COARSE-GRAINED MOLECULAR DYNAMICS SIMULATIONS OF THERMORESPONSIVE POLYMERS

by

Joshua E. Condon

A thesis submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Master of Chemical Engineering

Fall 2017

COARSE-GRAINED MOLECULAR DYNAMICS SIMULATIONS

OF THERMORESPONSIVE POLYMERS

by

Joshua E. Condon

Approved: _____

Arthi Jayaraman, Ph.D.
Professor in charge of thesis on behalf of the Advisory Committee

Approved: _____

Eric M. Furst, Ph.D.
Chair of the Department of Chemical and Biomolecular Engineering

Approved: _____

Babatunde A. Ogunnaike, Ph.D.
Dean of the College of Engineering

Approved: _____

Ann L. Ardis, Ph.D.
Senior Vice Provost for Graduate and Professional Education

ACKNOWLEDGMENTS

I would like to first and foremost thank my advisor, Professor Arthi Jayaraman, for her guidance, patience, and advice through this challenging but fruitful time in my life. I am most thankful for her commitment to helping me develop my communication abilities, skills that never lose their value over the course of a lifetime. Her passion for research is an inspiration to all of us that work in her lab to strive to be the best in our respective fields.

Next, I would like to thank my current and past lab group members for their continued friendship and advice. I especially want to thank Dr. Tyler Martin for his willingness to mentor me through the beginning stages of my research. He imprinted upon me a passion for software development that I hope to pursue as a career. I would like to thank Tom Gartner for his wisdom, both with research and interpersonal interactions, and for his assistance when I run into a wall with respect to research. When I needed to be cheered up, or needed emotional support, Michiel Wessels was sitting there next to me making me laugh. To all the postdocs in our group over the past two years: Dr. Francesca Stanzione, Dr. Ahmad Ghobadi, Dr. Daniel Beltran, Dr. Ivan Lyubimov, and our newest member Dr. Prhashanna Ammu, thank you for your assistance, guidance, and support. I would like to thank our collaborators and my thesis committee members Professor Kristi Kiick and Professor April Kloxin for their inspiring work, and to the graduate students that work with them on these projects, namely Tim Luo, Lucas Dunshee, and Amber Hilderbrand. I would also like to thank

my other thesis committee member, Professor Thomas Epps, III, for his guidance with respect to this thesis.

I am grateful for the friendships that I have made over the past few years in Delaware. I have met some amazing people, and cannot list them all in this document, but I want to mention the standouts, the people that have made this portion of my life enjoyable when things were going well and tolerable when things were not. David Phan and Matt Jouny, thanks so much for being great roommates and being even better friends. I am especially grateful for the many friends I have made at my church. Brad Dreibelbis, it has been an absolute pleasure serving with you in youth ministry, and without your belief and support, I do not believe I would have finished with my Master's. Additionally, I am forever grateful to my Wing Chun community and to Sifu Amy for their presence in my life. There were many days where I just needed to get away from it all, and Kung Fu was my solace. I hope to continue the tradition of excellence in Wing Chun teaching in the future. I would like to thank my roommate from University of Maryland, Connor Hart, for his consistent source of support over these past 6 years. Best of luck at Harvard!

Lastly, I want to thank my family. Mom, you are an inspiration and guiding light in my life, and I appreciate your honesty, though sometimes brutal, it has helped me through some very rough times. Dad, I cherish your belief in me, both in my abilities and in my character. It has given me the strength to press on and tackle the many challenges thrown at me. I would like to thank my siblings Jeffrey, Sarah, and John for the entertainment, love, and comradery that you provide whenever I get the opportunity to visit home. I look forward to the years to come.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Thermoresponsive polymers are present in many biomedical applications such as drug delivery and tissue engineering for their ability to change phase or morphology with changes in temperature. This thesis presents a coarse-grained (CG) molecular dynamics approach to understanding thermoresponsive behavior in three systems: elastin-like polypeptide (ELP), collagen-like polypeptide (CLP) and oligonucleic acid (ONA). CG molecular dynamics simulations allow for improved computational efficiency and facilitate study of macromolecular systems at length and time scales relevant to thermoresponsive behavior such as melting, aggregation, and assembly. The first system presented in this thesis is the elastin-like polypeptide (ELP)-collagen-like polypeptide (CLP) conjugate in an aqueous solution. ELP-CLP conjugates self-assemble above the lower critical solution temperature (LCST) of ELP into vesicles with drug delivery applications. An implicit solvent CG model of ELP-CLP conjugates was developed where ELP strands are considered as bead-spring polymers and CLP triple helices are considered as rigid rods. Through molecular dynamics simulations over a range of ELP-ELP interaction strengths, it was determined that ELP-CLP conjugates with multiple ELP strands conjugated to the CLP rigid body undergo the onset of aggregation at a lower ELP-ELP interaction strength than a corresponding system of ELP strands that are not conjugated to CLP. In addition, simulations of a system with only one ELP strand conjugated to a CLP rigid body did not show this decrease in the onset of interaction, thus supporting the hypothesis that

ELP crowding due to conjugation lowers the LCST of ELP-CLP conjugates in experiments with respect to the LCST of free ELP strands.

The second system presented in this thesis is the CLP triple helix in aqueous solution. This study primarily focuses on developing a phenomenological CG model of the CLP triple helix that mimics melting of CLP triple helix. By developing and implementing this CG model in the ELP-CLP conjugate system, analysis of ELP-CLP conjugates where CLP is in the melted state is possible. Using implicit solvent molecular dynamics simulations, the phenomenological CG model of CLP was able to qualitatively capture trends in melting temperature of the CLP triple helix with respect to CLP strand length and with substitution with charged amino acids. Further, this new model for CLP was substituted for the rigid body in the ELP-CLP conjugate system and simulations of this updated model demonstrated qualitative agreement with the original model with the added property of being able to undergo a melting transition.

The last system that was studied in this thesis is the oligonucleic acid (ONA)-star polymer conjugate system. A CG model for ONA-star polymer conjugates was developed where the ONA portion of the model was taken from a previous two site per nucleotide model that captures specific directional hydrogen bonding (H-bonding) interactions. The star polymer portion was modeled as a bead-spring polymer. Utilizing implicit solvent CG molecular dynamics simulations, both melting transition behavior and assembly transition behavior was studied with respect to changing ONA backbone flexibility, backbone charge, and number of star polymer arms. Increasing melting temperature was observed for ONA-star polymer conjugate systems with decreasing backbone flexibility and backbone charge. Assembly transition temperature

increases accordingly, and, in addition, increases with increasing number of star polymer arms per ONA-star polymer conjugates.

This thesis demonstrates the usefulness and importance of CG molecular dynamics simulations in screening large materials design parameter space, specifically for thermoresponsive polymers. These simulations provide a better understanding of the thermodynamics and driving forces for thermoresponsive behavior, and thus better inform scientists on how chemical and physical features of the polymer influences LCST, melting, and assembly behavior. In addition, using a CG framework, instead of experimental or atomistic framework, reduces the time and costs associated with testing a potentially vast design space of a thermoresponsive polymer. Future improvements in the models listed in this thesis may allow for better qualitative and quantitative characterization of thermoresponsive and self-assembly behavior in polymers.

Chapter 1

INTRODUCTION

Thermoresponsive polymers are a class of materials that undergo conformational changes and/or phase transitions with varying temperature making them attractive for use in a variety of application. In particular, biocompatible thermoresponsive polymers that undergo such changes in aqueous solutions and physiological relevant conditions find use in biological and pharmaceutical applications such as drug delivery, tissue engineering, molecular recognition/sensing, and imaging.[1-6] Limitations in the ability to quickly and inexpensively synthesize thermoresponsive polymers as well as an incomplete understanding of the thermodynamics associated with their phase transitions provide impetus for developing predictive computational models and methods to study the as a function of their chemistry. Atomistic molecular dynamics (MD) simulations are one such tool that enable linking the chemistry of these polymers to local structure of the polymer typically in the Angstrom to nanometer range in explicit solvent(s).[7-10] However, atomistic MD simulations are often not computationally efficient enough to probe phase transitions in these systems due to the macromolecular size of the polymer, the large number of solvent molecules necessary to represent concentrations of interest, and the number of timesteps necessary to run these simulations to achieve thermodynamic equilibrium or observe conformational transitions as a function of temperature. Alternatively, coarse-grained MD simulations with reduced degrees of freedom in the polymer and/or solvent are able to capture phase changes in polymers

at experimentally relevant length scales and time scales by reducing the level of detail in the models of these polymers and/or solvents while still retaining the key features of the systems and/or mimicking driving forces that are responsible for the phase transitions. Coarse-grained MD simulations thus enable the rapid screening of a broad design parameter space (e.g. polymer length, polymer flexibility, solvent conditions, etc.) for understanding how these design parameters impact polymer thermoresponsive behavior. In this thesis, under the advisement of my thesis advisor, I have focused on the development of coarse-grained models for simulations of three specific classes of biologically relevant thermoresponsive polymers to provide design rules and guidance for the synthesis of these polymers with properties that meet the criteria for their intended application.

The first class of polymers studied in this thesis are elastin-like polypeptides (ELPs) which are polymers that mimic elastin, a naturally occurring protein that imparts elasticity to biological tissues. ELPs consist of repeat units with the (Val-Pro-Gly-Xaa-Gly) amino acid sequence where Val (V) is valine, Pro (P) is proline, Gly (G) is glycine, and Xaa (X) can be any amino acid except proline. ELP strands undergo a lower critical solution temperature (LCST) transition in aqueous solutions where they are dispersed in solution below the LCST and aggregated in solution above the LCST. The LCST of ELP is tunable through ELP length, ELP concentration, solution quality, the amino acid in the X position of the repeat sequence.[10-13] The tunability of ELP's LCST transition in addition to its biocompatibility contribute to its use in biomedical applications such as drug delivery and tissue engineering.[1, 3, 14]

While experimental, theoretical, and computational efforts have elucidated the effect of ELP chemistry and solution quality on ELP LCST behavior in solution, less

is known about the effect of conjugating other polymers to ELP on the LCST of ELP. Recent work has shown that conjugating ELP strands to macromolecules has a significant impact on the LCST of ELP.[15-18] Luo and Kiick[16] have demonstrated that conjugating ELPs to collagen-like polypeptides (CLPs) significantly decreases the LCST of ELP in solution and leads to the self-assembly of ELP-CLP conjugates for applications in drug delivery. This leads to CLPs being the second class of bio-mimetic thermoresponsive polymers discussed in this thesis. CLPs are polymers that mimic the naturally occurring protein collagen which provides structural support in bodily tissue. CLP strands are polymers consisting of the (Gly-Xaa-Ybb) repeat sequence where Gly (G) is the amino acid glycine, and Xaa (X) and Ybb (Y) can be any amino acid. CLP strands form triple helices in solution at some temperatures, and undergo a melting transition as the temperature is increased. The melting temperature (T_m) is dependent on CLP strand length, the amino acids located in the X and Y positions of the CLP repeat unit, and solution quality.[9, 19-23] Additionally, stable CLP triple helices have a wide range of applications in the biomedical field in areas such as hydrogels, tissue engineering, and molecular recognition that rely on CLP melting transition to induce a change in phase, morphology, or assembly of the solution.[19, 24-27] Luo and Kiick[16] observed that ELP-CLP conjugates below the T_m of the CLP formed diblock structures where three ELP strands are connected to a rigid CLP triple helix. They observed a dramatic reduction in the LCST of the ELP strands in the ELP-CLP conjugate with respect to free ELP strands in solution. Chapter 2 of this thesis describes the utilization of a phenomenological coarse-grained model and MD simulations to present an understanding of the thermodynamics governing the differences in LCST between ELP-CLP conjugates and free ELP

strands for a given solution. The work presented in this Chapter 2 has been published as a peer-reviewed article in Soft Mater.[28]

While the coarse-grained model in Chapter 2 of this thesis mimics the ELP transition in free state and in ELP-CLP conjugated state, it is not designed to capture the melting behavior of CLP in ELP-CLP conjugates. While computational and experimental studies have elucidated the driving forces that govern the structure and thermal stability of CLP triple helices[7, 9, 19, 20, 22, 29], computational prediction of CLP T_m *a priori* from information about the CLP sequence and CLP length requires appropriate molecular models. Chapter 3 describes the development of an improved set of coarse-grained models for CLP that are capable of mimicking CLP triple helix melting with an increase in temperature, and qualitatively the correct trends in changing T_m with changing CLP design. Taking elements from a coarse-grained model for oligonucleic acids previously parameterized by Ghobadi and Jayaraman[30] that allows for specific and directional interactions between coarse-grained beads, we developed models for CLP that capture interstrand hydrogen bonding (H-bonding) necessary for stability of CLP triple helix formation[20]. Additionally, the impact of CLP strand length and CLP sequence is investigated by performing MD simulations with these newly developed coarse-grained models for CLP. The model for ELP-CLP conjugates described in Chapter 2 is then evaluated using the more recent CLP model for ELP-CLP conjugates that incorporates the possibility for CLP melting to determine if there are any qualitative/quantitative differences in LCST transition trends between the two models.

The third class of thermoresponsive polymers studied in this thesis are oligonucleic acids (ONAs). ONAs are short strands of nucleic acid repeat units that

are often conjugated to other polymers or nanoparticles for applications in drug and gene delivery, catalysis, electronics, and molecular recognition/sensing.[4-6, 31] Complementary ONA strands form double helices in solution that undergo a melting transition with increasing temperature. This melting behavior can be leveraged in the above applications by tuning the T_m of the ONA double helix through changing ONA length, ONA concentration, and ONA backbone chemistry.[32-38] Deoxyribonucleic acid (DNA) is the most commonly used ONA for ONA-polymer conjugates and has a melting transition that is well understood in scientific literature. One example of an DNA conjugate that has been recently synthesized is a DNA-star polymer conjugate which can form a self-assembled network formed from the hybridization of complementary DNA strands.[39, 40] However, DNA is expensive to synthesize and has a limited design space. Therefore, alternative backbone chemistries for ONA such as peptide nucleic acid (PNA)[38] and locked nucleic acid (LNA)[36, 37] strands have been synthesized that are different from DNA with respect to backbone charge and backbone flexibility.[41] In Chapter 4, coarse-grained MD simulation study of ONA-star polymer conjugates is presented to show how changing ONA chemistry through parameters such as ONA backbone flexibility and charge, impacts the melting and assembly of ONA-star polymer conjugates. The coarse-grained model for ONA-star polymer conjugates in Chapter 4 is based on the coarse-grained ONA model described in Ghobadi and Jayarman.[30] The work presented in Chapter 4 has been published as a peer-reviewed article in Soft Matter.[42]

Finally, in Chapter 5, I summarize the main conclusions of this thesis, highlighting the strengths and weaknesses of our coarse-grained approach, and present

a path of future work on coarse-grained model development and simulation of thermoresponsive polymers.

Chapter 2

COARSE-GRAINED SIMULATIONS OF THERMORESPONSIVE DIBLOCK CONJUGATES OF ELASTIN-LIKE POLYPEPTIDE (ELP) AND COLLAGEN-LIKE POLYPEPTIDE (CLP)

The text, figures, and tables in this chapter are adapted and/or reproduced, in part, from *Effect of conjugation on phase transitions in thermoresponsive polymers: an atomistic and coarse-grained simulation study* by Joshua E. Condon, Tyler B. Martin, and Arthi Jayaraman, published in *Soft Matter* volume 13, issue 16, pages 2907-2918, April 2017, with permission from the Royal Society of Chemistry.

2.1 Introduction

Thermoresponsive polymers are useful in a variety of applications such as chemical sensors, biological sensors, drug delivery, gene delivery, and tissue engineering because of their ability to undergo phase transitions and conformational changes with respect to a change in temperature.[43] A specific class of thermoresponsive polymers that undergo phase transitions or chain collapse with increasing temperature are lower critical solution temperature (LCST) polymers. The LCST phase behavior in solution of LCST polymers is thought to arise from a disruption in directional interactions, such as hydrogen bonding, between the polymer and solvent (in many cases water) as the temperature is increased.[44, 45] In one LCST polymer, poly(N-isopropylacrylamide) (PNIPAM), experimental[44] and computational[46-48] studies have suggested that with increasing temperature, chain collapse occurs as a result of a decrease in the number of polymer-water interactions,

and an increase in the number of polymer-polymer interactions. Further studies of PNIPAM in aqueous solutions have demonstrated that its LCST phase behavior involves a coil-to-globule chain collapse accompanied by aggregation as the temperature is raised above the LCST.[10, 44, 49, 50] Similar behavior was observed for poly[2-(dimethylamino)ethyl methacrylate] (PDMAEMA) in atomistic molecular dynamics simulations.[51]

A subclass of LCST polymers described above have physiologically relevant LCST transitions that make them useful in biomedical applications. [2, 43, 52] The focus of this chapter is on one of these polymers, elastin-like polypetide (ELP), which is nontoxic, biocompatible, and biodegradable allowing for its utilization in drug delivery and tissue engineering applications.[1, 3, 14] ELPs mimic the chemical structure of the natural polypeptide elastin and are polymers with the repeat unit Val-Pro-Gly-Xaa-Gly (VPGXG), where Val (V) is the amino acid valine, Pro (P) is the amino acid proline, Gly (G) is the amino acid glycine, and Xaa (X) can be any amino acid except proline. Similar to PNIPAM, ELP strands are dispersed in solution at temperatures below its LCST, and are aggregated at temperatures above its LCST.[13, 16, 53, 54] ELP LCST transitions are highly tunable with respect to ELP length, ELP concentration, salt concentration, solvent quality, and the amino acid in the X position of the VPGXG amino acid repeat unit sequence.[11-13, 54]

Conjugation of LCST polymer strands to macromolecules has been shown to lead to altered LCST behavior of these conjugates with respect to unconjugated “free” LCST polymer strands in solution.[1, 15-18, 55, 56] For example, Luo and Kiick[16] show that conjugating multiple ELP strands to a collagen-like polypeptide (CLP) triple helix reduces the LCST of ELP when compared to a solution of unconjugated ELP

strands of the same concentration, length, and composition. CLP triple helices are formed from the hybridization of three CLP strands which mimic the natural polypeptide collagen. CLP strands have the repeat sequence Gly-Xaa-Ybb (GXY) where Gly (G) is the amino acid glycine, and Xaa (X) and Ybb (Y) can be any amino acid though they are typically proline and hydroxyproline respectively. CLP triple helices undergo a melting transition with an increase in temperature, and this transition can be tuned with CLP length, the amino acids in the X and Y positions of its repeat sequence, and the solvent quality.[20, 21, 57, 58] Luo and Kiick showed that ELP-CLP conjugates, upon increasing the solution temperature above their LCST, self-assemble into bilayer vesicles which could be used for drug delivery applications.[16] Upon reaching the melting temperature of CLP, the CLP triple helices melt, and the ELP-CLP conjugates return to a dispersed state in solution. Understanding how the conjugation of ELP to other macromolecules such as CLP impacts the LCST of ELP in solution can inform the design of novel biocompatible thermoreponsive materials. In this chapter, we utilize coarse-grained molecular dynamic simulations to elucidate the thermodynamic driving forces governing the shift in ELP LCST transition upon conjugation to CLP.

In the past, atomistic molecular dynamics simulations of ELP have been performed by Zhao et al. that demonstrate the importance of changes in hydration of ELP strands with increasing temperature on the LCST phase behavior of ELP.[54] By characterizing the number of hydrogen bonds (H-bonds) between the ELP chain and water and comparing that value to the number of H-bonds between residues on the ELP chain, they showed that ELP strands have distinct hydration states corresponding to temperatures above and below the LCST of that ELP strand. Longer atomistic

molecular dynamics simulations performed by Jayaraman and coworkers[28] confirm this behavior for shorter ELP strands compared to the work of Zhao *et al.*[54] Furthermore, Jayaraman and coworkers also provide clustering algorithms that rigorously characterize the hydration states of ELP chains both above and below the LCST unlike the visual characterization of Zhao et al., Jayaraman and coworkers[28]use these simulations and the differences in LCST behavior between free ELP chains in solution and ELP chains tethered to a point to mimic the conjugation of ELP to a rigid macromolecule, and observed a decrease in LCST temperature in the tethered system with respect to the free system. Details about these atomistic molecular simulations and the conclusions drawn from their results can be found in Condon *et al.*[42] These atomistic studies highlight the importance of a decreasingly favorable polymer-solvent interaction with respect to the increase of temperature. Taking the understanding gained from the above atomistic molecular dynamics simulations, we develop a phenomenological coarse-grained (CG) model of ELP-CLP conjugates where ELP is represented as a bead-spring polymer in implicit solvent and the decrease in polymer-solvent interaction favorability is implicitly captured by increasing the strength of polymer-polymer interactions. Using the CG model, we simulate system sizes larger than what is feasible in atomistic simulations, and study how conjugating ELP to a rigid macromolecule such as CLP impacts the LCST-like assembly behavior marked by the onset of aggregation of the bead-spring polymer chains representing ELP. The local crowding of the model ELP chains upon conjugation to the model CLP rigid body (representative of a CLP triple helix below its melting temperature) shifts the onset of aggregation to lower ELP-ELP attraction strengths which is attributed to the reduction in the configurational entropic loss of the

ELP model polymer upon aggregation for the ELP-CLP conjugate system compared to the free ELP system. We also observe that in the absence of crowding, in a system where one ELP model strand is attached to one CLP model rigid body, the onset of aggregation shifts to higher ELP-ELP attraction strengths as compared to systems where multiple ELP model strands are conjugated to the CLP block which agrees with Luo and Kiick[16].

Chapter 2 is organized in the following manner. In Section 2.2, we detail the CG model of ELP-CLP conjugates, we present the simulation methods for studying aggregation of ELP model strands, we provide the analytical techniques used to characterize the onset of aggregation of ELP model strands, and finally, we present the parameters varied in this chapter. In Section 2.3, we present the results from the CG simulations of ELP-CLP conjugates. In Section 2.4, we conclude the chapter with a discussion of the implications of this work on thermoresponsive polymer conjugates.

2.2 Approach

We perform coarse-grained (CG) simulations of ELP-CLP conjugates at relevant length and time scales that are computationally unfeasible with an atomistic simulation framework. The phase transitions (coil-to-globule chain collapse[49, 50] and phase separation[13, 44, 59]) observed with regards to the LCST transition are mimicked with our CG framework through increasing the polymer-polymer pair-wise non-bonded interaction strength of the ELP model strand. This increase in pair-wise interaction strength captures the decreasingly favorable polymer-solvent interactions with respect to an increase in temperature in an ELP solution. The model, by being generic, coarsely represents ELP-CLP conjugates, and any diblock conjugate consisting of an LCST polymer block and a rigid macromolecule block.

2.2.1 Coarse-grained Model of ELP-CLP Conjugates

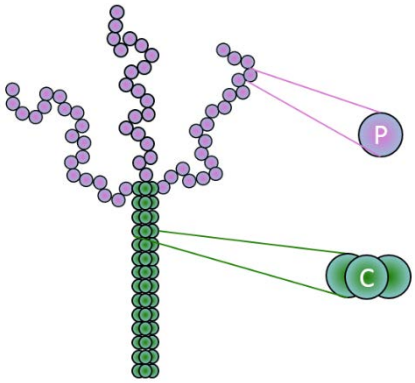

Figure 2.1: Schematic of the phenomenological coarse-grained (CG) model of three generic LCST polymer chains conjugated to a rod-like block in implicit solvent.

We model ELP strands as generic bead-spring polymers in an implicit solvent where the strands are composed of beads (denoted as P) of diameter 1.0σ and mass 1.0m with each bead representing 1-2 amino acids. The real values of σ and m can be calculated by directly mapping the CG bead to a specific group of atoms; given that we perform this study to capture the phenomena of LCST transition in generic diblock conjugate consisting of an LCST polymer block and a rigid macromolecule block we leave all distance and mass definitions in terms of σ and m. We model the CLP block as a rigid rod consisting of a triangular arrangement of three linear chains with beads (denoted as C) having diameter 1.0σ and mass 1.0m. The rigid rod behavior of the CLP model block mimics the stiff CLP triple-helix below the CLP melting temperature.[7, 9, 20, 23, 58, 60, 61] In Figure 2.1, we present a schematic of this ELP-CLP conjugate CG model.

The bonded interactions for C-C pairs are modeled with a rigid-body integrator defined in LAMMPS and are kept at a constant length of 1.0σ.[62, 63] The P-C and P-P bonded interactions are modeled with a harmonic potential defined in LAMMPS[62] with the equation

$$U_{bond}(r) = k_{bond}\left(r - r_{ij}^{0}\right)^{2} \tag{2.1}$$

where k_{bond} = 500ε is the bond constant and r_{ij}^{0} = 1.0σ is the equilibrium bond length between beads i and j. Similar to our comment above for σ and m, one can link ε to real energy units if we map the CG bead to a specific group of atoms. Keeping the model generic, we represent all energy values in terms of ε.

The non-bonded pair-wise P-P and C-C interactions are modeled with the 12-6 Lennard-Jones (LJ) potential[64]

$$U_{ij}(r) = 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r}\right)^{12} - \left(\frac{\sigma_{ij}}{r}\right)^{6}\right] \tag{2.2}$$

where ε_{ij} is the pair-wise interaction strength between beads i and j, and σ_{ij} is the arithmetic mean of the diameters of beads i and j. The interaction strength for P-P pair-wise interactions (ε_{PP}) is varied from 0.1ε to 1.2ε to mimic increasing hydrophobicity of the ELP model chain. We set the C-C interaction strength (ε_{CC}) to be constant at ε_{CC} = 0.1ε because we are primarily investigating the aggregation of ELP, hence we keep the C-C interaction strength low so the P-P interactions dominate in the system. All LJ interactions have a smooth cutoff at r = 2.5σ_{ij}. The P-C pairwise interaction potential is modeled with the Weeks-Chandler-Andersen (WCA) potential[65]

$$U_{ij}(r) = \begin{cases} 4\varepsilon_{PC}\left[\left(\frac{\sigma_{PC}}{r}\right)^{12} - \left(\frac{\sigma_{PC}}{r}\right)^{6}\right] + \varepsilon_{PC} & r < 2^{1/6}\sigma_{PC} \\ 0 & r \geq 2^{1/6}\sigma_{PC} \end{cases} \tag{2.3}$$

where the P-C interaction (ε_{PC}) is ε_{CC} = 1.0ε. The WCA potential is used, because experiments on ELP-CLP diblock conjugate self-assembled vesicles show that the ELP chains do not appear to interact with the CLP triple helices.[16]

2.2.2 Coarse-grained Simulation Details

We perform CG Langevin dynamics simulations in an NVT ensemble on the CG model of ELP-CLP conjugates from Section 2.2.1 using the LAMMPS[62] package. For free ELP systems, we randomly place ELP model chains in a periodic boundary condition cubic simulation box. For ELP-CLP conjugate systems, we randomly place ELP-CLP conjugate models in a periodic boundary condition cubic simulation box such that the number of ELP model strands in the ELP-CLP conjugate systems match the corresponding free ELP systems. Each simulation begins with a simulated annealing stage that begins at a reduced temperature of $T^* = 4.0$ and ends at $T^* = 1.0$ over 20 million timesteps where each timestep is $0.001(m\sigma^2/\varepsilon)^{0.5}$ in reduced units. The simulated annealing stage progresses in steps as described in Table 2.1. During annealing step 1, each simulation is initialized at $T^* = 4.0$ for 2 million timesteps before proceeding to step 2. At step 2, the temperature of the simulation is changed instantaneously to $T^* = 2.0$, and the simulation is then run for another 2 million timesteps. This process proceeds numerically through annealing steps 3 through 7 with the corresponding temperatures and timesteps for each step listed in Table 2.1.

Table 2.1: Simulation Protocol - Steps in the Annealing Schedule.

Step	Timesteps (in millions)	Temperature (T*)
1	2	4.0
2	2	2.0
3	2	1.5
4	3	1.25
5	3	1.1
6	4	1.01
7	4	1.0

After the simulated annealing stage, the system is simulated for 50 million timesteps at $T^* = 1.0$ and the last 10 million timesteps of this equilibration stage are used for analysis where data is taken every 100,000 timesteps. For every ε_{PP} value, we perform 5 independent simulations then calculate the ensemble averaged value of each of these simulations. The mean and standard deviation of these values taken from these 5 ensemble averaged values are then listed in the results section of this chapter.

2.2.3 Analyses for Onset of Aggregation of ELP Model Strands

To quantify chain aggregation, we calculate the ensemble average number of P-P pair-wise contacts (<# PP contacts>). For each P bead, we count the number of P beads within 2.5σ of its center, and take the ensemble average of this number for all

beads over all timesteps that are sampled. We take the averages from the 5 replicate simulations and report their mean as <# PP contacts>. We take the standard deviation directly from the averages of the 5 replicates. We plot <# PP contacts> as a function of ε_{PP} in the results section of this chapter with the mean and standard deviation represented by the symbol and error bar, respectively. To mark the ε_{PP} value at which chain aggregation occurs (ε_{Agg}), we fit a logistic curve to the <# PP contacts> versus ε_{PP} data series using non-linear regression:

$$y = \mathrm{b} + \frac{K-b}{1+A*e^{-k\varepsilon_{PP}}} \tag{2.4}$$

where y is <# PP contacts>, with K, b, A, and k being parameters that we fit using non-linear regression. The inflection point is calculated as $\left(ln\left(\frac{A}{k}\right), \frac{K+b}{2}\right)$ and in our simulations, y = <# PP contacts>, and $\varepsilon_{Agg} = \ln\left(\frac{A}{k}\right)$. For comparison, we also calculate ε_{Agg} by using the midpoint of a line between the minimum and maximum values of <# PP contacts>. Using linear interpolation, the ε_{PP} value where the midpoint located on the <# PP contacts> curve is recorded as ε_{Agg}.

2.2.4 Parameters Varied

For all simulations studied in this chapter, we vary ε_{PP} from 0.1ε to 1.2ε. For the first set of simulations, we study the differences between free ELP systems and ELP-CLP conjugates with the same volume fraction of ELP model strands. Table 2.2 summarizes the three systems studied in terms of number of free ELP model strands, number of rigid CLP blocks, and box size.

Table 2.2: Simulation details for the CG simulations. All ELP model strands in this table were 20 beads long, and all rigid CLP blocks were 14 beads long (see Figure 2.1 for a schematic of the model)

System	**Cubic simulation box length (in units of σ as defined in main paper)**	**Number of ELP model strands**	**Number of CLP rigid blocks**
Free ELP model strands	100	9	0
3 ELP model strands conjugated to rigid CLP block	100	9	3
1 ELP model strand conjugated to rigid CLP block	100	9	9

In all the above listed studies, we maintain the polymer chain length to be 20 P beads, and the CLP block length to be 14 C beads. To study the effect of the number of ELP model strands conjugated to the rigid CLP block on the onset of aggregation, we change the number of ELP chains on each CLP block from 3 to 1 while increasing the total number of ELP-CLP conjugates from 3 to 9, to maintain the same total number of polymer chains. To study the effect of length of the rigid CLP block on the onset of aggregation, we perform simulations with CLP block lengths of 1, 3, 7, and 14 (where length is defined as number of trimers of C beads). Lastly, to study the effect of CLP block flexibility on onset of aggregation, we do the following: In the system with 9 ELP-CLP conjugates, where each conjugate contains 1 ELP model strand conjugated to one rigid CLP block, we replace the CLP block with a flexible bead-spring polymer chain that has the same bonded interaction values as the ELP

model strand and the same non-bonded pair-wise interaction values as those in CLP blocks.

2.3 Results and Discussion

We present our findings from our CG molecular dynamic simulations of ELP-CLP conjugates below detailing how the onset of aggregation changes with respect to increasing P-P pair-wise interaction strength.

2.3.1 Impact of ELP-CLP Conjugation on ELP Onset of Aggregation

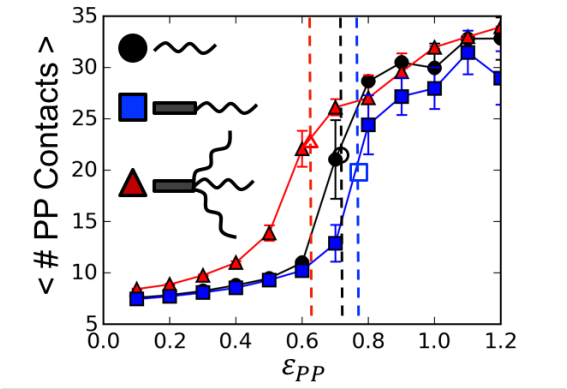

Figure 2.2: Average number of polymer-polymer (PP) contacts vs. ε_{PP}. The closed black circles represent the free ELP system with 9 ELP model chains with 20 ELP (P) beads per chain. The closed blue squares represent a system with 9 ELP-CLP conjugates, where in each conjugate, one 20mer ELP chain is conjugated to one CLP block 14 CLP (C) beads long. The closed red triangles represent a system with 3 ELP-CLP conjugates, where in each conjugate, three 20mer ELP chains are conjugated to one CLP block 14 C beads long. Simulation box length is 100σ in all three systems. The vertical dashed lines with open symbols denote the onset of aggregation as characterized by the logistic curve fit described in Section 2.2.3.

Table 2.3: Onset of aggregation transitions for free ELP model strands versus ELP model strand(s) conjugated to rigid CLP blocks

System	ε_{Agg} (logistic fit)	ε_{Agg} (linear interpolation)
9 free ELP model strands	0.706	0.709
3 ELP-CLP conjugates where each conjugate has 3 ELP model strands conjugated to 1 rigid CLP block	0.612	0.594
1 ELP-CLP conjugate where each conjugate has 1 ELP model strand conjugated to 1 rigid CLP block	0.760	0.766

In Figure 2.2, we present the ensemble average number of P-P contacts of the free ELP model strand system and the ELP-CLP conjugate systems, with increasing ε_{PP}, marking the onset of aggregation, ε_{Agg} with vertical dashed lines. Table 2.3 lists the values of ε_{Agg} for the systems presented in Figure 2.2. If we focus on the red triangles and black circles, we see that the onset of aggregation for the ELP-CLP conjugate system with 3 ELP model strands per rigid CLP block (red triangles) occurs at a lower ε_{PP} than in the free ELP model strand system (black circles). This result is consistent with the experimentally observed reduction in the LCST for ELP-CLP conjugates compared to free ELP in Luo and Kiick[16]. We also simulate systems of free ELP model strands and ELP-CLP conjugates with larger numbers of ELP model polymer chains while maintaining the same ELP volume fractions as the simulations described in Figure 2.2 and Table 2.3 to determine the effect of system size on our results. The larger system size simulations are described in Table 2.4, and the results are presented in Figure 2.3 and Table 2.5.

Table 2.4: Simulation details for the CG simulations at the same concentration but at higher number of ELP model strands than in the previous CG simulations described in Table 2.1. All ELP model strands in this table were 20 beads long, and all rigid CLP blocks were 14 beads long (see Figure 2.1 for the model)

System	Cubic simulation box length (in units of σ as defined in main paper)	Number of ELP model strands	Number of CLP rigid blocks
Free ELP model strands	208	81	0
3 ELP model strands conjugated to rigid CLP block	208	81	27

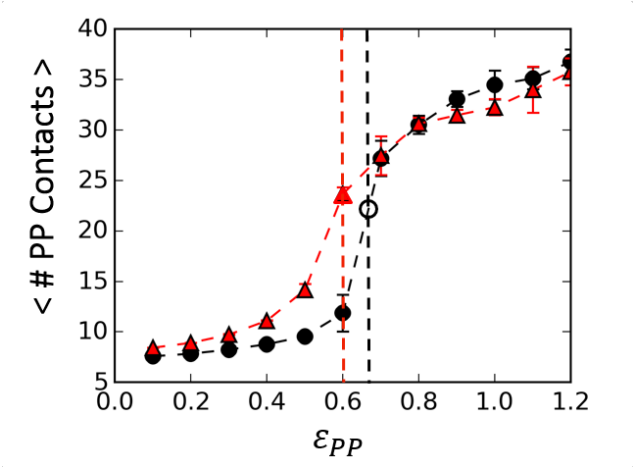

Figure 2.3: Average number of polymer-polymer contacts vs. ε_{PP}. The closed black circles represent the free ELP system with 81 ELP model chains with 20 ELP (P) beads per chain. The closed red triangles represent a system with 27 ELP-CLP conjugates, where in each conjugate, three 20mer ELP chains are conjugated to one CLP block 14 C beads long. Simulation box length is 208σ in all three systems. The vertical dashed lines with open symbols denote the onset of aggregation.

Table 2.5: Aggregation transitions for larger systems with 81 ELP model strands per simulation box.

System	ε_{Agg} (logistic fit)	ε_{Agg} (linear interpolation)
81 free ELP model strands	0.679	0.673
27 ELP-CLP conjugates where each conjugate has 3 ELP model strands conjugated to 1 rigid CLP block	0.601	0.588

We observe the same trend in the simulations with larger numbers of ELP model strands, namely the onset of aggregation for the ELP-CLP conjugate system occurs at a lower ε_{PP} than in the free ELP system, and conclude that system size has a negligible impact on the onset of aggregation for the systems that we study.

To understand why the ε_{Agg} value for the ELP-CLP conjugate system (3 ELP model strands per rigid block) is lower than the value for the corresponding free ELP conjugate system, we consider the enthalpic and entropic contributions to the change in free energy of the ELP model strands upon aggregation.

$$\Delta G_{Agg} = \Delta H_{Agg} - T\Delta S_{Agg} \tag{2.5}$$

For the ELP model strands to aggregate, the change in enthalpy upon aggregation (ΔH_{Agg}) and the change in entropy upon aggregation ($T\Delta S_{Agg}$) must be balanced such that the change in free energy upon aggregation (ΔG_{Agg}) ≤ 0 in Eq. 2.5. Since ε_{PP} is purely enthalpic in nature, the lower ε_{Agg} for the ELP-CLP conjugate system than the corresponding free ELP system suggests that a smaller ΔH_{Agg} is necessary to balance the $T\Delta S_{Agg}$. This further leads to our understanding that $T\Delta S_{Agg}$ for the ELP-CLP conjugate systems is smaller than the $T\Delta S_{Agg}$ for the free ELP systems. This agrees

with the hypothesis that conjugating multiple ELP strands to a CLP triple helix leads to increased crowding of the ELP strands, and thus decreases the loss in translational entropy upon polymer aggregation.

To further highlight that polymer crowding plays an important role in determining the onset of aggregation, we compare the 3 ELP model strand per rigid CLP block system to the 1 ELP model strand per rigid CLP block system shown in Figure 2.2 by the red triangles and blue squares, respectively. We observe that the 3 ELP model strand per rigid CLP block system has a lower ε_{Agg} than the corresponding 1 ELP model strand per rigid CLP block system. The latter system does not experience the same reduction in the loss of translational entropy upon polymer aggregation. The 1 ELP model strand per CLP block system also has a higher ε_{Agg} than the free ELP system which supports the understanding that increasing the length of the ELP chain without increasing the number of P-P interactions would increase the translational entropic loss upon aggregation.

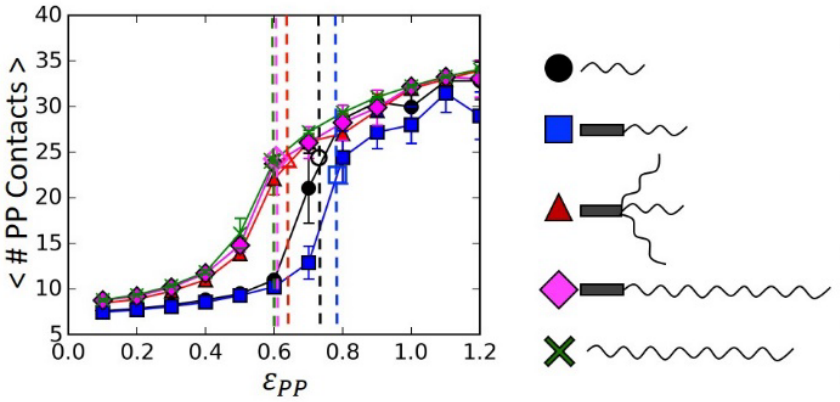

Figure 2.4: Average number of polymer-polymer contacts vs. ε_{PP}. The legend describes the systems presented in the plot. The vertical dashed lines denote the onset of aggregation also listed in two columns in the Table 2.6.

Table 2.6: Aggregation transitions listed for systems presented in Figure 2.4.

System	**ε_{Agg} (logistic fit)**	**ε_{Agg} (linear interpolation)**
9 20mer ELP model strands	0.706	0.709
9 ELP-CLP conjugates with 1 20mer ELP model strand per conjugate	0.760	0.766
3 ELP-CLP conjugates with 3 20mer ELP model strands per conjugate	0.612	0.594
3 ELP-CLP conjugates with 1 60mer ELP model strand per conjugate	0.593	0.577
3 60mer ELP model strands	0.581	0.579

Figure 2.4 lends additional support to the idea that increasing polymer bead crowding leads to decreasing loss in translational entropy of the polymer beads upon aggregation, and as a result, a lower energetic driving force (ε_{Agg}) for aggregation. All the systems in Figure 2.4 have the same total number of P beads and thus have the same total number of possible P-P contacts. First, we compare the system with three 20mer conjugated ELP model strands (red triangles) to the system with single 60mer conjugated ELP model strands (pink diamonds). Our simulations show that these two systems have a similar onset of aggregation. In both cases, the 60 LCST polymer (P) beads are crowded near the CLP block (i.e. the number of P beads effectively attached to each CLP block in the system is the same). Furthermore, the conjugation effect is similar, whether the three 20mer blocks are all tethered to a rod-like block (red triangles case) or sequentially tethered to a rod-like block and then each other (pink diamonds case). In both cases, the translational entropy of the monomers is restricted by being bound to one another and to the CLP block. Next, we compare the ELP-CLP conjugate system with one 60mer conjugated ELP model strand (pink diamonds in Figure 2.4) and the free ELP system with three 60mer ELP model strands (green crosses in Figure 2.4). Surprisingly, these two systems do not show a difference in ε_{Agg} while the analogous 20mer systems (blue squares and black circles) show a difference. We conjecture that the effect of the rigid CLP block on the ε_{Agg} diminishes as the polymer chain length increases (while the CLP block length remains constant) because the large number of P beads dominate the systems' $T\Delta S_{Agg}$ and the energetic driving force needed to balance that loss upon aggregation.

2.3.2 Impact of CLP Length on ELP-CLP Conjugate Onset of Aggregation

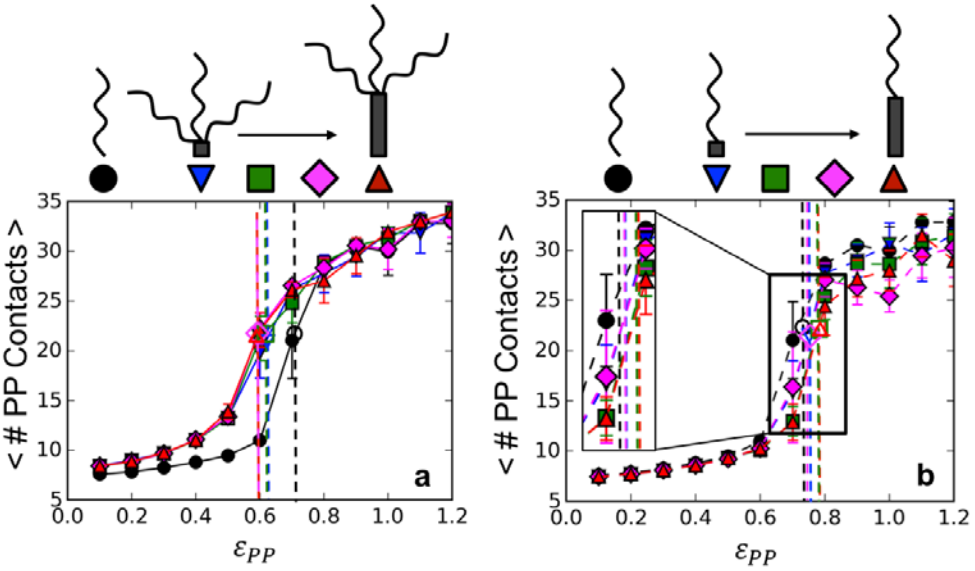

Figure 2.5: Effect of rigid CLP block length on average number of polymer-polymer contacts vs. ε_{PP}. (a) The closed black circles represent the free ELP system with 9 ELP model strands with 20 polymer (P) beads per chain. The closed blue upside-down triangles, green squares, magenta diamonds, and red triangles all represent a ELP-CLP conjugate system with 3 ELP-CLP conjugates, where the CLP block length is 1, 3, 7, and 14 CLP (C) beads, respectively, each with three 20mer ELP model strands conjugated to the CLP block. (b) Same as part a but with 9 ELP-CLP conjugates with one 20mer ELP model strand conjugated to each CLP block. The simulation box length is 100σ for all systems. The vertical dashed lines with open symbols denote the onset of aggregation.

Table 2.7: Aggregation transitions for varying the length of CLP in the ELP-CLP conjugate system where **three** ELP model strands are conjugated to each rigid CLP block

Length	**ε_{Agg} (logistic fit)**	**ε_{Agg} (linear interpolation)**
0	0.706	0.709
1	0.633	0.636
3	0.628	0.624
7	0.604	0.598
14	0.612	0.594

Table 2.8: Same as Table 2.7 but with **one** ELP model strand conjugated to each CLP block

Length	**ε_{Agg} (logistic fit)**	**ε_{Agg} (linear interpolation)**
0	0.706	0.709
1	0.737	0.748
3	0.757	0.760
7	0.736	0.733
14	0.760	0.765

To confirm that the local crowding of the ELP model strands rather than the rigid CLP blocks are causing the shift in ε_{Agg}, we vary the length of the CLP block in

Figure 2.5 and Tables 2.7 and 2.8. The results in Figure 2.5a along with Table 2.5, show that the ε_{Agg} is similar for all ELP-CLP conjugate systems where each conjugate has 3 ELP model strands attached to a CLP block, irrespective of the CLP block length. This indicates that the shift in ε_{Agg} is driven primarily by the local crowding of the ELP model strands on the grafting site. The results in Table 2.7 indicate that even at a CLP block length of 1, the shift in onset of aggregation is still apparent. This finding supports our earlier conclusion that local crowding is the dominant factor in shifting the onset of aggregation.

The results from Figure 2.5b along with Table 2.8, show that for ELP-CLP conjugate systems with 1 ELP model strand attached to each CLP block, for all CLP lengths we investigated, the ε_{Agg} is greater than the free ELP case. The trends in the inset graph in Figure 2.5b and Table 2.8 show that the onset of aggregation values decrease (with the exception of length 7) as the CLP block length is reduced. This trend is to be expected because by decreasing the CLP block length, we decrease the translational entropic loss with respect to the onset of aggregation.

2.3.3 Impact of CLP Flexibility on ELP-CLP Conjugate Onset of Aggregation

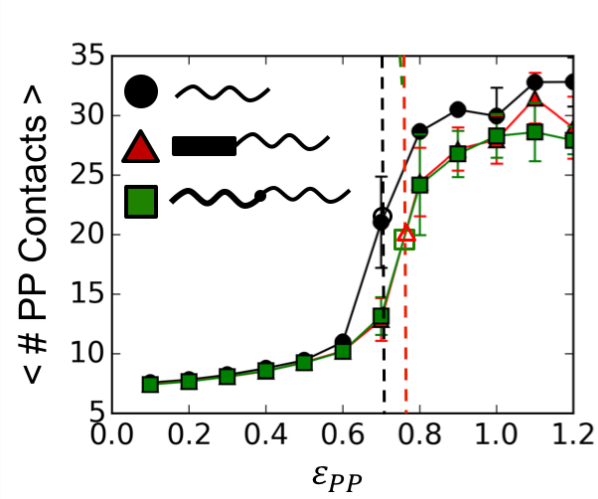

Figure 2.6: Average number of polymer-polymer contacts vs. ε_{PP} for the three systems depicted in the legend. In all simulations the total number of P beads is maintained same. Simulation box length is 100σ in all three systems. The vertical dashed lines with open symbols denote the onset of aggregation.

Table 2.9: Aggregation transitions for varying flexibility of CLP block to which the ELP model strand is conjugated

System	ε_{Agg} (logistic fit)	ε_{Agg} (linear interpolation)
9 Free ELP model strands	0.706	0.709
9 ELP-CLP conjugates with 1 ELP model strand per rigid CLP block	0.760	0.765
9 ELP-CLP conjugates with 1 ELP model strand per flexible CLP chain	0.753	0.752

To further understand the effect of the flexibility of the CLP block on $\varepsilon_{Agg,}$ we model the CLP block as a fully flexible linear polymer chain instead of a rigid body (Figure 2.6 and Table 2.9). This system of ELP-CLP conjugates with 1 ELP model strand conjugated to another fully flexible linear CLP block mimics the system of monomeric ELP-CLPs, i.e. single ELP chains attached to single (flexible) strands of CLP at temperatures above CLP triple helix melting temperature. In the experiments of Luo and Kiick, the monomeric state of ELP-CLPs had an LCST that was above the observable range of temperatures (i.e. LCST>80°C). In agreement with these experimental observations, in Figure 2.6, we find that the onset of aggregation ε_{Agg} for the ELP-CLP conjugate system where 1 ELP model strand is conjugated to a rigid CLP block (red triangle) or flexible CLP strand (green square) is higher than the onset of aggregation for free the corresponding free ELP system (black circle). We highlight that the green and red curves have significant overlap in Figure 2.6 which supports our hypothesis that it is the effective elongation of the ELP model strand with the additional block (rod-like or flexible) that increases the loss in configuration entropy of the diblocks, thus necessitating a larger energetic driving force for aggregation.

2.4 Conclusion

In this chapter, we performed coarse-grained simulations to understand recent experiments[15-17, 55] on conjugated thermoresponsive polymers that exhibit lower critical solution temperature behavior (LCST). In particular, Luo and Kiick[16] observed that conjugating three elastin-like polypeptides (ELP) to a collagen-like polypeptide (CLP) triple helix significantly reduced the LCST of the ELP-CLP conjugate with respect to a corresponding system of free ELP strands in solution. At higher temperatures upon melting of the CLP triple helix, the ELP-CLP conjugates

were found to once again become soluble. The reduction in the LCST was only found when multiple ELP strands one were conjugated CLP triple helix, thus increasing the crowding of the ELP strands. Motivated by these experiments, we have performed a coarse-grained (CG) molecular simulation study to better understand the molecular driving forces governing the effect of conjugation on the chain collapse and phase separation of ELP.

Atomistic molecular dynamics simulations have been performed to study ELP[28, 54], however a CG approach allows for the investigation of length and time scales relevant to aggregation of multiple ELP strands. These atomistic studies motivated our development of a phenomenological CG model, where we capture the LCST transition via varying the polymer-polymer attractive interaction strength, ε_{PP}. Using this model, we compared free ELP systems to systems with ELP model strands conjugated to a thermally-unresponsive, rigid CLP block which mimics ELP-CLP conjugates below the CLP melting temperature. In agreement with the experimental result, we observed a shift of transition ε_{PP} to lower values (i.e. smaller driving force needed to induce aggregation) for the ELP-CLP conjugate case as compared to the free ELP case. For the ELP-CLP conjugates, the increase in local concentration of ELP model strands due to its anchoring to a rigid CLP block likely decreases the polymer configurational entropy loss upon transition. This, in turn, reduces the enthalpic contributions needed to drive polymer aggregation and thus shifts the LCST to lower values in experiments. By reducing the length of the rigid CLP block to its minimum and still observing the shift in ε_{PP} compared to the free ELP system, we also confirmed that the local crowding of the ELP model strands due to anchoring of multiple ELP model strands in proximity is the primary cause for the shift in onset of

aggregation rather than the nature of the rigid CLP block. Lastly, we also confirmed that anchoring one ELP model strand to a flexible CLP strand, mimicking the monomeric ELP-CLP state at temperatures above the melting of CLP triple helix, shifts the onset of aggregation to slightly higher values than the free ELP system. This result matches the experimental observation of the ELP-CLP conjugates re-solubilizing above the CLP melting temperature.

We emphasize the generic nature of our ELP-CLP conjugate model that allows for it to predict trends in onset of aggregation for a variety of LCST polymers that are conjugated to rigid macromolecules. Our coarse-grained simulations presented in this chapter not only help explain the experimental results of Luo and Kiick, but also guide materials design for many synthetic (e.g. chemical sensors, heat-responsive adhesives) and biological applications (e.g. drug delivery vehicles, gene delivery carriers) where LCST polymers are often conjugated or tethered to other functional substrates/polymers.

Chapter 3

DEVELOPMENT OF AN IMPROVED CLP COARSE-GRAINED MODEL FOR STUDIES OF CLP MELTING AND DUAL-TRANSITIONS IN ELP-CLP DIBLOCK CONJUGATES

3.1 Introduction

Collagen-like peptides (CLPs) are polymer chains comprised of repeat units of amino acid triplets that often have the amino acid sequence Glycine (G)- Proline (P)-Hydroxyproline (O). Three CLP strands form a triple helix structures through inter-strand hydrogen bonding interactions.[19-22, 29] These CLP triple helices undergo a melting transition upon an increase in temperature; this meling transition can be tuned by the amino acid sequence, number of repeat units or length of CLP strands, and solvent quality.[19, 20, 57, 66, 67] The cell surface recognition/adhesion, structural, and fibrillar self-assembly of these CLP triple helices make them useful in many biologically relevant applications such as drug delivery[68-72], tissue engineering[24, 73-81], and nanoelectronics[82, 83]. For example, CLPs are useful components in hydrogels designed for tissue engineering to support the growth of cells by mimicking the extracellular matrix environment; this is due to the chemical similarity of CLPs to native collagen.[80] CLPs can also be tailored for biological applications where molecular recognition is key, through mutations of the canonical amino acids (P, O and G) or through addition of sequences of amino acids placed in the middle or at the ends of a canonical CLP strand. For example, Gly-Phe-Hyp-Gly-Glu-Arg is an integrin binding sequence that has been placed in the middle of two $(Gly\text{-}Pro\text{-}Hyp)_4$

sequences by Kiick and coworkers.[15, 16] Hartgerink and coworkers have demonstrated that CLP strands with modified ends having a $(Gly\text{-}Lys\text{-}Hyp)_4(Gly\text{-}Pro\text{-}Hyp)_4(Asp\text{-}Pro\text{-}Hyp)_4$ sequence form triple helices that assemble into fibrils which then form a hydrogel stabilized by electrostatic interactions between Lys and Asp.[84, 85] CLP strands are also conjugated to other thermoresponsive macromolecules (e.g. synthetic polymers[15, 80, 86] or polypeptides[16, 75, 87]) so that the resulting copolymer has structural and thermodynamic properties tailored for a specific application. For example, Luo and Kiick have achieved self-assembled vesicles for drug delivery through CLP conjugation to elastin like polypeptide (ELP) which exhibits inverse transition or lower critical solution transition behavior in water.[16] As described in Chapter 2, besides their use as delivery vehicles in the assembled state, this ELP-CLP conjugate is also fundamentally interesting as the conjugation of ELPs to an intact CLP triple helix reduces the LCST temperature of ELPs from that of unconjugated (or free) ELP in aqueous solution. Furthermore, upon melting of the attached CLP triple helix, the ELP's LCST shifts back to higher temperatures which in turn induces disassembly of the delivery vesicle held together by collapsed ELP, thus, releasing the cargo.[16] Therefore, one can envision designing vesicles that are functional near biologically relevant temperatures through tuning the ELP LCST and CLP melting transition temperatures.

Given that the melting behavior of CLP triple helices is critical to the stability and use of these biocompatible materials containing CLPs, it is important to understand how features of the CLP strand (e.g. sequence, length, mutations) impact melting temperature (T_m) of these triple helices *a priori*. Considerable effort has been placed on determining the structure of the triple helical region of collagen and CLP

triple helices, and how this structure affects CLP melting.[7, 9, 19-23, 29, 57, 88-92] CLP triple helices are right-handed helices of three $(XYG)_n$ strands in a left-handed polyproline II (PPII) type helical conformation where *n* is typically 6-12 repeats. These triple helices are stabilized by inter-strand 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 thus, exhibit a higher melting temperature than many other (XYG) combinations.[23] Additionally, increasing the length of $(POG)_n$ based triple helices improves the stability of the resulting triple helices by increasing the T_m.[20, 57, 66, 67] Another factor to consider in CLP triple helix stability is the network of water H-bonds that participate with the backbone and sidechains of the CLP triple helix strands. Diffraction studies have shown that water molecules are tightly bound to the surface of CLP triple helices in well-defined positions.[90] For example, the O-H hydrogen in hydroxyproline can form a H-bond with a water molecule, which in turn can form a hydrogen 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. Furthermore, 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.[84] 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.

Given that the CLP triple helix structural stability impacts the melting temperatures, both experimental[19, 20, 22, 57, 66, 89, 93-95] and computational[7, 9, 61, 96] methods have been used to show how CLP triple helix structure changes with respect to its sequence and its environment. For example, the length and radius of gyration for $(POG)_n$ triple helices for a range of CLP strand lengths in aqueous solution has been determined by small angle X-ray scattering experiments, and it is now known that CLP triple helices are not as rigid as once thought.[92] Circular dichroism experiments have determined the melting temperature of a limited selection of CLP triple helix sequences.[20, 57, 66, 67] Crystal structures of CLP triple helices obtained through X-ray diffraction experiments have also been instrumental in determining the structure of various CLP sequences and are useful in providing initial configurations for CLP atomistic simulations.[89, 90, 96, 97] However, observing melting behavior is still outside the practicality of atomistic simulations due to the length of simulation time required to observe melting in addition to the large number of atoms required for CLP and water in these simulations. Thus, the development of

coarse-grained models would be beneficial to the observation CLP triple helix melting on a reasonable simulation time scale for a range of CLP sequences and lengths. Another motivation of development of an improved coarse-grained model is to evaluate the validity of the results in the previous chapter where we used a rigid body CLP model[28]; this new model for CLP allows for a better understanding of ELP-CLP conjugates with respect to CLP melting.

In this chapter, we present two coarse-grained (CG) models: Model 1 and Model 2. We describe these in a pedagogical manner showing the limitations of the first model (Model 1) and the changes needed to improve the model to achieve the refined model (Model 2) for CLP. 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.[30] In these two phenomenological CG models, we capture the inter-strand H-bonding pattern experimentally observed in CLP triple helices[19] but do not 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. Using these 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 for 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 theory that repulsive electrostatic interactions of neighboring like charges in a CLP triple helix are destabilizing. Additionally, we compare our CLP coarse-grained models' end-to-end distance and diameter to that seen for CLP triple helices experimentally and computationally.[92, 96, 97] 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 1 does not correctly represent the aspect ratio or ratio of end-to-end distance to diameter of the experimental measurements of the CLP triple helix. Model 2 incorporates adjustments to bond lengths and dimensionless length of the model that attempt to match experimental length and diameter measurements. Lastly, we use Model 2 as an alternative model to represent CLP in ELP-CLP conjugates investigated in Condon *et al.*[28] We compare the results of our new ELP-CLP model against the results in Chapter 2 to ensure that the new ELP-CLP conjugate model captures similar qualitative behavior with respect to onset of aggregation of the ELP strands as was seen in Condon *et al*.[28]

3.2 Approach

We use molecular dynamics simulations with a phenomenological coarse-grained (CG) model inspired by the CG model developed by Ghobadi and Jayaraman[30] for oligomeric deoxyribonucleic acid (DNA). They describe the parameterization of a CG model with bonded and non-bonded interactions that mimic directional and specific H-bond interactions between donors and acceptors.[30] Ghobadi and Jayaraman have shown 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. Despite the success of that model for DNA melting, given the inherent differences between a DNA double helix and CLP triple helix, we modify that original DNA coarse-grained model extensively to represent CLP melting as a function of CLP design. We provide a detailed description of the development of this CLP coarse-grained model in two stages, Model 1 and Model 2, presenting the step by step process in improving its ability to capture melting behavior.

3.2.1 Coarse-grained CLP Model 1

In this model all values of length, mass, and energy in Model 1 are presented in reduced units of σ, m and ε which correspond to 0.3 nm, 42.0 g/mol, and 0.1 kcal/mol, respectively.[30] Each (XYG) triplet in the CLP sequence is represented by five interaction sites. Figure 3.1a 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. These length and mass parameters are set to approximately model the size and mass of individual amino acids where Gly has a molar mass of 75.07 g/mol, Pro has a molar mass of 115.13 g/mol, and Hyp has a molar mass of 131.13 g/mol. The length of these individual amino acids is approximately 0.4-0.5 nm, but can vary with solvent quality and temperature. The adjacent backbone beads in each triplet are connected through a harmonic bond potential defined in the LAMMPS[62] simulation package with bond length 0.84σ and prefactor 1000ε. Each hydrogen bonding bead is connected to its respective backbone bead with a harmonic bond potential with bond

length 0.37σ and prefactor 1000ε. In the case of CLP strands with a PKG triplet a 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. Figure 3.1b shows these types of triplet modifications modeled in this study. The lysine backbone (KB) bead has the same characteristics as the OB bead with the inclusion of a +1 charge valency. 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 valency 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 charge neutrality. A charge valency of +1 is used for positively charged ions and a charge valency of -1 is used for negatively charged ions. IN beads have a diameter of 0.7σ and a mass of 1.0m. The diameter and mass of the IN beads are taken directly from Ghobadi and Jayaraman to model monovalent ions and are not rescaled.[30]

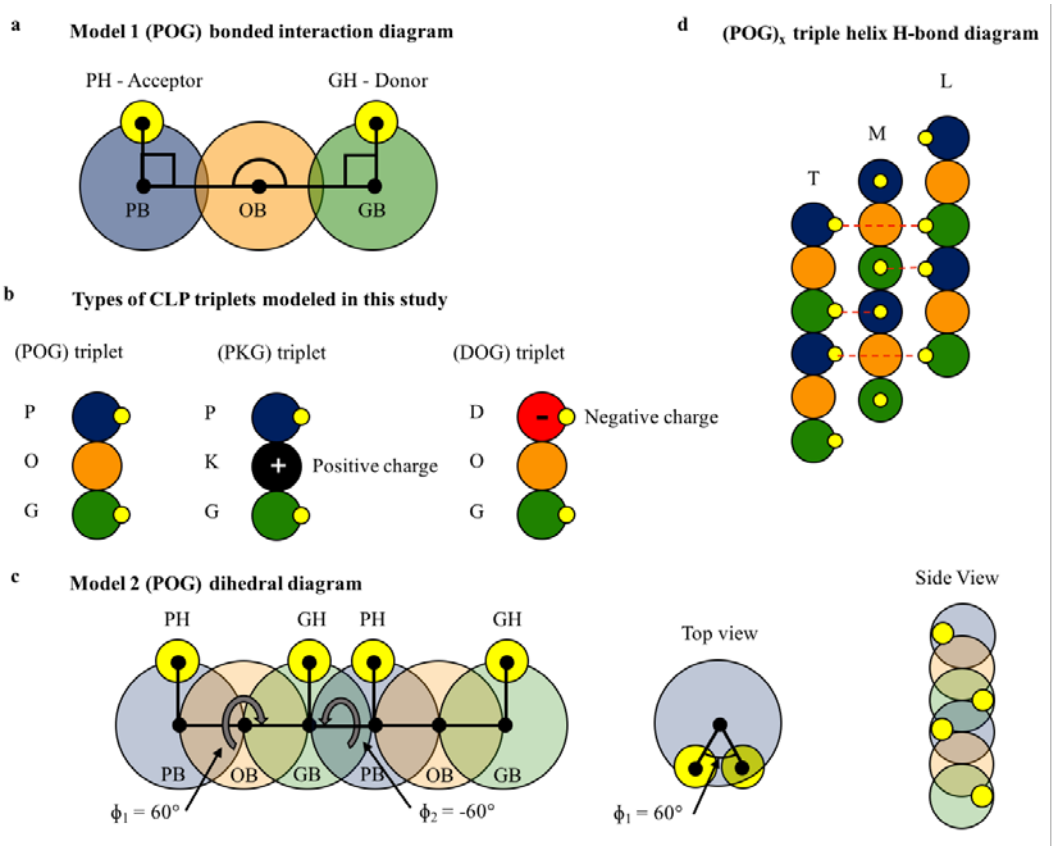

Figure 3.1: (a) Diagram of Model 1 coarse-grained (POG) triplet with bonded interactions shown by lines and angles connecting specific bead types. (b) Types of CLP triplets used in the simulations studied in this work. (c) Diagram of Model 2 coarse-grained CLP strand with the dihedral angles listed. (d) $(POG)_x$ 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.

Each CLP strand is a chain of XYG triplets where the G backbone bead of one triplet is bonded to the X backbone bead of the adjacent triplet via a harmonic bond potential with the parameters described above for bonded interaction between two adjacent backbone beads. 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 the angle constant $k_a^{BB} = 10\varepsilon/rad^2$. We note that the main phenomena we intend to mimic with this CG model are the trends in thermal transitions with varying CLP design, and not the structure. In this regard, the goals of our CLP CG model are similar to what was accomplished with the ONA model of Ghobadi and Jayaraman that quantitatively reproduce the trends of DNA melting with DNA strand design without modeling structural details such as helical twist. In the same vain, the backbone beads of individual CLP strands in this CLP model are not parameterized 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 isotropic bonded and non-bonded interactions involving the HB beads and their "parent" BB beads. First, we have a hydrogen bond-backbone bead-backbone bead angle potential with the angle set at 90° and the angle constant $k_a^{HB} = 300\varepsilon/rad^2$. 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 HB bead, XB is the X backbone bead, and YB is the Y backbone bead. The HB angle potential in the G position is defined as GH-GB-YB. Dihedral bonded interactions are not present in Model 1.

Non-bonded interactions between H-bond donors (D) and H-bond acceptors (A) (e.g. GH-PH interactions) are represented using the Lennard-Jones (LJ) potential as a function of interparticle distance (r)[98]

$$U_{D-A}^{HB}(r) = \varepsilon_{D-A}^{HB}\left[\left(\frac{\sigma^{HB}}{r}\right)^{12} - \left(\frac{\sigma^{HB}}{r}\right)^{6}\right] \tag{3.1}$$

in which ε_{D-A}^{HB} is the strength of interaction between the donor beads and the acceptor beads and has the value $\varepsilon_{D-A}^{HB} = 50\varepsilon$. $\sigma^{HB}$is the diameter of the hydrogen bonding beads. We use a force and potential cutoff of $1.9\sigma^{HB}$with a switching function from GROMACS[99] that smoothly ramps both to zero at $2.0\sigma^{HB}$. The potential for all other pairwise interactions besides the HB donor-acceptor interactions are purely repulsive Weeks-Chandler-Andersen (WCA) potential[65]

$$U_{ij}(r) = \begin{cases} 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r}\right)^{12} - \left(\frac{\sigma_{ij}}{r}\right)^{6}\right] + \varepsilon_{ij} & r < 2^{1/6}\sigma_{ij} \\ 0 & r \geq 2^{1/6}\sigma_{ij} \end{cases} \tag{3.2}$$

in which ε_{ij} = 1.0ε 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_{ii} 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. We use this pairwise interaction potential to simulate CLP strands in which the implicit solvent in this model has the same quality for all beads in the CLP model. Besides the above non-bonded potentials, the interactions between a backbone bead and its third neighbor (1-3 interaction) and fourth neighbor (1-4) pairwise interactions as defined in LAMMPS[62] are preserved in this model so that intra-chain bead overlap is prevented. All pairwise interaction values are listed in Table 3.1.

We note that the directionality of the hydrogen bond donor-acceptor attractive interactions arises from a combination of the repulsive nature of all pair-wise interactions besides donor-acceptor, the hydrogen bond bead-backbone bead-backbone bead angle potential and the relative sizes of the backbone bead and hydrogen bond

beads. Through this combination and the location of the hydrogen bond bead being within the repulsive shell of the parent backbone bead, there is only a small geometrical region extending from the hydrogen bond bead where hydrogen bond interactions occur.

Table 3.1: Parameters of the Lennard Jones potential* representing the non-bonded interactions described in Section 3.2 for the hydrogen 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*
HB (Acceptor)	HB (Acceptor)	WCA	1.0	0.7	0.786
HB (Donor)	HB (Donor)	WCA	1.0	0.7	0.786

** Potential and force smoothly ramped to zero by secondary cutoff of 0.600σ*

Lastly, for all electrostatic interactions among ionized beads (KB, DB, and IN), we use the Coloumb potential defined in LAMMPS[62]

$$U_{ij}^{coul}(r) = \frac{q_i q_j}{4\pi r \varepsilon_0 \varepsilon_r} \tag{3.3}$$

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 $\varepsilon_r = 80$ to be similar to that for water at room temperature.[100] The 1-3 and 1-4 electrostatic interactions in this model are set to zero to prevent strong intra-strand electrostatic repulsive or attractive forces from destabilizing the CLP strands.

3.2.2 Coarse-grained CLP Model 2

While Model 1 is an excellent first step towards 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 (POG)*n* with varying *n*, as we discuss in the results section. Additionally, there are other observations in the simulations (e.g. non-helical hybridized conformations) as described in further detail in the results section that motivated us improve this model. Model 2 retains many of the key aspects of Model 1 with a few notable differences. In Model 2, all values of length, mass, and energy are presented in reduced units of σ, m and ε which correspond to 0.5 nm, 42.0 g/mol, and 0.1 kcal/mol, respectively; the only difference between Model 2 and Model 1 is the value of σ. In conjunction with changing the unit of length in this model, we reduce the harmonic backbone bead-backbone bead bond length to 0.50σ. Together, these changes allow for a better match of the length to diameter ratio to that seen in experimental measurements of CLP. While Model 1 captures the directionality of hydrogen bond beads with respect to the backbone of CLP strands, it does not capture directionality of hydrogen bond beads with respect to neighboring hydrogen bond beads along the same strand. The directionality of sequential inter-strand H-bonds gives rise to specific CLP triple helix twist and pitch both of which are dependent on the CLP sequence of interest.[19] While neither model is intended to capture helicity, Model 2 does account for the directionality of hydrogen bond beads with respect to neighboring hydrogen bond beads along the same strand by introducing two hydrogen bond bead- backbone bead -adjacent backbone bead- hydrogen bond bead dihedral angle potentials defined in LAMMPS[62] as

$$U^{dih}(\phi) = k_d(1 + \cos(\phi - \phi_0)) \tag{3.4}$$

in which the dihedral constant $k_d = 15\varepsilon$ 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. Using a chain of (POG) triplets as an example, the angle for the first (PH-PB-GB-GH) dihedral is $\phi_0 = 120°$ where the reference plane is defined by the PH-PB-GB angle, and the angle for the second (GH-GB-PB-PH) dihedral $\phi_0 = -120°$ 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 $\phi_0 = 120°$ corresponds to all XH-XB-GB-GH dihedrals, and $\phi_0 = -120°$ corresponds to all GH-GB-XB-XH dihedrals. Figure 3.1c shows a Model 2 $(POG)_2$ CLP strand that demonstrates the above defined dihedral angles and provides the top and side view of this strand. Through this definition of dihedral angles, we achieve a staggered, triangular inter-strand H-bond network as shown in Figure 3.1d. The harmonic backbone beads angle potential in Model 2 has an angle constant of $k_a^{BB} = 20\varepsilon/rad^2$ compared to that of $10\varepsilon/rad^2$ in Model 1 to increase the stiffness of CLP triple helices with Model 2 compared to Model 1. All other bonded and non-bonded interaction parameters described in Section 3.2.1 remain the same for Model 2 with respect to the dimensionless units.

3.2.3 ELP-CLP Conjugate Coarse-grained Model

Building on the CG model of ELP-CLP from our previous work[28], in this work we replace the rigid body CLP portion of the previous model, where CLP is approximated as a collective rod of three CLP strands, with the current Model 2 described in Section 3.2.2. Each ELP-CLP conjugate strand consists of a CLP strand with all the same units, bonded and non-bonded 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 where the bond length is 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 similar to what is used in Chapter 2 and Condon *et al.*[28] where each ELP bead (EB) has a diameter of 1.0 and a mass of 1.0*m*, with σ defined as in Model 2 . Each of these EB beads could roughly represent a collection of atoms in a 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.

The atomistic simulations in our previous work showed that with increasing temperature the ELP chain undergoes reduction in hydration, quantified by the number of water molecules in the first shell around ELP and number of ELP-water H-bonds.[28] We mimic that decreasing solvophilicity or increasing solvophobicity implicitly using an effective attraction between non-bonded EB beads with the Lennard-Jones potential[98]

$$U^{EB}(r) = \varepsilon^{EB}\left[\left(\frac{\sigma^{EB}}{r}\right)^{12} - \left(\frac{\sigma^{EB}}{r}\right)^{6}\right] \tag{3.5}$$

where ε^{EB} is the effective attractive strength between EB beads that is gradually varied to mimic increasing temperature effects, $\sigma^{EB} = 1.0\sigma$, 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 non-bonded pairwise interactions of EB beads with any other bead in the CLP model are modeled using the WCA potential[65] (equation 3.2)**.**

3.2.4 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 periodic boundary condition cubic simulation box with sides of length 63.5σ to arrive at a concentration of ~2 mM of CLP triple helices. In the CG Model 2 simulations, we randomly place 10 CLP triple helices in a periodic boundary condition cubic simulation box with side lengths of 110σ such that the resulting concentration of CLP triple helices is ~0.1mM which is much closer to experimental conditions.[101] The triple helices are formed by placing three CLP strands in a triangular fashion lengthwise as detailed in Section 3.2.1 such that all possible H-bonds detailed in Figure 3.1d are formed and each strand is offset from one another by 1 bead similar to what is observed experimentally.[19] 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 experimental ELP-CLP conjugate concentration of 1mg/mL in Luo and Kiick.[16] 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 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 charge neutrality. We also insert an additional 1mM of monovalent salt to mimic a solution with salts added.

We use Langevin dynamics in the NVT ensemble as implemented within the LAMMPS[62] package. In all simulations, the friction coefficient is set to 10τ (where 0.001τ is approximately 6fs) as in Ghobadi and Jayaraman[30]. This was originally chosen to correspond to the diffusion of oligomeric DNA in water; since our focus is

on equilibrium results and not dynamics we do not optimize this friction coefficient in our simulation. We use a two-level RESPA[102] integrator for all simulations in which non-bonded interactions are integrated with a timestep of 0.001τ and bonded interactions are integrated with a timestep of 0.0005τ. In the case of charged systems, we use the particle-particle-particle-mesh (PPPM)[103] 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 timestep equilibration run followed by a 10^7 timestep 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 only in the random placement of the CLP helices in the box) and velocities.

3.2.5 Analyses

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 melting curves for DNA melting; here we choose 75% instead of 50% as done in DNA melting as it is possible that 2 strands can form a hybridized structure with almost two-thirds of the possible H-bonds in a triple helix formed without even being bonded to

the third strand. The standard deviation of $f_{75\%}$ is reported in the error bars in Section 3.3.

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. [96, 97] The ensemble average end-to-end distance and standard deviation is calculated by taking the average length of the vectors spanning the terminal residues of individual strands within a given helix followed by an average of this value and standard deviation over each strand for all simulation frames collected. The $\langle R_{EE} \rangle$ is computed from the collection of all simulation frames in all three replicates. The ensemble average diameter and standard deviation is calculated by taking 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 and taking the average and standard deviation over all simulation frames for three replicates for every triplet. 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 dependent 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 pair-wise contacts $\langle N_{EB\text{-}EB} \rangle$ with respect to the interaction strength between EB beads (ε^{EB}). This way a low value of $\langle N_{EB\text{-}EB} \rangle$ corresponds to a solvated ELP-like polymer and increasing value of $\langle N_{EB\text{-}EB} \rangle$ corresponds to decreasing hydration of the ELP-like polymer. For each

frame, $N_{EB\text{-}EB}$ 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 $N_{EP\text{-}EP}$ 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 $\langle N_{EB\text{-}EB} \rangle$ vs. ε^{EB} encounters an inflection point.

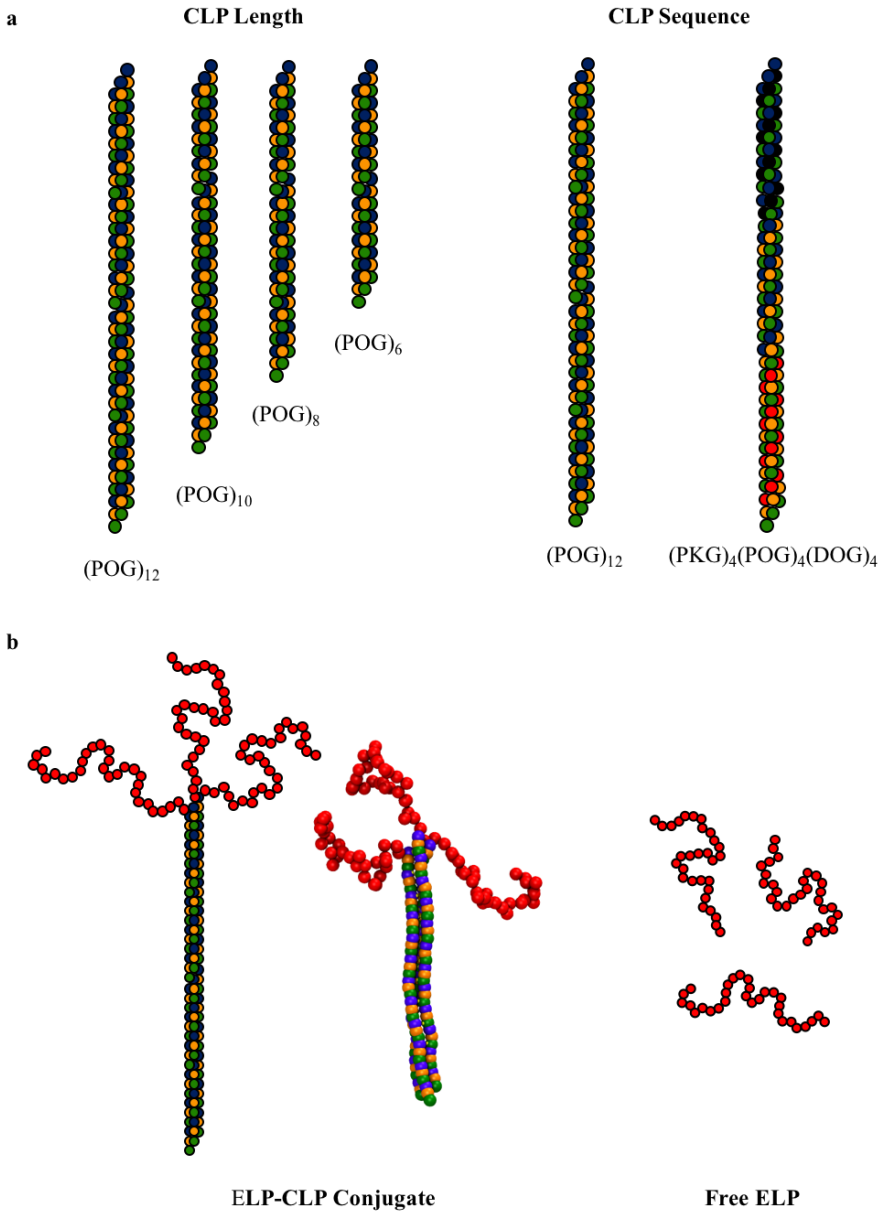

Figure 3.2: (a) Diagram of the systems studied to ascertain trends in CLP length and CLP modified sequence. (b) Visual rendering of the ELP-CLP conjugate model from simulation snapshot and diagram of the systems studied to determine how ELP aggregation is affected by conjugation to CLP.

3.2.6 Parameters Varied

To investigate whether either of the coarse-grained models (Model 1 and Model 2) capture the increase in CLP triple helix T_m observed experimentally[66] with increasing numbers of (POG) triplets in CLP chains comprising a CLP triple helix, we perform coarse-grained simulations of $(POG)_n$ where n = 6, 8, 10 and 12. To test the observation that T_m of CLP triple helices decreases with the double mutations (O to K and P to D) at both ends of CLP strands, we perform coarse-grained 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 3.2a we present visual representations of the different CLP helices studied in this work.

We also investigate whether modeling ELP-CLP conjugates with the CLP portion of the model as 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.[28] and experimentally in Luo and Kiick[16]. We perform a set of simulations for the ELP-CLP conjugate system as described by Section 3.2.4 with 10 ELP-CLP triple helix conjugates where the CLP triple helix is 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 where the ELP sequence used in the ELP-CLP conjugates was $(VPGFG)_6$ where V represents valine, P represents proline, G represents glycine and F represents phenylalanine. We also perform simulations of the free ELP system with 30 ELP strands also corresponding to 30 amino acids each to maintain the same volume fraction of ELP beads between

simulations. Snapshots of one of the ELP-CLP triple helices and one of the free ELP chains are shown in Figure 3.2b. The ELP-CLP conjugate and free ELP simulations are performed at a range of ε^{EB} values between $\varepsilon^{EB} = 0.1\varepsilon$ and $\varepsilon^{EB} = 1.3\varepsilon$ at a constant reduced temperature of T^{*}=2.0 that corresponds to a hybridized CLP triple helix.

3.3 Results and Discussion

3.3.1 Effect of CLP Length on Melting Transition Temperature

First, we explore how the number of repeat units impact the melting transition of $(POG)_n$ triple helices.

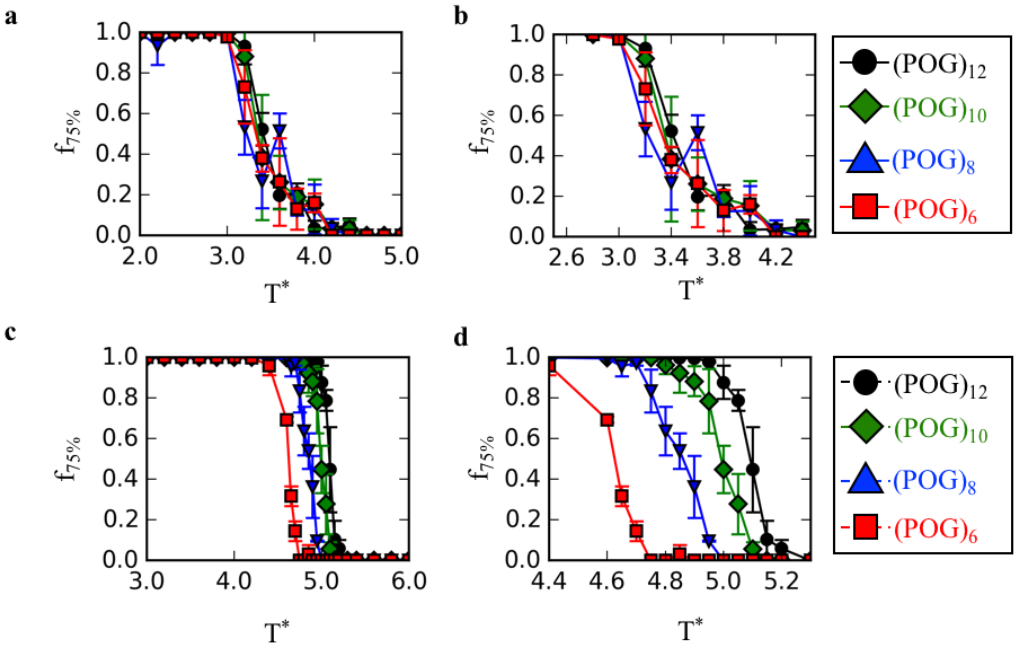

Figure 3.3: Ensemble average fraction of CLP triple helices that have at least 75% of possible H-bonds formed ($f_{75\%}$) as a function of reduced temperature (T^*) at a concentration of 0.1mM for $(POG)_n$ where n is equal to 12 (black circles), 10 (green diamonds), 8 (blue triangles), and 6 (red squares) using Model 1 (a and zoomed in version in b) and Model 2 (c and zoomed in version in d).

Figure 3.4: Simulation snapshot of Model 1 $(POG)_{12}$ system at T^{*}=3.6 that shows a non-helical hybridized structure

In Figure 3.3, we present the melting transition data for the simulations of $(POG)_{12}$, $(POG)_{10}$, $(POG)_{8}$, and $(POG)_{6}$ triple helix systems for both Model 1 (Figure 3.3a with zoomed in version in part b) and Model 2 (Figure 3.4c with a zoomed in version in part d). Based on circular dichroism melting temperature data[66], 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 3.3a and b), we

do not see a clear trend in melting temperature with respect to number of (POG) repeat units per CLP strand. This failure of Model 1 to capture the differences in melting temperatures with increase *n* in $(POG)_n$ can be 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 (say, ΔU_{hyb}) to balance out the entropic loss of the CLP single strands to be in a hybridized intact helix (say, ΔS_{hyb}) Therefore, for melting transition to increase with CLP strand length there should be an increase in the ΔU_{hyb} or 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, we find that the lack of stiffness (or restricted conformations) in the CLP strands leads to formation of non-triple helical hybridized structures as shown in Figure 3.4. Such structures 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 3.2.2 by introduction of dihedral angle potentials and more restricted angle potential that ensure the desired directionality of the H-bonds within the intact triple helix and increase stiffness of the CLP strands.

With Model 2, we see a distinct increase in the melting temperature transition with increase in CLP length (or increasing *n* in $(POG)_n$) confirming that the inclusion of the desired directionality in H-bonds described in Section 3.2.3 for Model 2 and the stiffening of the backbone are both important in capturing the experimental 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.[23] One may see this as analogous 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[30]; this is essential to capture the correct trends in DNA melting temperatures. Also, with Model 2, we do not see non-triple 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 reduced temperature 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*).

3.3.2 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.

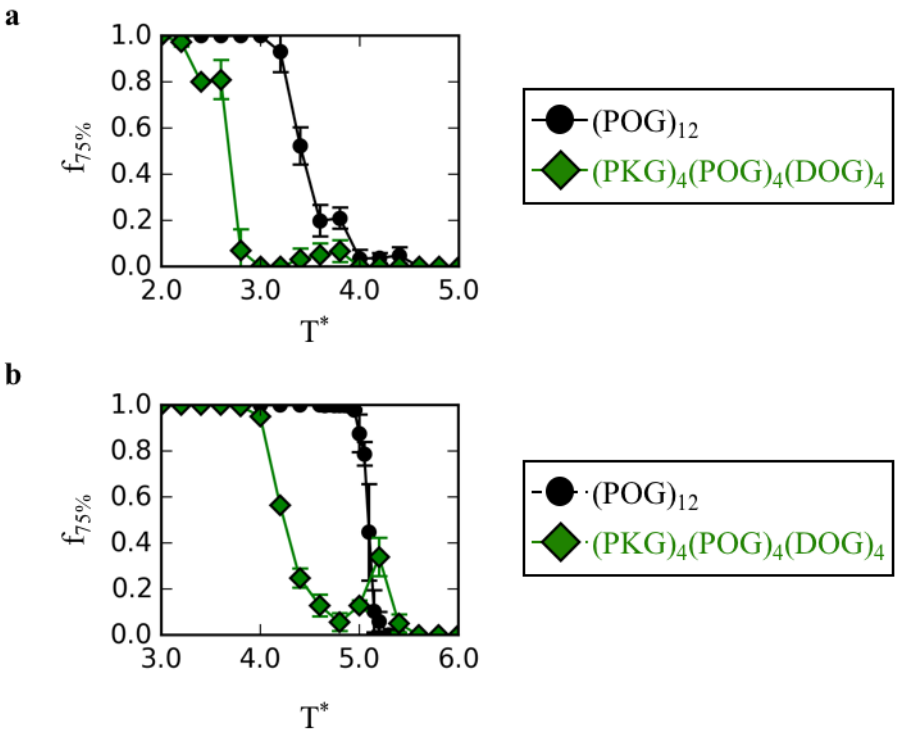

Figure 3.5: Fraction of CLP triple helices that have at least 75% of possible H-bonds formed ($f_{75\%}$) as a function of reduced temperature (T^*) at a concentration of 0.1mM for Model 1 (a) and Model 2 (b) for the sequence $(POG)_{12}$ (black circles) and the double mutation $(PKG)_4(POG)_4(DOG)_4$ (green diamonds).

Figure 3.6: Simulation snapshot of the $(PKG)_4(POG)_4(DOG)_4$ system for Model 2. Non-triple helical hybridized structures can form in such a way that they are stabilized by attractive electrostatic interactions.

In Figure 3.5a and Figure 3.5b, 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 experiments[84] 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 non-zero $f_{75\%}$ behavior of the $(PKG)_4(POG)_4(DOG)_4$ for Model 2 at higher temperatures is attributed to non-helical hybridized states. Favorable electrostatic interactions allow for rings of CLP strands to form as can be seen in Figure 3.6 and are seen in simulation snapshots ubiquitously at temperatures above the initial drop in $f_{75\%}$. For example, at $T^* = 4.2$ all of the strands in the simulation participate in H-bonds with one another, however very few have a triple helical

conformation and most participate in rings of CLP strands. At higher temperatures (T^* = 6.0), most strands are completely melted and do not form 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.

3.3.3 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 work[28] where CLP was modeled as a stiff rod in ELP-CLP conjugates. Given that the relative lengths of ELP and CLP blocks and CLP triple helix diameter could impact the crowding of the ELP chains and as a result, the onset of aggregation, we compare the end-to-end distance and diameter of the coarse-grained simulations of a single CLP triple helix to experimental values of maximum CLP length and diameter.

Table 3.2: 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

In Table 3.2, we show the average end-to-end distance for both Models as described in Section 3.2 for coarse-grained simulations of CLP triple helices at T^*=3.0. Setting σ to 0.3nm for Model 1 and σ to 0.5nm for Model 2 allows us to directly compare the two sets of simulations with experimental values for length and

diameter of CLP triple helices. Small angle X-ray scattering experiments list the length of $(POG)_{12}$ helices in solution to be 8.9nm.[92] X-ray diffraction data for the crystal structure of a similar CLP sequence $(POG)_4(POA)_1(POG)_5$ give a diameter of the CLP triple helix to be approximately 1nm with a resolution of 0.19nm.[104] Past atomistic simulations report the end to end distance of a $(POG)_9$ triple helix to be 7.24nm [96] or 7.97nm [97]with a diameter of 0.57nm [97]. If one wants to reproduce structural aspects of the CLP triple helix quantitatively right, then the unit of length and the backbone bond length in Model 2 could also be further adjusted to arrive at an accurate $(POG)_{12}$ aspect ratio based on experimental results and atomistic simulations. 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 ($<R_g^2>^{0.5}$) of a 30mer ELP-like polymer equal to 2.54 +/- 1.50σ at $\varepsilon^{EB} = 0.1\varepsilon$. Our past work[28], 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.

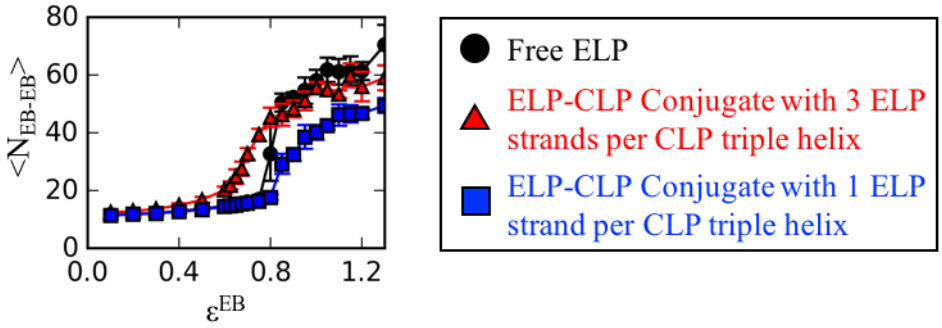

Figure 3.7: Ensemble average number of EB-EB contacts ($<N_{EB-EB}>$) vs. the strength of attractive pairwise interactions among EB beads (ε^{EB}) for both free ELP simulations (black circles) with 30mer ELP strands and for ELP-CLP conjugate simulations (red squares) with $(POG)_{12}$ CLP strands and three 30mer ELP strands per CLP helix at a reduced temperature of $T^* = 2.0$ and a concentration of ~1mg/mL. Note that $<N_{EB-EB}>$ is not 0 at low ε^{EB} values because the method of calculating this value includes nearest neighbors, in addition to the strands being fully flexible.

In Figure 3.7, we present ensemble average number of EB-EB contacts ($<N_{EB-EB}>$) vs. the strength of attractive pairwise 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 two curves, we see that the onset of aggregation is at a lower value 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 results[28] which was in agreement with experiments[16] in which the LCST of ELP-CLP conjugates is lowered by greater than 80 degrees Celsius compared to free ELP in solution. We showed in our previous CG molecular simulation results 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[28].

While one may think that given the remarkable agreement between the results using our more sophisticated CLP model (Model 2) and the rigid body model for CLP, it is sufficient to use the latter if one wants to study ELP-CLP conjugates. That may be the case for studying ELP-CLP systems at temperatures below the CLP melting temperature where CLP is an intact triple helix. However, if one intends 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 useful as it captures CLP helix melting. Also, if we wish 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 more valuable. These are future directions for this improved Model 2 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.

3.4 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 bonding (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. However, Model 1 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. Lastly, to demonstrate that the developed CLP models are extendable 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.*[28] where the onset of aggregation occurs at a lower ELP-ELP bead pair-wise interaction strength than for free ELP systems.

It is essential that we highlight some aspects of CLP Model 2 that can be further improved. First, with the implicit solvent representation, the water bridging stabilizing H-bond interactions that have been hypothesized to stabilize the CLP triple helix that are not captured explicitly. For example, experimental understanding of the atomistic structures of CLP triple helix show that the hydroxyl hydrogen on the

hydroxyproline side chain of $(POG)_x$ forms hydrogen bonds with water molecules which in turn form hydrogen bonds with a carboxyl oxygens of a neighboring CLP strand's backbone proline.[90] Future CG models of CLP could include explicit solvent beads in addition to CLP H-bond beads that are not represented in the current model that have the potential to capture these stabilizing interactions. However, the introduction of explicit incorporation of water-CLP interactions could significantly slow down the simulation and restrict length and time scales of such studies of CLP triple helix. Second, one could adjust the bond length between backbone beads and the dimensionless length of Model 2 to better quantitatively match the aspect ratio of the triple helices between the coarse-grained simulations and experiments. Third, to address the issue of the non-helical hybridized states of $(PKG)_4(POG)_4(DOG)_4$ triple helices at temperatures above T_m, one could increase the angle constant for the backbone beads to further stiffen the CLP strands and discourage the formation of these folded non-helical states. Lastly, given our focus on thermodynamics and not structure, neither model presented here differentiates between amino acids (e.g. proline vs. valine) that favor or do not favor the PPII helical conformation of individual strands; PPII helical conformation is necessary for the formation of a true triple helix This could be achieved 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.

Chapter 4

EFFECT OF OLIGONUCLEIC ACID (ONA) DESIGN ON ASSEMBLY OF ONA-STAR POLYMER CONJUGATES: A COARSE-GRAINED MOLECULAR SIMULATION STUDY

The text, figures, and tables in this chapter are adapted and/or reproduced, in part, from Effect of oligonucleic acid (ONA) backbone features on assembly of ONA-star polymer conjugates: a coarse-grained molecular simulation study by Joshua E. Condon and Arthi Jayaraman, published in Soft Matter volume 13, issue 38, pages 6770-6783, August 2017, with permission from the Royal Society of Chemistry.

4.1 Introduction

Soft materials comprised of polymers and/or nanoparticles conjugated to oligomers of nucleic acids (e.g. DNA, RNA, peptide nucleic acids or PNA) are useful in a wide range of nano- and bio-technologies such as electronics[105, 106], catalysis[107, 108], drug and gene delivery[109-119], shape-memory materials[120, 121], and molecule recognition/sensing[122-127]. The thermoreversible, specific, and directional nature of interactions between complementary oligomer strands of nucleic acids (or oligonucleic acids) are desirable properties that are also imparted to materials composed of polymer conjugated oligonucleic acids (ONAs). Though DNA is the most commonly used nucleic acid chemistry in many of the above applications, factors such as expensive synthesis and limited DNA design space have prompted the synthesis of other ONA chemistries that vary in backbone flexibility, nucleobase spacing, and backbone charge, while retaining hybridization and melting of the

bases.[41] For example, locked nucleic acid (LNA) strands are DNA mimics in which a methylene group links the 2'-O and 4'-C atoms in the backbone sugar moiety, locking the sugar in an 3'*endo*-conformation, and making the LNA backbone less flexible than DNA.[36, 37] The melting temperature (T_m) of hybridized LNA:LNA duplexes is higher than that of DNA:DNA duplexes with the same nucleotide sequences, suggesting that a decrease in flexibility corresponds to an increase in T_m. Threose nucleic acid (TNA) is another synthetic DNA mimic with a threose sugar in place of the deoxyribose moiety of DNA. In the TNA strands the adjacent phosphate groups are connected to the 3' and 2' positions of the sugar instead of the 5' and 3' positions as in DNA.[128] This altered sugar-phosphate placement and resulting torsional angles along the backbone in TNA should change its chain conformation, and as a result the thermodynamics of melting and hybridization compared to DNA. Peptide nucleic acids (PNA) are another class of DNA analogues comprised of N-(2-aminoethyl)-glycine units along the backbone to which bases are attached via a methyl carbonyl linker. This polyamide backbone in PNA instead of the sugar-phosphate backbone in DNA leads to the PNA backbone being neutral and more flexible than DNA.[38] These features lead to PNA:PNA duplexes having a higher T_m than DNA:DNA duplexes with the same nucleotide repeat sequence. Clickable nucleic acids (CNA) strands are synthesized through click chemistry between thiol groups and allylamine or acrylamide leading to a uncharged, more flexible backbone than DNA, and the possibility to controllably tune base-base spacing which is not possible with DNA.[129] Considering this diverse range of ONA chemistries, and the impact of these chemical modifications along the ONA backbone on the oligomer's physical features and melting, it is important to predict how these new chemistries impact

assembly of polymers conjugated with these ONA chemistries to envision their use in place of DNA.

Past work on assembly of soft materials driven by DNA hybridization has focused on DNA-star polymer conjugates[39, 40], DNA grafted nanoparticles [32, 33, 110-112, 130-133] and colloids[109, 134], and star-like structures composed of DNA strands (termed multivalent DNA)[135-139]. Extensive experimental and computational studies of DNA-nanoparticle conjugates have provided a thorough understanding of how DNA length, DNA sequence, DNA concentration, and solution properties such as pH and salt concentration affect structure (e.g. amorphous clusters versus crystals) and thermodynamics (e.g. melting temperature) of the assembled materials. For example, experiments and simulations have shown that increasing DNA grafting density on particle surface, increasing salt concentration, increasing G-C content in graft DNA, and increasing nanoparticle size at a constant DNA graft length stabilize the hybridized duplexes, and as a result, increase the melting temperature of the DNA functionalized nanoparticle assemblies.[32-35] Experimental studies of hydrogels formed by DNA hybridization show that increasing the flexibility of the DNA strands within the hydrogel by reducing the amount of hybridization that can occur between strands leads to an increase in modulus and stability of the resulting hydrogel.[140] On the computational front, coarse-grained simulations [8, 34, 35, 135-139, 141-146] have captured the assembly and percolation of DNA grafted particles and DNA dendrimers/stars. For example, Largo et al. have demonstrated that at low concentrations of 4-arm (tetravalent) DNA dendrimers and at low temperatures, the DNA dendrimers phase separate, whereas at intermediate concentrations, the DNA dendrimers percolate and form a network at intermediate temperatures, and become

kinetically arrested at low temperatures.[137] Simulations of tetravalent DNA nanostars, show that an increase in the tetravalent DNA nanostars concentration leads to an increase in the percolation transition temperature.[139] Phase diagrams have been developed using theory and simulation that predict phase behavior over a range of temperatures, concentrations, and nanostar architectures for DNA nanostar systems.[138, 139]

Like DNA grafted nanoparticles and DNA stars/dendrimers, studies have shown the ability of DNA-polymer conjugates[4, 147-152] to self-assemble into a range of morphologies with applications in drug delivery, gene delivery, and sensing. For example, short DNA strands have been conjugated to poly(N-isopropylacrylamide) (PNIPAM) linear and miktoarm star polymers to form thermoresponsive diblock copolymers that self-assemble into micelles in which the morphology of the micelle is tunable through architecture and size of the DNA-polymer conjugate.[152] DNA-linear polymer diblock conjugates do not require hybridization of the DNA strands for self-assembly in aqueous solutions and instead rely upon the relative hydrophilicity of the DNA strand and the hydrophobicity of the linear polymer to assemble. Experimental work has also been done to synthesize systems in which multiple DNA strands are grafted to a polymer backbone which then attaches to a DNA origami patterned surface to allow for single polymer patterning.[151] Lu *et al.* showed that by attaching DNA strands to the ends of brush polymers, hybridization of these DNA strands leads to the self-assembly of these DNA-brush polymers into wormlike structures.[150] Another example in which DNA hybridization stabilizes self-assembly is the conjugation of DNA to peptides where the resulting diblock polymer exhibits hybridization with complementary strands in

addition to forming β-sheet structures with the peptide block.[149] The self-assembly of these peptide-DNA subunits results in a toroidal structure.

While the above DNA based studies are valuable to the materials world, to design analogous materials with new ONA chemistries, it is critical to know *a priori* if and how the above trends hold when other ONA chemistries replace DNA. Molecular simulations play a valuable role in such a quest as they can predict quantitatively and/or qualitatively the behavior of the materials before time-intensive and costly materials synthesis and characterization. While atomistic simulations retain chemical detail of the new DNA chemistry and represent solvent effects explicitly, they require development/optimization of force fields to correctly represent the new ONA chemistries. Additionally, capturing the relevant length and time scales for studies of assembly in atomistic simulations is far from the reach of the current state of software and hardware. Coarse-grained models and simulations are the most appropriate choice for understanding assembly behavior as a function of a large ONA design space. In this regard, Ghobadi and Jayaraman developed a generic coarse-grained model for ONAs, to explore how ONA backbone flexibility, backbone charge, and nucleobase spacing impact the thermoresponsive behavior of both free (unconjugated) ONA systems[30] and ONA-linear polymer conjugate systems[153]. They found that decreasing backbone flexibility, nucleobase spacing, and backbone charge increased the melting temperature (T_m) of the ONA strands.[30] Further, they demonstrated that athermal polymers conjugated to ONA do not impact the T_m for stable complementary ONA strands, but shift the T_m to lower temperatures for less stable complementary ONA strands.[153] By evaluating physical properties such as ONA backbone charge and flexibility independently, instead of evaluating ONA chemistry on a case-by-case

basis, Ghobadi and Jayaraman established a framework to qualitatively and quantitatively evaluate thermoresponsive properties of conjugated ONA for a broad range of ONA chemistries. We extend this framework in this paper to evaluate structure and thermodynamics during assembly of ONA-star polymer conjugates.

In this work, we present coarse-grained (CG) simulations of the assembly of ONA-star polymer conjugates with varying ONA backbone flexibility, ONA backbone charge, ONA base sequence and composition, and ONA-star polymer architecture to elucidate how these design features impact ONA-star polymer conjugate strand hybridization, assembly, and thermodynamics. We use the CG model of ONA that has been parameterized in previous work by Ghobadi and Jayaraman to correctly predict melting behavior of DNA.[30] Our implicit solvent CG simulation approach allows us to investigate time and length scales of relevance to the experimental assembly of ONA-star polymer conjugates that would not be computationally feasible from an atomistic explicit solvent simulation framework. We find that for the 8 nucleotide long ONA sequences and the three star architectures considered here, increasing ONA backbone flexibility and increasing ONA backbone charge decrease the stability of the clusters formed by the ONA-star polymer conjugates shifting the ONA melting temperature and cluster assembly temperature to lower values; increasing the number of arms (3 to 6) in the ONA-star polymer architecture does not shift the thermal stability of the clusters ONA-star polymer conjugates. With regards to structure, going from neutral to charged ONA backbone, and increasing the number of arms per ONA-star polymer conjugate leads to better ordering of ONA-star polymer conjugates within the assembled systems, and increase the dispersion of ONA-star polymer conjugates in disassembled systems.

This chapter is organized in the following manner. In Section 4.2, we define the key attributes of our CG model of ONA-star polymer conjugates, provide our CG simulation details, present a description of the data analyses, and, finally, list all the parameters we vary in our simulations. In Section 4.3, we present simulation results for melting/hybridization, assembly transition as well as structure of ONA-star polymer conjugates for a range of temperatures both above and below melting and assembly transitions. In Section 4.4, we summarize our findings from this paper and include a discussion of the implications of our work on future studies involving ONA-star polymer conjugates.

4.2 Methods

4.2.1 Model

For this study, we use the coarse-grained (CG) model developed by Ghobadi and Jayaraman[30] that reproduces expected melting behavior of DNA oligomers and can be easily extended to other ONA chemistries with changing flexibility and charge. This model is described in detail in the original paper[30] and we only discuss the key aspects of the model for the purposes of this work. As shown in Figure 4.1a, each ONA strand is represented by a chain of nucleotides composed of a backbone (BB) bead and a hydrogen bond (HB) forming site, and the conjugated polymer is represented by a chain of polymer (PL) beads. The charge of the BB beads can be adjusted to mimic negatively charged ONA strands as in DNA, or mimic electrostatically neutral ONA strands as in PNA or CNA. The ONA backbone flexibility is tuned by changing the constant of the BB-BB-BB angle potential. The Watson-Crick H-bonds are mimicked through an attractive interaction between

complementary HB sites. The size of the HB site with respect to the BB site as well as repulsive interactions between non-complementary bases ensures directionality and specificity of the Watson Crick H-bonds. Stacking interactions between neighboring bases are modeled by a dihedral HB-BB-BB-HB interaction. Adjacent BB-BB, BB-PL, and PL-PL beads are connected through a harmonic bond potential with equilibrium bond length of 0.84σ and harmonic potential constant of 1000ε. While the polymer flexibility can be tuned through the PL-PL-PL angle potential, we set the angle constant (k_a^{PL}) in this paper to be $10\varepsilon/rad^2$ and angle to maintain a linear chain. The BB-BB-PL angle potential is of the same form as the BB-BB-BB angle potential; the angle constants used for this angle potential can be found in the parameters varied subsection. The BB-PL-PL angle potential is of the same form as the PL-PL-PL angle potential and has an angle constant of $10\varepsilon/rad^2$. We explicitly include monovalent counterions (IN) beads in which a charge valency of +1 is used for positively charged ions and –1 for negatively charged ions. Lastly, we denote the central PL bead in an ONA-star polymer conjugate as the core (CR) bead. To create an ONA-star polymer conjugate, we connect multiple ONA-polymer chains to a single CR bead as shown in Figure 4.1a using a harmonic bond potential for the CR-PL bonded interaction like the PL-PL bonded interaction. The CR bead has the same properties as the PL bead, however, we do not include any angle or dihedral potential for bonded interactions consisting of a CR bead with neighboring PL beads.

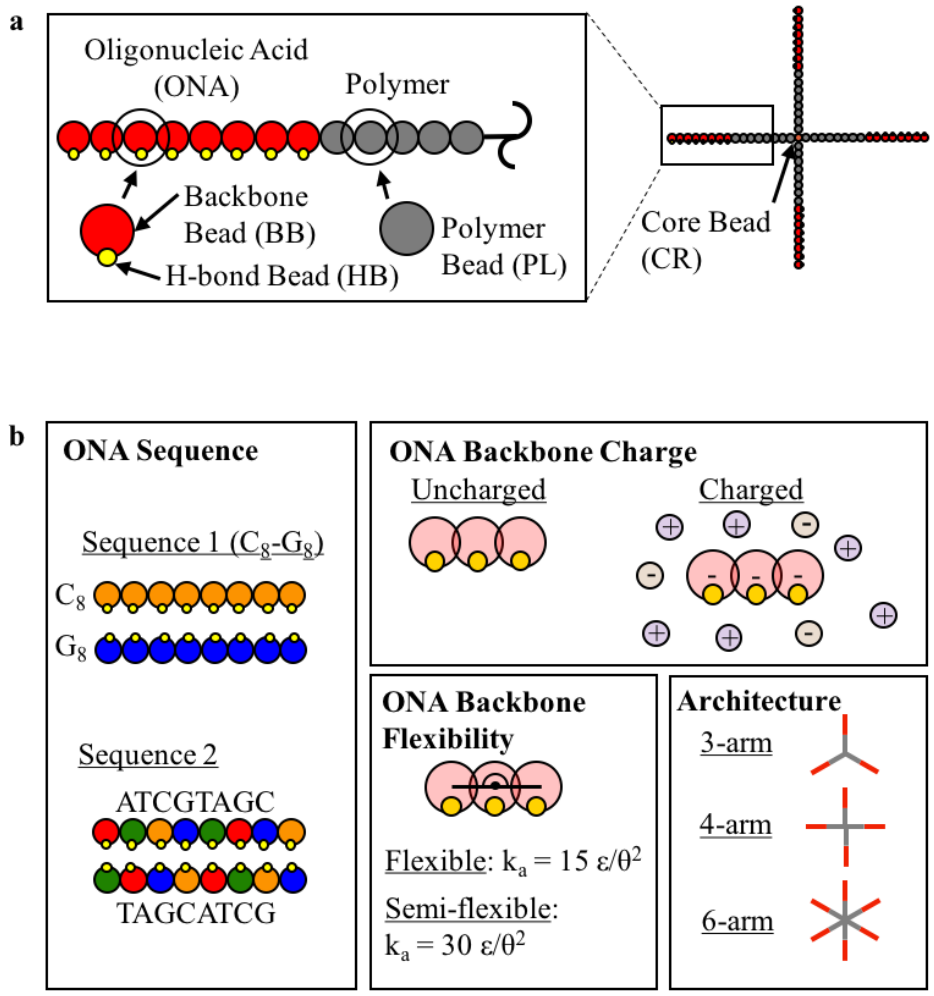

Figure 4.1: (a) Schematic of the coarse-grained (CG) model of a 4-arm oligonucleic acid (ONA)-star polymer conjugate. (b) Visual description of the design parameters varied in this paper.

The units of length, mass, and energy are set to be σ = 0.6 nm, m = 42.0 g/mol, and ε = 0.1 kcal/mol, respectively. The diameters of the BB, HB, PL, CR, and IN beads are 1.0σ, 0.3σ, 1.0σ, 1.0σ, and 0.7σ, respectively. While the mass of BB beads is set to 3.0m, the mass for the rest of the beads are set to 1.0m. The size and the mass selections for the BB and HB beads were originally chosen to represent the DNA

nucleotide backbone and nucleobase chemistries. The CG model used in this work was chosen to represent generic ONA, polymer, and monovalent ion chemistries and does not precisely capture the size or mass of every ONA, polymer, or ion.

The Lennard-Jones (LJ) potential[98] is used to model the directional Watson-Crick H-bond interactions between complementary HB sites:

$$U_{ij}^{HB-complementary}(r) = \varepsilon_{ij}^{HB}\left[\left(\frac{\sigma^{HB}}{r}\right)^{12} - \left(\frac{\sigma^{HB}}{r}\right)^{6}\right] \tag{4.1}$$

in which ε_{ij}^{HB} is the pairwise interaction strength as defined for HB sites i and j. σ^{HB} is the diameter of HB bead. The cutoff for the potential and force is $1.9\sigma^{HB}$ and a switching function taken from GROMACS[99] smoothly transitions both to zero before reaching $2\sigma^{HB}$. The LJ pairwise potential for all other bead types is purely repulsive and is modeled using the Weeks-Chandler-Andersen (WCA) potential[65]:

$$U_{ij}(r) = \begin{cases} 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r}\right)^{12} - \left(\frac{\sigma_{ij}}{r}\right)^{6}\right] + \varepsilon_{ij} & r < 2^{1/6}\sigma_{ij} \\ 0 & r \geq 2^{1/6}\sigma_{ij} \end{cases} \tag{4.2}$$

in which ε_{ij} is 1.0ε and σ_{ij} is the arithmetic mean of the diameters of bead i and bead j. The LJ potential for 1-3 and 1-4 pairs are retained within a given ONA-star polymer conjugate to avoid overlap. In the CG model, the solvent is implicitly represented where its quality is taken to be the same for both the ONA and polymer regions of the ONA-star polymer molecules. Additionally, the ONA region of the conjugate does not have any attractive interaction with the polymer region. The electrostatic interactions for beads that possess charge (BB and IN), are modeled using the Coulomb potential:

$$U_{ij}^{coul}(r) = \frac{q_i q_j}{4\pi r \varepsilon_0 \varepsilon_r(T)} \tag{4.3}$$

in which q_i and q_j are the charge valencies of bead i and j, respectively, and ε_0 is the permittivity of free space. $\varepsilon_r(T)$ is the relative permittivity of the solvent as a function

of temperature, represented with the following equation at low ion concentrations[100]:

$$\varepsilon_r(T) = 249.4 - 0.788T + 0.00072T^2 \quad (4.4)$$

where T is in units of Kelvin. The 1-3 and 1-4 BB electrostatic interactions are set to 0 in this model.

4.2.2 Simulation Details

We perform all coarse-grained (CG) molecular dynamics simulations with the LAMMPS[62] package in the NVT ensemble using Langevin dynamics. We prepare the initial configuration for each ONA-star polymer conjugate by randomly placing ONA-star polymer conjugates within a periodic cubic simulation box, avoiding any bead-bead overlap. In addition to the ONA-star polymer conjugates, we also prepare simulations of free ONA hybridized strands without polymer using the same initial configuration procedure. To achieve systems with a constant ONA strand volume fraction (ϕ_{ONA}) of 0.0032, we place the appropriate number of ONA-star polymer conjugates or ONA strands to arrive at a total of 96 ONA strands in a cubic simulation box with a side length of 50σ. For systems where the ONA backbone is uncharged, we perform the simulation without any added counterion (IN) beads. For systems where the ONA backbone is negatively charged, we include charged IN beads with a valency of +1 to achieve electrostatically neutrality, and an additional number of IN beads to obtain 1mM of monovalent salt; these IN beads are randomly placed in the simulation box while avoiding overlaps.

In the Langevin dynamics simulations, we set the friction coefficient to 10 (units of reduced time, τ) as chosen in Ghobadi and Jayaraman to agree with the diffusion of short nucleic acid chains in water.[30] Each simulation utilizes a two-level

RESPA[102] integrator in which the non-bonded interactions are calculated using a timestep of 0.001τ and the bonded interactions are calculated using a timestep of 0.0005τ (units of reduced time where 0.001τ corresponds to 6fs). All electrostatic interaction calculations are performed using the particle-particle-particle-mesh (PPPM) method.[103] All systems with uncharged ONA backbones undergo a 10^8 timesteps equilibration run. For all systems with charged ONA backbones, an initial 10^7 timestep run is performed with the electrostatic interactions turned off to allow for hybridization to occur among complementary ONA strands at reasonable time scales; then, the electrostatic interactions are turned on for the remainder of the 10^8 timesteps equilibration run. This protocol is chosen to achieve hybridization in charged systems with shorter simulation runs than if we started with electrostatics turned on with the initial configuration described earlier. In all systems, the equilibration run is followed by a 10^7 timestep sampling run in which configurations of the systems are recorded every 10^5 timesteps for analysis. Unless otherwise noted, for every temperature and system studied, we perform three replicate simulations with different initial configurations and velocities.

4.2.3 Analyses

To characterize the hybridization/melting behavior of ONA-star polymer conjugates and free ONA, we calculate the ensemble average fraction of ONA strands in a simulation box that are at least 50% hybridized ($f_{50\%}$) as a function of reduced temperature (T^*). We mimic the approach described in Ghobadi and Jayaraman.[30] We define the reduced melting temperature (T_m^*) as the T^* at which $f_{50\%} = 0.5$. We also define the width of the melting transition region (ΔT_m^*) to be the difference in the temperature values between points on the $f_{50\%}$ curve that correspond to $f_{50\%} = 0.9$ and

$f_{50\%}$ = 0.1. We utilize linear interpolation between data points to more precisely obtain temperatures that correspond to the above $f_{50\%}$ requirements for T_m^* and ΔT_m^*.

We also characterize the assembly behavior of ONA-star polymer conjugates. To do this we first identify the clusters in the assembly of ONA-star polymer conjugates. These clusters are defined as a collection of ONA-star polymer molecules that are connected to each other via hybridization of ONA strands where the strands are at least 50% hybridized. We calculate the ensemble average number of clusters ($<N_{clusters}>$) as a function of temperature, and normalize this value by the number of ONA-star polymer conjugates (N_{stars}) in the system so that we can observe relative clustering behavior for a range of ONA-star polymer conjugate architectures at the same ϕ_{ONA}. In the same way as we define T_m^*, we also define an assembly transition temperature (T_a^*) as the temperature at which $<N_{clusters}>/N_{stars}$ = 0.5. We define the width of the assembly transition region (ΔT_a^*) arbitrarily to be the difference in the values of the temperatures corresponding to $<N_{clusters}>/N_{stars}$ = 0.1 and $<N_{clusters}>/N_{stars}$ = 0.9, respectively. We utilize linear interpolation between data points to obtain temperatures that correspond to the above $<N_{clusters}>/N_{stars}$ requirements for T_a^* and ΔT_a^*.

In addition to the plots of average number of clusters versus temperature, we also show cluster size distributions for systems near T_a^*. The cluster size distributions at a given temperature are presented as the weighted probability distribution function, *P(s)*:

$$P(s) = \frac{N(s)}{N_{stars}} * s \tag{4.5}$$

where *s* is the number of ONA-star polymer conjugates in a cluster normalized by the number of stars in the simulation, *N(s)* is the number of occurrences of *s* in that

snapshot, and N_{stars} is the number of ONA-star polymer conjugates in the system. We choose to present the probability distribution in this manner so that we can a) fairly compare results among the different ONA-star polymer conjugate systems since the number of ONA-star polymer conjugates in a system decreases as arms in the star increases to keep ONA concentration constant) and b) more easily see occurrences of larger clusters that tend to have fewer occurrences (e.g. $N(s) = 1$ if all the star molecules come together to form one large cluster).

To quantify percolation in the system, we calculate the ensemble average fraction of configurations in a percolated state ($<\tau_{perc}>$) as a function of temperature. To evaluate whether a given configuration is percolated, we first count the number of clusters in that configuration, replicate the image of the configuration in all directions such that the resulting configuration is a cubic assembly of 27 periodic images of the original configuration, and count the number of ONA-star polymer conjugates participating in clusters in the larger configuration. If the size of any cluster in the new configuration is larger than the largest cluster in the original configuration, we list that system as being percolated which corresponds to $\tau_{perc} = 1$.

Lastly, we characterize the spacing of ONA-star polymer conjugates in a given system by calculating the pairwise distribution function between the CR beads ($g(r)_{core\text{-}core}$) in the system [154] :

$$g(r)_{core-core} = \frac{\rho(r)}{4\pi r^2 \Delta r \frac{N_{pairs}}{V}} \tag{4.6}$$

where $\rho(r)$ is the number density of CR bead pairs at a separation distance (r) between CR beads, N_{pairs} is the number of CR bead pairs in the system, and V is the volume of the cubic simulation box. We set the binning distance (Δr) to be 1σ. We calculate

$g(r)_{core\text{-}core}$ both at temperatures below T_a^* and above T_a^* to investigate how the ONA-star polymer systems arrange when assembled and when disassembled.

For all analyses presented in this paper, we obtain the ensemble average value of the property by calculating the value of that property for each individual configuration recorded in the sampling runs of three different trials, and then taking the average of these combined values across all configurations regardless of trial. The standard deviation is calculated for these values and is shown in the form of error bars in the figures in Section 4.3.

4.2.4 Parameters Varied

For all ONA-star polymer conjugate systems studied, each ONA-star polymer arm consists of a sequence of 8 PL beads connected to a sequence of 8 BB beads. We vary the number of arms in the ONA-star polymer arms from 3 to 6, with all arms in a star molecule connected to the same CR bead; schematics of these architectures are represented in Figure 4.1b. We study ONA-star polymer conjugates with the BB beads in the backbone being electrostatically neutral (i.e. neutral/uncharged ONA backbone) or with charge valency of -1 (i.e. charged ONA backbone). We also investigate the influence of ONA backbone flexibility by varying the harmonic angle constant for the BB-BB-BB angle (k_a^{BB}) from $15\varepsilon/rad^2$ (denoted as flexible ONA backbone) to $30\varepsilon/rad^2$ (denoted as semiflexible ONA backbone). We study the free (no polymer) ONA systems as a control case to understand how polymer conjugation in ONA-star polymer cases impact the structure and thermodynamics. For each combination of ONA-star polymer conjugate architecture, ONA backbone flexibility, and ONA backbone charge, we present data across a range of temperatures for systems where half of the ONA-star polymer conjugates (or free ONA strands) have ONA strands

with one sequence (e.g. C8 in the case of sequence 1) and the other half have the complementary sequence (e.g. G8 in the case of sequence 1). All simulations have an ONA strand volume fraction (ϕ_{ONA}) of 0.0032.

We also investigate ONA sequence and composition effects contrasting simulations of sequence 1 (C8 duplexed with G8) and sequence 2 (ATGCTACG duplexed with TACGATGC) as shown in Figure 4.1b. We must note here that our model does not distinguish between the 5' and 3' directions of DNA or other ONA strands. In both sequences hybridization (G8-C8) can occur in either direction. Sequence 1 (C8-G8) also allows for slip in which complementary strands can participate in hybridization with only a fraction of the possible H-bonds formed. For sequence 2, the possibility of partial hybridization is smaller than in sequence 1. Simulations for sequence 2 are performed for the same combination of ONA-star polymer conjugate architecture, ONA backbone flexibility, and ONA backbone charge as sequence 1. For sequence 1, the reduced temperature range for the uncharged ONA backbone systems is $T^* = 6.0\text{-}9.0$ and the reduced temperature range for the charged ONA backbone systems is $T^* = 4.0\text{-}7.0$. For sequence 2, the reduced temperature range for the uncharged ONA backbone systems is $T^* = 5.0\text{-}8.0$ and for charged ONA backbone systems is $T^* = 4.0\text{-}7.0$. Unless otherwise noted, 3 independent replicate simulations are performed at every T^* at interval of 0.2 within the range for each system.

4.3 Results and Discussion

4.3.1 Effect of ONA-Star Polymer Conjugate Design on Melting Transition

First, we investigate how the melting/hybridization and assembly behavior of ONA-star polymer conjugates is impacted by ONA backbone flexibility and charge, and number of arms in the star polymer for the C8-G8 ONA sequence.

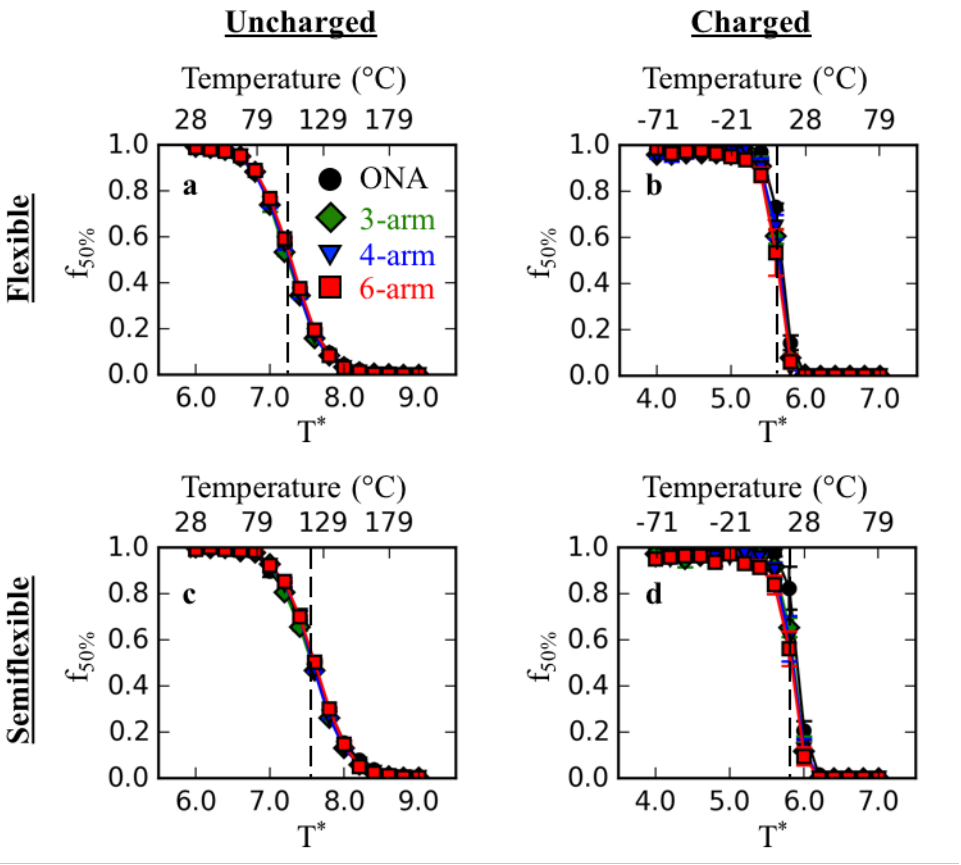

Figure 4.2: (a-d) Fraction of oligonucleic acid (ONA) strands that have at least fifty percent of possible H-bonds formed ($f_{50\%}$) as a function of reduced temperature (T^*) for free ONA (black circles), 3-arm stars (green diamonds), 4-arm stars (blue triangles), and 6-arm stars (red squares) for C8-G8 ONA systems at a volume fraction of ONA strands (ϕ_{ONA}) of 0.0032. The legend in Figure 4.2a holds for all plots in Figure 4.2. The titles along the top of the figure refer to the ONA backbone charge of the systems corresponding to the plots. The titles along the left side of the figure refer to the ONA backbone flexibility of the systems corresponding to the plots. In all plots, we show the reduced temperature in the x-axis and the corresponding temperature in degree Celsius on the top of the plot. The black dashed lines in all plots denote the reduced melting temperature, T_m^*, for the free ONA case. Melting temperatures for all systems are presented in Table 4.1.

Table 4.1: Melting temperature, T_m^* and width of the melting transition, ΔT_m^*, for free C_8-G_8 ONA and C_8-G_8 ONA-star polymer conjugate systems with varying star polymer architecture, ONA backbone charge and ONA backbone flexibility.

ONA Backbone Charge	ONA Backbone Flexibility	Architecture	T_m^*	ΔT_m^*
Uncharged	**Flexible**	Free ONA	7.26	1.06
		3-arm	7.24	1.01
		4-arm	7.26	1.01
		6-arm	7.29	1.00
	Semiflexible	Free ONA	7.60	1.15
		3-arm	7.56	1.04
		4-arm	7.57	1.03
		6-arm	7.60	1.03
Charged	**Flexible**	Free ONA	5.68	0.40
		3-arm	5.64	0.39
		4-arm	5.65	0.38
		6-arm	5.61	0.47
	Semiflexible	Free ONA	5.90	0.41
		3-arm	5.86	0.42
		4-arm	5.84	0.44
		6-arm	5.83	0.56

In Figure 4.2, we present the hybridization/melting of free ONA strands (no polymer attached to ONA) as well as for ONA conjugated with 3-arm to 6-arm star polymers, as a function of charge and flexibility of ONA backbones. For a given ONA flexibility and ONA charge, the reduced melting temperature (T_m^*) does not shift upon conjugation of ONA with athermal star polymers, at this ONA strand volume fraction (ϕ_{ONA}) of 0.0032. This is not surprising considering recent results using coarse-grained and atomistic simulations which showed that the ONA strand melting is not impacted by conjugation of short linear athermal polymer and conjugation to one arm of a 4-arm

star athermal polymer.[153] In our simulations, we conjecture that the negligible deviation in the thermoresponsive behavior with increasing number of arms is attributed to the negligible variation in configurational entropy loss of the ONA strands relative to the large enthalpic gain upon hybridization with varying number of arms, at the chosen low concentration of ONA strands.

Next, we explore how ONA backbone charge and flexibility impact melting of ONA strands in these ONA-star polymer conjugates. The T_m^* increases as ONA flexibility decreases, for both the charged and uncharged ONA backbones, and for a given ONA flexibility, the T_m^* decreases as we go from an uncharged/neutral to a charged ONA backbone. This behavior with varying ONA backbone flexibility and charge is also in agreement with the recent work on free ONA strands with the same CG model[30] where the effect of varying ONA design parameters on *free* ONA (no polymer conjugation) melting transition was explored. The rationale for these shifts in T_m^* with varying flexibility and charge is based on the energetic and entropic driving forces. The smaller loss in conformational entropy, ΔS_{conf}, upon hybridization for the stiffer ONA strands leads to a larger T_m^*. The presence of repulsive electrostatic interactions along the negatively charged ONA backbones decreases the gain in enthalpy upon hybridization compared to the uncharged ONA backbones, leading to a smaller T_m^* for charged ONA backbones versus uncharged ONA backbones. Going beyond the results of the work by Ghobadi and Jayaraman, here we note that the ONA-star polymer systems where (3-, 4- and 6- arm) star-polymers assemble via ONA hybridization, have similar dependence on ONA backbone flexibility and charge as the free ONA systems.

Besides the shift in T_m^* discussed above, as we go from the uncharged ONA backbone systems to the charged ONA backbone systems in Figure 4.2, we also see that the width of the melting transition (ΔT_m^*) becomes narrower. This can be explained by considering the competing interactions during hybridization. In the charged ONA backbone systems, the short-range attractive HB interactions that stabilize the hybridized state compensate for the repulsive electrostatic interactions along the ONA backbone and the thermal motion of the ONA strands in the system, both of which destabilize the hybridized state. Thus, partial melting of a few base pairs in the ONA duplex (also called bubbles) will not be favored in the charged systems as the electrostatic repulsions will no longer be compensated by the short-range attractive HB interactions. Due to these competing interactions, the charged system goes from a completely hybridized state to a completely melted state in a narrower range of temperatures than the uncharged ONA backbone systems where destabilization of the hybridized state due to electrostatic repulsion is absent. These uncharged ONA backbones systems can be thought to be analogous to the charged systems where screening of electrostatic interactions due to increased salt concentration (for example) among hybridized ONA backbones leads to the thermal motion being the only destabilizing force in the system, causing a broadening of the melting transition region. Such a sharpening/narrowing of the melting transition has been seen in DNA functionalized gold nanoparticles studied by Hurst *et al.*[33]. They observed an increase in melting transition temperature and melting transition width of the DNA-gold nanoparticles with an increase in salt concentration induced electrostatic screening.

4.3.2 Effect of ONA-Star Polymer Conjugate Design on Assembly Transition

Next, we characterize the impact of ONA backbone flexibility and charge, and star polymer architecture on the assembly of C8-G8 ONA-star polymer conjugates in Figure 4.3.

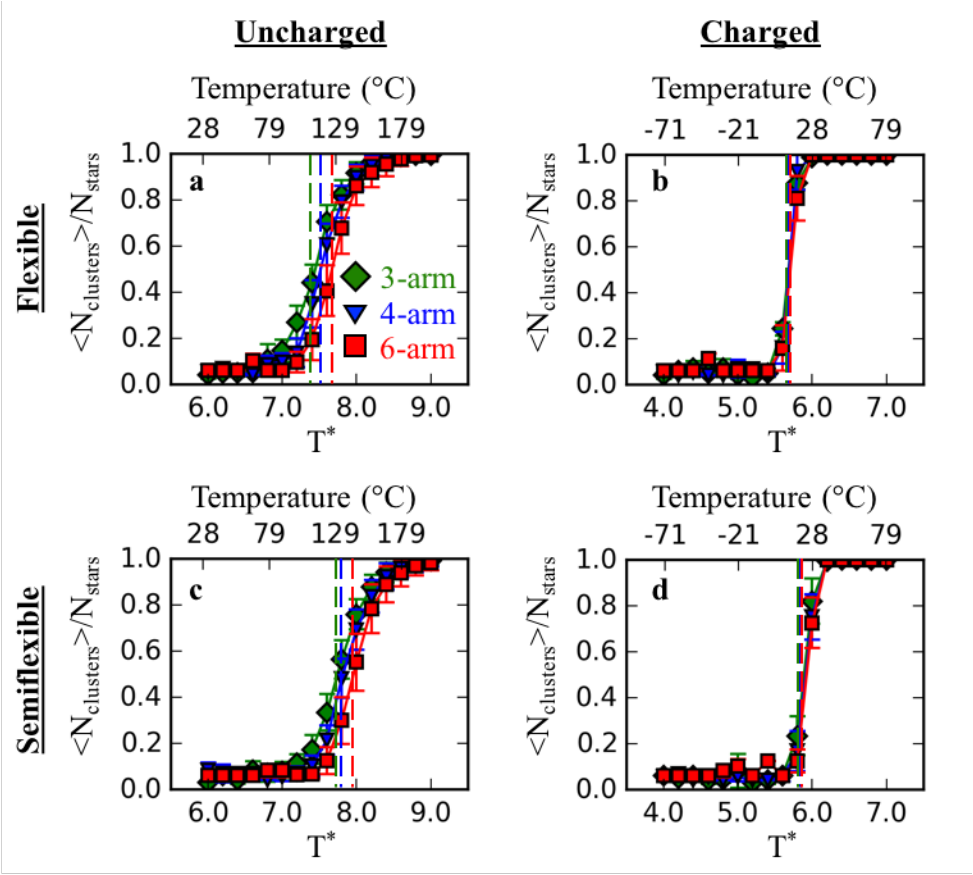

Figure 4.3: (a-d) Average number of clusters ($\langle N_{clusters} \rangle$) normalized by the total number of stars in each system (N_{stars}) as a function of temperature for 3-arm stars (green diamonds), 4-arm stars (blue triangles), and 6-arm stars (red squares) for C8-G8 ONA sequence at a volume fraction of ONA strands (ϕ_{ONA}) of 0.0032. The legend in Figure 4.3a holds for all plots in Figure 4.3. The titles along the top of the figure refer to the ONA backbone charge of the systems corresponding to the plots. The titles along the left side of the figure refer to the ONA backbone flexibility of the systems corresponding to the plots. The dashed lines in all plots denote the reduced assembly transition temperature, T_a^*, for the corresponding architecture.

Table 4.2: Assembly transition temperature, T_a^*, and assembly transition width, ΔT_a^*, for ONA-star polymer conjugates with C_8–G_8 sequence with respect to ONA backbone charge, ONA backbone flexibility, and ONA-star polymer conjugate architecture.

ONA Backbone Charge	**ONA Backbone Flexibility**	**Architecture**	**T_a^***	**ΔT_a^***
Uncharged	**Flexible**	3-arm	7.44	1.20
		4-arm	7.52	1.00
		6-arm	7.67	0.92
	Semiflexible	3-arm	7.74	1.15
		4-arm	7.82	0.97
		6-arm	7.96	0.93
Charged	**Flexible**	3-arm	5.68	0.39
		4-arm	5.69	0.30
		6-arm	5.70	0.42
	Semiflexible	3-arm	5.89	0.45
		4-arm	5.91	0.44
		6-arm	5.92	0.41

In Figure 4.3, the trends in assembly transition are similar to the melting behavior of ONA with varying backbone charge and flexibility described in Figure 4.2. In all plots, at high temperatures the system is in an unassembled state, and as a result, the average number of clusters *(*$<N_{clusters}>$*)* is the same as number of star-ONA conjugates (N_{stars}*)*. At low temperatures as the system assembles, the $<N_{clusters}>/N_{stars}$ value decreases and approaches the value of 1/ N_{stars} (i.e. a single cluster). In Figure 4, we provide representative simulation snapshots of the assembled states of ONA-star polymer conjugates for all the systems described in Figure 4.3. In Figure 4.3, we see that as ONA backbone flexibility increases, the assembly transition temperature (T_a^*) decreases, and as the ONA backbone charge is removed from a system, the assembly

transition temperature increases. We also see that the width of the assembly transition (ΔT_a^*) is larger for the uncharged systems than for the charged systems. These trends are expected based on hybridization/melting results in Figure 4.2 because the assembly of the star polymers is driven by the hybridization of the ONA strands.

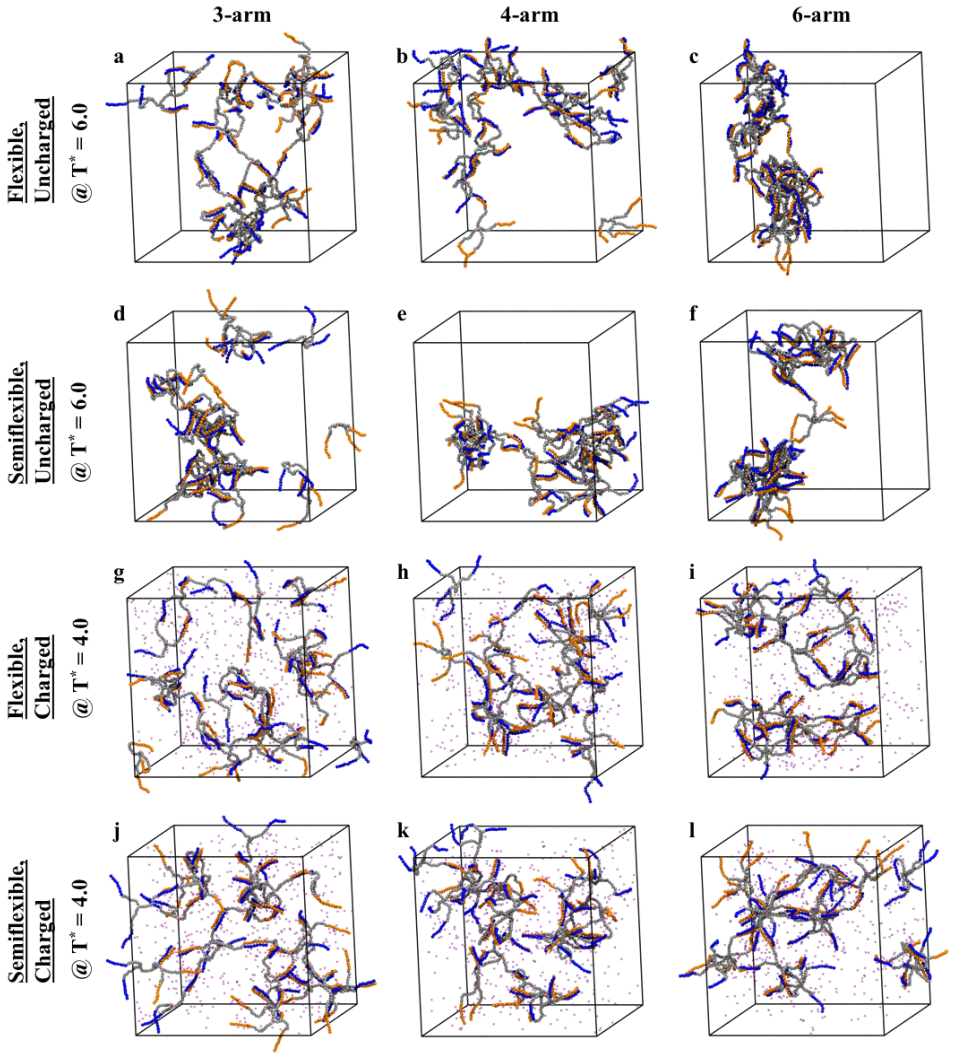

Figure 4.4: (a-l) Simulation snapshots of assembled ONA-star polymer conjugates with C_8–G_8 at a volume fraction of ONA strands (ϕ_{ONA}) of 0.0032. The titles along the top of the figure refer to the ONA-star polymer conjugate architecture. The titles along the left side of the figure refer to the ONA backbone flexibility, ONA backbone charge and the temperature of the systems corresponding to the snapshots. In each system, the backbone beads (BB) are represented by the colors orange and blue for type C and type G, respectively. The polymer (PL) beads are represented by the color grey, and the ion (IN) beads in the charged ONA backbone systems are represented by the colors purple and black for +1 and -1 valencies, respectively. In the charged systems, these ion beads include both the neutralizing +1 charge ions as well as +1 and -1 ions to represent 1mM of monovalent salt.

While the assembly transition and the melting transition of the system have similar thermodynamic driving forces and are inherently linked, there are some differences between the assembly/disassembly transition and ONA hybridization/melting behavior. Comparing Figure 4.3 with Figure 4.2, and Table 4.2 with Table 4.1, we see that the ONA melting transition (T_m^*) occurs at a slightly lower temperature than the assembly transition temperature (T_a^*) for the uncharged systems, whereas T_m^* and T_a^* are similar for the charged systems. We do not extrapolate much from this difference because we define the assembly transition temperature (T_a^*) as the temperature at which $<N_{clusters}>/N_{stars} = 0.5$, not as the temperature where at least half of the stars in the system participate in any cluster; the latter way would be analogous to the definition of T_m^* defined as temperature where at least half of the strands are at least 50% hybridized. For the uncharged ONA, we also see that T_a^* increases as the number of arms in the ONA-star polymers increase. While this is not apparent for the charged cases, the cluster distribution plots in Figure 4.5 indicate that the trend of T_a^* increasing with the number of arms in the ONA-star polymers holds for the charged systems as well by showing that within the assembly transition region, the 3-arm systems start forming small clusters at lower temperatures than 4-arm systems, and 4-arm systems start forming small clusters at lower temperatures than 6-arm systems. The likely reasoning for this is provided later in this section.

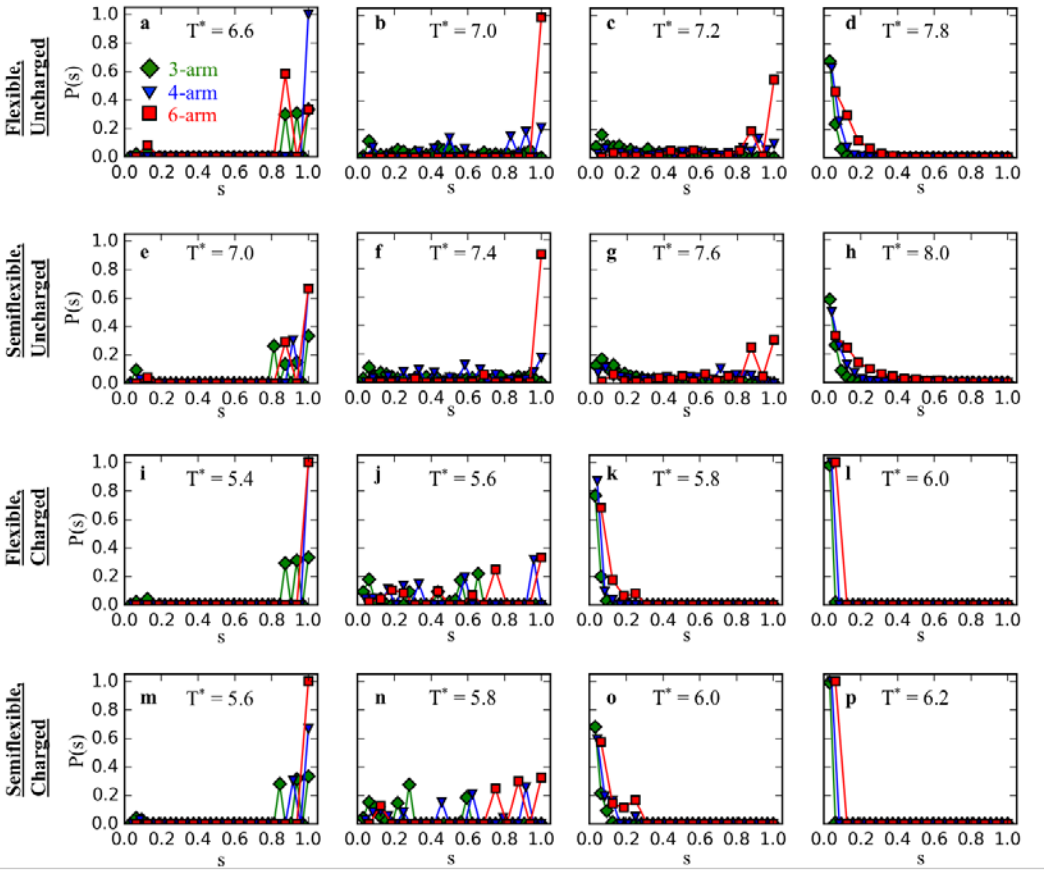

Figure 4.5: (a-p) Normalized probability distribution of clusters of size **s** denoted as P(s), for a range of temperatures within the assembly transition region for C8-G8 ONA sequence at a volume fraction of ONA strands (ϕ_{ONA}) of 0.0032. Data for 3-arm stars are denoted by green diamonds, 4-arm stars with blue triangles, and 6-arm stars with red squares. The legend in Figure 4.5a holds for all plots in Figure 4.5. The titles on the left refer to the ONA backbone flexibility and ONA backbone charge of the systems corresponding to the plots in that row.

In Figure 4.5, we plot the probability distribution of clusters of various sizes, *P(s) vs. s.* As described in the analyses section that *P(s)* is defined as $N(s)/N_{stars}$ times *s*, where *s* is the size of the cluster divided by number of ONA star-polymer molecules in the system, *N(s)* is number of occurrences of clusters of size *s*, and N_{stars} is the

number of ONA-star polymer conjugates in the system. We note again that the size *s* is depicted as the number of stars forming a cluster divided by the total number of stars in the system, to ensure that we can fairly compare systems with varying star architectures. Figure S1 in the Supplementary Information[42] provides the $N(s)/N_{stars}$ *vs. s* data. Without the weighting of the $N(s)/N_{stars}$ with the cluster size *s*, the occurrences of many small clusters can dominate the probability distribution over the occurrences of 1-2 large clusters, with the latter being too small to see. Upon comparing Figure 4.5 and Figure S1, we can see that a high value of P(s) at large s denotes the occurrence of a large cluster, with the probability of that occurrence weighted by the cluster size.

The results in Figure 4.5, especially the intermediate columns representing the temperatures that are close to the assembly transition, show that 6-arm stars statistically have a greater propensity to form larger clusters per ONA-star polymer conjugate than the 4-arm stars and the 3-arm stars at the same temperature, even though the number of ONA strands that could be hybridized remains the same in all cases. While the melting behavior of ONA-star polymer conjugate systems is largely independent of the number of the arms in the ONA-star polymer conjugate (at least among the systems we have studied), the assembly transition is impacted slightly by the star architecture (Figure 4.3). The driving forces for assembly are largely the same as the driving forces for melting transition, with the assembly transition additionally impacted by the translational entropic loss of the ONA-star polymer conjugates upon assembly. We have maintained the same number of ONA strands in the 6-arm star systems as the 4-arm and 3-arm star systems, to ensure the enthalpic driving force for assembly (ONA hybridization) is the same for these different polymer architectures.

As a result, the ONA melting behavior is similar for all architectures. However, the 6-arm systems have fewer molecules than the 4-arm or 3-arm system, leading to a reduction in the change in translational entropy of the molecules upon assembly which manifests itself as a higher T_a^* for the 6-arm system than the 4-arm or 3-arm systems. Knowing how the star polymer architecture impacts assembly at a given temperature provides valuable inferences about the stability of the assembled ONA-star polymer conjugate system for a range of star architectures.

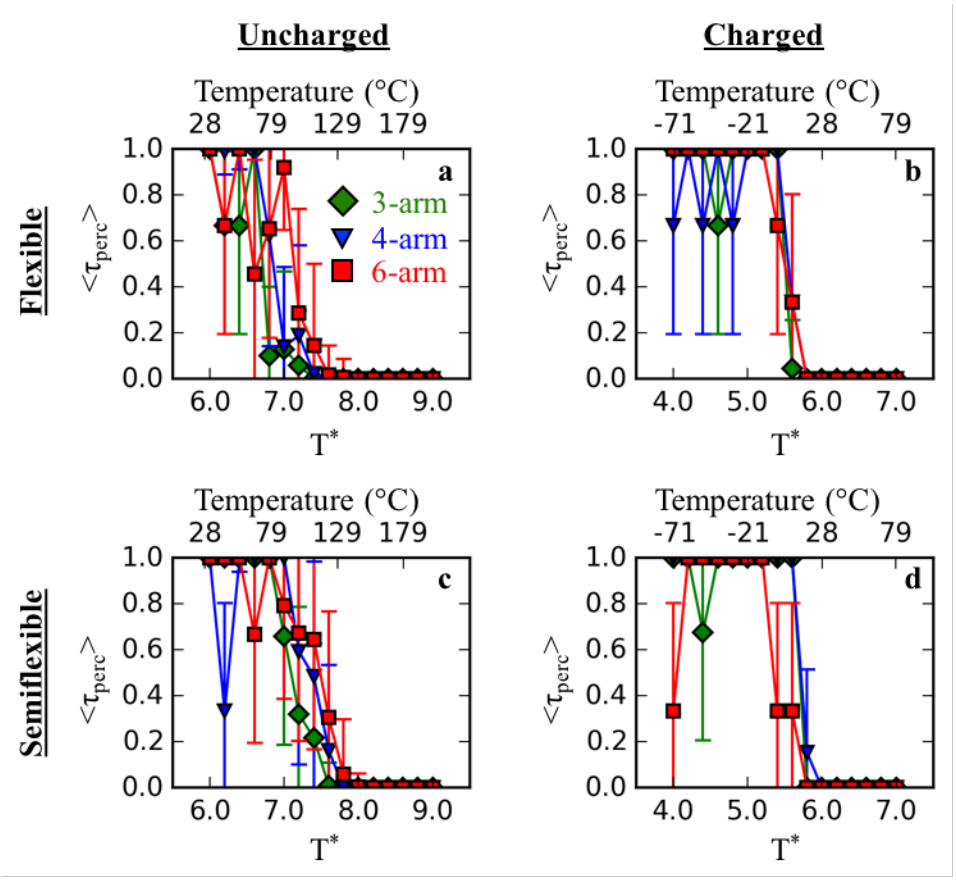

Figure 4.6: (a-d) Ensemble average fraction of simulation frames spent in a percolated state ($<\tau_{perc}>$) as a function of temperature for 3-arm stars (green diamonds), 4-arm stars (blue triangles), and 6-arm stars (red squares) for C8-G8 ONA sequence at a volume fraction of ONA strands (ϕ_{ONA}) of 0.0032. The legend in Figure 4.6a holds for all plots in Figure 4.6. The titles along the top of the figure refer to the ONA backbone charge of the systems corresponding to the plots. The titles along the left side of the figure refer to the ONA backbone flexibility of the systems corresponding to the plots.

Next, we present our findings on how ONA backbone charge and flexibility and number of arms in the star polymer impact the propensity of the assembled state to percolate. Figure 4.6 shows the percolation behavior of ONA-star polymer conjugates

through a parameter $<\tau_{perc}>$ defined in Section 4.2 as a fraction of the simulation frames where the percolation criterion is satisfied. Since τ_{perc} can be either 0 (not percolated) or 1 (percolated), the average ($<\tau_{perc}>$) value that is obtained from all the frames generated from 3 independent trials at each temperature, lies between 0 and 1. We arbitrarily define a percolation transition as the T* where the $<\tau_{perc}>$ reaches 0.5. Due to the large error bar in percolation values within the percolation transition region, we cannot mark a precise percolation transition temperature. From Figure 4.6, we observe that the width of the percolation transition, defined as the range of T* values corresponding to where $<\tau_{perc}>$ transitions from 1 to 0, is broad (and less precise) for the uncharged cases and sharp for charged systems. Based on the large error bars in the plots for the uncharged systems and through visualization, we see that the system visits percolated and unpercolated states in a dynamic manner leading to large variations in the amount of simulation frames spent in a percolated state. Figure S2 in the Supplementary Information[42] provides a visual representation of the distinction of a percolated and unpercolated state for the 6-arm ONA-star polymer conjugate system having charged, semiflexible ONA backbones at ϕ_{ONA} of 0.0032 and T^* of 4.0. The data point corresponding to Figure S2 in Figure 4.6 shows both percolated and unpercolated states at a low temperature, which is likely due to the low concentrations at which we perform these simulations. Based on past studies on analogous systems [137-139] and expectation from crowding effects, one would expect a higher propensity to percolate as we increase ONA strand concentration. By comparing Figure 4.6 to Figure 4.2, we see many of the trends in melting transition with varying ONA backbone flexibility and charge also hold for the percolation transition. As the flexibility of the ONA backbone increases, the transition region for

percolation shifts to lower temperatures. Also, the charged systems show a lower temperature for the percolation transition and a narrower/sharper percolation transition profile than the uncharged systems. This set of behaviors can again be linked to the interplay of the directional short-range HB interactions, the absence or presence of repulsive electrostatic interactions, and the thermal motion of the system. In the uncharged systems, due to the large error bars, it is difficult to conclude how increasing number of arms in the star polymer architecture impacts the percolation transition of the system. In the charged systems, the star architecture does not impact percolation transition in Figure 4.6 in agreement with melting behavior seen in Figure 4.2. However, the width of percolation transition region for the charged and flexible systems in Figure 4.6 is smaller/narrower than the width of the melting transition for the same system in Figure 4.2. This is more clearly seen for the second ONA sequence, described later in this paper.

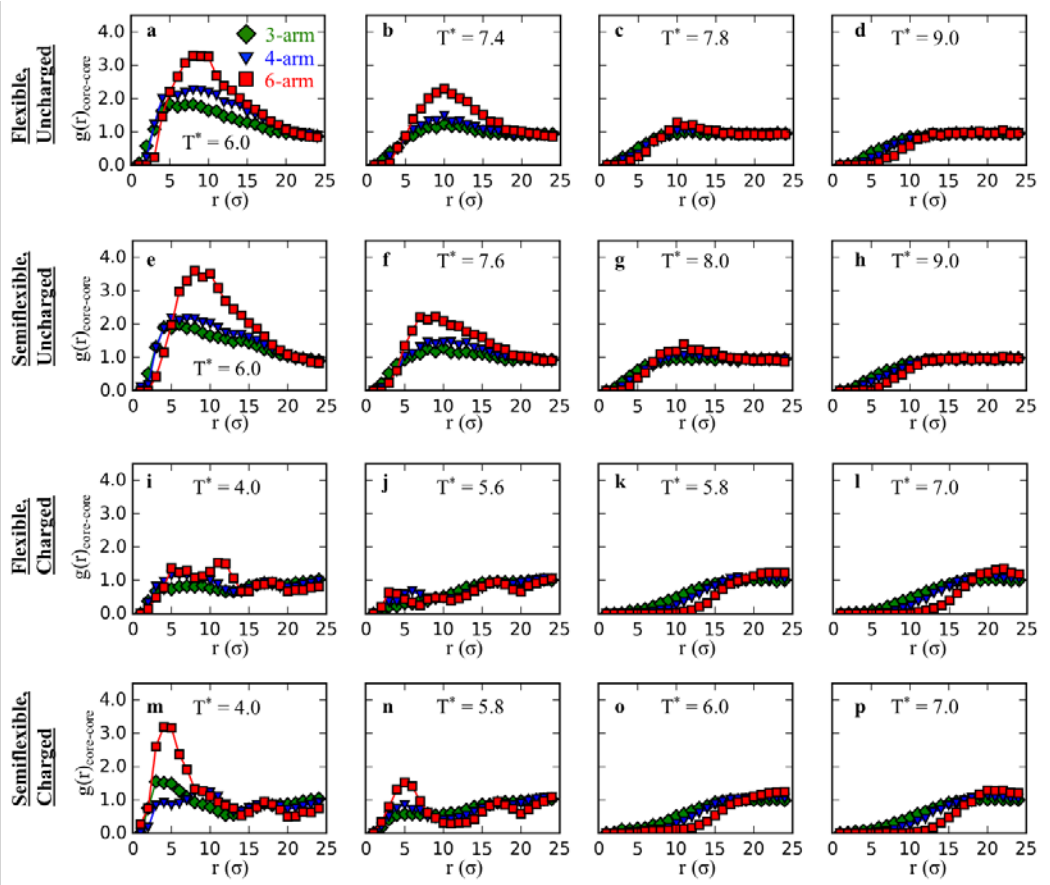

Figure 4.7: (a-p) Radial distribution function among core (CR) beads ($g(r)_{core\text{-}core}$) as a function of separation distance r (having units of σ) for C8-G8 ONA sequence at a volume fraction of ONA strands (ϕ_{ONA}) of 0.0032. Data for 3-arm stars are denoted by green diamonds, 4-arm stars with blue triangles, and 6-arm stars with red squares. The legend in Figure 4.7a holds for all plots in Figure 4.7. The titles on the left refer to the ONA backbone flexibility and ONA backbone charge of the systems corresponding to the plots in that row.

Within the percolated or unpercolated assembled states, we quantify the core-core distribution. In Figure 4.7, we show how ONA backbone charge, ONA backbone flexibility, and ONA-star polymer architecture impact the core-core pair correlation functions in the ONA-star polymer conjugates. Figure S2 in supplementary

information depicts these core beads in red. In Figure 4.7, for each combination of ONA backbone flexibility and charge, $g(r)_{core\text{-}core}$ is plotted for the lowest temperature of that system, the temperature just below the reduced assembly transition temperature (T_a^*), the temperature just above T_a^*, and the highest temperature of that system.

At T* below T_a* (first two columns), the charged ONA backbone systems have more ordered $g(r)_{core\text{-}core}$ than the uncharged ONA backbone systems. Additionally, the extent of ordering increases with increasing number of arms in the ONA-star polymer conjugate. The regular spacing between core beads in the charged ONA backbone clusters below T_a^* is to be expected because the net repulsive electrostatic interactions of the ONA-star polymer conjugate in the charged ONA backbone systems create an effective repulsive barrier that prevents core beads from entering within a certain distance of neighboring core beads. This effective repulsive electrostatic interaction between the cores of the star molecules increases as number of arms increases. The uncharged ONA-star polymer conjugates do not experience an electrostatic repulsive force acting on one another, allowing the core beads to sample a range of inter-core distances and form less ordered clusters than in the charged ONA backbone systems. While the $g(r)_{core\text{-}core}$ plots for uncharged ONA backbone systems differ slightly with respect to backbone flexibility at temperatures below T_a^*, for charged ONA backbone systems, in the case of semiflexible ONA backbone systems we see a large first peak and in the case of flexible ONA backbone systems a small first peak accompanied by a larger secondary peak. This implies that for charged ONAs, the semi-flexibility of the ONA is facilitating better ordering and periodic spacing among the cores. The visuals in Figure 4.4 confirm the trends presented in Figure 4.7.

At T* above T_a^*, we observe minimal differences in the size and position of the peaks in the $g(r)_{core\text{-}core}$ plots between the flexible and semiflexible ONA backbone systems. However, when accounting for ONA backbone charge at temperatures above T_a^*, we see that $g(r)_{core\text{-}core}$ approaches 1 at lower separation distances for the uncharged ONA backbone systems than the charged ONA backbone systems. Again, this can be explained by the presence of a net ONA-star polymer conjugate electrostatic repulsion in the charged ONA backbone system in addition to steric repulsion that prevents the core beads from moving within a certain distance of other neighboring core beads, whereas the uncharged ONA backbone system only encounters steric repulsion. With respect to ONA-star polymer architecture, the 6-arm star system $g(r)_{core\text{-}core}$ approaches 1 at higher separation distances than the 4-arm or 3-arm case. This is indicative of an increase in steric bulk of the ONA-star polymer molecule resulting in dispersive behavior at greater interparticle distances above T_a^*.

4.3.3 Effect of ONA Sequence on ONA-Star Polymer Conjugate Melting and Assembly Transitions

Please refer to the Supplementary Information section of Condon and Jayaraman[42] to view the data associated with this subsection.

Lastly, to prove that the above trends in structure and thermodynamics with varying ONA backbone charge, ONA backbone flexibility, and ONA-star polymer conjugate architecture shown (so far) for C8-G8 sequence, also hold for another ONA sequence, we also perform simulations with a second sequence (presented in Figure 4.1), which is both asymmetric in nucleobase sequence and has a lower G-C content of 50% compared to C8-G8. To ensure brevity in our scientific communication, all

results from simulations involving sequence 2 are found in the supplementary information and a brief discussion of these results is provided here.

Figure S3 and Table S1 in the Supplementary Information[42] for sequence 2 are analogous to Figure 4.2 and Table 4.1 for sequence 1. Comparing these figures and tables we see that the trends in T_m^* with varying ONA backbone charge and ONA flexibility hold for both sequences. The values of T_m^* for sequence 2 systems are lower than the corresponding sequence 1 (C_8-G_8) systems as expected since the G-C content is lower in sequence 2, and thus, the enthalpic driving force stabilizing the hybridized state is smaller for sequence 2. [35, 155] In addition, the width of the melting transition, ΔT_m^*, is smaller for sequence 2 as compared to sequence 1, due to higher occurrences of slip (i.e. partial hybridization of some Cs in one strand with some Gs in another strand) in sequence 1 than sequence 2. Lastly, we note that in charged systems with sequence 2 ONA, at low temperatures, we observe partial hybridization between arms of the same ONA-star polymer conjugate, as shown in Figure S4[42] which limits the system from complete hybridization and $f_{50\%} < 1$. These occurrences are seen primarily in the charged systems because they balance out the electrostatic repulsions between the ONA strands in the arms of the ONA-star polymer conjugate.

Figure S5, Table S2, and Figure S6 in the Supplementary Information[42] for sequence 2 are analogous to Figure 4.3, Table 4.2, and Figure 4.4, respectively, for sequence 1. Again, we observe for both sequences the same trends in T_a^* with respect to ONA backbone charge, ONA flexibility, and ONA-star polymer conjugate architecture. Additionally, like ΔT_m^*, we observe a reduction in ΔT_a^* across all systems when changing the sequence from sequence 1 to sequence 2. Because

assembly is inherently linked to hybridization, we expect that this reduction in ΔT_a^* is due to "slip" partial hybridization in sequence 1 systems. The simulation snapshots in Figure S6 and Figure 4.4 seem similar in all cases with the one noticeable difference for 6-arm charged systems. In the snapshots for sequence 2 (Figure S6f and S6h) the ONA-star polymer conjugates seem less percolated, and less regularly spaced apart as compared to the analogous systems with sequence 1 (Figure 4.4i and 4.4l). This is due to occurrences of intra-ONA-star polymer hybridization seen in charged sequence 2 systems (as shown in Figure S4) that we do not see in sequence 1 or in uncharged sequence 2 systems, which limits those sequence 2 strands from hybridizing with other ONA-star polymers in the systems and expanding the cluster. These visual observations are corroborated with results in Figures S7-S9[42].

Figure S7 shows the normalized cluster size distribution for a range of temperatures around T_a^* for sequence 2 systems. Comparing Figure S7 (sequence 2) to Figure 4.5 (sequence 1), we see that the trends in cluster size around T_a^* hold for both sequences studied. Again, the only exception is that in the case of the charged sequence 2 systems because of partial hybridization within the same ONA-star polymer conjugate we see several smaller clusters at low temperatures that are likely kinetically trapped. Comparing Figure S8 (sequence 2) to Figure 4.6 (sequence 1), we see that percolation does not occur at low temperatures for charged sequence 2 systems; the dominance of several small clusters that are likely kinetically trapped prevents percolation from happening. Percolation seems to occur with higher probability near the melting transition temperatures because of the increased likelihood of those intra-molecular partially hybridized bases melting and hybridizing with other star polymer molecules. Lastly, comparing Figure S9 (sequence 2) to

Figure 4.7 (sequence 1), we see that at temperatures below T_a^*, the first peak in the core-core pair correlation curve is shifted to higher interparticle distance for sequence 2 than sequence 1. This could be because there is inter-ONA star polymer partial hybridization in sequence 1 and the hybridized duplexes need not be directionally matched (i.e. 5' and 3' end are not distinguished in this CG model); these allow the CR core beads to approach one another at closer distances than if the hybridized state was directional as in sequence 2 systems. There does not appear to be any major difference in the dispersion of ONA-star polymer conjugates between sequence 1 and sequence 2 for temperatures above T_a^*. In addition, trends in $g(r)_{core\text{-}core}$ with respect to ONA charge, ONA flexibility, and ONA-star polymer conjugate architecture are similar for sequence 1 and sequence 2, except for the charged sequence 2 systems at low temperatures. As state above, at these low temperatures, the charged sequence 2 systems are likely kinetically trapped in a state of several small clusters rather than one or two large clusters and prevents longer range order from occurring.

Thus, based on the results for both sequences we conclude that the choice in ONA backbone flexibility, ONA backbone charge, and ONA-star polymer architecture all impact the ordering and the spacing of the assembled states of ONA-star polymer conjugates, whereas, ONA backbone charge, and ONA-star polymer architecture impact the dispersion of the disassembled states. Knowing the impact of each of these parameters on ONA-star polymer conjugate systems will allow for better control of both the assembly and the dispersion of ONA-star polymer conjugates in solution.

4.4 Conclusion

We have performed coarse-grained (CG) simulations that demonstrate how the selection of ONA chemistry and ONA-star polymer architecture impact the

melting/hybridization and assembly of ONA-star polymer conjugates. We have shown that upon increasing ONA backbone flexibility and ONA backbone charge, the melting temperature (T_m^*) of the ONA-star polymer conjugates decreases, while changing the number of arms (3 to 6) in the ONA-star polymer conjugate architecture has negligible impact on the melting behavior for the systems that we studied. In addition, with increasing ONA backbone flexibility, incorporation of ONA backbone charge, and decreasing the number of ONA-polymer arms per conjugate, the assembly transition temperature (T_a^*) decreases. The impact of ONA backbone flexibility and charge on ONA-star polymer conjugate melting behavior and assembly is directly linked to the varying contributions of driving forces associated with ONA strands hybridization (i.e. electrostatic forces, Watson-Crick specific direction H-bonding, and thermal motion) as discussed in the previous work using this CG model.[30] In the assembled state, ONA-star polymer conjugate systems with charged ONA backbones exhibit a higher level of structural order than systems with uncharged ONA backbones. We have demonstrated that ONA-star polymer conjugate architecture with larger number of ONA-polymer arms, show an increase in local ordering for systems that are in an assembled state. Lastly, for systems in a disassembled state, ONA-star polymer conjugates with increasing ONA backbone charge, and increasing number of ONA-polymer arms per conjugate are dispersed at larger inter particle distances which agrees well with our understanding of steric and electrostatic effects of star polymers.

Going beyond sequence 1 (C_8-G_8), we have repeated the structural and thermodynamic analyses for sequence 2 (ATGCTACG) in which the G-C content is smaller than sequence 1. For sequence 2, the qualitative trends in melting temperature and assembly transition with varying ONA flexibility, ONA charge, and number of

arms are the same as for sequence 1. Quantitatively, for a given ONA flexibility, ONA charge, and number of arms in the ONA-star polymer, the T_m^* and T_a^* for sequence 2 is smaller than sequence 1, primarily due to lower G-C content. Any additional small differences in structural characterization of the systems with the two sequences arise due to higher propensity for complementary strand slipping induced partial hybridization in sequence 1 (C_8-G_8) than sequence 2 (ATGCTACG), and kinetically trapped clusters in charged cases of sequence 2 systems arising from partial hybridization between arms in the same ONA star polymer molecule.

In summary, the findings we present in this paper inform experimentalists who work with ONA-star polymer conjugates on how the selection of ONA chemistry in ONA-star polymer conjugates impacts the overall thermoresponsive behavior and assembly of these systems. Future directions for this work include, but are not limited to, understanding behavior of ONA-star polymer conjugates at higher concentrations where percolation is unavoidable, and other polymer chemistries leading to semi-flexibility, non-athermal and/or charged interactions in the star polymers.

Chapter 5

CONCLUSION AND FUTURE DIRECTIONS

To conclude, through the development of coarse-grained (CG) modeling and utilization of molecular dynamics simulations we developed a better understanding of the thermodynamics driving thermoresponsive and assembly behavior. In this thesis, taking inspiration from previous CG models[30, 153], I developed coarse-grained models for a small selection of biopolymers, namely elastin-like polypeptide (ELP), collagen-like polypeptide (CLP), and oligonucleic acid (ONA)-star polymer conjugates for studying their individual thermoresponsive behavior. In Chapter 2, I presented the CG modeling and simulation of ELP-CLP conjugates to elucidate the effect of conjugation on the thermodynamics of ELP's aggregation behavior[28]. In their experiments, Luo and Kiick[16] observed that the conjugation of ELP strands to CLP triple helices significantly reduced the LCST of ELP with respect to unconjugated ELP strands in solution. Similarly, our simulation results indicated a reduction in the attractive pairwise interaction strength required for the onset of aggregation of ELP-CLP conjugate systems with respect to free ELP model strand systems for equivalent volume fractions of ELP model strands. Our reasoning for this behavior is that by conjugating multiple ELP strands to a single rigid body, due to crowding the loss in entropy upon aggregation is reduced, thus requiring less of an enthalpic driving force to balance the change in free energy upon aggregation. The simplicity of modeling ELP as a bead-spring polymer, CLP as a rigid body, and implicitly modeling solvent allows one flexibility in terms of being able to model not

only ELP-CLP conjugates, but any LCST-like polymer conjugated to a rigid macromolecule. Additionally, such a CG approach significantly reduces the amount of simulation time necessary to explore phase transitions as compared to atomistic simulations. Lastly, CG simulations provide access to a wide parameter space which may be difficult to screen experimentally or using atomistic simulations.

In Chapter 3, I presented our phenomenological model for CLP that adopted the directional, specific interaction sites from a recent CG model of oligonucleic acids developed by Ghobadi and Jayaraman [30] to represent interstrand H-bonds that stabilize triple helical conformations of CLP strands in solution. The sequential development of two CG models for CLP was presented in a pedagogical fashion to show how and why the second model was needed to improve upon the first. Both models for CLP were used to study the effect of CLP strand length and CLP sequence on the melting behavior of CLP triple helices in solution. While the first model did not qualitatively capture the experimentally observed trend of increasing melting temperature of CLP triple helices with increasing CLP strand length, the incorporation of addition of dihedral bonded potentials in the second model stiffened the CLP strands and stabilized the H-bonds that lead to triple helix formation, and captured the correct trend of melting temperature versus CLP length[67]. Next, following upon recent work by O'Leary *et al.*[84] that showed that CLP strands with the sequence $(PKG)_4(POG)_4(DOG)_4$ form triple helices with melting temperatures well below triple helices with the CLP sequence $(POG)_{12}$, we performed simulations with both CG CLP models. Both models corroborated this behavior by showing a reduction in the melting transition temperature of the $(PKG)_4(POG)_4(DOG)_4$ system with respect to the $(POG)_{12}$ system. The electrostatic repulsion of like charges within proximity of one

another in the $(PKG)_4(POG)_4(DOG)_4$ strand decrease the stability of the triple helices in comparison to $(POG)_{12}$ sequences. Lastly, the CLP model developed in Chapter 3 was substituted in the ELP-CLP conjugate model of Chapter 2 where CLP was represented with a rigid body. This substitution of rod-like rigid body CLP model with the improved second model of CLP enables us to study of ELP-CLP conjugate thermoresponsive behavior with CLP in melted or hybridized state. Results from the second model for ELP-CLP conjugates showed the same behavior as seen in Chapter 1 where the ELP-CLP conjugate system had a lower onset of aggregation than the corresponding unconjugated ELP system. This confirms that this newly developed CG model of CLP can also be utilized in conjugates with other inverse-transition polymers (e.g. ELP-CLP or PNIPAM-CLP) to capture the correct physics in this diblock conjugates as well. Future work with the ELP-CLP conjugate using the new CLP model will be aimed at studying the aggregation behavior of ELP-CLP conjugates above the CLP melting temperature as well as how conjugation of ELP to CLP impacts CLP melting. Dual thermoresponsive self-assembly of ELP-CLP conjugates has shown promise in drug delivery applications[16] and by using CG simulations one could investigate how ELP-CLP conjugate thermoresponsive behavior changes with the design of the ELP-CLP conjugate.

In Chapter 4, a CG model for another class of thermoresponsive polymer, oligonucleic acid (ONA), in conjugate with star polymers was simulated using the model developed by Ghobadi and Jayaraman[30, 153] to understand how melting and self-assembly of these ONA-star polymers change with ONA backbone charge, ONA backbone flexibility, and star polymer architecture (i.e. number of arms in the star polymer)[42]. Through hybridization analysis, a decrease in backbone charge and

backbone flexibility was shown to lead to an increase in the melting temperature of complementary ONA strands in the ONA-star polymer conjugate system, while star polymer architecture did not impact hybridization. As the charge on the ONA backbone decreases, the repulsion of like charges from complementary ONA strands is reduced, leading to a more enthalpically stable hybridized structure and higher melting temperature. As the flexibility of the ONA backbone decreases, the change in entropy upon melting is reduced, thus allowing a more stable ONA hybridized structure and a higher melting temperature. Furthermore, the assembly of ONA-star polymer conjugates was also probed by evaluating the average size of clusters present in ONA-star polymer simulations. The results of the assembly analysis demonstrated that decreasing both ONA backbone charge and ONA backbone flexibility led to an increase in the assembly temperature. Thermodynamic arguments for assembly behavior follow the above discussion of melting behavior since the two phenomena are closely linked. Additionally, increasing the number of arms of the star polymer in the ONA-star polymer conjugate increased the assembly temperature. This is because of an increase in the number of possible hybridization sites per ONA-star polymer conjugate with increasing number of arms in addition to a subsequent reduction in the change in entropy upon assembly due to fewer conjugates per system for higher number of arms (to maintain the same ONA strand volume fraction across various star polymer architectures). Lastly, percolation was studied with respect to ONA backbone charge, ONA backbone flexibility, and star polymer architecture and similar trends were observed, though the results were less clear than the previous two analytical studies due to a relatively low concentration of ONA-star polymer conjugates. Future work on ONA-star polymer conjugates will be aimed at studying these systems at

higher concentration. Additionally, mechanical properties of the assemblies formed in these systems could be investigated with respect to ONA backbone flexibility, ONA backbone charge, number of star polymer arms, ONA sequence, along with many other parameters of interest in this system. By evaluating both thermodynamic and mechanical properties, design of ONA-star polymer conjugate gels can be guided to provide the appropriate set of molecules for a given application while reducing the cost and time associated with screening an experimental design space.

Outside of the recommendations listed in the paragraphs above, the CG models developed throughout the course of this thesis can be used to study other questions within the same materials presented here or extended to new classes of soft materials with thermally responsive behavior. For example, the improved ELP-CLP conjugate model illustrated in chapter 3 can be used to study how ELP:CLP molecular weight ratios in ELP-CLP conjugates impact the self-assembly of these molecules with respect to morphology, curvature of interface, and permeability of small drug molecules. One could then inform experimentalists of the morphologies of ELP-CLP conjugate assemblies that selecting certain molecular weights of either polypeptide would achieve. Another example of expanding upon the usage of these CG models would be to perform simulations with a high concentration of CLP model triple helices to study fibril self-assembly behavior. With respect to the ONA-star polymer conjugate CG model, careful selection of sequence and assembly protocol in the simulation can allow for the study of ONA-star polymer conjugate nanogels. Similar to DNA dendrimers, these molecules could potentially be used as a biocompatible, nontoxic drug-delivery/drug-release mechanism with multiple temperature stimuli depending on the size and sequence of the ONA-star polymer conjugate nanogel.

REFERENCES

1. Despanie, J., J.P. Dhandhukia, S.F. Hamm-Alvarez, and J.A. MacKay, *Elastin-like polypeptides: Therapeutic applications for an emerging class of nanomedicines.* J Control Release, 2016. **240**: p. 93-108.

2. Gandhi, A., A. Paul, S.O. Sen, and K.K. Sen, *Studies on thermoresponsive polymers: Phase behaviour, drug delivery and biomedical applications.* Asian Journal of Pharmaceutical Sciences, 2015. **10**(2): p. 99-107.

3. Kyle, S., A. Aggeli, E. Ingham, and M.J. McPherson, *Production of self-assembling biomaterials for tissue engineering.* Trends in Biotechnology, 2009. **27**(7): p. 423-433.

4. Kedracki, D., I. Safir, N. Gour, K.X. Ngo, and C. Vebert-Nardin, *DNA–Polymer Conjugates: From Synthesis, Through Complex Formation and Self-assembly to Applications*, in *Bio-synthetic Polymer Conjugates*, H. Schlaad, Editor. 2013, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 115-149.

5. Kundu, A., S. Nandi, and A.K. Nandi, *Nucleic acid based polymer and nanoparticle conjugates: Synthesis, properties and applications.* Progress in Materials Science, 2017. **88**: p. 136-185.

6. Peterson, A.M. and J.M. Heemstra, *Controlling self-assembly of DNA-polymer conjugates for applications in imaging and drug delivery.* Wiley Interdisciplinary Reviews: Nanomedicine and Nanobiotechnology, 2015. **7**(3): p. 282-297.

7. Buehler, M.J., *Atomistic and continuum modeling of mechanical properties of collagen: Elasticity, fracture, and self-assembly.* Journal of Materials Research, 2006. **21**(8): p. 1947-1961.

8. Dans, P.D., J. Walther, H. Gómez, and M. Orozco, *Multiscale simulation of DNA.* Current Opinion in Structural Biology, 2016. **37**: p. 29-45.

9. Domene, C., C. Jorgensen, and S.W. Abbasi, *A perspective on structural and computational work on collagen.* Physical Chemistry Chemical Physics, 2016. **18**(36): p. 24802-24811.

10. Zhao, B., N.K. Li, Y.G. Yingling, and C.K. Hall, *LCST Behavior is Manifested in a Single Molecule: Elastin-like polypeptide (VPGVG) n.* Biomacromolecules, 2015. **17**(1): p. 111-118.

11. Cho, Y., Y. Zhang, T. Christensen, L.B. Sagle, A. Chilkoti, and P.S. Cremer, *Effects of Hofmeister Anions on the Phase Transition Temperature of Elastin-like Polypeptides.* The journal of physical chemistry. B, 2008. **112**(44): p. 13765-13771.

12. Meyer, D.E. and A. Chilkoti, *Quantification of the effects of chain length and concentration on the thermal behavior of elastin-like polypeptides.* Biomacromolecules, 2004. **5**(3): p. 846-851.

13. Urry, D.W., T.L. Trapane, and K.U. Prasad, *Phase-structure transitions of the elastin polypentapeptide–water system within the framework of composition–temperature studies.* Biopolymers, 1985. **24**(12): p. 2345-2356.

14. Ghoorchian, A., J.T. Cole, and N.B. Holland, *Thermoreversible micelle formation using a three-armed star elastin-like polypeptide.* Macromolecules, 2010. **43**(9): p. 4340-4345.

15. Luo, T., L. He, P. Theato, and K.L. Kiick, *Thermoresponsive Self-Assembly of Nanostructures from a Collagen-Like Peptide-Containing Diblock Copolymer.* Macromolecular Bioscience, 2015. **15**(1): p. 111-123.

16. Luo, T. and K.L. Kiick, *Noncovalent Modulation of the Inverse Temperature Transition and Self-Assembly of Elastin-b-Collagen-like Peptide Bioconjugates.* Journal of the American Chemical Society, 2015. **137**(49): p. 15362-15365.

17. Paik, B.A., M.A. Blanco, X. Jia, C.J. Roberts, and K.L. Kiick, *Aggregation of poly (acrylic acid)-containing elastin-mimetic copolymers.* Soft Matter, 2015. **11**(9): p. 1839-1850.

18. Hassouneh, W., E.B. Zhulina, A. Chilkoti, and M. Rubinstein, *Elastin-like polypeptide diblock copolymers self-assemble into weak micelles.* Macromolecules, 2015. **48**(12): p. 4183-4195.

19. Bella, J., *Collagen structure: new tricks from a very old dog.* Biochemical Journal, 2016. **473**(8): p. 1001.

20. Brodsky, B. and A.V. Persikov, *Molecular Structure of the Collagen Triple Helix*, in *Advances in Protein Chemistry*. 2005, Academic Press. p. 301-339.

21. Brodsky, B. and J.A.M. Ramshaw, *The collagen triple-helix structure.* Matrix Biology, 1997. **15**(8): p. 545-554.

22. Shoulders, M.D. and R.T. Raines, *Collagen Structure and Stability.* Annual Review of Biochemistry, 2009. **78**(1): p. 929-958.

23. Suárez, E., N. Díaz, and D. Suárez, *Entropic Control of the Relative Stability of Triple-helical Collagen Peptide Models.* The Journal of Physical Chemistry B, 2008. **112**(47): p. 15248-15255.

24. Koide, T., *Triple helical collagen-like peptides: engineering and applications in matrix biology*. Connective tissue research, 2005. **46**(3): p. 131-141.

25. Luo, T. and K.L. Kiick, *Collagen-like peptides and peptide–polymer conjugates in the design of assembled materials.* European Polymer Journal, 2013. **49**(10): p. 2998-3009.

26. Parmar, P.A., L.W. Chow, J.-P. St-Pierre, C.-M. Horejs, Y.Y. Peng, J.A. Werkmeister, J.A.M. Ramshaw, and M.M. Stevens, *Collagen-mimetic peptide-modifiable hydrogels for articular cartilage regeneration.* Biomaterials, 2015. **54**: p. 213-225.

27. Yamazaki, C.M., I. Nakase, H. Endo, S. Kishimoto, Y. Mashiyama, R. Masuda, S. Futaki, and T. Koide, *Collagen-like Cell-Penetrating Peptides.* Angewandte Chemie International Edition, 2013. **52**(21): p. 5497-5500.

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

29. Beck, K. and B. Brodsky, *Supercoiled Protein Motifs: The Collagen Triple-Helix and the α-Helical Coiled Coil.* Journal of Structural Biology, 1998. **122**(1): p. 17-29.

30. Ghobadi, A.F. and A. Jayaraman, *Effect of backbone chemistry on hybridization thermodynamics of oligonucleic acids: a coarse-grained molecular dynamics simulation study.* Soft matter, 2016. **12**(8): p. 2276-2287.

31. Sim, A.Y.L., *Nucleic acid polymeric properties and electrostatics: Directly comparing theory and simulation with experiment.* Advances in Colloid and Interface Science, 2016. **232**: p. 49-56.

32. Jin, R., G. Wu, Z. Li, C.A. Mirkin, and G.C. Schatz, *What Controls the Melting Properties of DNA-Linked Gold Nanoparticle Assemblies?* Journal of the American Chemical Society, 2003. **125**(6): p. 1643-1654.

33. Hurst, S.J., H.D. Hill, and C.A. Mirkin, *"Three-Dimensional Hybridization" with Polyvalent DNA−Gold Nanoparticle Conjugates.* Journal of the American Chemical Society, 2008. **130**(36): p. 12192-12200.

34. Seifpour, A., S.R. Dahl, and A. Jayaraman, *Molecular simulation study of assembly of DNA-grafted nanoparticles: effect of bidispersity in DNA strand length.* Molecular Simulation, 2014. **40**(14): p. 1085-1098.

35. Seifpour, A., S.R. Dahl, B. Lin, and A. Jayaraman, *Molecular simulation study of the assembly of DNA-functionalised nanoparticles: Effect of DNA strand sequence and composition.* Molecular Simulation, 2013. **39**(9): p. 741-753.

36. Koshkin, A.A., P. Nielsen, M. Meldgaard, V.K. Rajwanshi, S.K. Singh, and J. Wengel, *LNA (Locked Nucleic Acid): An RNA Mimic Forming Exceedingly Stable LNA:LNA Duplexes.* Journal of the American Chemical Society, 1998. **120**(50): p. 13252-13253.

37. Koshkin, A.A., S.K. Singh, P. Nielsen, V.K. Rajwanshi, R. Kumar, M. Meldgaard, C.E. Olsen, and J. Wengel, *LNA (Locked Nucleic Acids): Synthesis of the adenine, cytosine, guanine, 5-methylcytosine, thymine and uracil bicyclonucleoside monomers, oligomerisation, and unprecedented nucleic acid recognition.* Tetrahedron, 1998. **54**(14): p. 3607-3630.

38. Nielsen, P., M. Egholm, R. Berg, and O. Buchardt, *Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide.* Science, 1991. **254**(5037): p. 1497-1500.

39. Averick, S., E. Paredes, W. Li, K. Matyjaszewski, and S.R. Das, *Direct DNA Conjugation to Star Polymers for Controlled Reversible Assemblies.* Bioconjugate Chemistry, 2011. **22**(10): p. 2030-2037.

40. Okochi, K.D., L. Monfregola, S.M. Dickerson, R. McCaffrey, D.W. Domaille, C. Yu, G.R. Hafenstine, Y. Jin, J.N. Cha, R.D. Kuchta, M. Caruthers, and W. Zhang, *Synthesis of Small-Molecule/DNA Hybrids through On-Bead Amide-Coupling Approach.* The Journal of Organic Chemistry, 2017.

41. Leumann, C.J., *DNA Analogues: From Supramolecular Principles to Biological Properties.* Bioorganic & Medicinal Chemistry, 2002. **10**(4): p. 841-854.

42. Condon, J.E. and A. Jayaraman, *Effect of oligonucleic acid (ONA) backbone features on assembly of ONA-star polymer conjugates: a coarse-grained molecular simulation study.* Soft Matter, 2017. **13**(38): p. 6770-6783.

43. Ward, M.A. and T.K. Georgiou, *Thermoresponsive Polymers for Biomedical Applications.* Polymers, 2011. **3**(3): p. 1215-1242.

44. Schild, H.G., *POLY (N-ISOPROPYLACRYLAMIDE) - EXPERIMENT, THEORY AND APPLICATION.* Progress in Polymer Science, 1992. **17**(2): p. 163-249.

45. Schild, H.G. and D.A. Tirrell, *MICROCALORIMETRIC DETECTION OF LOWER CRITICAL SOLUTION TEMPERATURES IN AQUEOUS POLYMER-SOLUTIONS.* Journal of Physical Chemistry, 1990. **94**(10): p. 4352-4356.

46. Liu, M.S., C. Taylor, B. Chong, L.H. Liu, A. Bilic, N.S. Terefe, R. Stockmann, S.H. Thang, and K. De Silva, *Conformational transitions and dynamics of thermal responsive poly(N-isopropylacrylamide) polymers as revealed by molecular simulation.* European Polymer Journal, 2014. **55**: p. 153-159.

47. Abbott, L.J. and M.J. Stevens, *A temperature-dependent coarse-grained model for the thermoresponsive polymer poly(N-isopropylacrylamide).* Journal of Chemical Physics, 2015. **143**(24): p. 10.

48. Rissanou, A.N., S.H. Anastasiadis, and I.A. Bitsanis, *Monte Carlo study of the coil-to-globule transition of a model polymeric system.* Journal of Polymer Science Part B-Polymer Physics, 2006. **44**(24): p. 3651-3666.

49. Fujishige, S., K. Kubota, and I. Ando, *Phase transition of aqueous solutions of poly(N-isopropylacrylamide) and poly(N-isopropylmethacrylamide).* The Journal of Physical Chemistry, 1989. **93**(8): p. 3311-3313.

50. Kubota, K., S. Fujishige, and I. Ando, *SINGLE-CHAIN TRANSITION OF POLY(N-ISOPROPYLACRYLAMIDE) IN WATER.* Journal of Physical Chemistry, 1990. **94**(12): p. 5154-5158.

51. Min, S.H., S.K. Kwak, and B.S. Kim, *Atomistic simulation for coil-to-globule transition of poly(2-dimethylaminoethyl methacrylate).* Soft Matter, 2015. **11**(12): p. 2423-2433.

52. Matanović, M.R., J. Kristl, and P.A. Grabnar, *Thermoresponsive polymers: insights into decisive hydrogel characteristics, mechanisms of gelation, and*

promising biomedical applications. International Journal of Pharmaceutics, 2014. **472**(1): p. 262-275.

53. Li, N.K., F.G. Quiroz, C.K. Hall, A. Chilkoti, and Y.G. Yingling, *Molecular description of the LCST behavior of an elastin-like polypeptide.* Biomacromolecules, 2014. **15**(10): p. 3522-3530.

54. Zhao, B., N.K. Li, Y.G. Yingling, and C.K. Hall, *LCST Behavior is Manifested in a Single Molecule: Elastin-Like polypeptide (VPGVG)n.* Biomacromolecules, 2016. **17**(1): p. 111-118.

55. Lutz, J.-F. and A. Hoth, *Preparation of ideal PEG analogues with a tunable thermosensitivity by controlled radical copolymerization of 2-(2-methoxyethoxy) ethyl methacrylate and oligo (ethylene glycol) methacrylate.* Macromolecules, 2006. **39**(2): p. 893-896.

56. Yeboah, A., R.I. Cohen, C. Rabolli, M.L. Yarmush, and F. Berthiaume, *Elastin-like polypeptides: A strategic fusion partner for biologics.* Biotechnol Bioeng, 2016. **113**(8): p. 1617-27.

57. Brown, F.R., III, A.J. Hopfinger, and E.R. Blout, *The collagen-like triple helix to random-chain transition: Experiment and theory.* Journal of Molecular Biology, 1972. **63**(1): p. 101-115.

58. Engel, J. and H.P. Bächinger, *Structure, Stability and Folding of the Collagen Triple Helix,* in *Collagen: Primer in Structure, Processing and Assembly*, J. Brinckmann, H. Notbohm, and P.K. Müller, Editors. 2005, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 7-33.

59. Heskins, M. and J.E. Guillet, *Solution Properties of Poly(N-isopropylacrylamide).* Journal of Macromolecular Science: Part A - Chemistry, 1968. **2**(8): p. 1441-1455.

60. Lorenzo, A.C. and E.R. Caffarena, *Elastic properties, Young's modulus determination and structural stability of the tropocollagen molecule: a computational study by steered molecular dynamics.* Journal of Biomechanics, 2005. **38**(7): p. 1527-1533.

61. Raman, S.S., R. Parthasarathi, V. Subramanian, and T. Ramasami, *Role of Aspartic Acid in Collagen Structure and Stability: A Molecular Dynamics Investigation.* The Journal of Physical Chemistry B, 2006. **110**(41): p. 20678-20685.

62. Plimpton, S., *Fast parallel algorithms for short-range molecular dynamics.* Journal of Computational Physics, 1995. **117**(1): p. 1-19.

63. Miller, T.F., III, M. Eleftheriou, P. Pattnaik, A. Ndirango, D. Newns, and G.J. Martyna, *Symplectic quaternion scheme for biophysical molecular dynamics.* The Journal of Chemical Physics, 2002. **116**(20): p. 8649-8659.

64. Frenkel, D. and B. Smit, *Understanding Molecular Simulations*. 2nd ed. Computational Science Series. 2002, San Diego: Academic Press.

65. Weeks, J.D., D. Chandler, and H.C. Andersen, *Role of repulsive forces in determining the equilibrium structure of simple liquids.* The Journal of Chemical Physics, 1971. **54**(12): p. 5237-5247.

66. Sakakibara, S., K. Inouye, K. Shudo, Y. Kishida, Y. Kobayashi, and D.J. Prockop, *Synthesis of (Pro-Hyp-Gly)n of defined molecular weights Evidence for the stabilization of collagen triple helix by hydroxypyroline.* Biochimica et Biophysica Acta (BBA) - Protein Structure, 1973. **303**(1): p. 198-202.

67. Persikov, A.V., J.A.M. Ramshaw, and B. Brodsky, *Prediction of Collagen Stability from Amino Acid Sequence.* Journal of Biological Chemistry, 2005. **280**(19): p. 19343-19349.

68. An, B., Y.-S. Lin, and B. Brodsky, *Collagen interactions: Drug design and delivery.* Advanced Drug Delivery Reviews, 2016. **97**(Supplement C): p. 69-84.

69. Kojima, C., S. Tsumura, A. Harada, and K. Kono, *A Collagen-Mimic Dendrimer Capable of Controlled Release.* Journal of the American Chemical Society, 2009. **131**(17): p. 6052-6053.

70. Yamazaki, C.M., I. Nakase, H. Endo, S. Kishimoto, Y. Mashiyama, R. Masuda, S. Futaki, and T. Koide, *Collagen-like Cell-Penetrating Peptides.* Angewandte Chemie International Edition, 2013. **52**(21): p. 5497-5500.

71. Yasui, H., C.M. Yamazaki, H. Nose, C. Awada, T. Takao, and T. Koide, *Potential of collagen-like triple helical peptides as drug carriers: Their in vivo distribution, metabolism, and excretion profiles in rodents.* Peptide Science, 2013. **100**(6): p. 705-713.

72. Koide, T., *Collagen-like triple helical peptides: applications in drug development and regenerative medicine.* Yakugaku Zasshi, 2013. **133**(3): p. 387-92.

73. Chattopadhyay, S. and R.T. Raines, *Review collagen-based biomaterials for wound healing.* Biopolymers, 2014. **101**(8): p. 821-833.

74. Fields, G.B., *Synthesis and biological applications of collagen-model triple-helical peptides.* Organic & Biomolecular Chemistry, 2010. **8**(6): p. 1237-1258.

75. Golinska, M.D., M.K. Włodarczyk-Biegun, M.W.T. Werten, M.A.C. Stuart, F.A. de Wolf, and R. de Vries, *Dilute Self-Healing Hydrogels of Silk-Collagen-Like Block Copolypeptides at Neutral pH.* Biomacromolecules, 2014. **15**(3): p. 699-706.

76. Hernandez-Gordillo, V. and J. Chmielewski, *Mimicking the extracellular matrix with functionalized, metal-assembled collagen peptide scaffolds.* Biomaterials, 2014. **35**(26): p. 7363-7373.

77. Islam, M.M., R. Ravichandran, D. Olsen, M.K. Ljunggren, P. Fagerholm, C.J. Lee, M. Griffith, and J. Phopase, *Self-assembled collagen-like-peptide implants as alternatives to human donor corneal transplantation.* RSC Advances, 2016. **6**(61): p. 55745-55749.

78. Pires, M.M., D.E. Przybyla, and J. Chmielewski, *A Metal–Collagen Peptide Framework for Three-Dimensional Cell Culture.* Angewandte Chemie International Edition, 2009. **48**(42): p. 7813-7817.

79. Reyes, C.D. and A.J. García, *Engineering integrin-specific surfaces with a triple-helical collagen-mimetic peptide.* Journal of Biomedical Materials Research Part A, 2003. **65A**(4): p. 511-523.

80. Rubert Pérez, C.M., A. Panitch, and J. Chmielewski, *A Collagen Peptide-Based Physical Hydrogel for Cell Encapsulation.* Macromolecular Bioscience, 2011. **11**(10): p. 1426-1431.

81. Woolfson, D.N., *Building fibrous biomaterials from α-helical and collagen-like coiled-coil peptides.* Peptide Science, 2010. **94**(1): p. 118-127.

82. Bai, H., K. Xu, Y. Xu, and H. Matsui, *Fabrication of Au Nanowires of Uniform Length and Diameter Using a Monodisperse and Rigid Biomolecular Template: Collagen-like Triple Helix.* Angewandte Chemie International Edition, 2007. **46**(18): p. 3319-3322.

83. Gottlieb, D., S.A. Morin, S. Jin, and R.T. Raines, *Self-assembled collagen-like peptide fibers as templates for metallic nanowires.* Journal of Materials Chemistry, 2008. **18**(32): p. 3865-3870.

84. O'Leary, L.E.R., J.A. Fallas, E.L. Bakota, M.K. Kang, and J.D. Hartgerink, *Multi-hierarchical self-assembly of a collagen mimetic peptide from triple helix to nanofibre and hydrogel.* Nat Chem, 2011. **3**(10): p. 821-828.

85. Sarkar, B., L.E.R. O'Leary, and J.D. Hartgerink, *Self-Assembly of Fiber-Forming Collagen Mimetic Peptides Controlled by Triple-Helical Nucleation.* Journal of the American Chemical Society, 2014. **136**(41): p. 14417-14424.

86. Rubert Perez, C.M., L.A. Rank, and J. Chmielewski, *Tuning the thermosensitive properties of hybrid collagen peptide-polymer hydrogels.* Chemical Communications, 2014. **50**(60): p. 8174-8176.

87. Jin, H.-E., J. Jang, J. Chung, H.J. Lee, E. Wang, S.-W. Lee, and W.-J. Chung, *Biomimetic Self-Templated Hierarchical Structures of Collagen-Like Peptide Amphiphiles.* Nano Letters, 2015. **15**(10): p. 7138-7145.

88. Bella, J., *A new method for describing the helical conformation of collagen: Dependence of the triple helical twist on amino acid sequence.* Journal of Structural Biology, 2010. **170**(2): p. 377-391.

89. Berisio, R., L. Vitagliano, L. Mazzarella, and A. Zagari, *Crystal structure of the collagen triple helix model [(Pro-Pro-Gly)(10)](3).* Protein Science : A Publication of the Protein Society, 2002. **11**(2): p. 262-270.

90. Okuyama, K., C. Hongo, G. Wu, K. Mizuno, K. Noguchi, S. Ebisuzaki, Y. Tanaka, N. Nishino, and H.P. Bächinger, *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): p. 361-372.

91. Persikov, A.V., J.A.M. Ramshaw, A. Kirkpatrick, and B. Brodsky, *Amino Acid Propensities for the Collagen Triple-Helix.* Biochemistry, 2000. **39**(48): p. 14960-14967.

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

93. Shoulders, M.D. and R.T. Raines, *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*, S.D. Valle, E. Escher, and W.D. Lubell, Editors. 2009, Springer New York: New York, NY. p. 251-252.

94. Ying Chow, W., D. Bihan, C.J. Forman, D.A. Slatter, D.G. Reid, D.J. Wales, R.W. Farndale, and M.J. Duer, *Hydroxyproline Ring Pucker Causes Frustration of Helix Parameters in the Collagen Triple Helix.* 2015. **5**: p. 12556.

95. Khew, S.T. and Y.W. Tong, *Characterization of triple-helical conformations and melting analyses of synthetic collagen-like peptides by reversed-phase HPLC.* Journal of Chromatography B, 2007. **858**(1): p. 79-90.

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

97. Gopalakrishnan, R., E.R. Azhagiya Singam, J. Vijaya Sundar, and V. Subramanian, *Interaction of collagen like peptides with gold nanosurfaces: a molecular dynamics investigation.* Physical Chemistry Chemical Physics, 2015. **17**(7): p. 5172-5186.

98. Jones, J.E., *On the Determination of Molecular Fields. II. From the Equation of State of a Gas.* Proceedings of the Royal Society of London. Series A, 1924. **106**(738): p. 463-477.

99. Abraham, M.J., T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, and E. Lindahl, *GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers.* SoftwareX, 2015. **1–2**: p. 19-25.

100. Stogryn, A., *Equations for Calculating the Dielectric Constant of Saline Water (Correspondence).* IEEE Transactions on Microwave Theory and Techniques, 1971. **19**(8): p. 733-736.

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

102. Tuckerman, M., B.J. Berne, and G.J. Martyna, *Reversible multiple time scale molecular dynamics.* The Journal of Chemical Physics, 1992. **97**(3): p. 1990-2001.

103. Hockney, R.W. and J.W. Eastwood, *Computer Simulation Using Particles*. 1988: CRC Press.

104. Berisio, R., L. Vitagliano, L. Mazzarella, and A. Zagari, *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): p. 8-13.

105. Van Nguyen, K. and S.D. Minteer, *Investigating DNA hydrogels as a new biomaterial for enzyme immobilization in biobatteries.* Chemical Communications, 2015. **51**(66): p. 13071-13073.

106. Vura-Weis, J., M.R. Wasielewski, A.K. Thazhathveetil, and F.D. Lewis, *Efficient Charge Transport in DNA Diblock Oligomers.* Journal of the American Chemical Society, 2009. **131**(28): p. 9722-9727.

107. Zinchenko, A., Y. Che, S. Taniguchi, L.I. Lopatina, V. G. Sergeyev, and S. Murata, *Metallization of DNA hydrogel: application of soft matter host for preparation and nesting of catalytic nanoparticles.* Journal of Nanoparticle Research, 2016. **18**(7): p. 179.

108. Zinchenko, A., Y. Miwa, L.I. Lopatina, V.G. Sergeyev, and S. Murata, *DNA Hydrogel as a Template for Synthesis of Ultrasmall Gold Nanoparticles for Catalytic Applications.* ACS Applied Materials & Interfaces, 2014. **6**(5): p. 3226-3232.

109. Banga, R.J., S.A. Krovi, S.P. Narayan, A.J. Sprangers, G. Liu, C.A. Mirkin, and S.T. Nguyen, *Drug-Loaded Polymeric Spherical Nucleic Acids: Enhancing Colloidal Stability and Cellular Uptake of Polymeric Nanoparticles through DNA Surface-Functionalization.* Biomacromolecules, 2017. **18**(2): p. 483-489.

110. Chen, C., J. Geng, F. Pu, X. Yang, J. Ren, and X. Qu, *Polyvalent Nucleic Acid/Mesoporous Silica Nanoparticle Conjugates: Dual Stimuli-Responsive Vehicles for Intracellular Drug Delivery.* Angewandte Chemie International Edition, 2011. **50**(4): p. 882-886.

111. Ding, Y., Z. Jiang, K. Saha, C.S. Kim, S.T. Kim, R.F. Landis, and V.M. Rotello, *Gold Nanoparticles for Nucleic Acid Delivery.* Molecular Therapy, 2014. **22**(6): p. 1075-1083.

112. Jensen, S.A., E.S. Day, C.H. Ko, L.A. Hurley, J.P. Luciano, F.M. Kouri, T.J. Merkel, A.J. Luthi, P.C. Patel, J.I. Cutler, W.L. Daniel, A.W. Scott, M.W. Rotz, T.J. Meade, D.A. Giljohann, C.A. Mirkin, and A.H. Stegh, *Spherical Nucleic Acid Nanoparticle Conjugates as an RNAi-Based Therapy for Glioblastoma.* Science Translational Medicine, 2013. **5**(209): p. 209ra152.

113. Kim, M.-G., Y. Shon, W. Miao, J. Lee, and Y.-K. Oh, *Biodegradable graphene oxide and polyaptamer DNA hybrid hydrogels for implantable drug delivery.* Carbon, 2016. **105**: p. 14-22.

114. Moroz, E., S.H. Lee, K. Yamada, F. Halloy, S. Martínez-Montero, H. Jahns, J. Hall, M.J. Damha, B. Castagner, and J.-C. Leroux, *Carrier-free Gene Silencing by Amphiphilic Nucleic Acid Conjugates in Differentiated Intestinal Cells.* Molecular Therapy - Nucleic Acids, 2016. **5**: p. e364.

115. Nishida, Y., S. Ohtsuki, Y. Araie, Y. Umeki, M. Endo, T. Emura, K. Hidaka, H. Sugiyama, Y. Takahashi, Y. Takakura, and M. Nishikawa, *Self-assembling DNA hydrogel-based delivery of immunoinhibitory nucleic acids to immune cells.* Nanomedicine: Nanotechnology, Biology and Medicine, 2016. **12**(1): p. 123-130.

116. Nishikawa, M., Y. Mizuno, K. Mohri, N. Matsuoka, S. Rattanakiat, Y. Takahashi, H. Funabashi, D. Luo, and Y. Takakura, *Biodegradable CpG DNA hydrogels for sustained delivery of doxorubicin and immunostimulatory signals in tumor-bearing mice.* Biomaterials, 2011. **32**(2): p. 488-494.

117. Nishikawa, M., K. Ogawa, Y. Umeki, K. Mohri, Y. Kawasaki, H. Watanabe, N. Takahashi, E. Kusuki, R. Takahashi, Y. Takahashi, and Y. Takakura, *Injectable, self-gelling, biodegradable, and immunomodulatory DNA hydrogel for antigen delivery.* Journal of Controlled Release, 2014. **180**: p. 25-32.

118. Sharma, A., T. Garg, A. Aman, K. Panchal, R. Sharma, S. Kumar, and T. Markandeywar, *Nanogel—an advanced drug delivery tool: Current and future.* Artificial Cells, Nanomedicine, and Biotechnology, 2016. **44**(1): p. 165-177.

119. Tan, X., B.B. Li, X. Lu, F. Jia, C. Santori, P. Menon, H. Li, B. Zhang, J.J. Zhao, and K. Zhang, *Light-Triggered, Self-Immolative Nucleic Acid-Drug*

Nanostructures. Journal of the American Chemical Society, 2015. **137**(19): p. 6112-6115.

120. Lu, C.-H., W. Guo, Y. Hu, X.-J. Qi, and I. Willner, *Multitriggered Shape-Memory Acrylamide–DNA Hydrogels.* Journal of the American Chemical Society, 2015. **137**(50): p. 15723-15731.

121. Guo, W., C.-H. Lu, R. Orbach, F. Wang, X.-J. Qi, A. Cecconello, D. Seliktar, and I. Willner, *pH-Stimulated DNA Hydrogels Exhibiting Shape-Memory Properties.* Advanced Materials, 2015. **27**(1): p. 73-78.

122. Du, Y., B.J. Lim, B. Li, Y.S. Jiang, J.L. Sessler, and A.D. Ellington, *Reagentless, Ratiometric Electrochemical DNA Sensors with Improved Robustness and Reproducibility.* Analytical Chemistry, 2014. **86**(15): p. 8010-8016.

123. Freeman, R., J. Girsh, and I. Willner, *Nucleic Acid/Quantum Dots (QDs) Hybrid Systems for Optical and Photoelectrochemical Sensing.* ACS Applied Materials & Interfaces, 2013. **5**(8): p. 2815-2834.

124. Gaylord, B.S., A.J. Heeger, and G.C. Bazan, *DNA detection using water-soluble conjugated polymers and peptide nucleic acid probes.* Proceedings of the National Academy of Sciences, 2002. **99**(17): p. 10954-10957.

125. Riahi, R., S. Wang, M. Long, N. Li, P.-Y. Chiou, D.D. Zhang, and P.K. Wong, *Mapping Photothermally Induced Gene Expression in Living Cells and Tissues by Nanorod-Locked Nucleic Acid Complexes.* ACS Nano, 2014. **8**(4): p. 3597-3605.

126. Wang, H., R. Yang, L. Yang, and W. Tan, *Nucleic Acid Conjugated Nanomaterials for Enhanced Molecular Recognition.* ACS Nano, 2009. **3**(9): p. 2451-2460.

127. Wu, Z., G.-Q. Liu, X.-L. Yang, and J.-H. Jiang, *Electrostatic Nucleic Acid Nanoassembly Enables Hybridization Chain Reaction in Living Cells for Ultrasensitive mRNA Imaging.* Journal of the American Chemical Society, 2015. **137**(21): p. 6829-6836.

128. Schöning, K.U., P. Scholz, S. Guntha, X. Wu, R. Krishnamurthy, and A. Eschenmoser, *Chemical Etiology of Nucleic Acid Structure: The α-Threofuranosyl-(3'→2') Oligonucleotide System.* Science, 2000. **290**(5495): p. 1347-1351.

129. Xi, W., S. Pattanayak, C. Wang, B. Fairbanks, T. Gong, J. Wagner, C.J. Kloxin, and C.N. Bowman, *Clickable Nucleic Acids: Sequence-Controlled Periodic Copolymer/Oligomer Synthesis by Orthogonal Thiol-X Reactions.* Angewandte Chemie International Edition, 2015. **54**(48): p. 14462-14467.

130. Li, T.I.N.G., R. Sknepnek, R.J. Macfarlane, C.A. Mirkin, and M. Olvera de la Cruz, *Modeling the Crystallization of Spherical Nucleic Acid Nanoparticle Conjugates with Molecular Dynamics Simulations.* Nano Letters, 2012. **12**(5): p. 2509-2514.

131. Morris, W., W.E. Briley, E. Auyeung, M.D. Cabezas, and C.A. Mirkin, *Nucleic Acid–Metal Organic Framework (MOF) Nanoparticle Conjugates.* Journal of the American Chemical Society, 2014. **136**(20): p. 7261-7264.

132. Pu, F., Z. Liu, J. Ren, and X. Qu, *Nucleic acid-mesoporous silica nanoparticle conjugates for keypad lock security operation.* Chemical Communications, 2013. **49**(23): p. 2305-2307.

133. Zheng, D., D.A. Giljohann, D.L. Chen, M.D. Massich, X.-Q. Wang, H. Iordanov, C.A. Mirkin, and A.S. Paller, *Topical delivery of siRNA-based spherical nucleic acid nanoparticle conjugates for gene regulation.* Proceedings of the National Academy of Sciences, 2012. **109**(30): p. 11975-11980.

134. Angioletti-Uberti, S., B.M. Mognetti, and D. Frenkel, *Theory and simulation of DNA-coated colloids: a guide for rational design.* Physical Chemistry Chemical Physics, 2016. **18**(9): p. 6373-6393.

135. Fernandez-Castanon, J., F. Bomboi, L. Rovigatti, M. Zanatta, A. Paciaroni, L. Comez, L. Porcar, C.J. Jafta, G.C. Fadda, T. Bellini, and F. Sciortino, *Small-angle neutron scattering and molecular dynamics structural study of gelling DNA nanostars.* The Journal of Chemical Physics, 2016. **145**(8): p. 084910.

136. Starr, F.W. and F. Sciortino, *Model for assembly and gelation of four-armed DNA dendrimers.* Journal of Physics: Condensed Matter, 2006. **18**(26): p. L347-L353.

137. Largo, J., F.W. Starr, and F. Sciortino, *Self-Assembling DNA Dendrimers: A Numerical Study.* Langmuir, 2007. **23**(11): p. 5896-5905.

138. Locatelli, E., P.H. Handle, C.N. Likos, F. Sciortino, and L. Rovigatti, *Condensation and Demixing in Solutions of DNA Nanostars and Their Mixtures.* ACS Nano, 2017. **11**(2): p. 2094-2102.

139. Rovigatti, L., F. Bomboi, and F. Sciortino, *Accurate phase diagram of tetravalent DNA nanostars.* The Journal of Chemical Physics, 2014. **140**(15): p. 154903.

140. Pan, W., H. Wen, L. Niu, C. Su, C. Liu, J. Zhao, C. Mao, and D. Liang, *Effects of chain flexibility on the properties of DNA hydrogels.* Soft Matter, 2016. **12**(25): p. 5537-5541.

141. Douglas, J.F., F.W. Starr, F. Vargas–Lara, A. D'Amore, D. Acierno, and L. Grassia, *Conformational nature of DNA–grafted chains on spherical gold nanoparticles.* AIP Conference Proceedings, 2016. **1736**(1): p. 020081.

142. Huang, Y., I. Szleifer, and N.A. Peppas, *A Molecular Theory of Polymer Gels.* Macromolecules, 2002. **35**(4): p. 1373-1380.

143. Lequieu, J.P., D.M. Hinckley, and J.J. de Pablo, *A molecular view of DNA-conjugated nanoparticle association energies.* Soft Matter, 2015. **11**(10): p. 1919-1929.

144. Li, T.I.N.G., R. Sknepnek, and M. Olvera de la Cruz, *Thermally Active Hybridization Drives the Crystallization of DNA-Functionalized Nanoparticles.* Journal of the American Chemical Society, 2013. **135**(23): p. 8535-8541.

145. Prytkova, T.R., I. Eryazici, B. Stepp, S.-B. Nguyen, and G.C. Schatz, *DNA Melting in Small-Molecule−DNA-Hybrid Dimer Structures: Experimental Characterization and Coarse-Grained Molecular Dynamics Simulations.* The Journal of Physical Chemistry B, 2010. **114**(8): p. 2627-2634.

146. Rovigatti, L., F. Smallenburg, F. Romano, and F. Sciortino, *Gels of DNA Nanostars Never Crystallize.* ACS Nano, 2014. **8**(4): p. 3567-3574.

147. Edwardson, T.G.W., K.M.M. Carneiro, C.J. Serpell, and H.F. Sleiman, *An Efficient and Modular Route to Sequence-Defined Polymers Appended to DNA.* Angewandte Chemie International Edition, 2014. **53**(18): p. 4567-4571.

148. Jia, Y., X. Zuo, X. Lou, M. Miao, Y. Cheng, X. Min, X. Li, and F. Xia, *Rational Designed Bipolar, Conjugated Polymer-DNA Composite Beacon for the Sensitive Detection of Proteins and Ions.* Analytical Chemistry, 2015. **87**(7): p. 3890-3894.

149. Kye, M. and Y.-b. Lim, *Reciprocal Self-Assembly of Peptide–DNA Conjugates into a Programmable Sub-10-nm Supramolecular Deoxyribonucleoprotein.* Angewandte Chemie International Edition, 2016. **55**(39): p. 12003-12007.

150. Lu, X., E. Watts, F. Jia, X. Tan, and K. Zhang, *Polycondensation of Polymer Brushes via DNA Hybridization.* Journal of the American Chemical Society, 2014. **136**(29): p. 10214-10217.

151. Madsen, M., R.S. Christensen, A. Krissanaprasit, M.R. Bakke, C.F. Riber, K.S. Nielsen, A. Zelikin, and K.V. Gothelf, *Preparation, single-molecule manipulation and energy transfer investigation of a polyfluorene-graft-DNA polymer.* Chemistry, 2017.

152. Pan, P., M. Fujita, W.-Y. Ooi, K. Sudesh, T. Takarada, A. Goto, and M. Maeda, *Thermoresponsive Micellization and Micellar Stability of Poly(N-isopropylacrylamide)-b-DNA Diblock and Miktoarm Star Polymers.* Langmuir, 2012. **28**(40): p. 14347-14356.

153. Ghobadi, A.F. and A. Jayaraman, *Effects of Polymer Conjugation on Hybridization Thermodynamics of Oligonucleic Acids.* The Journal of Physical Chemistry B, 2016. **120**(36): p. 9788-9799.

154. Levine, B.G., J.E. Stone, and A. Kohlmeyer, *Fast Analysis of Molecular Dynamics Trajectories with Graphics Processing Units—Radial Distribution Function Histogramming.* Journal of Computational Physics, 2011. **230**(9): p. 3556-3569.

155. Yakovchuk, P., E. Protozanova, and M.D. Frank-Kamenetskii, *Base-stacking and base-pairing contributions into thermal stability of the DNA double helix.* Nucleic Acids Research, 2006. **34**(2): p. 564-574.

Appendix

REPRINT PERMISSIONS

Note from RSC regarding reuse of work:

Authors contributing to RSC publications (journal articles, books or book chapters) do not need to formally request permission to reproduce material contained in this article provided that the correct acknowledgement is given with the reproduced material.

If the material has been adapted instead of reproduced from the original RSC publication "Reproduced from" can be substituted with "Adapted from".