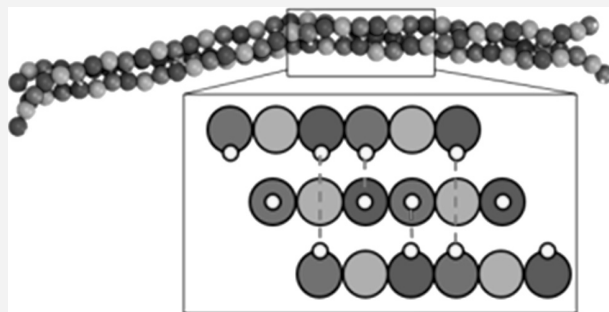


Development of a Coarse-Grained Model of Collagen-Like Peptide (CLP) for Studies of CLP Triple Helix Melting

Joshua E. Condon[†] and Arthi Jayaraman^{*,†,‡,§}[†]Colburn Laboratory, Department of Chemical and Biomolecular Engineering, University of Delaware, 150 Academy Street, Newark Delaware 19716, United States[‡]Department of Materials Science and Engineering, University of Delaware, 201 Dupont Hall, Newark Delaware 19716, United States

S Supporting Information

ABSTRACT: In this paper, we present the development of a phenomenological coarse-grained model that represents single strands of collagen-like peptides (CLPs) as well as CLP triple helices. The goal of this model development is to enable coarse-grained molecular simulations of solutions of CLPs and conjugates of CLPs with other macromolecules and to predict trends in the CLP melting temperature with varying CLP design, namely CLP length and composition. Since the CLP triple helix is stabilized primarily by hydrogen bonds between amino acids in adjacent strands, for modeling CLP melting we get inspiration from a recent coarse-grained (CG) model that was used to capture specific and directional hydrogen-bonding interactions in base-pair hybridization within oligonucleotides and reproduced known DNA melting trends with DNA sequence and composition in implicit water. In this paper, we systematically describe the changes we make to this original CG model and then show that these improvements reproduce the known melting trends of CLPs seen in past experiments. Specifically, the CG simulations of CLP solutions at experimentally relevant concentrations show increasing melting temperature with increasing CLP length and decreasing melting temperature with incorporation of charged residues in place of uncharged residues in the CLP, in agreement with past experimental observations. Finally, results from simulations of CLP triple helices conjugated with elastin like peptides (ELPs), using this new CG model of CLP, reproduce the same trends in ELP aggregation as seen in past experiments.



I. INTRODUCTION

Collagen-like peptides (CLPs) are polymer chains comprised of repeat units of amino acid triplets that most commonly have the amino acid sequence glycine (G)–proline (P)–hydroxyproline (O). Three such CLP strands form a stable triple helix structure through hydrogen bond (H-bond) interactions between strands and with water.^{1–5} With increasing temperature of the solution, the CLP triple helices undergo a melting transition with the value of the melting temperature tuned by the amino acid sequence, number of repeat units or length of CLP strands, and solvent quality.^{1,3,6,7} The motivation to understand how CLP melting transition changes with CLP design lies in the use of intact CLP triple helices in many biologically relevant applications, such as drug delivery,^{8–12} tissue engineering,^{13–22} and nanoelectronics.^{23,24} CLP triple helices are attractive for such applications because of their ability for cell surface recognition/adhesion, suitable structural properties and ability to self-assemble into fibers. Given its chemical similarity to native collagen, CLP is incorporated in hydrogels designed for tissue engineering to support the growth of cells by mimicking the extracellular matrix environment.²¹ CLPs are also incorporated into materials tailored for molecular recognition through strategic mutations of the amino acids in the CLP repeat units (usually P, O, and G) or through addition

of sequences of amino acids in the middle or at the ends of a POG based CLP strand. For example, in past work Gly-Phe-Hyp-Gly-Glu-Arg, an integrin binding sequence, has been placed in the middle of two (Gly-Pro-Hyp)_n sequences.^{25,26} CLP strands have also been modified at their ends with (Gly-Lys-Hyp)₄(Gly-Pro-Hyp)₄(Asp-Pro-Hyp)₄ sequences to form triple helices that assemble into fibrils in a hydrogel through electrostatic interactions between Lys and Asp.^{27,28}

CLP strands have also been conjugated to other thermoresponsive synthetic polymers^{21,25,29} and polypeptides^{15,26,30} so that the resulting copolymer has structural and thermodynamic properties tailored for a specific application. For example, vesicles for drug delivery have been achieved through self-assembly of macromolecules created by CLP conjugation to elastin like polypeptide (ELP).²⁶ Besides their use as delivery vehicles in the assembled state, systems of ELP–CLP conjugates are also fundamentally interesting as aqueous solutions of ELP exhibit inverse transition or lower critical solution transition (LCST) phase behavior.²⁶ Experiments show that upon the conjugation of ELPs to an intact CLP triple

Received: November 5, 2017

Revised: December 13, 2017

Published: February 7, 2018



helix, the LCST temperature of ELPs is lowered by greater than 80 °C compared to the LCST transition of unconjugated (or free) ELP in aqueous solution.²⁶ Atomistic and coarse-grained simulations explain these results by showing that it is the local crowding of the three ELP chains conjugated to the CLP triple helix that reduces the loss in conformational entropy as ELP goes through LCST transition, and as a result shifts the LCST transition to lower temperatures.³¹ Experiments of ELP–CLP conjugates also show that at higher temperatures, the conjugated CLP triple helix melts, separating the CLP strands returning the ELP's LCST to higher temperatures. This CLP melting and subsequent ELP LCST transition shifting to higher temperature leads to destabilization of ELP chain aggregate and disassembly of the delivery vesicle held together by the collapsed ELP.²⁶ Therefore, by tuning the ELP LCST and CLP melting transition temperatures, one can design customized vesicles that are functional (i.e., intact and disassembled) within desired temperature range.

Given that the melting behavior of CLP triple helices is critical to the stability and use of biocompatible thermoresponsive materials containing CLPs, for designing these materials it is important to understand which features of the CLP strand (e.g., sequence, length, mutations) impact melting temperature (T_m) of these triple helices and how/why. Past studies have described the structure of the triple helical region of collagen and CLP triple helices and how this structure impacts CLP melting.^{1–6,32–39} CLP triple helices are right-handed helices of three (XYG)_n strands in a left-handed polypyrrolone II (PPII) type helical conformation where n is typically 6–12 repeats. These triple helices are stabilized by interstrand hydrogen bonds (H-bonds) between the N–H hydrogen of glycine and the C=O oxygen of the X amino acid in a neighboring strand. As stated above, often the X amino acid in the XYG triplet is proline (Pro or P) and the Y amino acid is hydroxyproline (Hyp or O) because these two amino acids in these positions stabilize the PPII conformation of the individual CLP strands in the helix through their preferred ϕ and ψ backbone dihedral angles. Thus, CLP chains with (POG) triplets show a smaller change in entropy upon triple helical formation, and exhibit a higher melting temperature than many other (XYG) combinations.³⁸ Additionally, increasing the length of (POG)_n based triple helices improves the stability of the resulting triple helices, increasing the T_m .^{3,7,40} Another factor to consider in CLP triple helix stability is the network of water H-bonds that participate with the backbone and side chains of the CLP triple helix strands. Diffraction studies have shown that water molecules are bound to the surface of CLP triple helices in well-defined positions.³⁶ The O–H hydrogen in hydroxyproline can form a H-bond with a water molecule, which in turn can form a H-bond with a backbone C=O oxygen of a neighboring chain. Experimentally, (POG)_n is postulated to have a higher melting temperature than (PPG)_n for two reasons, one, the conformation of the (POG) PPII type helix in the triple helix is more entropically favorable than (PPG) and two, the ability of the hydroxyproline in (POG) to participate in the network of H-bonds surrounding and bridging the CLP triple helix through water is more enthalpically favorable than (PPG) which cannot participate in this network. Last, it has been shown that CLP strands with triplets that possess amino acids with like charges in proximity with one another destabilize the CLP triple helix. For example, the sequence (PKG)₄(POG)₄(DOG)₄ (where K is lysine or Lys and D is aspartic acid or Asp) forms triple helices with a lower

T_m than (POG)₁₂ triple helices.²⁷ The combination of the proximally located like charges that result in repulsive electrostatic forces in combination with the less favorable range of ϕ and ψ dihedral angles for K and D grants the (PKG)₄(POG)₄(DOG)₄ triple helix less stability with respect to (POG)₁₂ at the same temperature.

Additional experimental^{1,3,5–7,33,41–43} and computational^{34,35,44,45} studies have been done to quantify sequence effects on melting and other structural features of the CLP triple helices. Small angle X-ray scattering experiments have also been used to determine the length and radius of gyration for (POG)_n triple helices for a range of CLP strand lengths in aqueous solution.³⁹ Crystal structures of CLP triple helices obtained through X-ray diffraction experiments serve as initial configurations for CLP atomistic simulations^{33,36,45,46} that study stability of CLP triple helices. While atomistic simulations of CLP triple helices explicitly capture all atom details of both the CLP and water, capturing melting curves over a series of temperature is still outside the practicality of atomistic simulations. This is due to the unfeasible simulation time required to observe melting in addition to the large number of atoms depicting CLP and explicit water in these simulations. Thus, the development of CLP coarse-grained models with implicit solvent representation would be beneficial if one wishes to study trends in CLP helix melting transition for a broad range of CLP design parameters and for varying CLP-macromolecular conjugate design.

In this paper, we present two coarse-grained (CG) models: model 1 and model 2. Both these models adopt the H-bond functionality from a recent coarse-grained model developed by Ghobadi and Jayaraman aimed at reproducing oligomeric DNA melting trends.⁴⁷ In these two phenomenological CG models, we capture the interstrand H-bonding pattern experimentally observed in CLP triple helices but do not *explicitly* capture the water mediated H-bonding due to implicit solvent representation. Model 1 and model 2 are sequential models where model 2 addresses some of the shortcomings of model 1 with respect to correctly reproducing trends in CLP melting temperature, length, and diameter. We present these two models in a pedagogical manner showing the limitations of the first model (model 1) and the changes needed to improve the model to achieve a more refined model (model 2) for CLP.

Using these two CG models, we investigate melting behavior for varying length of (POG)_n triple helices. We do not see the experimentally observed increase in T_m with respect to (POG)_n length in model 1, but we are able to capture this phenomenon qualitatively with model 2. Next, we perform simulations that compare the melting transition of (POG)₁₂ without charged amino acids and (PKG)₄(POG)₄(DOG)₄ with charged amino acids using both models. Our CG simulations for both models show that (PKG)₄(POG)₄(DOG)₄ has a melting transition at lower temperatures than (POG)₁₂ which qualitatively agrees with experimental T_m data for these two sequences and supports the hypothesis that repulsive electrostatic interactions of neighboring like charges in a CLP triple helix are destabilizing. Additionally, we compare the end-to-end distance and diameter of CLP triple helix using these two coarse-grained models to that seen experimentally and computationally for CLP triple helices.^{39,45,46} We show that either the experimental length or diameter of CLP (but not both) can be matched by model 1 by scaling the dimensionless length of the model and/or adjusting the bond length of the backbone. Model 2 incorporates adjustments to bond lengths and dimensionless

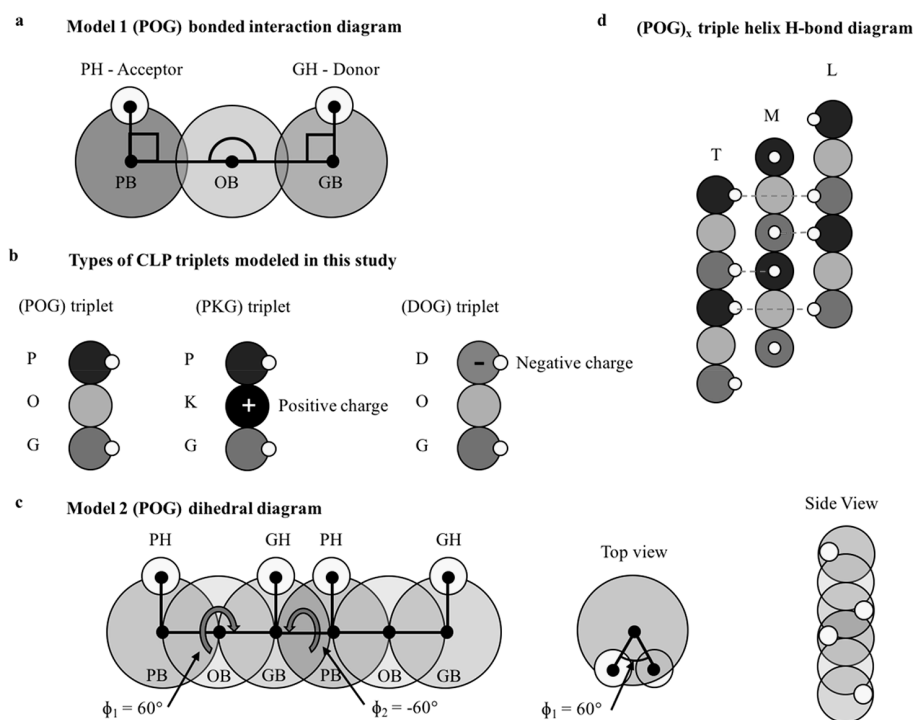


Figure 1. (a) Diagram of model 1 coarse-grained (POG) triplet with bonded interactions shown by lines and angles connecting specific bead types. (b) Types of CLP triplets used in the simulations studied in this work. (c) Diagram of model 2 coarse-grained CLP strand with the dihedral angles listed. (d) $(\text{POG})_x$ triple helical H-bond diagram highlighting the donor–acceptor interaction and the offset of the individual strands where T represents the trailing strand, M represents the middle strand, and L represents the leading strand.

length of the model to better match experimental length and diameter measurements. Last, we use model 2 as an alternative model to represent CLP in ELP–CLP conjugates investigated in simulation studies of Condon et al.³¹ and experiments of Luo and Kiick.²⁶ We compare the results of our new ELP–CLP model against the results in Condon et al.³¹ to ensure that the new ELP–CLP conjugate model captures similar behavior with respect to onset of aggregation of the ELP strands as was seen in the previous work of Condon et al.³¹

The paper is organized as follows. In section II, we describe the simulation approach and present all the details of both coarse-grained models. In section III, we describe the results from simulations with one or both CG models, of CLP melting with varying CLP length and composition as well as results of ELP aggregation within ELP–CLP conjugate systems. We end the paper with a Conclusion, where we summarize these results and present some limitations of the model which could be directions for future model development/improvement.

II. APPROACH

We use molecular dynamics simulations in the LAMMPS⁴⁸ simulation package with a newly developed phenomenological coarse-grained (CG) model of CLP inspired by the recently published CG model developed by Ghobadi and Jayaraman⁴⁷ for oligomeric deoxyribonucleic acid (DNA). Ghobadi and Jayaraman parametrized the bonded and nonbonded interactions in their oligonucleotide CG model to mimic directional and specific H-bond interactions between donors and acceptors within complementary base pairs.⁴⁷ They showed that their CG model is capable of qualitatively capturing the trends in melting temperatures of DNA as a function of DNA sequence, concentration and salt concentration. In this paper, we present

our efforts to modify that original DNA coarse-grained model to represent our system of CLP triple helices and mimic CLP melting as a function of CLP design. In the next two subsections, we present the development of this CLP coarse-grained model in two stages, model 1 and model 2, sharing the justification for the improvements in model 1 that lead to model 2.

II.A. Coarse-Grained CLP Model 1. Each CLP strand is a chain of XYG triplets, where each (XYG) triplet in the CLP sequence is represented by five interaction sites. Figure 1a shows the (POG) triplet which has 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 backbone (PB, OB and GB) beads have a diameter of 1.0σ and a mass of $3.0m$, and all H-bond beads (PH and GH) have a diameter of 0.3σ and a mass of $1.0m$. In this model, σ and ϵ correspond to 0.3 nm and 0.1 kcal/mol, respectively, as was stated in the original model of Ghobadi and Jayaraman. As for the value of m , since the goal of this model is to capture melting thermodynamics and not to mimic the dynamics of these CLP–water solutions for which the masses of the CG beads should be correct, the backbone and H-bond bead masses here are chosen arbitrarily for simplicity and do not reproduce the masses of the residues they represent. The relative sizes of the backbone and H-bond bead (1.0σ and 0.3σ) are chosen specifically to enable modeling of directionality in the H-bond interaction. The details of how the directionality of H-bond interaction is achieved is described in more detail later in this section. The adjacent backbone beads in each triplet are connected through a harmonic bond potential with bond length 0.84σ and force constant 1000ϵ . Each H-bond bead is connected to its respective backbone bead

with a harmonic bond potential with bond length 0.37σ and force constant 1000ϵ . Figure 1b shows the CLP composition modifications studied in this paper. In the case of CLP strands with a PKG triplet, the positively charged lysine (K) replaces hydroxyproline (O) in the (POG) triplet, and in the case of (DOG) triplet, a negatively charged aspartic acid (D) replaces proline (P) in the (POG) triplet. In these modified CLP sequences, the lysine backbone (KB) bead has the same characteristics as the OB bead along with a +1 charge that is missing in OB. The aspartic acid backbone (DB) bead and H-bond (DH) bead are identical to the PB and PH beads respectively with the exception that there is a -1 charge on the DB bead. For systems with charged residues in the CLP strand, explicit monovalent counterion (IN) beads are included in the system to maintain net charge neutrality. A charge of +1 is used for positively charged ions and a charge of -1 is used for negatively charged ions. IN beads have a diameter of 0.7σ and a mass of 1.0m. The diameter of the IN beads are taken directly from Ghobadi and Jayaraman to model monovalent ions.⁴⁷

To account for the rigidity of the PPII type helical conformation of individual CLP strands in a triple helix, we introduce a harmonic angle potential between three adjacent backbone beads with the angle set at 180° and constant $k_a^{BB} = 10\epsilon/\text{rad}^2$. Since the main goal of developing this CLP CG model is to capture correctly the trends in thermal transitions rather than the structure with varying CLP design, our CLP CG model is like the ONA model of Ghobadi and Jayaraman⁴⁷ that quantitatively reproduced the trends of DNA melting with DNA strand design without accurately reproducing structural details of DNA such as helical twist, minor/major grooves etc. In the same vein, the backbone beads of individual CLP strands in this CLP model are not parametrized to reproduce the experimentally observed helicity for these strands. The exciting aspect of this CLP CG model, like the ONA model of Ghobadi and Jayaraman, is that it captures the directionality and specificity of the H-bond interactions using a combination of bead sizes and isotropic bonded and nonbonded interactions involving the H-bond beads and their "parent" backbone beads. This is described next.

First, we have a H-bond bead–backbone bead–backbone bead angle potential with the angle set at 90° (as seen in Figure 1a) and the angle constant $k_a^{HB} = 300\epsilon/\text{rad}^2$. This ensures that H-bond formation occurs perpendicular to the backbone. For the (POG) triplet, this angle is defined as PH–PB–OB for the PH bead and GH–GB–OB for the GH bead. This angle definition holds for all triplets studied in this paper such that the angle potential in the X position in (XYG) is defined as XH–XB–YB where XH is the X H-bond bead, XB is the X backbone bead, and YB is the Y backbone bead. The angle potential in the G position is defined as GH–GB–YB. Dihedral bonded interactions are not present in model 1.

Next, nonbonded interactions between H-bond donors (D) and H-bond acceptors (A) (e.g., GH–PH interactions)⁴⁹ are represented using the Lennard-Jones (LJ) potential⁴⁹

$$U_{D-A}^{HB}(r) = \epsilon_{D-A}^{HB} \left[\left(\frac{\sigma^{HB}}{r} \right)^{12} - \left(\frac{\sigma^{HB}}{r} \right)^6 \right] \quad (1)$$

in which ϵ_{D-A}^{HB} is the strength of interaction between the donor beads and the acceptor beads and has the value $\epsilon_{D-A}^{HB} = 50.4\epsilon$. σ^{HB} is the diameter of the H-bond bead and r is the distance between the centers of two H-bond beads. We use a force and

potential cutoff of $1.9\sigma^{HB}$ with a switching function taken from GROMACS⁵⁰ and implemented in LAMMPS package that smoothly ramps both to zero at $2.0\sigma^{HB}$.

The potential for all other pairwise interactions besides the H-bond donor–acceptor interactions is the purely repulsive Weeks–Chandler–Andersen (WCA) potential⁵¹

$$U_{ij}(r) = \begin{cases} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r} \right)^{12} - \left(\frac{\sigma_{ij}}{r} \right)^6 \right] + \epsilon_{ij} & r < 2^{1/6}\sigma_{ij} \\ 0 & r \geq 2^{1/6}\sigma_{ij} \end{cases} \quad (2)$$

in which $\epsilon_{ij} = 1.0\epsilon$ and σ_{ij} is taken to be the arithmetic mean of the diameters of beads i and j . In the case of donor–donor and acceptor–acceptor H-bond interactions, σ_{ij} is 0.7σ instead of 0.3σ to prevent multiple donor HB beads from H-bonding to a single HB acceptor bead and to prevent multiple acceptor HB beads from H-bonding to a single donor HB bead. By using this same purely repulsive pairwise interaction potential for all backbone beads we are essentially simulating CLP strands in an implicit solvent which has the same good quality for all residues in the CLP. Besides the above nonbonded potentials, the pairwise interactions between a backbone bead and its third neighbor (1–3 interaction) and fourth neighbor (1–4 interaction) are included in this model so that intrachain bead overlap is prevented. All pairwise interaction values are listed in Table 1.

Table 1. Parameters of the Lennard-Jones Potentials^a Representing the Non-Bonded Interactions for the H-Bond Beads (PH and GH Generally Noted as HB Here) and Backbone Beads (PB, OB, and GB Generally Noted as BB) in the Coarse-Grained CLP model

bead type (i)	bead type (j)	pairwise interaction	ϵ_{ij} (ϵ)	σ_{ij} (σ)	cutoff (σ)
BB	BB	WCA	1.0	1.0	1.225
BB	HB (acceptor or donor)	WCA	1.0	0.65	0.729
HB (acceptor)	HB (donor)	LJ	50.4	0.3	0.570 ^a
HB (acceptor)	HB (acceptor)	WCA	1.0	0.7	0.786
HB (donor)	HB (donor)	WCA	1.0	0.7	0.786

^aPotential and force smoothly ramped to zero by secondary cutoff of 0.600σ

As stated earlier, the directionality of the H-bond donor–acceptor attractive interactions arises from a combination of the relative H-bond and backbone bead sizes, the repulsive nature of all pairwise interactions besides H-bond donor–acceptor, and H-bond bead–backbone bead–backbone bead angle potential. Through this combination and the location of the H-bond bead within the repulsive shell of the parent backbone bead, there is only a small geometrical region extending from the H-bond bead where attractive H-bond interactions can occur.

Finally, for all electrostatic interactions among ionized beads (KB, DB, and IN), we use the Coloumb potential

$$U_{ij}^{coul}(r) = \frac{q_i q_j}{4\pi r \epsilon_0 \epsilon_r} \quad (3)$$

in which q_i is the charge valency of bead i , q_j is the charge valency of bead j , and ϵ_0 is the permittivity of free space. The

relative permittivity of water in this model is set to $\epsilon_r = 80$ to be similar to that for water at room temperature.⁵² The 1–3 and 1–4 electrostatic interactions in this model are set to zero to prevent strong intrastrand electrostatic repulsive or attractive forces from destabilizing the CLP strands. These electrostatic interactions are present only in the systems where there are charged residues in the CLP strands.

II.B. Coarse-Grained CLP Model 2. While model 1 is a first step toward mimicking CLP triple helix melting, due to some missing interactions (e.g., no dihedral angles in model 1), model 1 is incapable of distinguishing the melting temperature of $(\text{POG})_n$ with varying n , as we discuss in the Results. Additionally, there are other observations in the simulations (e.g., nonhelical hybridized conformations) described in further detail in the Results that motivated us to edit model 1. In model 2, the length and energy are presented in reduced units of σ and ϵ which correspond to 0.5 nm and 0.1 kcal/mol, respectively, and masses chosen arbitrarily like model 1. The only difference between model 2 and model 1 is the value of σ . We also reduce the harmonic backbone bead-backbone bead bond length in model 2 to 0.50σ . Together, these changes bring the CLP length to diameter ratio seen in simulations closer to that seen in experimental measurements of CLP, as described in the Results. The harmonic backbone beads angle potential in model 2 has an angle constant of $k_a^{\text{BB}} = 20\epsilon/\text{rad}^2$ compared to that of $10\epsilon/\text{rad}^2$ in model 1 to increase the stiffness of CLP triple helices with model 2 compared to model 1. While model 1 captures the directionality of H-bond beads with respect to the backbone of CLP strands, the directionality of sequential interstrand H-bonds gives rise to specific CLP triple helix twist and pitch, both of which are dependent on the CLP sequence of interest.¹ While neither model is intended to capture helicity correctly, in order to correctly obtain the melting transition trends with increasing CLP length and to preserve the cylindrical nature of the CLP triple helix in the hybridized state, model 2 incorporates the directionality of H-bond beads with respect to neighboring H-bond beads along the same strand. This is achieved by introducing two H-bond bead-backbone bead-adjacent backbone bead-H-bond bead dihedral angle potentials

$$U^{\text{dih}}(\phi) = k_d(1 + \cos(\phi - \phi_0)) \quad (4)$$

in which the dihedral constant $k_d = 15\epsilon$ and the dihedral angle ϕ is with reference to the plane defined by the angle made by the first three beads participating in the dihedral angle. For example, for CLP chain of (POG) triplets, the angle for the first (PH-PB-GB-GH) dihedral is $\phi_0 = 120^\circ$ where the reference plane is defined by the PH-PB-GB angle, and the angle for the second (GH-GB-PB-PH) dihedral $\phi_0 = -120^\circ$ where the reference plane is defined by the GH-GB-PB angle. The definition of these two dihedral angles holds for all model 2 (XYG) triplets modeled in this paper where $\phi_0 = 120^\circ$ corresponds to all XH-XB-GB-GH dihedrals, and $\phi_0 = -120^\circ$ corresponds to all GH-GB-XB-XH dihedrals. Figure 1c shows a $(\text{POG})_2$ CLP strand using model 2 that demonstrates the above-defined dihedral angles and provides the top and side view of this strand. Through this definition of dihedral angle potential, we achieve a staggered, triangular interstrand H-bond network as shown in Figure 1d. All other bonded and nonbonded interaction parameters remain the same for model 1 and model 2 with respect to the dimensionless units.

II.C. ELP–CLP Conjugate Coarse-Grained Model. To test if the CLP model developed in this work is able to reproduce the expected trends seen in systems of ELP–CLP conjugates, we replace the rigid rod model of CLP used in our previous work³¹ with the newly developed model 2. Each ELP–CLP conjugate strand consists of a CLP strand with all the same units, bonded and nonbonded potentials, and types of beads as model 2, along with one bead–spring ELP polymer model attached terminally to each CLP strand by means of a harmonic bond potential with bond length 0.84σ and the bond constant is 1000ϵ , with σ and ϵ as defined in model 2. The ELP chain is represented as a generic bead–spring polymer model (as done in our previous work³¹) where each ELP bead (EB) has a diameter of 1.0σ and a mass of $1.0m$. ELPs are polymers that mimic the natural polypeptide elastin with the repeat unit Val-Pro-Gly-Xaa-Gly (VPGXG), where Val is valine, Pro is proline, Gly is glycine, and Xaa can be any amino acid except proline. Each of these CG EB beads could roughly represent a collection of atoms in one amino acid/residue in the ELP. The neighboring bonded EB beads are modeled via a harmonic bond potential where the bond length is 0.84σ and the bond constant is 1000ϵ . EB beads do not have any other bonded interactions.

Our recent atomistic simulations showed that with increasing temperature, the ELP chains of sequence $(\text{VPGFG})_6$, in aqueous solutions undergoes an LCST transition because of decreasing chain hydration, quantified by the number of water molecules in the first shell around ELP and number of ELP–water H-bonds.³¹ In the CG model of ELP, we mimic that decreasing hydration or increasing solvophobicity implicitly using an effective attraction between nonbonded EB beads with the Lennard-Jones potential⁴⁹

$$U^{\text{EB}}(r) = \epsilon^{\text{EB}} \left[\left(\frac{\sigma^{\text{EB}}}{r} \right)^{12} - \left(\frac{\sigma^{\text{EB}}}{r} \right)^6 \right] \quad (5)$$

where ϵ^{EB} is the effective attractive strength between EB beads that is gradually varied to mimic increasing temperature effects, $\sigma^{\text{EB}} = 1.0\sigma$, and the cutoff distance is 2.5σ . Thus, with this CG model of a generic (LCST or inverse transition) ELP-like polymer, we observe the collapse of the ELP polymer with increasing ϵ^{EB} , mimicking the ELP chain in water undergoing decreasing hydration with increasing temperature.

All other nonbonded pairwise interactions of EB beads with any other bead in the CLP model are captured using the WCA potential.⁵¹

II.D. Simulation Details. In the simulations of CG model 1 CLP triple helices, we build the initial configuration by randomly placing 10 CLP triple helices in a cubic simulation box of size 63.5σ with periodic boundary conditions. This system represents a concentration of ~ 2 mM for CLP triple helices. In the CG model 2 simulations, we randomly place 10 CLP triple helices in a cubic simulation box with size 110σ with periodic boundary conditions, achieving experimentally relevant concentration range.⁵³ The triple helices are formed by placing three individual CLP single strands where the backbone beads from each of the strands arrange in a triangular fashion, all along the CLP length, such that all possible H-bonds detailed in Figure 1d are formed, and each strand is staggered from one another by 1 bead similar to what is observed experimentally.¹ We only include IN beads in systems where the CLP helices have charged beads to achieve charge neutrality in the system.

For every charged CLP backbone bead, we randomly place a counterion IN bead of the opposite charge in the simulation box without overlap to achieve net charge neutrality in the simulation box. We also insert an additional 1 mM of monovalent salt to represent a solution with salts added, to mimic common experimental conditions where salts are present in the aqueous solution. For the ELP–CLP conjugate system, we place 10 ELP–CLP triple helix conjugates in a periodic boundary condition box with side lengths of 140σ to mimic the experimental ELP–CLP conjugate concentration regime of Luo and Kiick.²⁶ The initial configuration of the ELP–CLP conjugate in the simulation has 3 ELP strands connected to a CLP triple helix on the same end.

We use Langevin dynamics in the NVT ensemble as implemented within the LAMMPS⁴⁸ package. In all simulations, the friction coefficient is set to 10τ (where 0.001τ is approximately 6 fs) as in Ghobadi and Jayaraman.⁴⁷ This was originally chosen to correspond to the diffusion of oligomeric DNA in water; since our focus is on results at equilibrium and not dynamics we arbitrarily use the same friction coefficient in our simulation. We use a two-level RESPA⁵⁴ integrator for all simulations in which nonbonded interactions are integrated with a time step of 0.001τ and bonded interactions are integrated with a time step of 0.0005τ . In the case of charged systems, we use the particle–particle–particle–mesh (PPPM)⁵⁵ method for electrostatics with a force tolerance of 10^{-2} , the interpolation order is 2, and the real-space cutoff distance is 20σ . All simulations are performed with a 10^8 time step equilibration run followed by a 10^7 time step sampling run in which data is collected every 100,000 timesteps. For the simulations with charged beads in the CLP triple helices, the first 10^7 timesteps of the equilibration run are performed without electrostatic interactions, after which the electrostatic interactions are turned on for the remainder of the simulation. This protocol is chosen to allow for equilibration of the triple helical structure before introduction of electrostatic forces. Given the long equilibration run, this choice of initial protocol should not impact the equilibrium sampling and ensemble averages.

For every system studied, we perform three replicate simulations with differing initial configurations, differing in the random placement of the CLP helices in the box and in velocities.

II.E. Analyses. As shown in the simulation snapshots in the Supporting Information, Figure S1, the CLP triple helices are stable at low temperatures and melt at higher temperature, separating into individual strands. To quantify the hybridization/melting transition, we define the ensemble average fraction of CLP triple helices that have at least 75% H-bonds formed as $f_{75\%}$ for the system at a range of reduced temperatures (T^*). This $f_{75\%}$ versus T^* is like the DNA melting curves ($f_{50\%}$ versus T^*) but we choose 75% instead of 50% as done in DNA melting because we see that two CLP strands can hybridize with almost two-thirds of the H-bonds that are possible in a CLP triple helix, without involvement of a third CLP strand. Thus, $f_{50\%}$ is not a good metric if we wish to distinguish a fully hybridized triple helix structure with three CLP strands from partially hybridized triple helix structure with three CLP strands, while $f_{75\%}$ is. We also report in the Supporting Information another metric to quantify the hybridization/melting that produces the same trends in melting curves as $f_{75\%}$ versus T^* . All melting curves reported in the

Results in this paper are $f_{75\%}$ versus T^* with standard deviation of $f_{75\%}$ reported in the error bars.

Even though we do not intend to mimic the true structure of CLP triple helix with either CG models, we compare the ensemble average end-to-end distance ($\langle R_{EE} \rangle$) and diameter of the CLP triple helices for both models to that from atomistic simulations.^{45,46} The end-to-end distance is the length of the vectors spanning the terminal residues of individual strands within a given helix. The $\langle R_{EE} \rangle$ and standard deviation is computed from the collection of all simulation frames in all three replicates. The diameter of the CLP triple helices is the diameter of a circle that contains the three central-most (Y-position) beads of a given triplet from three CLP strands participating in a triple helix. Since both the $\langle R_{EE} \rangle$ and the average diameter of our CG simulations of both models are reported in reduced units of σ , to convert $\langle R_{EE} \rangle$ and average diameter to real units for direct comparison with experimental values, we multiply $\langle R_{EE} \rangle$ and average diameter by the value of σ depending on the model used.

In the ELP–CLP conjugate simulations, given that our primary goal is to distinguish the solvated states of ELP-like polymer before and after the inverse transition, we quantify the ensemble average number of EB pairwise contacts ($\langle N_{EB-EB} \rangle$) with respect to the interaction strength between EB beads (ϵ^{EB}). A low value of $\langle N_{EB-EB} \rangle$ corresponds to a hydrated ELP-like polymer and increasing value of $\langle N_{EB-EB} \rangle$ corresponds to decreasing hydration of the ELP-like polymer. For each frame, N_{EB-EB} is computed by counting the number of EB beads within 2.5σ of each EB bead in the simulation then dividing the total by the number of EB beads in the simulation. Then for all frames over all three replicates, the ensemble average and standard deviation of N_{EB-EB} is recorded and presented. We mark the onset of aggregation or the inverse temperature transition as the ϵ^{EB} value at which the graph of $\langle N_{EB-EB} \rangle$ vs ϵ^{EB} encounters an inflection point.

II.F. Parameters Varied. To investigate whether the coarse-grained models (model 1 and model 2) capture the increase in CLP triple helix T_m observed experimentally^{7,40} with increasing numbers of (POG) triplets in CLP chains comprising a CLP triple helix, we perform simulations of $(POG)_n$ where $n = 6, 8, 10$, and 12 . To test if we reproduce the experimental observation²⁷ of decreasing T_m of CLP triple helices with the double mutations (O to K and P to D) at both ends of CLP strands, we perform simulations of $(POG)_{12}$ and $(PKG)_4(POG)_4(DOG)_4$ with both models. Model 1 simulations are performed for a range of reduced temperatures between $T^* = 2.0$ and $T^* = 5.0$ and model 2 simulations are performed for a range of reduced temperatures between $T^* = 3.0$ and $T^* = 6.0$, as the modifications to the model shifts the temperature scale where the CLPs are intact/melted. In Figure 2a, we present visual representations of the different CLP helices studied in this work.

We also investigate whether modeling ELP–CLP conjugates using CLP model 2 described in section II.B qualitatively captures the decrease in ϵ^{EB} of the onset of aggregation of ELP–CLP conjugates with respect to free ELP strands in solution seen computationally in Condon et al.³¹ and experimentally in Luo and Kiick.²⁶ We perform a set of simulations for the ELP–CLP conjugate system with the CLP triple helix modeled as $(POG)_{12}$ and ELP as strands of 30 amino acids each which is like the ELP–CLP conjugates studied in Luo and Kiick with ELP being $(VPFGF)_6$ where V represents valine, P represents proline, G represents glycine,

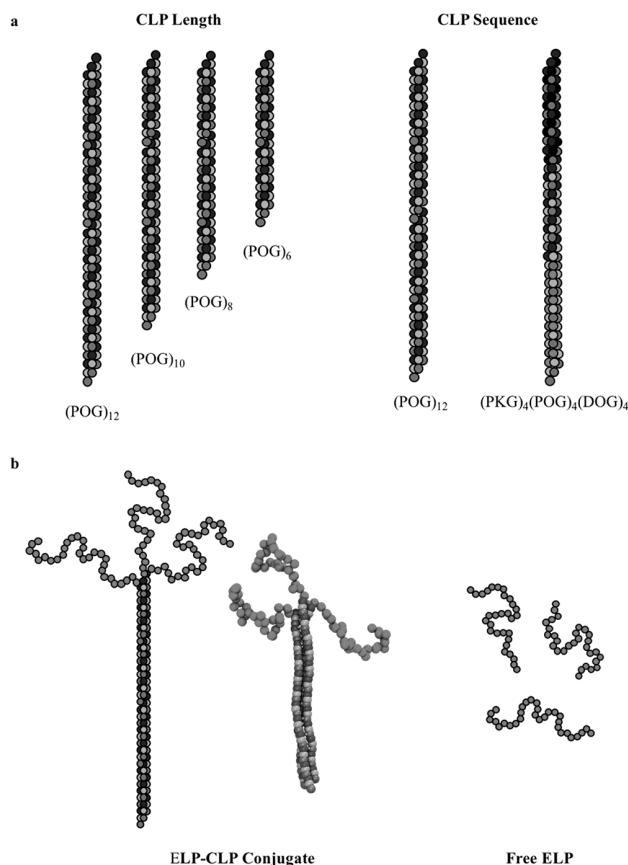


Figure 2. (a) Schematic of the CLP length and sequences studied. (b) Schematic of ELP–CLP conjugate and free ELPs along with a simulation snapshot of the ELP–CLP conjugate.

and F represents phenylalanine. We also perform simulations of the free ELP system with 30 ELP strands, each strand with 30 amino acids, to maintain the same volume fraction of ELP beads between simulations. A snapshot of one of the ELP–CLP triple helices is shown in Figure 2b. The ELP–CLP conjugate and free ELP simulations are performed at a range of ϵ^{EB} values between $\epsilon^{EB} = 0.1\epsilon$ and $\epsilon^{EB} = 1.3\epsilon$ at a constant reduced temperature of $T^* = 2.0$ that corresponds to a hybridized CLP triple helix.

III. RESULTS AND DISCUSSION

III.A. Effect of CLP Length on CLP Melting Transition Temperature. First, we explore how the number of repeat units impacts the melting transition of $(\text{POG})_n$ triple helices.

In Figure 3, we present the melting transition data for the simulations of $(\text{POG})_{12}$, $(\text{POG})_{10}$, $(\text{POG})_8$, and $(\text{POG})_6$ triple helix systems using both model 1 (Figure 3a with zoomed in version in part b) and model 2 (Figure 3c with a zoomed in version in part d). In these plots in Figure 3, a value of $f_{75\%} = 1.0$ corresponds to a completely hybridized CLP triple helix and a value of 0 corresponds to a completely melted CLP triple helix (see Figure S.1 for representative simulation snapshots of hybridized and melted states). An approximate way to identify the melting temperature would be to mark the T^* that corresponds to $f_{75\%} = 0.5$. On the basis of circular dichroism melting temperature data,^{7,40} we expect that the melting transition temperature should increase as the numbers of (POG) repeat units per CLP strand is increased. With model 1,

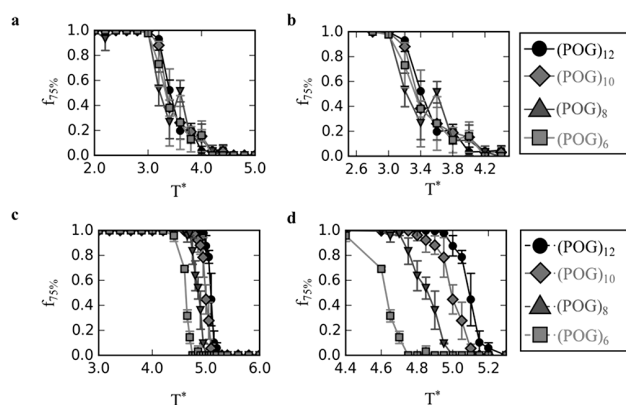


Figure 3. Ensemble average fraction of CLP triple helices that have at least 75% of possible H-bonds formed ($f_{75\%}$) as a function of reduced temperature (T^*) at a concentration of 0.1 mM for $(\text{POG})_n$ where n is equal to 12 (black circles), 10 (green diamonds), 8 (blue triangles), and 6 (red squares) using model 1 (part a and zoomed in version in part b) and model 2 (part c and zoomed in version in part d).

(Figure 3, parts a and b), we do not see any significant differences in the melting temperature as number of (POG) repeat units per CLP strand increases. We hypothesize that this failure of model 1 to capture the differences in melting temperatures with increase n in $(\text{POG})_n$ is due to the following reasons: (i) We know the melting temperature of CLP triple helix is dictated by the temperature where we expect overall enthalpic gain for helix to be hybridized (e.g., ΔU_{hyb}) to balance out the entropic loss of the CLP single strands to be in a hybridized intact helix (e.g., ΔS_{hyb}). Therefore, for melting temperature to increase with CLP strand length there should either be an increase in the ΔU_{hyb} or a decrease in the ΔS_{hyb} with increasing CLP length. The primary contribution to the ΔU_{hyb} term in model 1 is the attractive H-bond interactions and the bonded interactions that facilitate the H-bond interactions. The terms that contribute to the ΔS_{hyb} are the loss in translational entropy which decreases with increasing CLP length, and the loss in conformational entropy which decreases with increasing stiffness of the CLP strand. Upon closer visualization of the snapshots in the partially hybridized region for model 1 simulations, we find that the lack of stiffness (or restricted conformations) in the CLP strands leads to formation of nontriple helical hybridized structures as shown in Figure 4. Such structures as seen in Figure 4 would be avoided if bonded interactions and stiffness of the strands are tuned to maintain the desired directionality of the interstrand H-bond network among CLP strands that form triple helices. If the stiffness of the CLP is increased through bonded interactions (e.g., dihedrals) along the length of the CLP strand, both the ΔU_{hyb} and the ΔS_{hyb} terms described above would be affected with increasing CLP strand length. Thus, these aspects are addressed in model 2 as described in section II.B. by introduction of dihedral angle potentials and more restricted angle potentials that ensure the desired directionality of the H-bonds within the intact triple helix and increase stiffness of the CLP strands.

In parts c and d of Figure 3, with model 2, we see a distinct increase in the melting temperature transition with increase in CLP length (or increasing n in $(\text{POG})_n$) confirming that the inclusion of the desired directionality in H-bonds and the stiffening of the backbone, incorporated in model 2 but missing in model 1, are both important in capturing the experimental



Figure 4. Simulation snapshot of model 1 (POG)₁₂ system at $T^* = 3.6$ that shows a nonhelical hybridized structure.

melting trends of CLP triple helices with increasing CLP length. As suggested earlier, the interstrand H-bond directionality impacts the stability of the CLP triple helix.³⁸ One may see this as being analogous to behavior in DNA double helix where the directionality of H-bond is stabilized via stacking interactions between adjacent bases along the strand, mimicked using the dihedral interactions in the ONA model of Ghobadi and Jayaraman;⁴⁷ this is essential to capture the correct trends in DNA melting temperatures. Also, with model 2, we do not see nontriple helical hybridized structures (as in Figure 4) when we look through the simulation snapshots. One may note that the melting transition for model 2 occurs at a higher T^* for all systems studied than the corresponding system for model 1. This again is not surprising given that we intended to increase the ΔU_{hyb} and/or decrease the ΔS_{hyb} with the intended changes to model 1, and as a result increase the stability of the triple helix (or increase melting transition T^*).

III.B. Effect of CLP Sequence on Melting Transition Temperature. Next, we investigate how the substitution of charged subunits in place of (POG) impacts the melting temperature of CLP triple helices.

In parts a and b of Figure 5, we see that for both models the melting transitions for (PKG)₄(POG)₄(DOG)₄ triple helices are at a lower temperature than the (POG)₁₂ triple helix system. This behavior is observed in experiments²⁷ and is expected because the hybridization of (PKG)₄(POG)₄(DOG)₄ triple helices places like charges within proximity of one another leading to repulsive electrostatic forces that destabilize the ends of triple helix, and in turn, reduces the melting temperature of the CLP triple helix. The nonzero $f_{75\%}$ behavior of the (PKG)₄(POG)₄(DOG)₄ for model 2 at higher temperatures is attributed to nonhelical hybridized states. Favorable electrostatic interactions between the negatively charged residues in DOG triple and PKG triplets along the CLP strand cause rings of CLP strands to form as can be seen in Figure 6 at temperatures above the initial drop in $f_{75\%}$. For example, at $T^* = 4.2$ all of the strands in the simulation are participating in H-bonds with one another, however very few have a triple helical conformation, and many participate in rings of CLP strands. At higher temperatures ($T^* = 6.0$), most

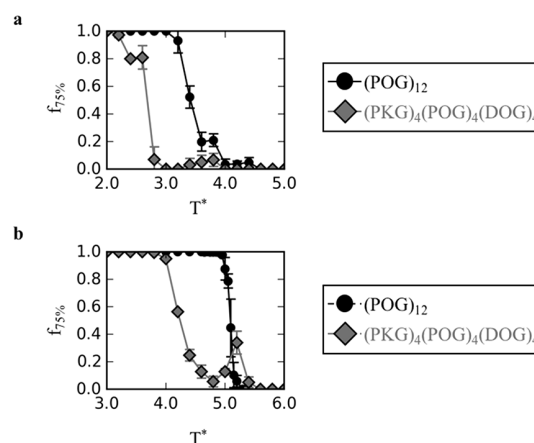


Figure 5. Fraction of CLP triple helices that have at least 75% of possible H-bonds formed ($f_{75\%}$) as a function of reduced temperature (T^*) and a concentration of 0.1 mM for model 1 (a) and model 2 (b) for the sequence (POG)₁₂ (black circles) and the double mutation (PKG)₄(POG)₄(DOG)₄ (green diamonds).



Figure 6. Simulation snapshot of the (PKG)₄(POG)₄(DOG)₄ system for model 2. Nontriple helical hybridized structures can form in such a way that they are stabilized by attractive electrostatic interactions.

strands are completely melted and are no longer forming H-bonds. A possible way to eliminate this behavior is to increase the backbone angle coefficient even further than what is done in model 2. Alternate method to quantify melting curves, like the one described in Supporting Information, result in cleaner melting curves (Figures S.2 and S.3) by better distinguishing free unhybridized/melted strands, hybridized nonhelical conformations, and hybridized helical conformations; the trends in melting temperature do not change with such alternate metrics.

III.C. Use of Model 2 in ELP–CLP CG Model. Next, we investigate how the use of model 2 in ELP–CLP conjugates impact the trends in the onset of aggregation behavior for ELP–CLP conjugates and free ELP, and compare these results to that in our recent work³¹ where CLP was modeled as a stiff rod in ELP–CLP conjugates. To understand the relative lengths of ELP and CLP blocks, we calculate the end-to-end distance of a single CLP triple helix in our coarse-grained simulations and compare that to experimental values of maximum CLP length. We also calculate the diameter of the CLP triple helix to better understand the size of the substrate that the ELP chains are attached to.

In Table 2, we show the average end-to-end distance for both models as described in section II for coarse-grained simulations of CLP triple helices at $T^* = 3.0$. Setting σ to 0.3 nm for model 1 and σ to 0.5 nm for model 2 allows us to directly compare the results from two sets of simulations with experimental values

Table 2. Ensemble Average End-to-End Distance ($\langle R_{EE} \rangle$) and Diameter of CLP Helices in Simulations Using Model 1 and Model 2 at a Reduced Temperature $T^* = 3.0$ Where the Triple Helix Is Intact

model	sequence	$\langle R_{EE} \rangle$ (nm)	diameter (nm)
model 1	(POG) ₁₂	7.71 ± 0.30	0.77 ± 0.03
model 2	(POG) ₁₂	8.44 ± 0.12	1.16 ± 0.02

for length and diameter of CLP triple helices. Small angle X-ray scattering experiments report the length of (POG)₁₂ helices in solution to be 8.9 nm.³⁹ X-ray diffraction data for the crystal structure of a similar CLP sequence (POG)₄(POA)₁(POG)₅ report the diameter of the CLP triple helix to be approximately 1 nm with a resolution of 0.19 nm.⁵⁶ Atomistic simulations in literature show that the end to end distance of triple helix with nine POG is 7.24 nm,⁴⁵ and in another study, the end to end distance of (POG)₉ is 7.97 nm with a diameter of 0.57 nm.⁴⁶ Our simulation results for (POG)₁₂ presented in Table 2, in particular from model 2, are in good agreement with the experimental characterization of length ~ 8.9 nm³⁹ and diameter ~ 1 nm of (POG)₁₂ helices.⁵⁶ For the ELP–CLP systems in terms of relative sizes, the (POG)₁₂ CLP triple helix represented with model 2 has an end-end distance of 16.88σ in contrast to the root-mean-square radius of gyration ($\langle R_g^2 \rangle^{0.5}$) of a 30mer ELP-like polymer equal to $2.54 \pm 1.50\sigma$ at $\epsilon^{EB} = 0.1\epsilon$. Our past work³¹ showed that the length of CLP triple helix varied from 1 triplet to 14 triplets did not alter the inverse transition of the conjugated ELP strands. Therefore, we do not explore the CLP length effect on ELP transition here.

In Figure 7, we present ensemble average number of EB–EB contacts ($\langle N_{EB-EB} \rangle$) vs the strength of attractive pairwise

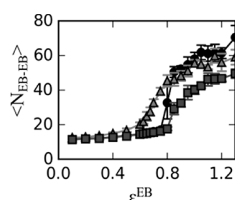


Figure 7. Ensemble average number of EB–EB contacts ($\langle N_{EB-EB} \rangle$) vs the strength of attractive pairwise interactions among EB beads (ϵ^{EB}) for free ELP (black circles) with 30mer ELP strands, for ELP–CLP conjugates with (POG)₁₂ CLP strands and three 30mer ELP strands per CLP helix (red triangles), and for ELP–CLP conjugate with (POG)₁₂ CLP strands and one 30mer ELP strand per CLP helix (blue squares) at a reduced temperature of $T^* = 2.0$. The number of molecules and the simulation box size are reported in section II.

interactions among EB beads for both free ELP and for ELP–CLP conjugates at a reduced temperature of $T^* = 2.0$. We note that $T^* = 2.0$ is chosen because it is well below the melting temperature for the (POG)₁₂ triple helix, therefore the CLP portions of the ELP–CLP conjugates in these simulations are in a triple helical conformation for the duration of the simulations. If we take the inflection point of these curves in Figure 7, we see that the onset of aggregation is at a lower value of attractive pairwise interactions among EB beads (i.e., lower ϵ^{EB}) for the ELP–CLP conjugate system than it is for the free ELP system. This result follows the trends of our previous coarse-grained molecular simulation results³¹ which was in agreement with experiments²⁶ in which the LCST of ELP–CLP conjugates is lowered by greater than 80 °C compared to

free ELP in solution. The work in Condon et al.³¹ has shown that the conjugation of multiple ELP strands to CLP reduces the loss in the change in entropy upon aggregation, and thus, requires a lower interaction strength between EB beads for aggregation to occur. If we eliminate this crowding by attaching only a single ELP to the CLP triple helix, indeed we shift the onset of aggregation to higher ϵ^{EB} than the free ELP, in agreement with our previous results.³¹

The rigid body model for CLP in Condon et al.³¹ performs as well as the more sophisticated CLP model 2, but if the goal is to understand how ELP conjugation impacts CLP melting, a rigid body model of CLP will be insufficient and our significantly more sophisticated CLP model 2 will be the appropriate model as it captures CLP helix melting. Also, to explore the impact of CLP conjugation to ELP aggregation at higher T^* where CLP helix is at the cusp of melting, model 2 will be appropriate. The future directions for this improved CLP model 2 are to investigate how CLP melting impacts the onset of aggregation of ELP–CLP conjugates, how ELP aggregation impacts ELP–CLP conjugate melting, and how the ratio of ELP length to CLP length in these conjugates impacts both melting and aggregation.

IV. CONCLUSION

In this study, we present our sequential development of two coarse-grained models—model 1 and model 2—for CLP to investigate how the length of CLP strands or the n in (POG) _{n} impact the resulting melting transition of the CLP triple helix. Both models are capable of capturing the directional hydrogen bond (H-bond) interactions between the H-bond donors and acceptors along the CLP strands that lead to a stable triple helix structure at low temperature. Model 1, however, was not able to show the experimentally observed increase in T_m with increasing n in (POG) _{n} , whereas model 2 was able to show a distinct increase in T_m with respect to increasing n in (POG) _{n} . This was achieved through the incorporation of additional bonded interactions in model 2 that effectively stiffened the CLP strands and maintains the H-bond interactions between the correct donor and acceptor pairs along the helix. We also evaluated the melting behavior of (PKG)₄(POG)₄(DOG)₄ with respect to (POG)₁₂ and showed that simulations with both models of (PKG)₄(POG)₄(DOG)₄ shows a lower temperature for the melting transition than (POG)₁₂ system. The reduction in melting temperature for the (PKG)₄(POG)₄(DOG)₄ versus (POG)₁₂ system is likely due to the presence of destabilizing electrostatic repulsion among like charged beads at the ends of the (PKG)₄(POG)₄(DOG)₄ strands. Finally, to demonstrate that the developed CLP models can be extended to capture thermodynamics in systems of CLPs conjugated to other macromolecules, we showed that simulations of ELP–CLP conjugates using model 2 to represent the CLP triple helix portion of the conjugate qualitatively captures the same behavior seen in Condon et al.³¹ where the onset of aggregation occurs at a lower ELP–ELP bead pairwise interaction strength than for free ELP systems.

Next, we discuss some of the limitations and possible improvements for CLP model 2. With the implicit solvent representation in our model, the water bridging stabilizing H-bond interactions that have been hypothesized to stabilize the POG-based CLP triple helix are not captured explicitly. Experimental understanding of the atomistic structures of POG based CLP triple helix is that the hydroxyl hydrogen on the hydroxyproline (O) side chain of (POG) _{x} forms H-bonds

with water molecules which in turn form H-bonds with a carboxyl oxygens of a neighboring CLP strand's backbone proline (P).³⁶ As stated earlier, some compositions and sequences of CLP (e.g., PPG triplets) that do not have this hydroxyproline will not have such water mediated H-bond interactions. Given that our CG model (model 2) seems to capture the effect of POG length on CLP helix melting temperature correctly, we chose not to incorporate explicit solvent CG beads capable of H-bond interactions with the CLP H-bond beads. We also note that while this explicit representation of water-CLP interactions may seem more realistic, it comes at a cost of significantly slowing down the simulation speed and significantly restricting the length and time scales of such studies of CLP triple helix. Finally, given our focus on thermodynamics and not structure, neither model presented here differentiates between amino acids that favor or do not favor the PPII helical conformation of individual strands (e.g., proline vs glycine); PPII helical conformation is necessary for the formation of a true triple helix. If these structural features are critical for an application, this model could be improved by distinguishing the identity of the backbone beads, and incorporating specific bonded potentials around/on those backbone beads to capture the differences in the amino acids to favor PPII helical conformation.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcb.7b10916.

Simulation snapshots and an additional method to quantify melting curves (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*(A.J.) E-mail: arthij@udel.edu.

ORCID

Arthi Jayaraman: 0000-0002-5295-4581

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank the National Science Foundation (NSF) for Grant 1703402 for financial support. The authors thank the information technologies resources program at the University of Delaware, specifically the Farber high-performance computing resources program, for supporting the computational needs of this project.

■ REFERENCES

- (1) Bella, J. Collagen structure: New tricks from a very old dog. *Biochem. J.* **2016**, *473* (8), 1001.
- (2) Beck, K.; Brodsky, B. Supercoiled protein motifs: The collagen triple-helix and the α -helical coiled coil. *J. Struct. Biol.* **1998**, *122* (1), 17–29.
- (3) Brodsky, B.; Persikov, A. V. Molecular structure of the collagen triple helix. *Adv. Protein Chem.* **2005**, *70*, 301–339.
- (4) Brodsky, B.; Ramshaw, J. A. M. The collagen triple-helix structure. *Matrix Biol.* **1997**, *15* (8), 545–554.
- (5) Shoulders, M. D.; Raines, R. T. Collagen structure and stability. *Annu. Rev. Biochem.* **2009**, *78* (1), 929–958.
- (6) Brown, F. R., III; Hopfinger, A. J.; Blout, E. R. The collagen-like triple helix to random-chain transition: Experiment and theory. *J. Mol. Biol.* **1972**, *63* (1), 101–115.
- (7) Sakakibara, S.; Inouye, K.; Shudo, K.; Kishida, Y.; Kobayashi, Y.; Prockop, D. J. Synthesis of (pro-hyp-gly)_n of defined molecular weights evidence for the stabilization of collagen triple helix by hydroxyproline. *Biochim. Biophys. Acta, Protein Struct.* **1973**, *303* (1), 198–202.
- (8) An, B.; Lin, Y.-S.; Brodsky, B. Collagen interactions: Drug design and delivery. *Adv. Drug Delivery Rev.* **2016**, *97*, 69–84.
- (9) Kojima, C.; Tsumura, S.; Harada, A.; Kono, K. A collagen-mimic dendrimer capable of controlled release. *J. Am. Chem. Soc.* **2009**, *131* (17), 6052–6053.
- (10) Yamazaki, C. M.; Nakase, I.; Endo, H.; Kishimoto, S.; Mashiyama, Y.; Masuda, R.; Futaki, S.; Koide, T. Collagen-like cell-penetrating peptides. *Angew. Chem., Int. Ed.* **2013**, *52* (21), 5497–5500.
- (11) Yasui, H.; Yamazaki, C. M.; Nose, H.; Awada, C.; Takao, T.; Koide, T. Potential of collagen-like triple helical peptides as drug carriers: Their in vivo distribution, metabolism, and excretion profiles in rodents. *Biopolymers* **2013**, *100* (6), 705–713.
- (12) Koide, T. Collagen-like triple helical peptides: Applications in drug development and regenerative medicine. *Yakugaku Zasshi* **2013**, *133* (3), 387–92.
- (13) Chattopadhyay, S.; Raines, R. T. Review collagen-based biomaterials for wound healing. *Biopolymers* **2014**, *101* (8), 821–833.
- (14) Fields, G. B. Synthesis and biological applications of collagen-model triple-helical peptides. *Org. Biomol. Chem.* **2010**, *8* (6), 1237–1258.
- (15) Golinska, M. D.; Włodarczyk-Biegun, M. K.; Werten, M. W. T.; Stuart, M. A. C.; de Wolf, F. A.; de Vries, R. Dilute self-healing hydrogels of silk-collagen-like block copolypeptides at neutral pH. *Biomacromolecules* **2014**, *15* (3), 699–706.
- (16) Hernandez-Gordillo, V.; Chmielewski, J. Mimicking the extracellular matrix with functionalized, metal-assembled collagen peptide scaffolds. *Biomaterials* **2014**, *35* (26), 7363–7373.
- (17) Islam, M. M.; Ravichandran, R.; Olsen, D.; Ljunggren, M. K.; Fagerholm, P.; Lee, C. J.; Griffith, M.; Phopase, J. Self-assembled collagen-like-peptide implants as alternatives to human donor corneal transplantation. *RSC Adv.* **2016**, *6* (61), 55745–55749.
- (18) Koide, T. Triple helical collagen-like peptides: Engineering and applications in matrix biology. *Connect. Tissue Res.* **2005**, *46* (3), 131–141.
- (19) Pires, M. M.; Przybyla, D. E.; Chmielewski, J. A metal–collagen peptide framework for three-dimensional cell culture. *Angew. Chem., Int. Ed.* **2009**, *48* (42), 7813–7817.
- (20) Reyes, C. D.; Garcia, A. J. Engineering integrin-specific surfaces with a triple-helical collagen-mimetic peptide. *J. Biomed. Mater. Res., Part A* **2003**, *65A* (4), 511–523.
- (21) Rubert Pérez, C. M.; Panitch, A.; Chmielewski, J. A collagen peptide-based physical hydrogel for cell encapsulation. *Macromol. Biosci.* **2011**, *11* (10), 1426–1431.
- (22) Woolfson, D. N. Building fibrous biomaterials from α -helical and collagen-like coiled-coil peptides. *Biopolymers* **2010**, *94* (1), 118–127.
- (23) Bai, H.; Xu, K.; Xu, Y.; Matsui, H. Fabrication of au nanowires of uniform length and diameter using a monodisperse and rigid biomolecular template: Collagen-like triple helix. *Angew. Chem., Int. Ed.* **2007**, *46* (18), 3319–3322.
- (24) Gottlieb, D.; Morin, S. A.; Jin, S.; Raines, R. T. Self-assembled collagen-like peptide fibers as templates for metallic nanowires. *J. Mater. Chem.* **2008**, *18* (32), 3865–3870.
- (25) Luo, T.; He, L.; Theato, P.; Kiick, K. L. Thermoresponsive self-assembly of nanostructures from a collagen-like peptide-containing diblock copolymer. *Macromol. Biosci.* **2015**, *15* (1), 111–123.
- (26) Luo, T.; Kiick, K. L. Noncovalent modulation of the inverse temperature transition and self-assembly of elastin-b-collagen-like peptide bioconjugates. *J. Am. Chem. Soc.* **2015**, *137* (49), 15362–15365.
- (27) O'Leary, L. E. R.; Fallas, J. A.; Bakota, E. L.; Kang, M. K.; Hartgerink, J. D. Multi-hierarchical self-assembly of a collagen mimetic

peptide from triple helix to nanofibre and hydrogel. *Nat. Chem.* **2011**, *3* (10), 821–828.

(28) Sarkar, B.; O'Leary, L. E. R.; Hartgerink, J. D. Self-assembly of fiber-forming collagen mimetic peptides controlled by triple-helical nucleation. *J. Am. Chem. Soc.* **2014**, *136* (41), 14417–14424.

(29) Rubert Perez, C. M.; Rank, L. A.; Chmielewski, J. Tuning the thermosensitive properties of hybrid collagen peptide-polymer hydrogels. *Chem. Commun.* **2014**, *50* (60), 8174–8176.

(30) Jin, H.-E.; Jang, J.; Chung, J.; Lee, H. J.; Wang, E.; Lee, S.-W.; Chung, W.-J. Biomimetic self-templated hierarchical structures of collagen-like peptide amphiphiles. *Nano Lett.* **2015**, *15* (10), 7138–7145.

(31) Condon, J. E.; Martin, T. B.; Jayaraman, A. Effect of conjugation on phase transitions in thermoresponsive polymers: An atomistic and coarse-grained simulation study. *Soft Matter* **2017**, *13* (16), 2907–2918.

(32) Bella, J. A new method for describing the helical conformation of collagen: Dependence of the triple helical twist on amino acid sequence. *J. Struct. Biol.* **2010**, *170* (2), 377–391.

(33) Berisio, R.; Vitagliano, L.; Mazzarella, L.; Zagari, A. Crystal structure of the collagen triple helix model [(pro-pro-gly)(10)](3). *Protein Science: A Publication of the Protein Society* **2002**, *11* (2), 262–270.

(34) Buehler, M. J. Atomistic and continuum modeling of mechanical properties of collagen: Elasticity, fracture, and self-assembly. *J. Mater. Res.* **2006**, *21* (8), 1947–1961.

(35) Domene, C.; Jorgensen, C.; Abbasi, S. W. A perspective on structural and computational work on collagen. *Phys. Chem. Chem. Phys.* **2016**, *18* (36), 24802–24811.

(36) Okuyama, K.; Hongo, C.; Wu, G.; Mizuno, K.; Noguchi, K.; Ebisuzaki, S.; Tanaka, Y.; Nishino, N.; Bächinger, H. P. High-resolution structures of collagen-like peptides [(pro-pro-gly)₄-xaa-*yaa*-gly-(pro-pro-gly)₄]: Implications for triple-helix hydration and hyp(x) puckering. *Biopolymers* **2009**, *91* (5), 361–372.

(37) Persikov, A. V.; Ramshaw, J. A. M.; Kirkpatrick, A.; Brodsky, B. Amino acid propensities for the collagen triple-helix. *Biochemistry* **2000**, *39* (48), 14960–14967.

(38) Suárez, E.; Díaz, N.; Suárez, D. Entropic control of the relative stability of triple-helical collagen peptide models. *J. Phys. Chem. B* **2008**, *112* (47), 15248–15255.

(39) Walker, K. T.; Nan, R.; Wright, D. W.; Gor, J.; Bishop, A. C.; Makhatadze, G. I.; Brodsky, B.; Perkins, S. J. Non-linearity of the collagen triple helix in solution and implications for collagen function. *Biochem. J.* **2017**, *474* (13), 2203.

(40) Persikov, A. V.; Ramshaw, J. A. M.; Brodsky, B. Prediction of collagen stability from amino acid sequence. *J. Biol. Chem.* **2005**, *280* (19), 19343–19349.

(41) Shoulders, M. D.; Raines, R. T. Modulating collagen triple-helix stability with 4-chloro, 4-fluoro, and 4-methylprolines. In *Peptides for youth: The proceedings of the 20th American peptide symposium*, Valle, S. D., Escher, E., Lubell, W. D., Eds.; Springer New York: New York, 2009; pp 251–252.

(42) Ying Chow, W.; Bihan, D.; Forman, C. J.; Slatter, D. A.; Reid, D. G.; Wales, D. J.; Farndale, R. W.; Duer, M. J. Hydroxyproline ring pucker causes frustration of helix parameters in the collagen triple helix. *Sci. Rep.* **2015**, *5*, 12556.

(43) Khew, S. T.; Tong, Y. W. Characterization of triple-helical conformations and melting analyses of synthetic collagen-like peptides by reversed-phase hplc. *J. Chromatogr. B: Anal. Technol. Biomed. Life Sci.* **2007**, *858* (1), 79–90.

(44) Raman, S. S.; Parthasarathi, R.; Subramanian, V.; Ramasami, T. Role of aspartic acid in collagen structure and stability: A molecular dynamics investigation. *J. Phys. Chem. B* **2006**, *110* (41), 20678–20685.

(45) Klein, T. E.; Huang, C. C. Computational investigations of structural changes resulting from point mutations in a collagen-like peptide. *Biopolymers* **1999**, *49* (2), 167–183.

(46) Gopalakrishnan, R.; Azhagiya Singam, E. R.; Vijaya Sundar, J.; Subramanian, V. Interaction of collagen like peptides with gold

nanosurfaces: A molecular dynamics investigation. *Phys. Chem. Chem. Phys.* **2015**, *17* (7), 5172–5186.

(47) Ghobadi, A. F.; Jayaraman, A. Effect of backbone chemistry on hybridization thermodynamics of oligonucleic acids: A coarse-grained molecular dynamics simulation study. *Soft Matter* **2016**, *12* (8), 2276–2287.

(48) Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.* **1995**, *117* (1), 1–19.

(49) Jones, J. E. On the determination of molecular fields. II. From the equation of state of a gas. *Proc. R. Soc. London, Ser. A* **1924**, *106* (738), 463–477.

(50) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. Gromacs: High performance molecular simulations through multi-level parallelism from laptops to super-computers. *SoftwareX* **2015**, *1*–2, 19–25.

(51) Weeks, J. D.; Chandler, D.; Andersen, H. C. Role of repulsive forces in determining the equilibrium structure of simple liquids. *J. Chem. Phys.* **1971**, *54* (12), 5237–5247.

(52) Stogryn, A. Equations for calculating the dielectric constant of saline water (correspondence). *IEEE Trans. Microwave Theory Tech.* **1971**, *19* (8), 733–736.

(53) Hilderbrand, A.; Stanzione, F.; Condon, J.; Larue, M.; Jayaraman, A.; Kloxin, A. *Understanding the impact of non-natural amino acid incorporation on the assembly of multifunctional collagen mimetic peptides*. Presented at the ACS National Meeting and Exposition, Washington, D.C., 2017.

(54) Tuckerman, M.; Berne, B. J.; Martyna, G. J. Reversible multiple time scale molecular dynamics. *J. Chem. Phys.* **1992**, *97* (3), 1990–2001.

(55) Hockney, R. W.; Eastwood, J. W. *Computer simulation using particles*; CRC Press: 1988.

(56) Berisio, R.; Vitagliano, L.; Mazzarella, L.; Zagari, A. Crystal structure of a collagen-like polypeptide with repeating sequence pro-hyp-gly at 1.4 Å resolution: Implications for collagen hydration. *Biopolymers* **2000**, *56* (1), 8–13.