peptides that form on the TEM grid during drying. As the temperature increases to 80 °C, which is above

the CLP T_m , the vesicles are completely dissociated as shown in Fig. 4D, although some aggregates were observed (indicated by red arrows), which likely formed during TEM sample preparation.

In Table 1, for both the F_6 and F_5 series, a single F-to-Y substitution does not cause any change in T_t , while for two or more F-to-Y substitutions, the T_t shifts to lower values. The observed behavior is consistent with the work of Urry *et al.*,⁴⁹ indicating a reduction in T_t with the inclusion of Y, rather than the hydrophobicity scales of Kyte and Doolittle,⁴⁸ which suggest that Y should be more hydrophilic than F (and thus raise the T_t). The atomistic and CG molecular simulation work described later describes the molecular interactions that explain these experimental trends.

Next, we switched the position of the Y-containing penta-d(s) from the C-terminus to N-terminus of the ELP domain in both the F_6 and F_5 series to investigate the impact of the position of F-to-Y substitution (*i.e.*, sequence order) on the T_t . According to previous studies, the T_t of ELP is altered by the ELP length, the choice of guest residue, as well as conjugation to other peptides (*e.g.*, CLP, linear poly-L-lysine, polylysine dendrimer, polyethylene glycol, *etc.*).^{24,27,49,71–73} In Fig. 5, we keep the composition of the ELP constant and focus on the effects of the placement of Y-containing penta-ds in the F_6 series with either two or three F-to-Y substitutions. As illustrated in Fig. 5A, for Y $_2$ F $_4$ -G $_8$, the onset temperature is ~ 3 °C lower than that observed for F_4 Y $_2$ -G $_8$. As the temperature increases above the T_m of the CLP domain, the diameter decreases sharply indicating the unfolding of the CLP triple helix (confirmed by circular dichroism measurement in Fig. S17†). Similarly, in Fig. 5B, the onset temperature of Y $_3$ F $_3$ -G $_8$ is also ~ 3 °C lower than that observed for F_3 Y $_3$ -G $_8$. As illustrated in both figures, despite the compositional similarity of the ELP, when the F-to-Y substitution is made in pentads at the N-terminus (away from the CLP), the T_t is shifted to lower values. This trend is also observed for the F_5 series (Fig. 6); Table 2 presents the T_t for all of these sequences and shows quantitatively that the T_t is reduced by ~ 3 °C in all cases in which the Y-substituted pentad is shifted to the N-terminus of the ELP.

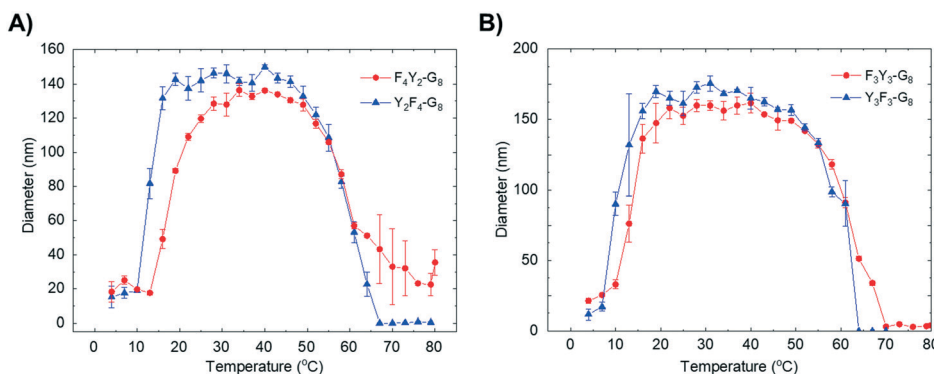


Fig. 5 Average hydrodynamic diameter as a function of temperature obtained from DLS measurements for A) F_4 Y $_2$ -G $_8$ & Y $_2$ F $_4$ -G $_8$; B) F_3 Y $_3$ -G $_8$ & Y $_3$ F $_3$ -G $_8$.

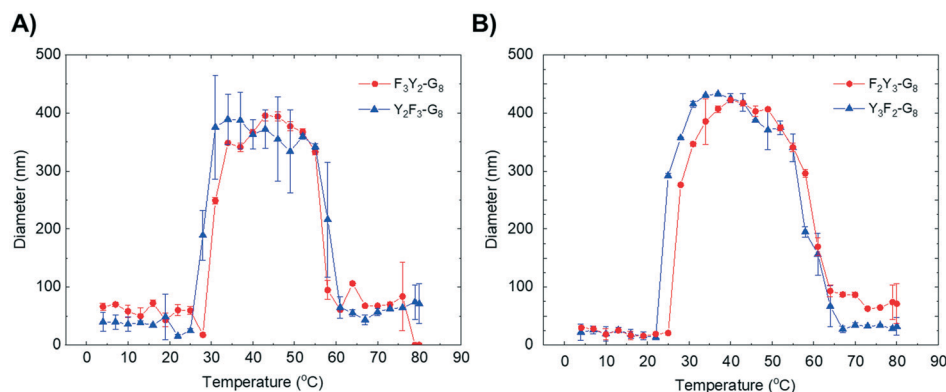


Fig. 6 Average hydrodynamic diameter as a function of temperature obtained from DLS measurements for A) $F_3Y_2-G_8$ & $Y_2F_3-G_8$; B) $F_2Y_3-G_8$ & $Y_3F_2-G_8$.

In all of the experimental results presented, simply by making two (or three) F-to-Y substitutions in an ELP chain comprising 30 total amino acids, or by simply changing the placement of the substitution relative to placement of the substitution relative to the CLP tethering point, statistically significant changes are observed in the T_t ($p < 0.02$). As an increased number of Y guest residues are included in the ELP sequence and the position of Y-containing pentads is nearer to the tethering point between the ELP and CLP (*i.e.*, $F_3Y_3-G_8$ compared to $Y_3F_3-G_8$ and $F_2Y_3-G_8$ compared to $Y_3F_2-G_8$), the T_t difference between the pairs decreases, as the impact of the increasing number of Y-containing pentads becomes more dominant in affecting the T_t .

Our experimental observations of the reduction in T_t with increasing Y-content are consistent with previous reports. Early work by Urry⁴⁹ as well as that by Meyer *et al.*²⁴ showed that the T_t of an ELP can be lowered by increasing the length of the ELP and/or the hydrophobicity of the guest residue. Additional experimental studies have also illustrated low T_t (< 4 °C) for long, Y-containing ELP sequences.⁷⁴ Hathorne *et al.*'s work⁷⁵ showed that increasing the incubation temperature for Y-containing short ELP sequences resulted in an increase in type I/III β -turns and hydrogen bonds. Our empirical observation of the sensitivity of the T_t to the position of the Y-containing pentads in the short ELP-CLPs highlights the advantages of employing short ELPs and illustrates the tunability of the transitions of these ELP-CLP

conjugates near physiologically relevant temperatures at which their temperature sensitivity could be used to trigger the release of drugs or imaging agents. The conjugation of the ELP to the triple-helix-forming CLP results in observable T_t transitions, similar to reports for short ELPs conjugated to dendrimers and synthetic polymers^{72,73} and short ELP-polyethylene glycol conjugates,⁷¹ but the inclusion of the CLP domain offers strategies to target collagen-containing matrices and tissues, therefore allowing for collagen-targeted delivery strategies.^{10,46}

To provide detailed molecular insight and understanding of the driving forces governing the impact of F-to-Y substitution on the T_t of these ELP-CLP conjugates, atomistic and coarse-grained MD simulations were performed. First, we present atomistic simulation results for the 'tethered' F_6 (*i.e.*, $(VPGFG)_6$ tethered at its C-terminus) at various temperatures ranging from below experimentally observed T_t to above T_t . We note that these atomistic simulations included only water molecules, salt molecules, and three ELP chains that were all tethered to a point in space (Fig. 2A and B). Although the distance between the three tethering points was the same as that expected for three 'conjugation' sites at the N-terminus of the folded CLP, the explicit role of a given CLP and any concentration effects observed in experiments are not accounted for in these atomistic simulations.

In Fig. 7, we plot the number of hydrogen bonds (total peptide-peptide, inter-peptide, intra-peptide, and turn structures) observed in simulations as a function of temperature as tethered F_6 undergoes the LCST-like transition. An increase in the total number of peptide-peptide hydrogen bonds is observed with increasing temperature for tethered F_6 . The intra-peptide hydrogen bonds and the number of turn structures (a subset of the former) follow a similar trend where there is an initial increase as the ELP undergoes its transition between 5 °C and 25 °C, followed by a decrease (for intra-peptide) and a plateau (secondary structure) at higher temperatures. Interestingly, at those same high temperatures, the number of inter-peptide hydrogen bonds increases. These trends

Table 2 Transition temperatures (T_t) obtained from DLS measurement of samples with various substitutions of Y in the F_6 and F_5 series ELP-CLPs. The error represents the standard error of the T_t obtained from three different measurement replicates of each of three synthetic repeats

ELP-CLP sequences	T_t
$F_4Y_2-(GPO)_8GG$	16.4 ± 0.4 °C
$Y_2F_4-(GPO)_8GG$	12.9 ± 0.2 °C
$F_3Y_3-(GPO)_8GG$	13.0 ± 0.0 °C
$Y_3F_3-(GPO)_8GG$	11.0 ± 0.9 °C
$F_3Y_2-(GPO)_8GG$	31.3 ± 0.3 °C
$Y_2F_3-(GPO)_8GG$	27.8 ± 0.3 °C
$F_2Y_3-(GPO)_8GG$	28.0 ± 0.5 °C
$Y_3F_2-(GPO)_8GG$	25.3 ± 0.2 °C

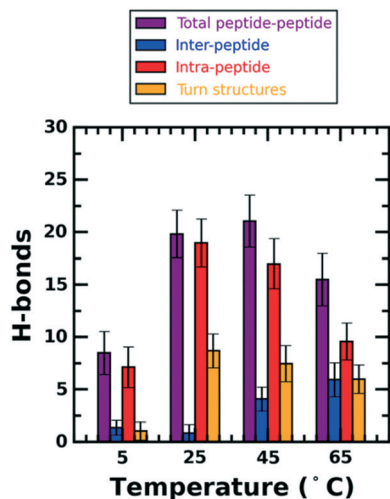


Fig. 7 Number of hydrogen bonds (H-bonds) vs. temperature obtained from atomistic simulations of tethered F_6 (i.e., (VPGFG) $_6$ tethered at the C-terminus of the ELP).

indicate a two-step mechanism: each ELP chain, at low temperatures, likely undergoes an intra-peptide collapse stabilized by intra-peptide hydrogen bonds, some of which qualify as turn structures. Similarly, reports from Matt *et al.*⁷⁶ showed that as the temperature increases, the hydrogen bonding network will be interrupted and reorganized. In our work, as the temperature is increased, the three ELP chains undergo inter-peptide aggregation stabilized by inter-molecular hydrogen bonds. Unlike the coarse-grained simulations discussed later, there were not multiple sets of tethered ELPs included in the atomistic simulations, so we cannot comment on the aggregation of multiple sets of tethered ELPs. We also would not expect to capture signatures of this two-step mechanism using experimental

techniques such as DLS given the difference in length scales of the original ELP chain folding *versus* the larger dimension observable in DLS.

Considering the aromatic nature of F and Y residues, the π - π stacking interactions were then quantified both below and above the T_t as the ELP chains undergo the LCST-like transition (Fig. 8). For tethered F_6 (without any Y) at temperatures below the T_t (5 °C), there are weak and heterogeneous stacking interactions involving stacking with both parallel and perpendicular orientations, as shown in Fig. 8A (and Fig. S19†). Between 5 °C and 25 °C, near the experimentally observed T_t , the simulation results show a shift to primarily parallel stacking interactions as observed with a peak in the distribution near $\theta = 0^\circ$ (Fig. 8B). Then, at higher temperatures (Fig. 8C and D) there are weak stacking interactions that continue to decrease with increasing temperature. The decrease in the number of stacking interactions occurs as a result of the increasing entropic penalty that the aromatic rings would incur to maintain a stacked configuration at higher temperatures.

As we have shown that H-bonding plays a pivotal role during the LCST-like transition of F-containing ELPs, we first investigate the impact of F-to-Y substitutions on the number of hydrogen bonds in Fig. 9. For the results in Fig. 9, two-sided *t*-tests were performed using the stats module in SciPy.⁷⁷ “****” indicates $p < 0.001$, “***” indicates $p < 0.01$, and “**” indicates $p < 0.05$. All pairwise *t*-tests shown are for comparisons of results for F_6 with F_4Y_2 at each temperature. It is not clear *a priori* whether the presence of the additional hydroxyl (OH) group on Y vs. F would facilitate the formation of more inter- and intra-peptide hydrogen bonds thus leading to a reduction in T_t relative to unsubstituted ELP-CLP conjugates, or if the OH group would facilitate additional interactions with the surrounding solvent (water) and lead to

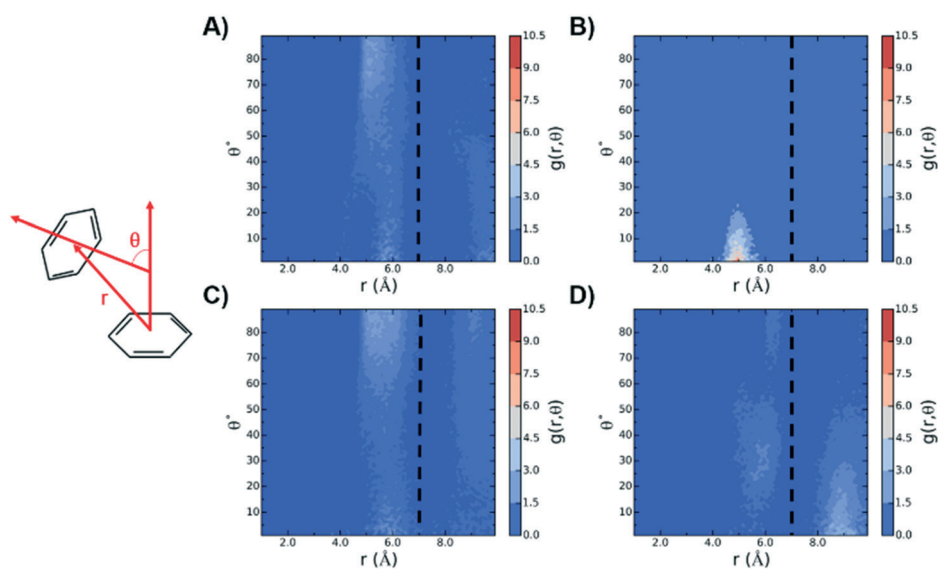


Fig. 8 Effect of temperature on the π - π stacking of tethered F_6 obtained from atomistic simulations at temperatures of A) 5 °C, B) 25 °C, C) 45 °C, D) 65 °C. Dashed lines show the radial cut off for a π - π stacking interaction (7 Å).

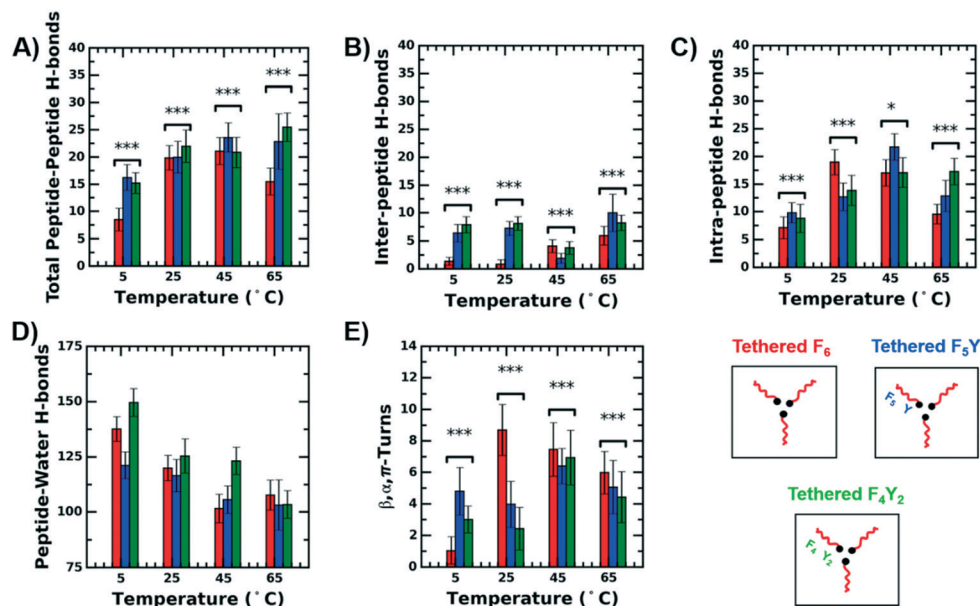


Fig. 9 Hydrogen bonding analyses for tethered F_6 , F_5Y , and F_4Y_2 (shown in red, blue, and green, respectively) obtained from atomistic simulations: A) total peptide–peptide hydrogen bonds, B) inter-peptide hydrogen bonds, C) intra-peptide hydrogen bonds, D) peptide–water hydrogen bonds, and E) β , α , and π turns. Two-sided t -tests were performed using the stats module in SciPy.⁷¹ “***” indicates $p < 0.001$, “**” indicates $p < 0.01$, and “*” indicates $p < 0.05$. All pairwise t -tests shown are for comparisons of results for F_6 with F_4Y_2 at each temperature.

an increase in the T_t relative to unsubstituted ELP–CLP conjugates. With increasing F-to-Y substitutions, the total peptide–peptide hydrogen bonds increase (Fig. 9A). The number of inter-peptide hydrogen bonds (Fig. 9B) around the LCST-like transition (5 °C and 25 °C) is significantly higher for sequences with F-to-Y substitutions as compared to F_6 . In contrast there is no consistent trend in the number of intra-peptide hydrogen bonds (Fig. 9C). All ELP chains, regardless of composition, become dehydrated with increasing temperature as shown by a decrease in the number of peptide–water hydrogen bonds (Fig. 9D). We note that a hydration analysis in terms of the number of peptide–water hydrogen bonds and number of water molecules in the 1st hydration shell of the ELP, as used in the work of Condon *et al.*,¹⁷ also confirms dehydration but does not provide a single LCST-like transition from a well-hydrated state at low temperature to dehydrated states at high temperature for these short ELPs (Fig. S20†). At 5 °C, below the LCST-like transition, the F-to-Y substitution increases the number of turn (β , α , π) structures adopted by the ELP chain, but at 25 °C, above the LCST-like transition, we see an opposing trend in which the F-to-Y substitutions leads to a reduction in the number of turn structures (Fig. 9E). This is unlike the trends in our previous studies with tryptophan, W, where an increased propensity to form turn structures could be correlated with observation of a lower T_t .^{38,40}

In Fig. 10, we present the same results as in Fig. 9 but for the F_5 series. The LCST-like transition for all sequences in the F_5 series occurs between 25 °C and 45 °C in the experiments. In the computational results shown in Fig. 10, at all temperatures, with an increasing number of F-to-Y

substitutions, there is an increase in the total peptide–peptide hydrogen bonds (Fig. 10A). The number of inter-peptide hydrogen bonds is higher for the Y-containing ELP sequences than F_5 at 45 °C; we, however, observe the opposite trend at 5 °C and 25 °C where there is a mild (but statistically significant) decrease in the number of inter-peptide hydrogen bonds for the Y-containing ELP sequences *versus* F_5 (Fig. 10B). The number of intra-peptide hydrogen bonds is higher for only the case where we include two Y guest residue substitutions (F_3Y_2) in the F_5 at all temperatures, but we do not see consistent trends for F_4Y from 5 °C to 65 °C (Fig. 10C). Similar to the F_6 series (Fig. 9D), we observe that all ELP sequences, regardless of composition, become dehydrated with increasing temperature as shown by a decrease in the number of peptide–water hydrogen bonds (Fig. 10D). Although the existence of β -turn secondary structures for ELPs in the collapsed state has been reported previously,⁷⁸ we show no conclusive evidence in experiments for their presence in the CD spectra of the conjugates, most likely due to the considerably weaker absorption of β -turns relative to the absorption of triple helices, β -sheets, and α -helices⁷⁷ and because spectral signatures can deviate significantly for different sequences.⁷⁹ In simulations, we observe the presence of turn structures but we do not see a consistent effect of F-to-Y substitution on the number of β , α , and π turn structures. Therefore, our simulation data shows that ELP sequences containing F or Y guest residues have similar propensity to form secondary structures as these ELP chains undergo the LCST-like transition, suggesting that turn structures are an important driving force for intramolecular

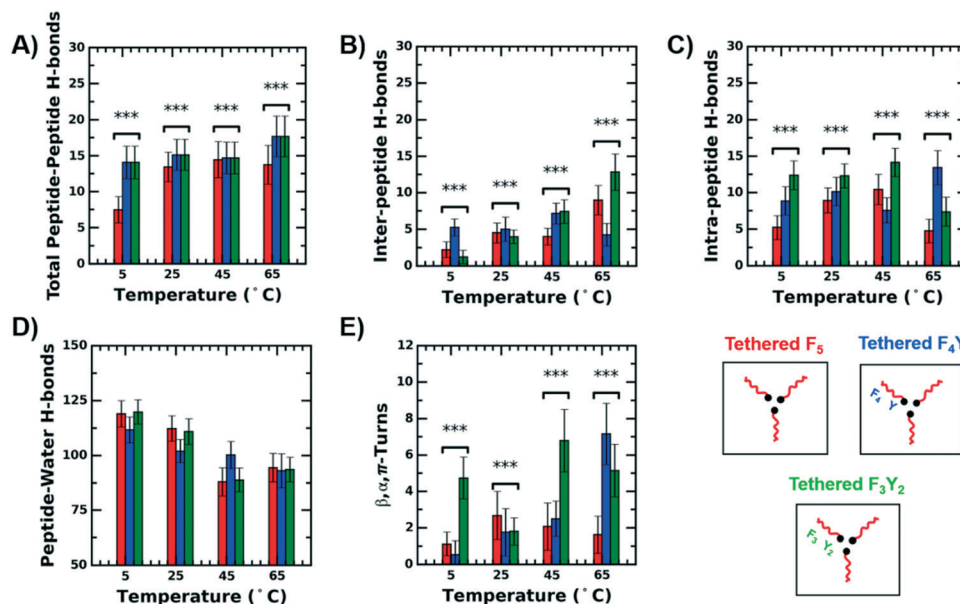


Fig. 10 Hydrogen bonding analyses for tethered F_5 , F_4Y , and F_3Y_2 (shown red, blue, and green, respectively) obtained from atomistic simulations: A) total peptide–peptide hydrogen bonds, B) inter-peptide hydrogen bonds, C) intra-peptide hydrogen bonds, D) peptide–water hydrogen bonds, and E) β , α , and π turns. Two-sided t -tests were performed using the stats module in SciPy.⁷¹ “***” indicates $p < 0.001$, “**” indicates $p < 0.01$, and “*” indicates $p < 0.05$. All pairwise t -tests shown are for comparisons of results for F_5 with F_3Y_2 at each temperature.

collapse for ELPs with either Y or F guest residues, in agreement with our past studies on ELP sequences containing W guest residues.³⁸ We also confirm that the choice of a β -spiral structure as the starting conformation for these atomistic ELP simulations does not impact these results. Previous work from Condon *et al.* has shown that the initial β -structure does not bias the simulation trajectory.¹⁷ To prove that this is the case in this study as well, snapshots of the initial and final conformations for tethered F_4Y_2 were collected, and the number of turn structures at the end of the MD simulation was calculated for all ELP sequences (Fig. S21†). We observe that the final configuration does not resemble a β -spiral structure, and that a collection of different turn structures (β , α , and π) are present at the end of the simulation.

The trends observed in the atomistic simulation results for both the F_6 and F_5 series are the following: i) for all ELP sequences there is increasing dehydration (*i.e.*, decreasing peptide–water hydrogen bonds) with increasing temperature; ii) the total peptide–peptide hydrogen bonds (including inter- and intra-peptide hydrogen bonds) increase with F-to-Y substitutions. However, unlike the experiments where at least two F-to-Y substitutions are required to bring about a shift in the T_t , in these atomistic simulations we observe increases in the total peptide–peptide hydrogen bonds even with one F-to-Y substitution and also observe comparable numbers of hydrogen bonds for both single and double Y substitutions. We attribute this disagreement to the limitations associated with these atomistic simulations. As noted earlier, the atomistic simulations, due to computational intensity, are restricted to having only three ELP chains

tethered to three points in space and do not represent ELP–CLP conjugates at experimentally relevant concentration. We conjecture that the experimental observation of needing more than one F-to-Y substitution to observe any shift is linked to thermodynamic driving forces that change only after a certain number of F-to-Y substitutions at finite ELP–CLP concentration. To test this, CG MD simulations were conducted that explicitly include the CLP block using the CLP model of Condon and Jayaraman and include multiple ELP–CLP conjugates at the same concentration of 1 mg mL^{-1} (0.1 mM) as in experiments.

In the CG MD simulations, in place of temperature, we increase the attraction strength ϵ_{EB} . As ϵ_{EB} is increased, an increase in the number of ELP CG bead contacts is also observed. A lower ‘onset’ ϵ_{EB} in these CG MD simulations corresponds to a lower T_t in experiments.^{17,38,40} Analogous to the T_t temperature in experiments, we mark the ϵ_{EB} at which an inflection point is observed in the plot of the number of contacts *vs.* ϵ_{EB} . In our atomistically-informed CG model, every residue is able to form hydrogen bonds with nearby residues, in which tyrosine residues have the greatest affinity to form hydrogen bonds. Inspired by previous studies from the Jayaraman group,^{38,64,66,67} each residue comprises a backbone bead which represents the peptide backbone and an H-bond bead which represents the corresponding hydrogen bonding donor/acceptor. The combinations of bead sizes, placement, and isotropic, interaction potentials allow our CG model to capture the directionality and specificity needed to model hydrogen bonds. Details of parameterization are described in the Methods section and the ESI.† In Fig. 11A, it is illustrated that when there is only

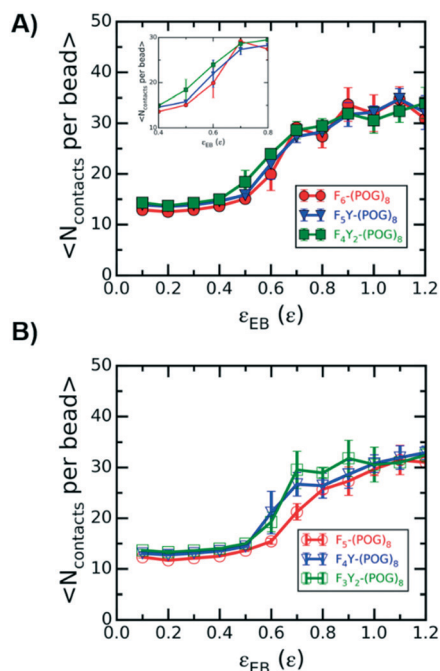


Fig. 11 Ensemble average number of pairwise contacts per ELP coarse-grained (CG) bead vs. the strength of attractive pairwise interactions among ELP beads for (A) F_6 series and (B) F_5 series, obtained from CG simulations.

one F-to-Y substitution (*i.e.*, $F_5Y-(POG)_8$) our CG model shows a mild shift in the onset point relative to F_6 . Whereas, for two F-to-Y substitutions (*i.e.*, $F_4Y_2-(POG)_8$), a larger, more distinguishable shift in the onset point is observed, indicating that two substitutions are required to observe a significant shift in T_t for the F_6 series. On the other hand, for the F_5 series, $F_5-(POG)_8$, $F_4Y-(POG)_8$, and $F_3Y_2-(POG)_8$, both the $F_3Y_2-(POG)_8$ and $F_4Y-(POG)_8$ systems exhibit a shift in T_t compared to $F_5-(POG)_8$ (Fig. 11B).

Finally, we explain the experimental observations of the effect of sequence order (*i.e.* changing the position of Y-containing pentad from C-terminus to N-terminus) on the LCST-like transition of ELP–CLP conjugates. Our hypothesis is that a combination of hydrogen bonding and π – π stacking could be inhibited or promoted as a result of the positioning of Y guest residues relative to the tethering point (C-terminus) of ELP to CLP. To test this hypothesis, two ELP sequences each containing six repeat units with the same composition: tethered F_4Y_2 and tethered Y_2F_4 were simulated atomistically. The differences in hydrogen bonding (total peptide–peptide, inter, intra, and turn structures) between the two sequences is equivalent to error for both sequences (ESI† Fig. S22). In contrast, the effect of sequence order is observed in the π – π stacking interactions for the ELP chains. As shown in Fig. 12, when the ELP is conjugated to CLP at its F end (*i.e.* the F_4 block is located at the C-terminal conjugation point) the ELP chains show strong stacking interactions *vs.* the case where the ELP is conjugated to CLP at its Y end (*i.e.* the Y_2 block is located at the C-terminal conjugation point). At higher temperatures above 45 °C, we do not observe significant effects of the tethering point on the stacking interactions, possibly due to the entropic penalty incurred by the stacking of aromatic groups at high temperatures (Fig. S23†).

Moreover, although our stacking analyses do not show a consistent effect of Y substitutions on the stacking interactions between and within ELP chains (Fig. S24†), past studies have shown that Y forms stronger stacking interactions than F.⁸⁰ Therefore, when the ELP chain is tethered at the Y end, the Y guest residues are constrained as a result of the tethering and cannot participate in favorable stacking interactions. The distance between tethered termini in atomistic simulations is ~ 1.03 nm, which is greater than the cut-off distance for π – π stacking interactions of approximately 0.7 nm (black dashed line in Fig. 12). So, Y

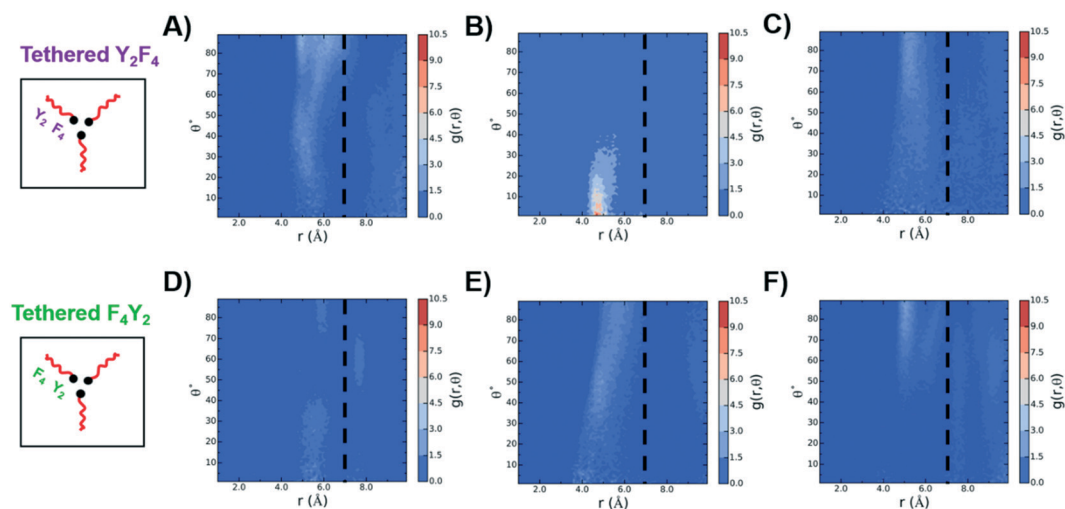


Fig. 12 Effect of the tethering point/terminus on the π – π stacking of ELP 6-mers (F_6 series) at low temperatures: (A and D) 5 °C, (B and E) 25 °C, and (C and F) 45 °C. Dashed lines show the radial cut off for a π – π stacking interaction (7 Å).

guest residues are constrained such that they are unable to make close contacts corresponding to a centroid-centroid distance (r) of less than 0.7 nm and show fewer stacking interactions when the ELP chain is tethered at the Y end. On the other hand, when the ELP chain is tethered at the F end, the Y guest residues are less constrained and therefore can form more stacking interactions leading to a lower T_t , as observed in our experimental results.

Conclusions

We have shown using a combination of experiments, atomistic and coarse-grained molecular dynamics (MD) simulations that we can tune the LCST-like transition of ELP-CLP conjugates by varying both the length and composition of the ELP block *via* the substitution of phenylalanine (F) guest residues with tyrosine (Y). Previously published hydrophobicity scales disagree on the relative hydrophobicities of F and Y, where some scales (such as that of Kyte and Doolittle⁴⁸) rank tyrosine to be more hydrophilic than phenylalanine while our results, although for short ELPs, follow the trends observed by Urry and co-workers⁴⁹ for high molecular weight ELPs. Past studies have also noted that the collapse of the ELP domain occurs due to a combination of hydrophobic collapse and intra-pentamer H-bonds. Our simulations showed that with an increase in temperature, the number of total peptide-peptide hydrogen bonds increases (with increases in both inter- and intra-peptide hydrogen bonds), which is consistent with previous studies by Hathorne *et al.* and Matt *et al.*^{75,76} We also observed that the sequence directionality of the ELP block when conjugated to CLP significantly impacts the LCST-like transition. When Y guest residues were placed near the C-terminus of the ELP chain (*i.e.* near the tethering point to CLP), there was an increase in T_t as compared to cases where the Y guest residues were placed near the untethered N-terminus of the ELP. The increase in T_t when tyrosine was placed near the tethering point is suggested by our studies to result from reduced π - π stacking interactions due to the spatial constraints placed on Y as a result of its proximity to the tethering point; there were no significant differences in terms of hydrogen bonding for these sequences based on variations in sequence order. While previous studies have mainly focused on hydrogen bonding and hydrophobic effects as driving forces for the LCST-like transition, this study highlights the importance of also considering π - π stacking interactions when designing ELPs and ELP-CLP conjugates as responsive materials for applications in sensing, electronics, and medicine.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

We thank the National Science Foundation (NSF; Grant CBET1703402) for financially supporting this work. Additionally, HH and KK want to thank the National Institutes

of Health (NIH; R21 AR069778A, 1 P30 GM110758, 1 P30 GM103519 and 1 P20 GM104316) for partial financial support and instrument resources. The computational work in this paper was supported through the use of information technologies resources at the University of Delaware, specifically the Farber high-performance computing resources. The views expressed in this work do not necessarily reflect the views of the funding agencies.

Notes and references

- 1 Y. Cho, Y. Zhang, T. Christensen, L. B. Sagle, A. Chilkoti and P. S. Cremer, *J. Phys. Chem. B*, 2008, **112**, 13765–13771.
- 2 J. C. Rodríguez-Cabello, F. J. Arias, M. A. Rodrigo and A. Girotti, *Adv. Drug Delivery Rev.*, 2016, **97**, 85–100.
- 3 M. E. Roth-Konforti, M. Comune, M. Halperin-Sternfeld, I. Grigoriants, D. Shabat and L. Adler-Abramovich, *Macromol. Rapid Commun.*, 2018, **39**, 1800588.
- 4 D. Schmaljohann, *Adv. Drug Delivery Rev.*, 2006, **58**, 1655–1670.
- 5 A. M. Hilderbrand, E. M. Ford, C. Guo, J. D. Sloppy and A. M. Kloxin, *Biomater. Sci.*, 2020, **8**, 1256–1269.
- 6 M. R. Dreher, A. J. Simnick, K. Fischer, R. J. Smith, A. Patel, M. Schmidt and A. Chilkoti, *J. Am. Chem. Soc.*, 2008, **130**, 687–694.
- 7 L. Martín, E. Castro, A. Ribeiro, M. Alonso and J. C. Rodríguez-Cabello, *Biomacromolecules*, 2012, **13**, 293–298.
- 8 T. Dvir, M. R. Banghart, B. P. Timko, R. Langer and D. S. Kohane, *Nano Lett.*, 2010, **10**, 250–254.
- 9 T. Jiang, Z. Zhang, Y. Zhang, H. Lv, J. Zhou, C. Li, L. Hou and Q. Zhang, *Biomaterials*, 2012, **33**, 9246–9258.
- 10 T. Luo, M. A. David, L. C. Dunshee, R. A. Scott, M. A. Urello, C. Price and K. L. Kiick, *Biomacromolecules*, 2017, **18**, 2539–2551.
- 11 Y. Shamay, L. Adar, G. Ashkenasy and A. David, *Biomaterials*, 2011, **32**, 1377–1386.
- 12 E. Meco and K. J. Lampe, *Biomacromolecules*, 2019, **20**, 1914–1925.
- 13 Y. Zhang, M. S. Desai, T. Wang and S.-W. Lee, *Biomacromolecules*, 2020, **21**, 1149–1156.
- 14 D. W. Urry, *Prog. Biophys. Mol. Biol.*, 1992, **57**, 23–57.
- 15 M. Martino, T. Perri and A. M. Tamburro, *Macromol. Biosci.*, 2002, **2**, 319–328.
- 16 D. W. Urry, T. Trapane and K. Prasad, *Biopolymers*, 1985, **24**, 2345–2356.
- 17 J. E. Condon, T. B. Martin and A. Jayaraman, *Soft Matter*, 2017, **13**, 2907–2918.
- 18 S. R. MacEwan and A. Chilkoti, *Nano Lett.*, 2012, **12**, 3322–3328.
- 19 J. A. MacKay, D. J. Callahan, K. N. FitzGerald and A. Chilkoti, *Biomacromolecules*, 2010, **11**, 2873–2879.
- 20 D. E. Meyer and A. Chilkoti, *Nat. Biotechnol.*, 1999, **17**, 1112–1115.
- 21 J. Reguera, D. W. Urry, T. M. Parker, D. T. McPherson and J. C. Rodríguez-Cabello, *Biomacromolecules*, 2007, **8**, 354–358.
- 22 S. Roberts, M. Dzuricky and A. Chilkoti, *FEBS Lett.*, 2015, **589**, 2477–2486.

- 23 R. L. Teeuwen, F. A. de Wolf, H. Zuilhof and J. C. van Hest, *Soft Matter*, 2009, **5**, 4305–4310.
- 24 D. E. Meyer and A. Chilkoti, *Biomacromolecules*, 2004, **5**, 846–851.
- 25 H. Reiersen, A. R. Clarke and A. R. Rees, *J. Mol. Biol.*, 1998, **283**, 255–264.
- 26 P. Paricaud, A. Galindo and G. Jackson, *Mol. Phys.*, 2003, **101**, 2575–2600.
- 27 M. Pechar, J. Brus, L. Kostka, Č. Koňák, M. Urbanová and M. Šlouf, *Macromol. Biosci.*, 2007, **7**, 56–69.
- 28 R. Saxena and M. J. Nanjan, *Drug Delivery*, 2015, **22**, 156–167.
- 29 L. Ayres, K. Koch and J. C. M. van Hest, *Macromolecules*, 2005, **38**, 1699–1704.
- 30 Y. Navon, M. Zhou, J. B. Matson and R. Bitton, *Biomacromolecules*, 2016, **17**, 262–270.
- 31 G. L. Dignon, W. Zheng, Y. C. Kim and J. Mittal, *ACS Cent. Sci.*, 2019, **5**, 821–830.
- 32 N. K. Li, F. G. Quiroz, C. K. Hall, A. Chilkoti and Y. G. Yingling, *Biomacromolecules*, 2014, **15**, 3522–3530.
- 33 R. Rousseau, E. Schreiner, A. Kohlmeyer and D. Marx, *Biophys. J.*, 2004, **86**, 1393–1407.
- 34 E. Schreiner, C. Nicolini, B. Ludolph, R. Ravindra, N. Otte, A. Kohlmeyer, R. Rousseau, R. Winter and D. Marx, *Phys. Rev. Lett.*, 2004, **92**, 148101.
- 35 J. D. Tang, C. E. McAnany, C. Mura and K. J. Lampe, *Biomacromolecules*, 2016, **17**, 3222–3233.
- 36 B. Zhao, N. K. Li, Y. G. Yingling and C. K. Hall, *Biomacromolecules*, 2016, **17**, 111–118.
- 37 B. Zhao, T. Lindeboom, S. Benner, G. Jackson, A. Galindo and C. K. Hall, *Langmuir*, 2017, **33**, 11733–11745.
- 38 A. Prhashanna, P. A. Taylor, J. Qin, K. L. Kiick and A. Jayaraman, *Biomacromolecules*, 2019, **20**, 1178–1189.
- 39 B. Li, D. O. Alonso and V. Daggett, *J. Mol. Biol.*, 2001, **305**, 581–592.
- 40 H. Nuhn and H.-A. Klok, *Biomacromolecules*, 2008, **9**, 2755–2763.
- 41 J. Qin, T. Luo and K. L. Kiick, *Biomacromolecules*, 2019, **20**, 1514–1521.
- 42 T. Luo and K. L. Kiick, *J. Am. Chem. Soc.*, 2015, **137**, 15362–15365.
- 43 M. D. Shoulders and R. T. Raines, *Annu. Rev. Biochem.*, 2009, **78**, 929–958.
- 44 S. Chattopadhyay and R. T. Raines, *Biopolymers*, 2014, **101**, 821–833.
- 45 S. D. Kim, H. Y. Lee, J. W. Shim, H. J. Kim, Y. H. Yoo, J. S. Park, S.-H. Baek, B. A. Zabel and Y.-S. Bae, *Am. J. Respir. Crit. Care Med.*, 2011, **184**, 243–251.
- 46 Y. Li, C. A. Foss, D. D. Summerfield, J. J. Doyle, C. M. Torok, H. C. Dietz, M. G. Pomper and S. M. Yu, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 14767–14772.
- 47 L. C. Dunshee, M. O. Sullivan and K. L. Kiick, *Bioeng. Transl. Med.*, 2020, **5**, e10145.
- 48 J. Kyte and R. F. Doolittle, *J. Mol. Biol.*, 1982, **157**, 105–132.
- 49 D. W. Urry, C. H. Luan, T. M. Parker, D. C. Gowda, K. U. Prasad, M. C. Reid and A. Safavy, *J. Am. Chem. Soc.*, 1991, **113**, 4346–4348.
- 50 L. Zhou, T. Qiu, F. Lv, L. Liu, J. Ying and S. Wang, *Adv. Healthcare Mater.*, 2018, **7**, 1800670.
- 51 S. R. MacEwan, W. Hassounah and A. Chilkoti, *J. Visualized Exp.*, 2014, e51583, DOI: 10.3791/51583.
- 52 T. Christensen, W. Hassounah, K. Trabbic-Carlson and A. Chilkoti, *Biomacromolecules*, 2013, **14**, 1514–1519.
- 53 V. Sarangthem, E. A. Cho, A. Yi, S. K. Kim, B.-H. Lee and R.-W. Park, *Sci. Rep.*, 2018, **8**, 3892.
- 54 M. Shah, P.-Y. Hsueh, G. Sun, H. Y. Chang, S. M. Janib and J. A. MacKay, *Protein Sci.*, 2012, **21**, 743–750.
- 55 B. Hess, C. Kutzner, D. Van Der Spoel and E. Lindahl, *J. Chem. Theory Comput.*, 2008, **4**, 435–447.
- 56 W. L. DeLano, <http://www.pymol.org>, 2002.
- 57 G. A. Kaminski, R. A. Friesner, J. Tirado-Rives and W. L. Jorgensen, *J. Phys. Chem. B*, 2001, **105**, 6474–6487.
- 58 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.
- 59 B. Hess, H. Bekker, H. J. Berendsen and J. G. Fraaije, *J. Comput. Chem.*, 1997, **18**, 1463–1472.
- 60 G. Bussi, D. Donadio and M. Parrinello, *J. Chem. Phys.*, 2007, **126**, 014101.
- 61 H. J. Berendsen, J. V. Postma, W. F. van Gunsteren, A. DiNola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684–3690.
- 62 T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089–10092.
- 63 T. F. Headen, C. A. Howard, N. T. Skipper, M. A. Wilkinson, D. T. Bowron and A. K. Soper, *J. Am. Chem. Soc.*, 2010, **132**, 5735–5742.
- 64 J. E. Condon and A. Jayaraman, *J. Phys. Chem. B*, 2018, **122**, 1929–1939.
- 65 A. F. Ghobadi and A. Jayaraman, *Soft Matter*, 2016, **12**, 2276–2287.
- 66 A. Kulshreshtha, K. J. Modica and A. Jayaraman, *Macromolecules*, 2019, **52**, 2725–2735.
- 67 D. J. Beltran-Villegas, D. Intriago, K. H. Kim, N. Behabtu, J. D. Londono and A. Jayaraman, *Soft Matter*, 2019, **15**, 4669–4681.
- 68 J. E. Jones, *Proc. R. Soc. London, Ser. A*, 1924, **106**, 463–477.
- 69 S. Plimpton, *J. Comput. Phys.*, 1995, **117**, 1–19.
- 70 C. Huang, D. Quinn, Y. Sadovsky, S. Suresh and K. J. Hsia, *Proc. Natl. Acad. Sci. U. S. A.*, 2017, **114**, 2910–2915.
- 71 A. Araújo, B. D. Olsen and A. V. Machado, *Biomacromolecules*, 2018, **19**, 329–339.
- 72 C. Kojima, U. H. Sk, D. Fukushima, K. Irie, N. Akazawa, M. Umeda and T. Niidome, *RSC Adv.*, 2015, **5**, 104900–104906.
- 73 M. Zhou, Y. Shmidov, J. B. Matson and R. Bitton, *Colloids Surf., B*, 2017, **153**, 141–151.
- 74 B. M. Seifried, J. Cao and B. D. Olsen, *Bioconjugate Chem.*, 2018, **29**, 1876–1884.
- 75 A. P. Hathorne and H. Bermudez, *Biotechnol. Bioeng.*, 2013, **110**, 1822–1830.

- 76 A. Matt, B. Kuttich, I. Grillo, S. Weißheit, C. M. Thiele and B. Stühn, *Soft Matter*, 2019, **15**, 4192–4199.
- 77 P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser and J. Bright, *Nat. Methods*, 2020, 1–12.
- 78 J. Heyda, H. I. Okur, J. Hladílková, K. B. Rembert, W. Hunn, T. Yang, J. Dzubiella, P. Jungwirth and P. S. Cremer, *J. Am. Chem. Soc.*, 2017, **139**, 863–870.
- 79 W. Kim and V. P. Conticello, *Polym. Rev.*, 2007, **47**, 93–119.
- 80 R. Chelli, F. L. Gervasio, P. Procacci and V. Schettino, *J. Am. Chem. Soc.*, 2002, **124**, 6133–6143.