Computational exploration of helical polymer geometry, folding, and structural stability.

by

Matthew John Williams

(Under the direction of Michael Bachmann)

Abstract

This computational study aims at a better understanding of structure formation processes in helical polymers. For this purpose, parallel tempering Monte Carlo simulations were preformed using a generic coarse-grained model. Structural conformations exhibit helical order with tertiary structural organizations including single helices, multiple helical segments organized into bundles, and disorganized helical arrangements. We compare helical structure formation for flexible and semiflexible polymers. For each model, we analyze low-energy structural geometries, structural stability, and folding dynamics. This exploration lends insight into the restricted flexibility of biological polymers such as double-stranded DNA and proteins. Additionally, the influence of substrate adsorption on helical structures and stability is investigated in detail.

Index words: Biopolymers, Protein Folding, Helix Bundles, Simulation, Adsorption, Stability

Computational exploration of helical polymer geometry, folding, and structural stability.

by

Matthew John Williams

B.A., University of Louisiana at Lafayette, 2008

A Dissertation Submitted to the Graduate Faculty
of The University of Georgia in Partial Fulfillment
of the
Requirements for the Degree

Doctor of Philosophy

Athens, Georgia

2015

Computational exploration of helical polymer geometry, folding, and structural stability.

by

Matthew John Williams

Approved:

Major Professor: Michael Bachmann

Committee: William M. Dennis
Steven P. Lewis
Heinz-Bernd Schüttler

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
December 2015

Acknowledgments

Graduate school has been filled with wonderful individuals who have had enormous impact on me both in my work and my personal life. I would like to first thank my advisor Dr. Michael Bachmann. Throughout my time working with Dr. Bachmann, I have felt both challenged and supported. In my research, his creativity in problem finding and solving has kept my work interesting and enjoyable. I appreciate the high standard to which he has held me and feel that this habit will serve me well for many years. Most of all I appreciate Dr. Bachmann's excitement for physics both in our direct line of study and throughout the discipline.

I would also like to thank my former advisor, Dr. Chad Fertig. His constant support and guidance during the early years of my graduate studies taught me many things which remained useful throughout my time at the University of Georgia. I am also grateful to Dr. Dennis and Dr. Lewis for the many conversations both personal and work related in which you have mentored me and guided me through this process. To my entire committee, Dr. Bachmann, Dr. Dennis, Dr. Lewis and Schüttler, I appreciate the time you spent reading my dissertation and your insightful comments.

I am also grateful to have had the opportunity to work alongside Kai Qi, Tomas Koci, Dilina Perera, and Manoj Manjare. Through coffee breaks, conference trips, and shared meals, I have enjoyed the time we spent together both as colleagues and friends. During my final semester, I have been splitting my time between my family's new home in Kentucky

and my work in Athens. I want to say a special thank you to the people who have hosted me in their homes during this time, Larry, Rebecca, Jeff, Jocelyn, Judah, Shenming, Susannah, Jonathan, Kristy, David, Susana, Luke, and Natalie.

Throughout graduate school and the rest of my life, my parents have been a constant source of encouragement and wisdom. Without them, I would not be the person I have become. I only hope that I can guide Liam, in whatever path he chooses, as well as they have guided me.

Finally, I want to thank my wife and son. Liam, thank you for the much needed breaks. Thank you for playing legos, and reading, and exploring outside with me. Rebecca, your love and support has held me up and brought joy to my life during every stage of my work. I look forward to the years ahead as we continue our journey together.

Contents

List of Figures

Chapter 1

Introduction

Many modern problems in physics are difficult to study using the traditional mathematical tools available in theoretical physics. Additionally experimental explorations are often limited in terms of scope and flexibility. Due to the cooperative effects in complex systems it has become essential that we embrace computation as a methodology. The algorithmic approach allows us to study systems which are too large to be fully modeled theoretically and lack convenient opportunity for simplification. As our ability to study complex systems experimentally expands, interesting problems which can be approached computationally emerge. Some examples of such a problems are my recent work on tertiary structure formation in helical polymers [1] and evaporative cooling of ultracold atoms, which is the basis of my early graduate work [2]. Computation has also proved extremely useful in many areas involving phase transitions such as magnetism [3], networks [4], and cosmology [5]. Polymer systems are often too complex to be fully solvable and are too small to be treated as infinitely large. For these reasons the study of polymers has benefitted greatly from computational exploration [6, 7].

Polymers are macromolecules consisting of many repeated units called monomers. They occur frequently in biological systems and are an essential element in the function of all

biological material. Biological polymers include polypeptides or proteins, polynucleotides such as DNA and RNA, and polysaccharides. An understating of the structure and thus the function of these biopolymers gives insight into how biological systems work and how they evolved. This insight is critical because biopolymers are at the heart of many disease-causing mechanisms such as viruses, prion diseases, and *Alzheimer's disease* [8, 9].

Several disease causing mechanisms are linked to the relationship between biopolymer sequence, structure, and function. Prion diseases such as *bovine spongiform encephalopathy* and *Creutzfeld-Jakob disease* are caused by structural changes to proteins in the body caused by foreign molecules. While the cause is less clear, *Alzheimer's disease* is also caused by the misfolding of proteins in the neuronal cellular network. In this dissertation, we study the relationship between biopolymer sequence, structure, and function in hopes of contributing to the better understanding of the biological systems relevant in cellular functions, evolution, and disease.

1.1 Polypeptides

The most relevant systems to the work presented in this dissertation are polypeptide systems. Polypeptides are biopolymers made up of a sequence of amino acids joined together by peptide bonds. Each amino acid consists of an amine group ($-NH_2$), a side chain unique to each amino acid, and a carboxylic acid group ($-COOH$). Two individual amino acids can join together by forming a peptide bond between the amine group of one amino acid and the carboxylic group of the other. Many amino acids can join together into polypeptide chains called proteins. Almost all proteins found in human biology are made from a basis of 20 different amino acids. The length of proteins ranges from approximately 20 to 20,000 amino acids. Each polypeptide folds into a geometrical structure which depends on the sequence of the amino acids. This structure dictates the protein's function.

1.2 Simulation and thermodynamics

Simulation allows for the study of thermodynamic quantities which describe the behavior of a system. Over the course of a simulation, an ensemble of structures is generated according to the canonical distribution which corresponds to a specific temperature. To understand the behavior of the system in temperature space, we generate an array of canonical ensembles each corresponding to a different temperature.

We can describe the ensembles by calculating averages of structural quantities. The most obvious quantity is the energy. By plotting the canonical average of the energy, $\langle E \rangle$, versus temperature we learn how the system evolves as the temperature varies. The rate of change of the canonical energy, the specific heat, gives insight into the presence of phase transitions. The specific heat can be calculated using

$$C_v = d\langle E \rangle / dT = (\langle E^2 \rangle - \langle E \rangle^2)/(k_B T^2). \tag{1.1}$$

Here, and for the remainder of this dissertation $k_B \equiv 1$.

In the thermodynamic limit, an effectively infinite system may exhibit discontinuities in thermodynamic quantities such as entropy or a divergence in C_v at a phase transition temperature. However, in many cases considering the thermodynamic limit for the study of polymers does not make sense. In polymers the finite-size effects are an essential element of the behavior of the physical system. Although the system does not exhibit phase transitions in the strict thermodynamic sense [10, 11], qualitative changes in the system behavior are interesting and can be thought of as a finite analogue to phase transitions [12]. In this dissertation, these transitions are referred to as structural transitions, and often correspond to peaks or shoulders in the specific heat curve.

It can sometimes be useful to measure other canonical quantities to find structural transitions without a strong signal in energy. For example, the end-to-end length L of a polymer

reliably shows a dramatic change between the liquid like structures and gas like random-coil structures. This change presents itself as a peak in the temperature derivative of the length, $d\langle L\rangle/dT$. The temperature derivative of some canonical quantity, O, can be calculated using the formula

$$d\langle O\rangle/dT = (\langle OE\rangle - \langle O\rangle\langle E\rangle)/(k_B T^2). \tag{1.2}$$

In addition to canonical averages, it can also be useful to examine histograms of canonical ensembles across various structural quantities. The negative logarithm of the histogram can be interpreted as a free-energy landscape in the space of system energy. The evolution of the free-energy landscape can provide insight into the folding pathways [13–19] and structure stability [20–23].

Another way to think about the evolution of a system is according to its microcanonical properties [24]. The microcanonical ensemble is the collection of structures observed in the system at a single energy. The microcanonical entropy is a relative measure of the logarithm of the number of available states at a given energy. This quantity along with its derivatives gives unparalleled insight into the phase transitions present in a system [25].

1.3 Study of biopolymers

Nuclear magnetic resonance spectroscopy (NMR) is a frequently used experimental technique used to determine structures of proteins [26]. Due to the cooperative effects in long polymers, the behavior of these structures can only be explored theoretically for a very narrow set of circumstances . Because of these restraints, many systems can only be studied computationally . All-atom models are commonly used to simulate proteins on the basis of the fully atomic configuration [27]. In these simulations, polymers studied must either be small or the scope of the experiment must be limited.

A coarse-grained approach simplifies the problem by replacing the system with an analogous system with reduced complexity but similar behavior [28]. For polymers, this is done by defining an arbitrary repeating unit called a monomer. Monomers interact according to a set of potentials chosen to reproduce the behavior of the system in question. Each potential does not have a direct one-to-one relationship with a physical force derivable from first principles but is rather an effective representation of an observable behavior of a system.

While coarse-grained models do not model a single specific protein sequence, their merit lies in the ability to discern more broadly applicable effects. The generality of the model allows for direct manipulation of a system's behavior without regard for the microscopic details responsible for it. Because of the increased efficiency gained from a dramatic decrease in the degrees of freedom, it is possible to explore a much wider array of conditions, which generates results that are more broadly applicable. The results of such a simulation can be used in the construction of a hyper-phase diagram which describes the behavior of a system across a multi-dimensional model parameter space.

Coarse-grained polymer models have been used successfully to study many different systems. These include both lattice models [29] and off-lattice models [30]. Systems include flexible heteropolymers [31], flexible polymers [32–34], and semiflexible polymers [35–37]. The relationship between a collection of polymers has been explored using aggregation models [38, 39]. There has also been much interest in the effect of an adsorption substrate on polymer structures [40–44]. Most importantly for the purposes of this dissertation, is the research into helical polymer structures and transitions [45–48].

We employ a coarse-grained approach in order to study helical polymer systems. An illustration of the difference between an all-atom representation of a 3-helix bundle and a coarse-grained 3-helix bundle is given in Fig. 1.1. In the top panel, the atomic configuration of VPS4A (vacuolar protein sorting factor 4A) is presented with the helical structures highlighted as green ribbons. An all-atom simulation would require that all 1285 atoms be

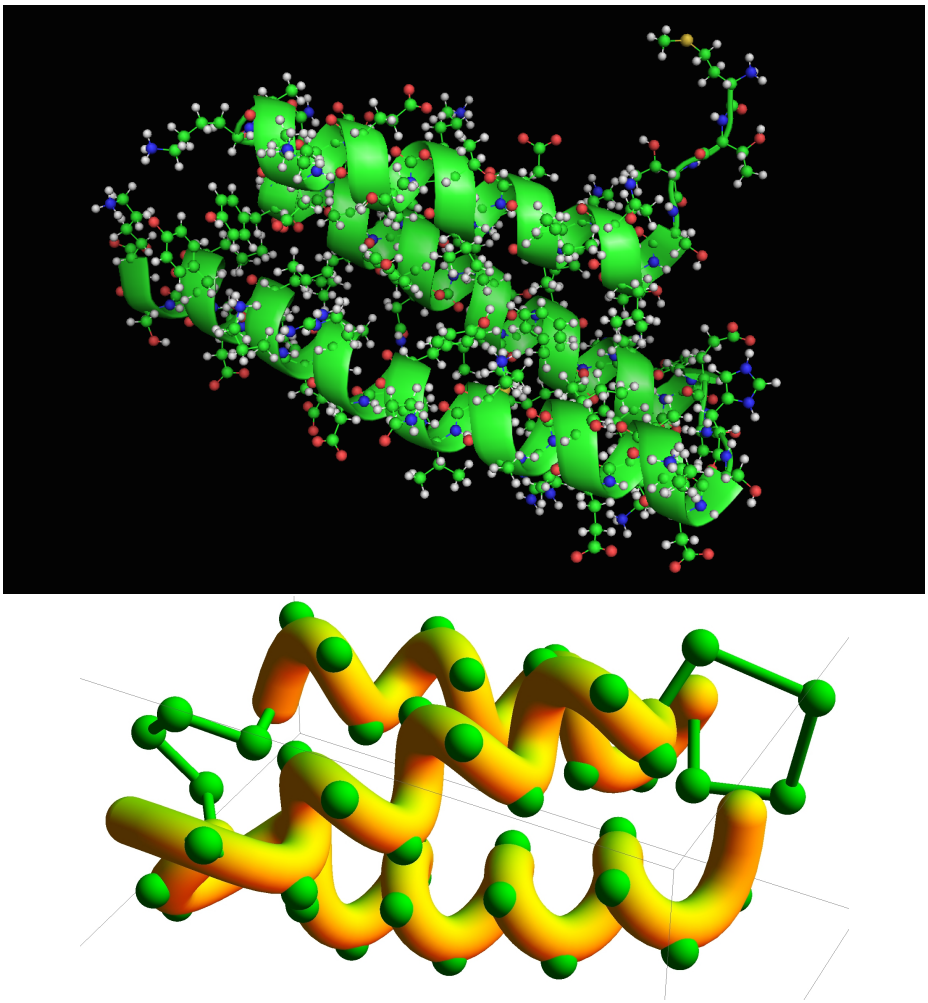

Figure 1.1: Top: Structure of vacuolar protein sorting factor 4A, a protein with 77 residues comprising a total of 1285 atoms. The structure was observed in Nuclear Magnetic Resonance experiments [49,50]. Bottom: Coarse-grained 3-helix bundle generated by parallel tempering simulation of polymer with 60 monomers.

considered. In the bottom panel, a coarse-grained approach is taken to model a generic 3-helix bundle similar to the protein above. The coarse-grained approach only requires the interactions between the 60 monomers included to be considered. Not including the medium, the coarse-grained system is a factor of 20 smaller than the all-atom simulation. Additionally, the computational requirements are approximately proportional to the square system size, meaning that the computational time for a coarse-grained system is reduced by a factor of 500 from the simulation of an all-atom model. Despite this decrease in complexity, the coarse-grained system still captures the cooperative interactions involved in the formation of a structure such as VPS4A.

1.4 Coarse-grained helical polymers

Helical structures are among the most commonly found structures exhibited by polypeptide systems [51]. In proteins, helices can occur in several different contexts. While peptide structures consisting of only a single helical rod are found, often helical segments are part of a larger tertiary structure consisting of many different elements such as helix bundles made up of several helical segments.

A propensity for formation of helical structures in biological systems can be explained by hydrogen bonding between backbone atoms of the polymer or as the result of an ordering principle such as many-body constraint [52–54]. As polymers with a propensity for helical order cool, they undergo a structural transition from random-coil structures to helical structures [55–61]. The inclusion of non-bonded interaction causes helical segments of sufficient length to fold into helical bundles [62–70]. These helical structures vary greatly depending on the particular interactions present [71, 72]. In this dissertation we expand on the previous work done in helical biopolymers using an effective potential model which represents a generic protein chain with a propensity for helix formation. We simulate these biopolymers

under an array of conditions in order to gain insight into the various influences which dictate structure formation.

In Chapter 2, we describe in detail the effective potential model used to represent the biopolymer. We include a discussion of the bond potential, non-bonded interaction, the potential which facilitates helix formation, the bending angle interaction, and the adsorption potential. We describe the computational techniques used in our simulations of these systems. Optimization strategies used to improve simulation efficiency are also discussed.

In physical polypeptide systems, it can be observed that the angle between successive monomer-monomer bonds is restrained to a fixed value. This property is typical for a semiflexible polymer [46]. Chapter 3 explores the effect of these bending-angle restraints and considers the reason for this restraint in nature. This is done by comparing the simulation results for flexible and semiflexible systems under otherwise identical conditions.

Adsorption surfaces have been shown to influence conformation geometry, transition dynamics, and stability in many different systems [22, 73–75]. The effect of an adsorption surface on the helix-coil transition has been explored experimentally [76]. Adsorbed helical polymers in the regime of very large systems have also been studied using an exactly solvable model [77].

Based on the experimental observation that helical polymers are stabilized by adsorption onto the surface of silica nanoparticles, it has been hypothesized that the presence of these particles during the formation and evolution of early life on earth may have played an important role in stabilizing early biological polymers [78]. Chapter 4 is an investigation of this claim. We accomplish this by simulating helical polymers in the presence of an adsorbing surface to determine its impact on structure formation and stability.

Both research questions under consideration in this dissertation are studied using replica-exchange Monte Carlo simulations. The systems are modeled using coarse-grained homopolymers which include a torsion potential that facilitates the formation of helical order. These

investigations provide insight into helical polymer ground state structures, stability, folding pathways, and transitions in both temperature and torsion strength [79, 80].

Chapter 2

Simulation

2.1 Helical polymer model

The behavior of a simulated system is a direct result of the model used. In coarse-grained models, effective potentials are used to represent essential system features observed in nature. We use effective potentials to model the behavior of single-chain linear polymer systems. Polymers are macromolecules with some number N of monomers which are bonded in a linear chain. These potentials are first presented as unitless quantities which will be multiplied by relative energy scales presented in Equation 2.6.

We model the relationship between successive bonded monomers according to the FENE (finitely extensible nonlinear elastic) potential [81–83]. For neighboring monomers separated by a distance r the FENE potential is given by

$$v_{\mathrm{FENE}}(r) = -\frac{1}{2}R^2 \log\{1 - [(r - r_0)/R]^2\}. \tag{2.1}$$

The bond length associated with the potential minimum is given by $r = r_0 \equiv 1$ and the maximum deviation from this value is $r - r_0 = R \equiv 3/7$. The shape of the potential is

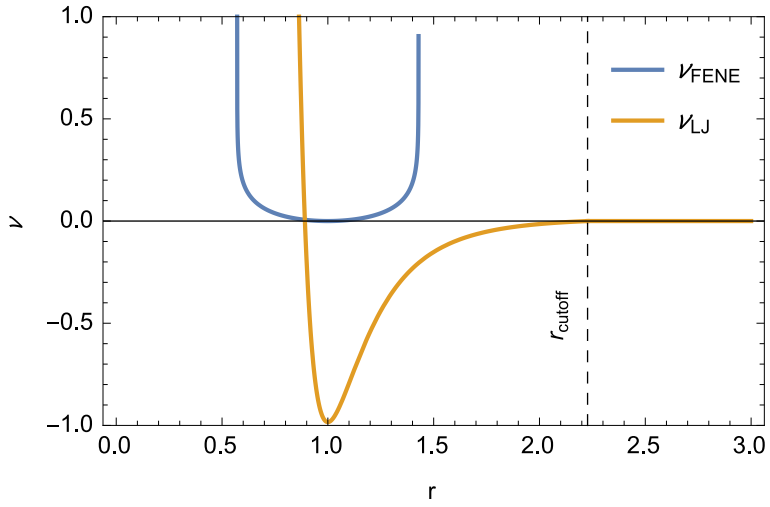

Figure 2.1: The blue curve represents the FENE potential used for all bonded interactions and orange is the Lennard-Jones which is used for non-bonded interactions. The vertical dashed line is the cutoff distance for the LJ potential, any monomers beyond this distance do not interact with each other.

shown in Fig. 2.1. If a monomer in a Monte Carlo update is shifted beyond the maximum deviation, $r > r_0 + R$ or $r < r_0 - R$, the potential is considered infinite and the move is rejected. In effect, the FENE potential maintains the polymer as a single unbroken chain of monomers.

The second interaction observed in polymer systems is an attractive interaction between non-bonded monomers in close proximity to one another. Between monomers separated by a distance r, the non-bonded interaction is calculated by the Lennard-Jones (LJ) potential [84]. In our simulations, we use it in the form

$$v_{\mathrm{LJ}}(r) = \begin{cases} 4[(\sigma/r)^{12} - (\sigma/r)^{6}] - v_{\mathrm{shift}}, & \text{if } r < r_{\mathrm{cutoff}}, \\ 0, & \text{otherwise.} \end{cases} \tag{2.2}$$

We define $\sigma = 2^{-1/6} r_0$, where $r_0 \equiv 1$ is the separation distance which yields an energetic minimum. The LJ potential acts as a repulsive potential if $r < r_0$ and it is attractive for $r > r_0$. The LJ potential is only calculated between monomers which are separated by $r < r_{\text{cutoff}}$; this improves the efficiency of the simulation by greatly reducing the number of times in which v_{LJ} needs to be calculated. This improvement in efficiency comes at very little cost since the gradient of the potential becomes very small at large r. We use a cutoff distance of $r_{\text{cutoff}} = 2.5\sigma$. A shift to the LJ potential must be included to avoid a discontinuity in v_{LJ} at $r = r_{\text{cutoff}}$. This shift is given by $v_{\text{shift}} = 4[(\sigma/r_{\text{cutoff}})^{12} - (\sigma/r_{\text{cutoff}})^{6}]$. Monomers separated by $r < r_{\text{cutoff}}$ are considered to be in contact. Each contact between non-bonded monomers provides a net energetic benefit. Because of this benefit polymers have a tendency to maximize the number of monomer-monomer contacts by forming collapsed dense structures.

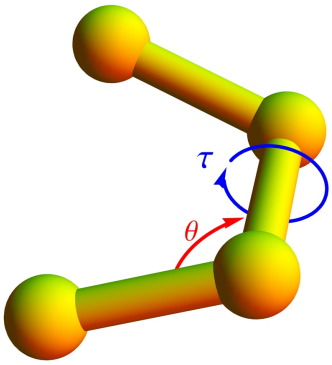

Figure 2.2: Four monomers forming three bonds are required to define a torsion angle, τ, shown in blue. Any two adjacent bonds define a bending angle represented by θ in red.

Beyond the standard effective potentials, we include a torsion potential to introduce helical order into the system. The torsion potential is associated with the twisting of the bonds as shown in Fig. 2.2. The form of the torsion potential is used as suggested by Rapaport [46],

$$v_{\text{tor}}(\tau) = 1 - \cos(\tau - \tau_0), \tag{2.3}$$

where τ represents the dihedral angle formed by the two surfaces which can be constructed using four successive monomers. Any deviation in τ away from the reference torsion angle, $\tau_0 = 0.873$, results in an increase in the torsion potential. The value of τ_0 is chosen such that a resultant helix will have approximately 4 monomers per turn.

In biological polymers, which often exhibit helical structures, there is nearly always an additional restraint placed on successive bending angles. To include this restraint in our model we apply a potential similar to the torsion potential which acts on the bending angle between neighboring bonds. The form of this potential is given by

$$v_{\text{bend}}(\theta) = 1 - \cos(\theta - \theta_0). \tag{2.4}$$

Here, θ is the bending angle and the reference angle is chosen to be $\theta_0 = 1.4$. Our implementation of the bending angle restraint is identical to the semiflexible polymer model [35, 85] with a non-zero reference angle.

In Chapter 4, we discuss the effect of an adsorbing substrate on helical polymers. The potential of a monomer at a distance h from an adsorption surface is calculated by integrating the Lennard-Jones potential over the entire half-space of the surface. This gives a potential of the form

$$v_{\text{A}}(h) = \begin{cases} \frac{2}{15}\left(\frac{\sigma}{h}\right)^9 - \left(\frac{\sigma}{h}\right)^3 - v_{\text{A shift}}, & \text{if } h < h_{\text{cutoff}} \\ 0, & \text{otherwise.} \end{cases} \tag{2.5}$$

We include a shift in the adsorption potential in the same way that a shift term was included in the LJ potential with $h_{\text{cutoff}} = 2.5\sigma$. To avoid a discontinuity, $v_{\text{A shift}} = \frac{2}{15}\left(\frac{\sigma}{h_{\text{cutoff}}}\right)^9 - \left(\frac{\sigma}{h_{\text{cutoff}}}\right)^3$. To keep the polymer from drifting excessively far from the adsorption surface, we impose a steric impenetrable boundary at a distance of $h_{\text{max}} = 200$ from the surface. This decreases the time taken for a free polymer to attempt adsorption to the surface while having minimal effect on adsorbed and desorbed structures.

To calculate an energy associated with a particular polymer structure we combine all five of these potentials with each having its own associated energy scale. The energy can be calculated by

$$\begin{aligned} E(\mathbf{X}) \quad &= S_{\mathrm{FENE}} \sum_i v_{\mathrm{FENE}}(r_{i\,i+1}) + S_{\mathrm{LJ}} \sum_{i>j+1} v_{\mathrm{LJ}}(r_{ij}) \\ &+ S_\tau \sum_l v_{\mathrm{tor}}(\tau_l) + S_\theta \sum_k v_{\mathrm{bend}}(\theta_k) \\ &+ S_{\mathrm{A}} \sum_i v_{\mathrm{A}}(h_i). \end{aligned} \tag{2.6}$$

The energy scales for the FENE and LJ potentials take the standard values $S_{\mathrm{FENE}} = 98/5$ and $S_{\mathrm{LJ}} = 1$. Throughout this paper, we will vary the torsion energy scale S_τ to explore helical structures formed at different levels of helix forming propensity. To restrict the model parameter space, we do not explore an array of values for the bending potential energy scale, but rather consider it to be either $S_\theta = 0$ or 200, representing the flexible and semiflexible cases, respectively. Finally, the adsorption energy scale is considered to be $S_{\mathrm{A}} = 0$ for Chapter 3 in which adsorption is not considered. For Chapter 4 values between $S_{\mathrm{A}} = 0$ and 2 are chosen to explore the effect of the adsorption surface on helical polymer structures. For the cases in which $S_{\mathrm{A}} = 0$, the steric boundary is not present.

2.2 Metropolis sampling

In this study, polymers of length between 30 and 60 are simulated using replica-exchange Monte Carlo simulation (parallel tempering) [86–90]. Here, the initially random configuration is continually modified by iterative random updates to its configuration. Each change to the configuration will result in a change in the polymer's energy by an amount ΔE. The modification will be accepted with probability P_{accept} according to the Metropolis crite-

rion [91],

$$P_{\text{accept}} = \begin{cases} e^{-\beta\ \Delta E}, & \text{if } \Delta E > 0 \\ 1, & \text{otherwise,} \end{cases} \tag{2.7}$$

which depends on the inverse temperature $\beta = 1/(k_B T)$.

There are many different types of updates which can be used in Metropolis sampling of off-lattice polymers. The choice of update types can have a strong influence on the correlation time and therefore the simulation efficiency. It is important when choosing updates that ergodicity and detailed balance are satisfied. Ergodicity is the requirement that it is possible that all possible configurations of the system are accessible from all other configurations by making many repeated updates. This does not have to be satisfied for each update type independently but only for the combination of all updates. Detailed balance is the requirement that any update is equally likely to be attempted in both directions. For example, generating the state X_2 starting with X_1 should have the same probability as generating X_1 from X_2 for all updates used. The most elementary update type is the local displacement update, in which a single monomer is chosen at random a to be moved to a new location within a box of size r_d surrounding its original location. The displacement update is particularly efficient because the entire energy need not be calculated at each energy step. Only those components influenced directly by that monomer are recalculated (2 FENE interactions, $N-3$ LJ interactions, 4 torsion angles, 3 bending angles, and 1 adsorption interaction).

For the displacement update, optimization of the box size can lead to gains in simulation efficiency [92]. For large values of r_d the acceptance rate decreases. In the case that acceptance becomes rare, the simulation efficiency decreases as it will take many attempted updates to change the polymer configuration significantly. Conversely, for very small values of r_d, changes in energy will be quite small, leading to a high acceptance rate. Although the acceptance rate is high, the relative change to the polymers configuration in each move will

be very small. In this case, it will take many updates to change the polymer's configuration, appreciably also decreasing efficiency. The optimal r_d will lie somewhere in between these two extremes.

We choose r_d such that the acceptance rate is approximately $R_{\text{accept}} = 0.5$. This is done during an initialization period before data collection begins. After every 100 updates, the displacement size is changed to $r'_d = r_d + p(R_{\text{accept}} - 0.5)$. The parameter p is a factor determining the size of each iterative adjustment of r_d which we choose to be 0.04. A lower limit for displacement size is also implemented to avoid having displacements grow unreasonably small in regimes where very few moves are accepted.

For polymer systems which include both bending and torsion potentials, the strongest variation between unique structure types comes from variation in the torsion angles. For example, the difference between a single helix and two-helix bundle is a disturbance in the torsion angles in the joint between the two helical segments. Allowing for random variation of the torsion angles of low-temperature systems, leads to shorter autocorrelation time due to the increased ease by which the system can transition between structure types. Random variation of the torsion angles can be achieved by selecting a single bond at random, and rotating all monomers on one side of the bond around a rotation axis collinear with the bond. Monomers are rotated by the same randomly chosen angle between $+\tau_r$ and $-\tau_r$. The torsion update can also be optimized in the same way as the displacement update was optimized, by changing τ_r dynamically during the initialization period based on the torsion update acceptance rate. Because of the global nature of this update the entire energy must be calculated to find the energy difference. Because of this, the torsion update takes significantly longer than the single monomer displacement update.

With the inclusion of an adsorption substrate, adsorption attempts can be drastically increased by including a translational update of the entire polymer helping it to drift through the simulation region more rapidly. This update chooses a random displacement which is

applied to every monomer at once. This effectively translates the entire polymer relative to the adsorption surface changing only the adsorption energy and leaving the monomer-monomer interactions unchanged. While this update is not important for low temperatures at which the polymer will be adsorbed, it can greatly increase the speed at which the simulation can overcome the adsorption transition.

We choose between the various update types at random with weights applied to each update type depending on its desired frequency. We apply $2N$ single monomer displacement updates for each torsion update and polymer displacement update.

2.3 Parallel tempering

Metropolis sampling has the fundamental problem that for systems with multi-welled free-energy landscapes, the simulation is prone to becoming stuck in one of the free-energy minima, taking very many updates to eventually overcome the entropic barrier. While a physical system exhibits some fraction of the population in each well, a system being simulated with Metropolis sampling will likely fall into one of the wells and continue sampling only that well. Because every move has some probability of acceptance, there is always a non-zero probability of transitioning between the wells. This transition, though, may take far longer than the length of the simulation.

Replica-exchange parallel tempering nicely mitigates this problem while making only a minor change to the simulation. Instead of generating an array of temperature ensembles independently, all ensembles are simulated in parallel. Each simulation is performed on a separate computational thread and is propagated using the original Metropolis algorithm. Periodically neighboring temperature threads compare polymer conformations and attempt to exchange. This allows a single polymer to travel through multiple simulation threads at different temperatures. Ideally, a single replica (polymer structure) stuck in a free-energy

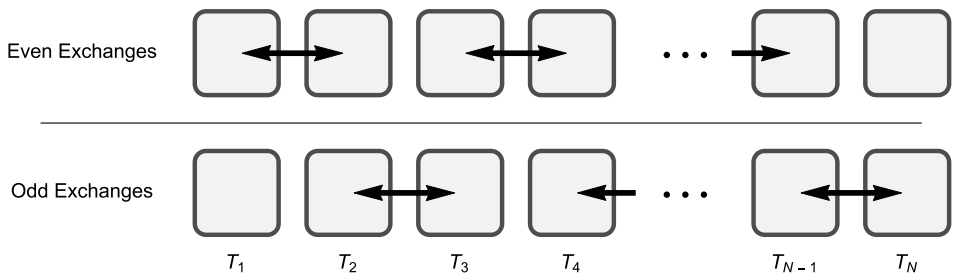

Figure 2.3: Exchanges represented by arrows are attempted between neighboring temperature threads every 400 updates. Each thread alternates between attempting exchange with its higher and lower temperature neighbor. The first and last threads only participate in half of the exchange attempts.

minimum will eventually transition to a higher temperature thread, melt, and then eventually transition back into a low-temperature thread possibly cooling into a different free-energy minimum. Each simulation thread will sample many different replicas and explore structure space much more efficiently.

In the studies discussed here, we perform a single parallel tempering simulation for each model we explore. In each simulation, an array of N temperatures $\{T_1, T_2, \ldots, T_N\}$ is chosen. Metropolis simulations are performed in each thread attempting to exchange replicas every 400 updates. The ith temperature thread alternates between attempting exchange with the $i+1$ thread and the $i-1$ thread. Because both T_1 and T_N have only one neighbor, they will each remain idle during half of the attempts. This pattern is explained further in Fig. 2.3. A single exchange attempt between thread i and thread j is accepted with probability

$$P_{\text{accept}} = \min(1, e^{-(\beta_i - \beta_j)(E_j - E_i)}). \tag{2.8}$$

Ideally, the thread temperature arrangement is chosen in a way which allows exchanges and rejections to both occur frequently. For this to happen, the histograms of threads i and j

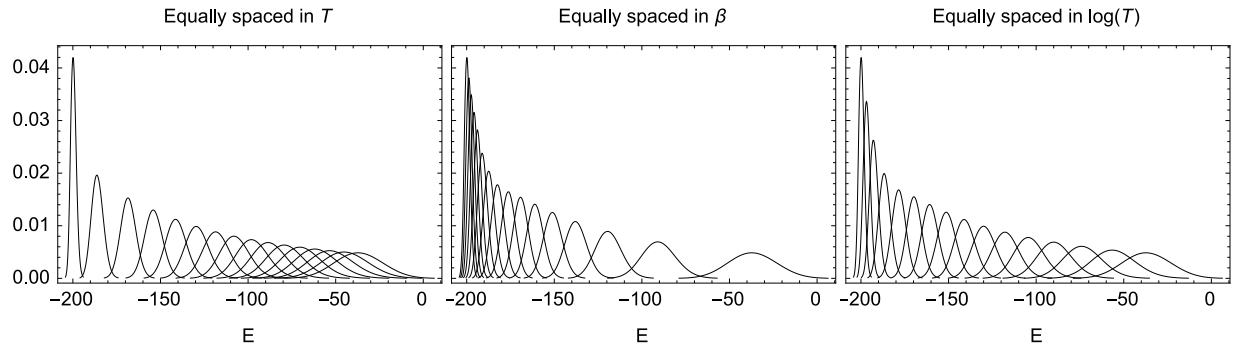

Figure 2.4: Energy histograms for each temperature thread in three different 16 thread simulations. Compared are simulations in which temperature are chosen linearly, inverse temperatures are chosen linearly, and temperatures are chosen exponentially. Exponential temperature spacing leads to the most consistent histogram overlap.

must overlap by approximately 50%. The obvious choice might be to choose either evenly spaced temperatures or temperatures which are evenly spaced in β. We find that evenly spaced temperatures results in histograms with excessive overlap at high-temperature and very little overlap at low-temperature. Conversely, evenly spacing the inverse temperature leaves very little overlap in the histograms of high-temperature threads. We find that an even spacing of $\log(T)$ results in the most consistent overlap as shown in Fig. 2.4. It is also important to the effectiveness of the parallel tempering simulation that the largest temperature thread, T_N, is large enough that structures are fully melted.

For systems exhibiting strong first-order phase transitions, the ideal spacing of temperatures cannot be generated so simply in the region of the transition. While the exponential choice of temperatures is a good starting point each specific model may benefit from a unique choice of temperatures. To dynamically tune the temperature spacing of simulation threads we have tried implementing a proportional control system during the system equilibration period, considering each thread starting with an initial random configuration. A few thousand sweeps are allowed for the system to stabilize before dynamically tuning the individual

temperature threads. After this time, a dynamic control of each temperature is implemented every 20 exchanges. The update to each temperature is calculated using

$$
\begin{aligned}
c_i &= p(R_t - R_i)(T_i - T_{i-1}) \\
T'_i &= \max(\ \ T_i + (T'_{i-1} - T_{i-1}) + c_i \ \ ,\ \ T'_{i-1} + T_{sp\min}),
\end{aligned}
\tag{2.9}
$$

where T_i represents the original temperature of the i^{th} thread, T'_i represents the updated temperature of the i^{th} thread, R_i is the exchange rejection rate in the i^{th} thread, R_t is the target rejection rate, and p is the proportional factor. In our simulations we have used values of $R_t = 0.6$ and $p = 0.02$. It is also important that we impose a minimum temperature spacing ($T_{sp\min}$) so that all replicas do not move into a region of a particularly strong first-order phase transition. In practice we find that dynamic updating of temperatures, while useful, often introduces more problems than it fixes and was therefore not included in the simulations used to generate the data presented in this dissertation.

Chapter 3

Flexible and semiflexible helical macromolecules

In this chapter we present the results of simulations of flexible and semiflexible polymers with a varied propensity for helical order. We begin by explaining our methodology for classifying structures. This methodology is then used to explain changes in structure formation as S_τ is varied. We further our understanding of each system by examining its thermodynamic behavior across temperature space. With information about transitions in both S_τ and T, we construct a hyper-phase diagram which lays out structure types in this two-dimensional parameter space. Finally, we take a more detailed look at low-temperature structures and structure stability.

3.1 Structure classification

A variety of unique structures are generated in the simulation of flexible and semiflexible helical polymers across torsion strength and temperature. These structures include single helical strands, bundles of two or more helical segments, globular clusters with or without lo-

cal helical order, amorphous solids with or without helical order, and random-coil structures. Descriptive parameters must be used to distinguish between structures types. Traditional parameters, including energy, end-to-end length, radius of gyration, and number of contacts, can provide some insight into the structures generated but lack the specific property of distinguishing helices and helix bundles. For a more targeted approach to structure classification, we introduce several unique parameters.

Helical disorder, q_τ, gives a measure of a structure's torsional deviation away from the idealized helical polymer. It can be calculated using the equation

$$q_\tau = \sum_l v_{\mathrm{tor}}(\tau_l). \tag{3.1}$$

A structure in which all torsion angles are near to the reference torsion angle will have a small value for q_τ. While this parameter provides useful information about the polymer structure, it is directly correlated with a change in the torsion strength with any change in q_τ is attributable to a change in S_τ. This parameter is not a good indication of tertiary structural information such as helix bundle formation.

To distinguish between helix bundles with different numbers of helical segments, we introduce the pair of parameters q_1 and q_2. These parameters describe the local and global Lennard-Jones interaction, respectively. For a particular polymer conformation, the average over all monomers of their LJ interaction with other monomers separated by less than seven bonds is given by the equation

$$q_1 = \frac{1}{N} \sum_{i=1}^{N-2} \sum_{j=i+2}^{N} \Theta_{6,j-i}\, v_{\mathrm{LJ}}(r_{ij}). \tag{3.2}$$

Here the Heaviside function is presented in a discrete form such that $\Theta_{m,n}$ is 1 for $m \geq n$ and 0 otherwise. The average over all monomers of interaction with monomers separated by

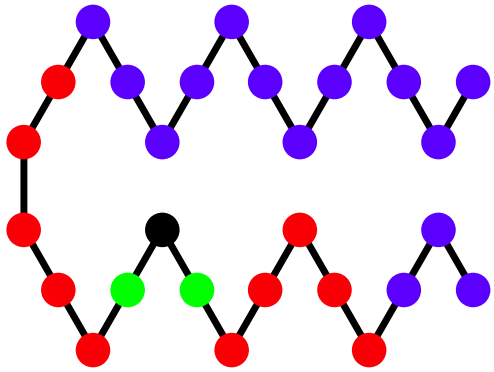

Figure 3.1: Definition of the order parameters q_1 and q_2. The black monomer interacts with the green monomers via the FENE potential and with the blue and red monomers via the LJ potential. The total potential of the LJ interactions between nonbonded monomers separated from the black monomer by 6 or fewer bonds, as represented by red monomers, contributes to q_1. Consequently, q_2, accounts for the LJ contributions from the monomers more than 6 bonds away (blue monomers).

more than six bonds is given by

$$q_2 = \frac{1}{N} \sum_{i=1}^{N-2} \sum_{j=i+2}^{N} \Theta_{j-i,7}\, v_{\mathrm{LJ}}(r_{ij}). \tag{3.3}$$

The distinction is explained visually in Fig. 3.1, where the interactions between the black monomer and all red monomers contribute to q_1 and the black monomers interaction with blue monomers contributes to q_2.

To understand the usefulness of this pair of parameters, consider two example structures, a single helix and a two-helix bundle. For a single helix, all monomers are in contact with other monomers which are either, in its own helix turn, the turn before its own, or the turn after its own. All of the monomers in these three turns are separated by no more than 6 bonds and will, therefore, contribute a negative potential to q_1. All monomers more than 6 bonds away from a particular monomer are too far away to be in contact. For this reason, in a single helix q_1 is minimal and q_2 maximal. In contrast, for a two-helix bundle local

LJ interaction at the joint between the two helix segments is sacrificed, causing an increase in q_1. This joint allows monomers from different helix segments, separated by more than 6 bonds, to come into contact with each other. These additional contacts give a negative contribution to q_2. Similarly, higher-order helix bundles can also be distinguished by the relationship between q_1 and q_2. If we plot a collection of structures onto a space defined by q_1 and q_2, we find that qualitatively similar structures form distinct clusters. This clustering allows us to quantitatively distinguish between different helix bundle types.

3.2 Distribution in structure parameter space

In both the flexible and semiflexible cases, we explore polymers at an array of S_τ values. Each simulation consists of 24 temperature threads arranged exponentially between $T = 0.1$ and 2. From each temperature thread 1000 polymer structures are saved. Structures from all temperature threads for a single simulation are compiled and distributed in $q_1 - q_2$ space. The black areas shown in each panel of Fig. 3.2 depict the regions in which structures are formed for a selection of S_τ values. The left column (a)-(d) represents semiflexible models ($S_\theta = 200$), and the right, flexible models ($S_\theta = 0$). Red regions represent only structures formed at $T = 0.1$, and the gray regions are simply a background which shows the combined region for all values of S_τ with either semiflexible or flexible. In panel (a) we see that the low-temperature structures collect at low q_1 and as the temperature increases so does q_1. It is apparent that, at low temperature, there are several distinct structural clusters formed which correspond to a multi-welled free-energy landscape near the ground state. In panel (b), the larger S_τ creates helical segments which organize into bundles. We still see multiple low-temperature clusters corresponding to unique bundling configurations. Further increasing S_τ stiffens the helical segments, leading to production of two-helix bundles in panel (c). The two-helix bundle offers fewer possible orientations and variations, therefore

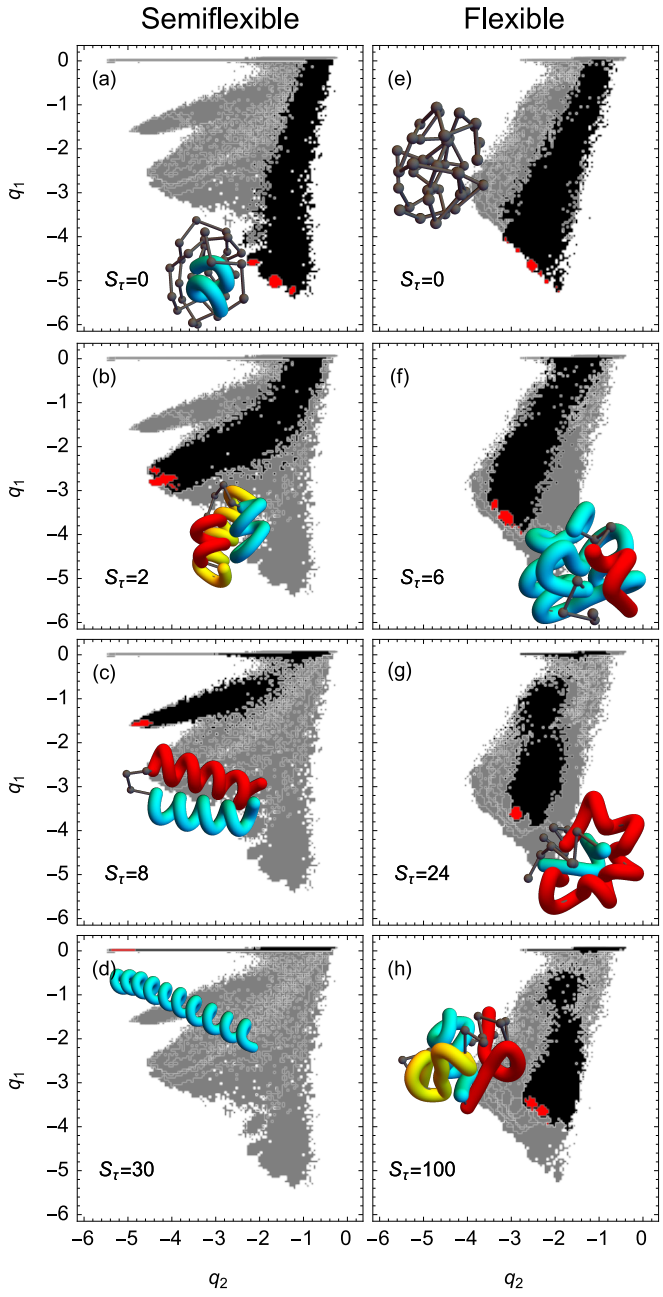

Figure 3.2: Regions of structure formation in $(q_1 - q_2)$ space for the (a)-(d) semiflexible (bending restrained) and (e)-(h) flexible (bending unrestrained) polymers with 40 monomers. Light gray regions represent the generalized ensemble of all conformations found at all temperatures T and torsion strengths S_τ simulated. Black regions correspond to the most populated states at given S_τ values. Red regions represent only the states populated for $T \leq 0.1$. Representative conformations for each low-temperature ensemble are shown.

exhibiting a single well-defined low-temperature cluster corresponding to a single structure type. For $S_\tau = 30$, we see that a single stiff helix is consistently produced. All of the panels in the bending restrained case have high-temperature structures in the upper-right hand region. When cooled these ensembles move toward the lower-left into one of several branches depending on the native folded state for the particular model being explored.

In contrast to the bending restrained case, Fig. 3.2 panels (e)-(h) show flexible polymer structures ($S_\theta = 0$). For nearly every value of S_τ, we see multiple low-temperature clusters. While helical order emerges as S_τ increases from 0, the lack of bending restraint lends no stiffness to the helical segments. Because of the lack of stiffness there is far less predictability and organization in the evolution of global structures as S_τ varies.

3.3 Free-energy minima folding trajectories

A more detailed analysis of the free-energy landscape in structure parameter space $q_1 - q_2$ can give further insight into the folding pathways and their dependance on S_τ. We determine the free-energy for each canonical ensemble from the inverse frequency of states in each bin of a partitioned space of $q_1 - q_2$. The free-energy can be calculated by

$$F_{S_\tau,T}(q_1, q_2) = -k_\mathrm{B}T \log Z_{S_\tau,T}(q_1, q_2), \tag{3.4}$$

where

$$Z_{S_\tau,T}(q_1', q_2') = \int \mathcal{D}X \delta(q_1' - q_1(\mathbf{X}))\delta(q_2' - q_2(\mathbf{X}))e^{-E(\mathbf{X})/k_\mathrm{B}T} \tag{3.5}$$

is the restricted partition function in the space of all structures present in the ensemble.

For a given ensemble, there is a global free-energy minimum corresponding to the dominant structural configuration present. Fig. 3.3 is divided between bending-restrained (left) and bending-unrestrained (right). Each black point represents the global free-energy min-

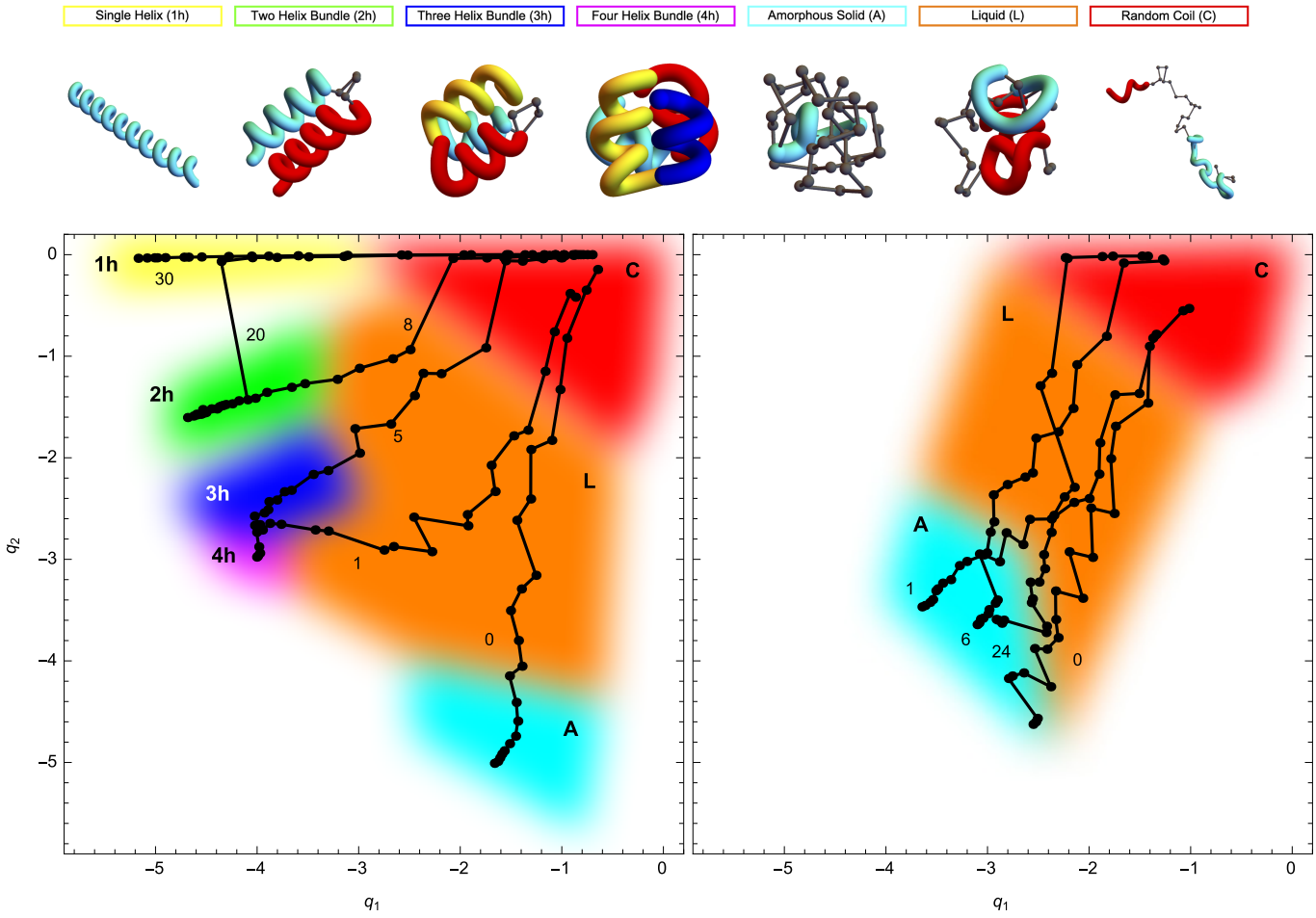

Figure 3.3: Structural phase diagrams for bending-restrained semiflexible (left) and unrestrained flexible polymers (right) in $(q_1 - q_2)$ order parameter space for the temperature and torsion strength space (T, S_τ) covered in our simulations. Colored regions represent structural phases. Black dots locate free-energy minima at given T and S_τ values. Trajectories show the helical folding pathways at fixed torsion strengths S_τ by decreasing the temperature.

imum for a single canonical ensemble, and lines connecting free-energy minima connect different temperatures for the same set of model parameters. Regions are marked in the space according to the dominant structure types found in the region, as well as the locations of transitions in specific heat plots.

The lines signify the folding pathway for fixed torsion strength. For all pathways the high-temperature ensembles which are in the random-coil phase, begin in the upper-right hand corner of the panel. As the temperature decreases, the free-energy minima move toward the lower left into one of the folding branches depending on the value of S_τ. For specific values such as the bending restrained $S_\tau = 20$ case, it is possible for the dominant phase to jump from one structural branch to another as temperature decreases. In this particular case, as the system is cooled structures which begin in the random-coil phase transition predominantly to the single-helix phase. Upon further cooling, the system exhibits coexistence of the single-helix and 2-helix phases. Eventually, the single-helix phase dies out completely leaving structures in only the 2-helix phase at the lowest temperatures.

While the folding trajectories in the bending restrained case show clear behavior with predictable evolution as S_τ is varied. The unrestrained case exhibits none of these properties. In contrast to the distinct solid structural phases present on the left, bending unrestrained folding trajectories all transition from random-coil configurations through a liquid phase into an amorphous solid phase. In this solid phase the local helical order increases with increasing S_τ as seen in the bending unrestrained curve of Fig. 3.4. While the helical order is increasing, the helical segments are very short allowing them to contort into less structured phases which do not show a drastic qualitative global change in organization across S_τ. The global structure type changes are the cause of the stair-step events seen in the bending restrained curve in Fig. 3.4.

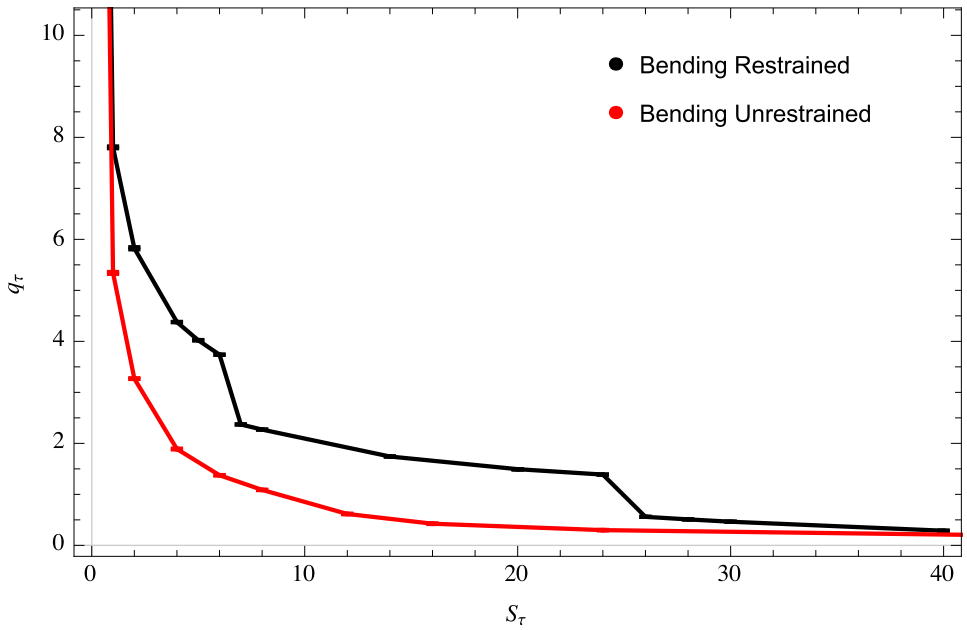

Figure 3.4: Torsion disorder plotted as a function of S_τ for canonical ensembles at $T = 0.1$. While both systems show decreasing torsion disorder with increasing torsion restraint the structure transitions present in the bending restrained case cause jumps in the torsion disorder.

3.4 Transitions in temperature

Transitions in S_τ are nicely revealed by the discrete branches seen in $q_1 - q_2$ space but transitions in temperature are not nearly as apparent using the methods explored up to this point. To detect these transitions we consider how the specific heat behaves in temperature. Fig. 3.5 shows the specific heat for the semiflexible model on the left and the flexible model on the right across a variety of S_τ values. For the semiflexible case, starting from the top panel where the torsion potential is quite strong, we see a single peak in the specific heat curve corresponding to the freezing transition between the solid single-helix phase and the random-coil phase. As S_τ decreases, a second sharper transition corresponding to the a solid-solid transition between the single-helix phase and two-helix bundle phase comes in from the left starting at $S_\tau = 24$. At $S_\tau = 14$, the solid-solid transition merges with the freezing

transition. As we continue to decrease S_τ, the transition splits into a Θ transition and a freezing transition which spread apart to form an increasingly large liquid phase between them. For $S_\tau \leq 8$, the region to the right of both transitions corresponds to the random-coil phase. The region between the transition has a more liquid-like behavior, with no defined single organization. And to the left of the lower-temperature transition is a solid phase corresponding to either 2-helix, 3-helix, 4-helix, or amorphous solid phases distinguishable by the $q_1 - q_2$ branch in which they lie. For $S_\tau < 4$, transitions in temperature become increasingly complex with the introduction of multi-welled low-temperature ensembles. We do not find it useful to analyze the canonical specific heat quantity for these cases.

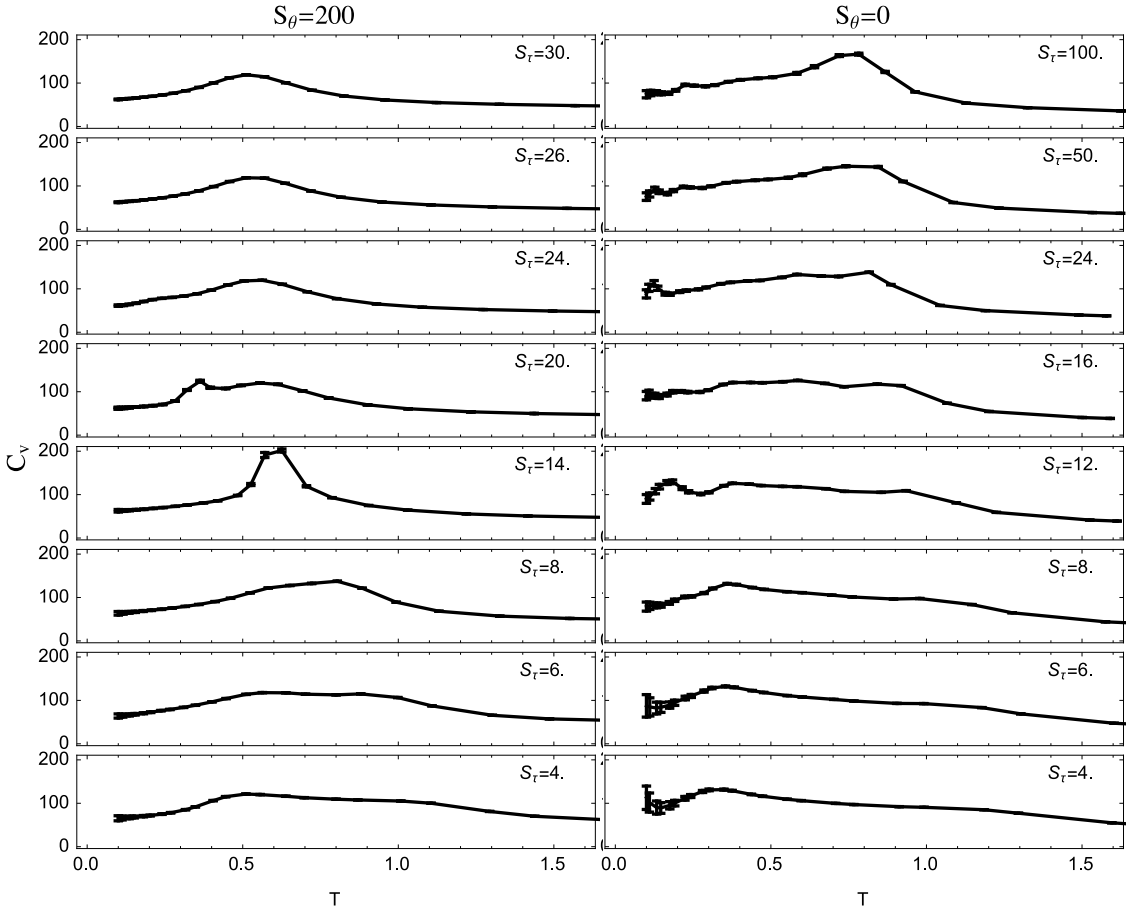

Figure 3.5: Specific heat vs. temperature for S_τ between 4 and 100, with bending restrained cases on the left and bending unrestrained on the right. The peaks and shoulders in these curves give insight into the location of structural transitions in temperature.

Transitions in the bending restrained case exhibit far fewer systematic features which can be observed across S_τ. While we still in general observe a Θ and freezing transition, we see additional transitions which appear and disappear for each unique value of S_τ. For a more detailed analysis of the phase transitions present in the flexible model, a microcanonical analysis would be required.

Analysis of transitions recognizable in the specific heat curves gives insight into not only the temperatures at which each transition occurs but also the region of $q_1 - q_2$ space shown in Fig. 3.3 in which we find solid, liquid, and random-coil phases.

3.5 Hyper-phase diagram in system parameter space

Using information gathered in both the folding trajectory study and the specific heat vs. temperature curves, we are able to construct a hyper-phase diagram parameterized by torsion strength and temperature. In Fig. 3.6 we see the dominant phase present for all values of S_τ and T, in both the semiflexible and flexible models. In the semiflexible case (left) we observe much more robust organization of unique structural phases. For $S_\tau \geq 7$, there is clear distinction between random-coil, liquid, single-helix, and 2-helix phases. In each of the regions specified, the dominant structure is qualitatively distinguishable and recognizable. Moving to lower values of S_τ the dominant phase becomes increasingly less clear. For $S_\tau = 6$ low-temperature structures are clearly dominated by 3-helix bundles but as we move closer to the 4-helix and amorphous solid regime we begin to have variability in the organization of the three-helix bundles, with helical segments of differing lengths and orientations. The strong presence of distinct clustering also begins in this region. These features become more pronounced as S_τ decreases further into the 4-helix and amorphous solid regions.

Included in Fig. 3.7 are the hyper-phase diagrams for semiflexible 30, 40, and 60 monomer systems. We find that for a 60-mer the instability in the 3-helix bundle reduces, with the

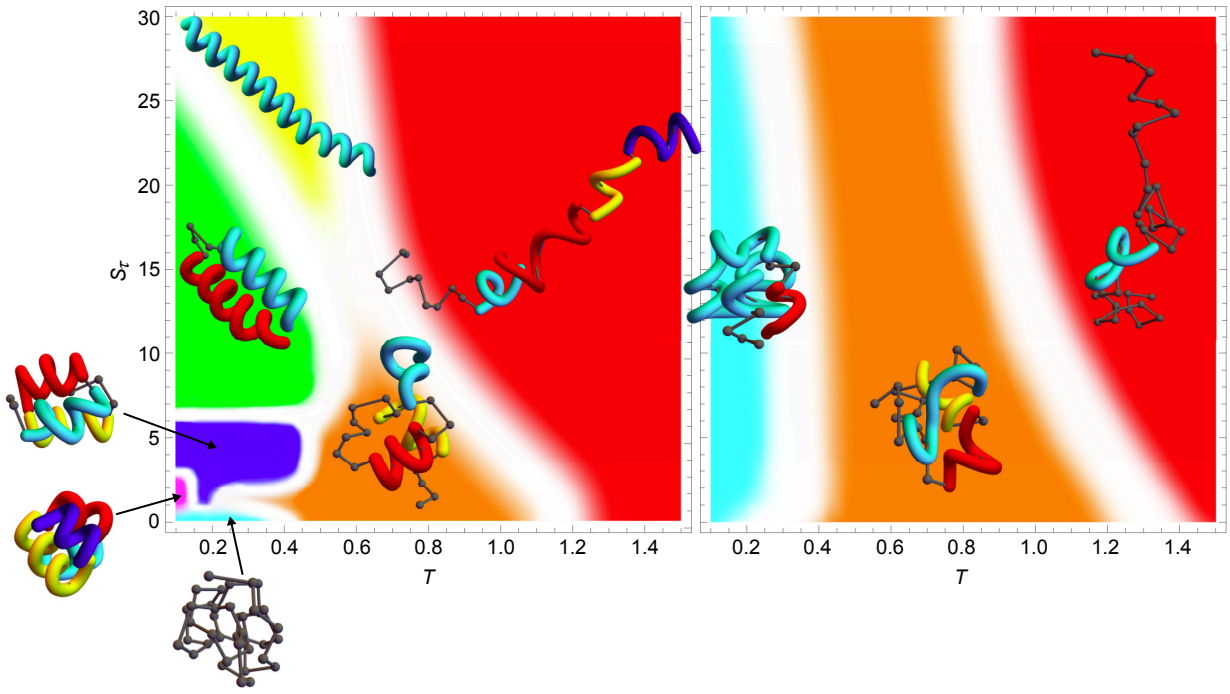

Figure 3.6: Hyper-phase diagrams of bending-restrained semiflexible (left) and unrestrained flexible polymers (right) with 40 monomers. Regions are represented in the space of the torsion strength S_τ as a material parameter distinguishing classes of polymers and the temperature T as an external control parameter for the formation of structural phases. The color code is the same as in Fig. 3.3.

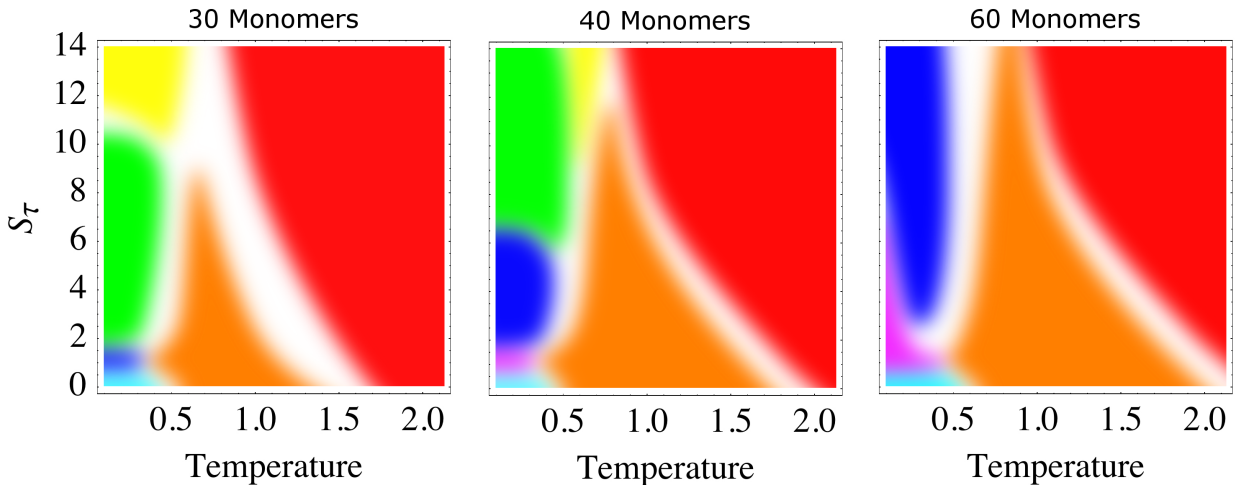

Figure 3.7: Hyper-phase diagrams of semiflexible polymers with varied system size. The color code is the same as in Fig. 3.3.

3-helix branch forming at larger values of S_τ and separating more obviously from the 4-helix branch. The bundling for the 3-helix bundle becomes less variable in orientation and helix-segment length. The variability of the 4-helix bundle reduces as well, but to a lesser extent. Conversely, in the case of the 30-mer we see disappearance of the 4-helix phase. The parameter space for which 3-helix bundles are formed is greatly reduced from the 40 monomer case and the bundles are highly unstable.

As shown in the right side of Fig. 3.6, the folding process is not influenced strongly by the torsion strength. Again, while helical order emerges for increased S_τ, there is no organization of helical segments. While structures in this region are not all of a single structure type, they are not well separated in $q_1 - q_2$ space and are therefore projected into a single phase. Also, in contrast to the bending restrained case, we no longer see the disappearance of the liquid phase at high S_τ.

3.6 Low-temperature structure analysis

Here we introduce a new parameter q_2^{frac}, which gives the fraction of Lennard-Jones interaction which occurs between monomers separated by more than 6 bonds. The value of q_2^{frac} can be calculated using the equation

$$q_2^{\text{frac}} = q_2/(q_1 + q_2). \tag{3.6}$$

In Fig. 3.8, we plot the canonical averages of q_2^{frac} at each value of S_τ considering only the ensembles for a single temperature. For each panel, a different temperature is chosen shown in the upper left. The behavior of q_2^{frac} highlights the structural transitions in S_τ at low temperature.

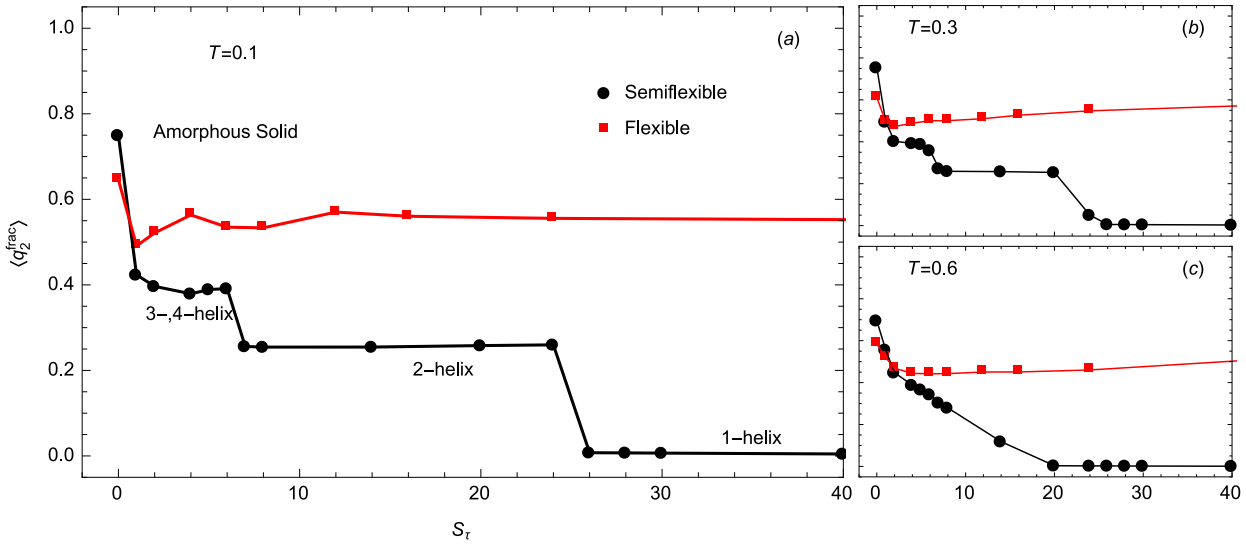

Figure 3.8: Structural parameter $\langle q_2^{\text{frac}} \rangle$ plotted for a single temperature for each value S_τ. Regions of constant $\langle q_2^{\text{frac}} \rangle$ represent consistent phases over a range of S_τ. (a) For the bending restrained case at $T = 0.1$, there is strong division between distinct states. This behavior is not present in the unrestrained case. At higher temperature (b) we see the sharpness of the transitions decrease as the structural variability increases. In (c) the distinct states are no longer discernible.

In panel (a), we have plotted the behavior of q_2^{frac} for $T = 0.1$. We see progression, in the bending restrained case (black), from most of the energy in q_2, which corresponds to the amorphous solid phase, to just less the half for the 3- and 4-helix bundles, to less than 1/4 in the 2-helix bundle, and almost 0 in the single helix. We can see in the 3- and 4-helix bundle region that the there is no well-defined separation between the two for the parameter q_2^{frac}. The inconsistency of q_2^{frac} in the 3- and 4- helix region is evidence of some instability in the system, this will be more obvious later when we look at the ensembles from which these averages are taken. The single helix and 2-helix regions are both highly consistent over the extent of their domain.

In contrast to the constrained case, the unconstrained case is highly unstable. This is evident based on the erratic behavior of q_2^{frac} along S_τ.

As the temperature is increased to $T = 0.3$, we see softening of the transitions due to the increased variability within each structure type. For $S_\tau = 0.6$, the transition between stable unique states has completely disappeared in favor of a continuous evolution between globule-like $S_\tau = 0$ structures and the stiffer helical segments of $S_\tau = 0.6$. No notable behavior is observed in the bending unrestrained case at higher temperature.

To give further insight into the ensemble averages of q_2^{frac} shown in Fig. 3.8, we present the histograms from which these averages are taken. Fig. 3.9 shows the histograms of the q_2^{frac} from structures generated in each $T = 0.1$ ensemble. It is apparent that for the fully flexible polymer many of the ensembles present spread our multi-peaked histograms when projected into the q_2^{frac} parameter space. The free-energy is the negative logarithm of the histogram, implying that the multi-peaked histograms presented here correspond to multi-welled free-energy landscapes. Multi-peaked histograms also occur in the semiflexible helical polymer, but only for low torsion potential in the three-helix, four-helix, and amorphous solid structures and in higher temperature ensembles which are not shown here.

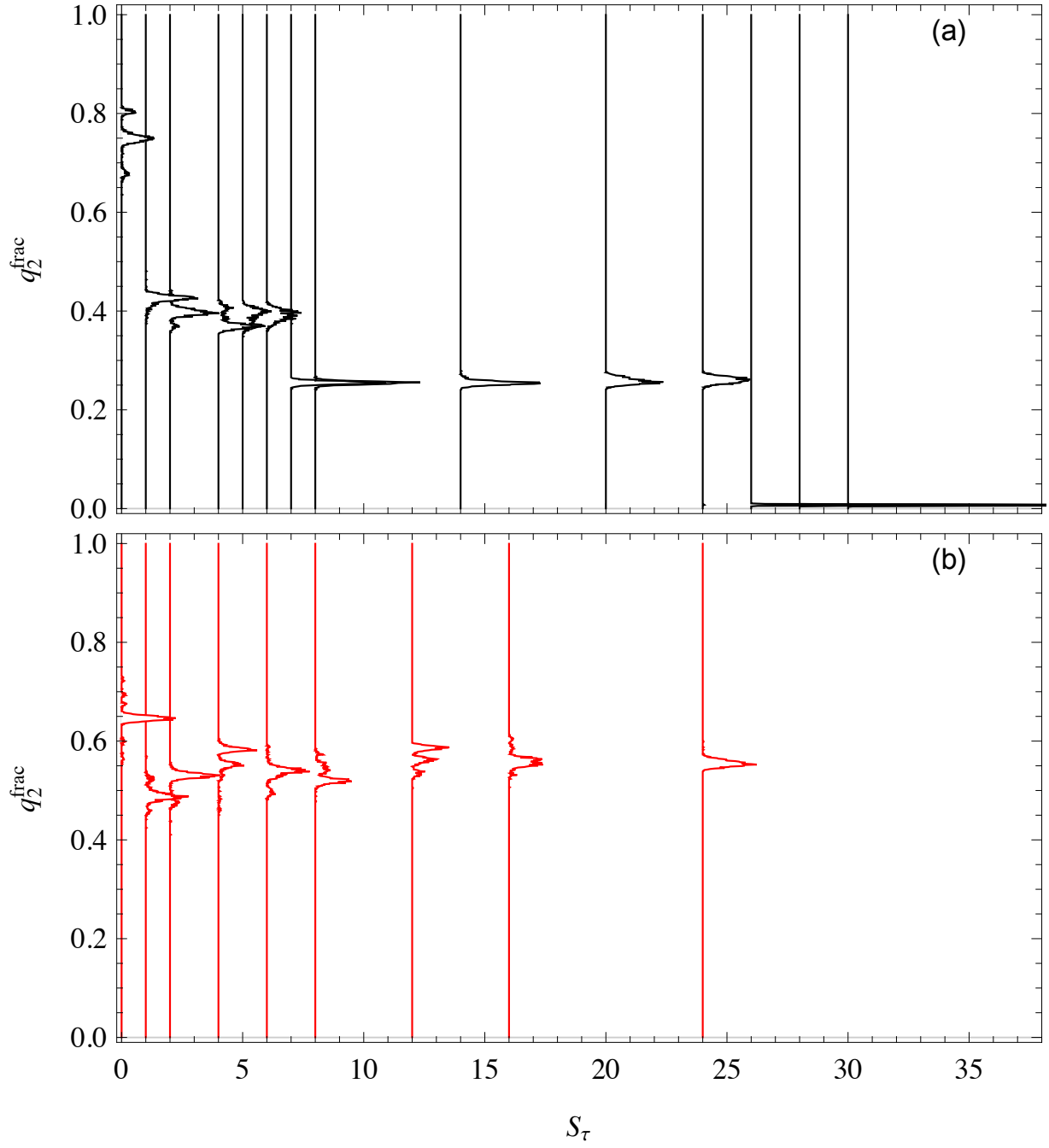

Figure 3.9: q_2^{frac} histograms for canonical ensembles at $T = 0.1$. The base line for each histogram aligns with the value of S_τ at which it was generated. Panel (a) gives the histograms generated in the semiflexible case, and panel (b) gives the histograms generated for fully flexible polymers.

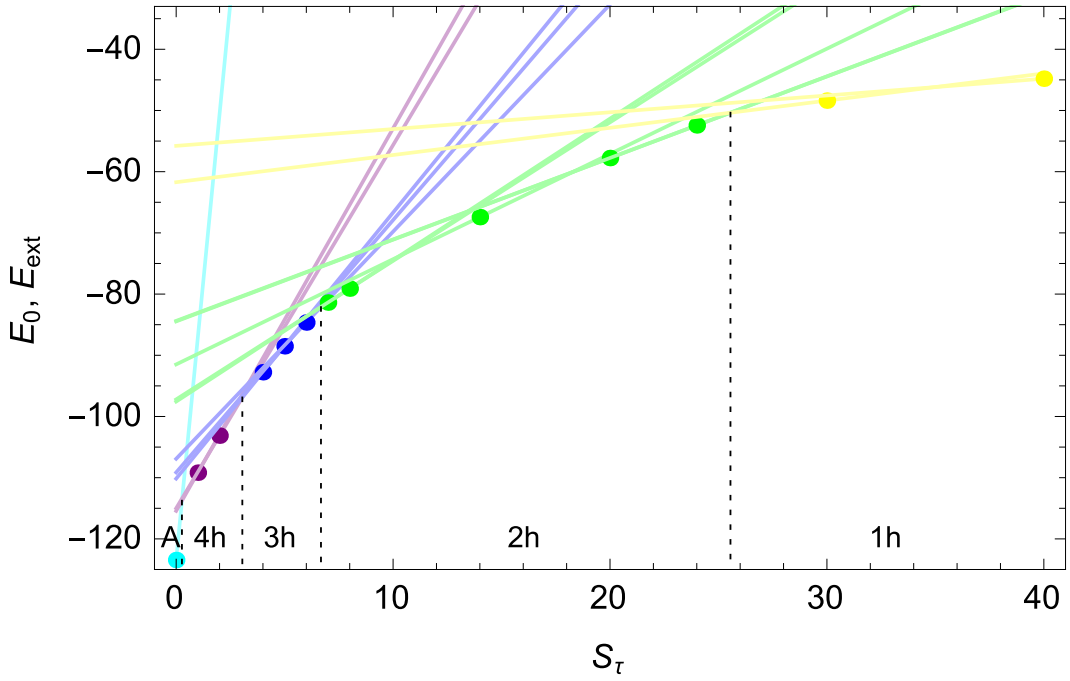

Figure 3.10: Energies E_0 of putative ground-state structures at different values of torsion strength S_τ (dots) for bending-restrained polymers ($N = 40$) with torsional barriers. The color of the dots and curves is consistent with the key in Fig. 3.3. The solid lines are hypothetic extrapolations of the energy $E_{\rm ext}(S_\tau)$ if the torsion strength in the torsion potential of a given ground-state structure is changed. The intersection points of lines with different color mark the crossover between different structure types of ground-state conformations. The thus identified S_τ threshold values agree with the zero-temperature transition points in the hyper-phase diagram shown in Fig. 3.6.

To understand the location of structure transition in S_τ, we consider only the single lowest-energy structure for each value of S_τ in the bending restrained case to be the putative ground state for that particular model. In Fig. 3.10 we plot a point for each structure's energy E_0, as well as a line for its extended energy

$$E_{\text{ext}}(S_\tau) = E_0 + (S_\tau - S_\tau^0) \sum_l v_{\text{tor}}(\tau_l), \tag{3.7}$$

where S_τ^0 is the torsion strength at which the structure was originally formed. E_{ext} effectively gives that structure's total energy in an alternate environment. Because $\sum_l v_{\text{tor}}(\tau_l)$ is a property of the structure and is not dependent on the environment, E_{ext} is a linear function with a slope proportional to the torsion disorder. A system in its ground state will always form the structure type with the lowest possible energy and will therefore form a structure along the E_{ext} line which is the lowest for a particular value of S_τ. The clustering of the lines is due to the nonlinear nature of the change in the torsion disorder. For example the three 3-helix blue curves are more similar to each other than they are to any of the 2-helix or 4-helix curves. It must be noted that there are three distinct clusterings for 2-helix structures, and this is due to the way in which the joint between the helices is constructed.

It can be inferred that the point at which lines from one structure type cross with lines from another structure type will be the point at which ground state structures will transition between the two. We can see good agreement between the transition locations on this plot and in Fig. 3.8.

3.7 Summary

In this chapter, we simulated bead-spring homopolymers with propensity for helical order. We performed a rigorous comparison of structures, transitions, and stability for flexible and

semiflexible models. The flexible polymers have no potential associated with its bending angles, and the semiflexible models have bending angles restrained to maintain a near fixed value. The tendency to form helical segments was controlled via a torsion potential. The strength of the torsion potential is tuned to explore a range of models spanning the space from no torsion potential to a very strong helix potential.

We find that stiff helical segments are formed when both a torsion and bending potential are included. The helical segments can vary in length and will align into bundles. The stiffness of the helical segments and, consequently, the number of helices per bundle are determined by the strength of the torsion potential.

Polymer chains lack helical segment stiffness and do not form stable organized helix bundles when an effective bending restraint is not included. Ensembles produced without bending-stabilization exhibit clustering at low temperature. This indicates instability in the structures formed. Without bending-stabilization, we also observe unpredictable sensitivity to a change in environmental factors such as torsion strength and temperature.

The lack of stability and tolerance to environmental variability provides insight into the observed preference in semiflexible biopolymers for the restriction of bond angles. This observation can be made in both DNA and most protein structures. The reduced degrees of freedom obtained by fixed bond angles in polypeptides is essential for functional structures to behave predictably and consistently.

Chapter 4

Adsorption of helix bundles

In this chapter, we further explore the structures and stability of helical bundles by considering the effect of an adsorption surface. Like the introduction of a bending restraint, adsorption to a substrate provides a restriction on the configuration space available to helical polymers. In this chapter we first study the influence of surfaces with many different adsorption strengths on a system with a single torsion strength. This will aid in our understanding of the adsorption transition for helical polymers and show for a single system the effect of varied surface adsorption strengths representing different substrate materials. We then consider only a single adsorption strength and look the effect of adsorption at an array of torsion strengths to understand more broadly the effect of adsorption for classes of polymers.

4.1 Adsorption of helical polymers on different surfaces

For simulations performed with the torsion energy scale $S_\tau = 6$, we consider an array of values for adsorption strengths in the interval $S_A \in [0, 6]$. Examples of the structures generated in

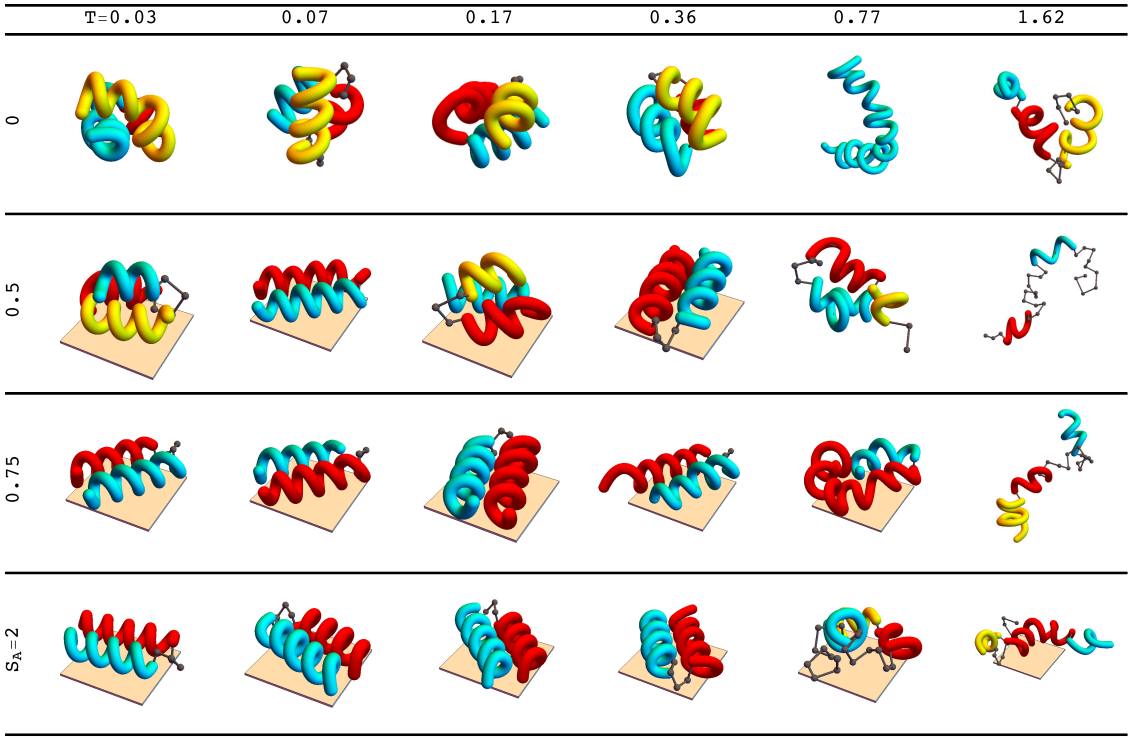

Figure 4.1: Representative polymer conformations formed under varied conditions. Each row shows structures for a single value of S_A along an array of temperatures between $T = 0.03$ and 1.62. The adsorption strength increases from top to bottom from a value of $S_A = 0$ to 2.

this series of simulations are shown in Fig. 4.1. Considering the top row ($S_A = 0$), polymer structures at low temperature are dominated by three-helix bundles and exhibit variability in terms of alignment and helix segment length. As the adsorption strength is increased we see that the polymers adsorb onto the surface at low temperatures and form two-helix bundles instead of three-helix bundles. We also find qualitatively that the adsorbed structures exhibit far less variability over a wider range of temperatures.

4.1.1 Variation of structures with temperature

It can be useful to calculate canonical averages for structural quantities to observe their variation as temperature and model parameters are changed. The parameter $\langle q_2 \rangle / \langle q_1 \rangle$ is the ratio of the global to local Lennard-Jones energy, and it is a good parameter to distinguish between various helical structure types. For three-helix bundles observed at low temperature in the case of $S_A = 0$, $\langle q_2 \rangle / \langle q_1 \rangle \approx 0.64$. In the case of the two-helix bundles which dominate at low temperature when $S_A \geq 0.75$ we find that $\langle q_2 \rangle / \langle q_1 \rangle \approx 0.33$.

Fig. 4.2 (a) gives $\langle q_2 \rangle / \langle q_1 \rangle$ as a function of T for each value of S_A. By comparing the different values S_A, we can see splitting at low temperature between two- and three-helix bundle branches. For higher temperature it is apparent that the different simulation ensembles become more similar. Fig 4.2 (b) shows the thermal fluctuation of the center of mass distance from the substrate (dh_{cm}/dT). The temperature at which polymers desorb from the surface can be determined by noting the peak in this plot as it corresponds to the temperature at which there the most even mixture between adsorbed and desorbed structures. There is agreement between the desorption temperature noted in panel (b) and the convergence to the free case ($S_A = 0$) case for the curves in panel (a). We can also see that the desorption temperature increases with increasing adsorption strength as is expected. More interestingly, in the range from $S_A = 0.75$ to 2 the low-temperature structure and structural stability to perturbation in temperature changes very little.

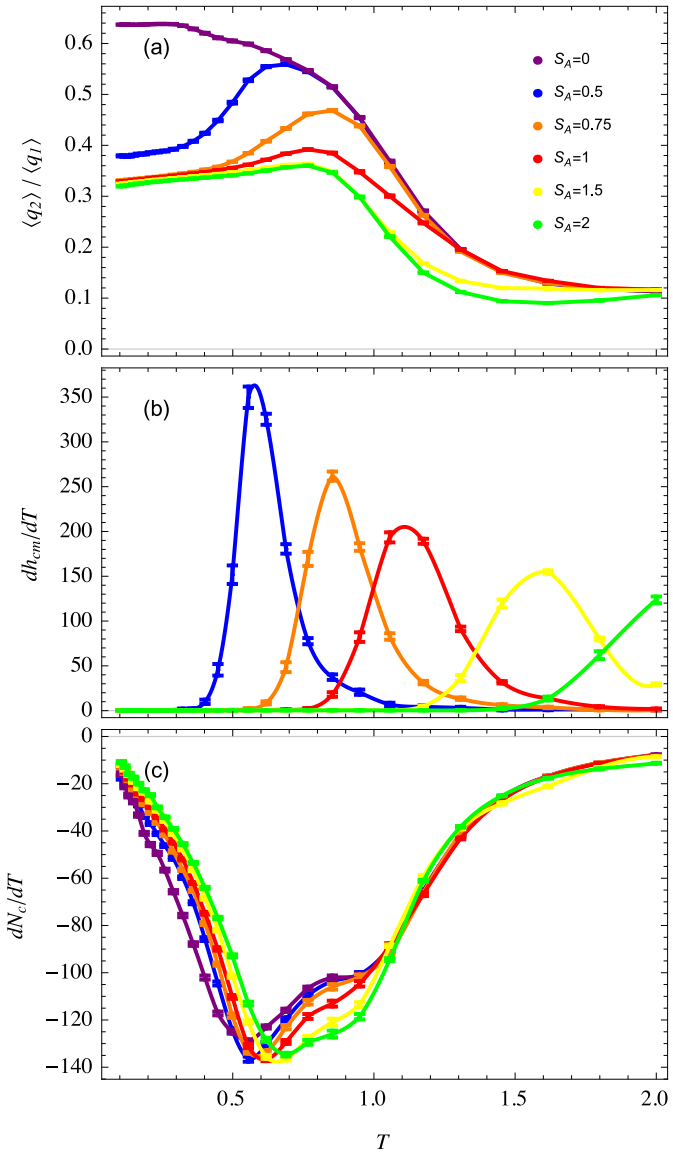

Figure 4.2: (a) $\langle q_2\rangle/\langle q_1\rangle$ as a function of temperature for an array of different values for S_A. (b) Temperature variation of the center of mass distance of polymers at several different adsorption strengths. Note that the peak locations correspond to the temperature at which the polymer desorbs from the surface. (c) Temperature variation of the number of monomer-monomer contacts vs temperature. N_c decreases most rapidly during structural transitions between solid, liquid and gas exhibiting inverted peaks or shoulders in the plot of dN_c/dT.

Fig. 4.2 (c) shows the thermal fluctuation of the number of monomer-monomer contacts as a function of T. The gas-liquid transition and the liquid-solid transition are both marked by a rapid rate of change in the number of monomer-monomer contacts and can be located in this plot by finding inverted peaks and shoulders. We can see that as the adsorption strength increases the temperature range over which the liquid phase occurs narrows.

4.1.2 Structure clustering in $q_1 - q_2$ space

We can better understand the relationship between the ensembles formed at each value of S_A by considering the structures produced in each of these simulations in $q_1 - q_2$ space, as seen in Fig. 4.3. The low-energy structures (produced at low temperature) must in general have lower q_1 and q_2 values, and therefore lie towards lower left.

For $S_A = 0$ or 0.5, low-temperature structures are spread out over several unique clusters. Each of these clusters represents a unique structure type. The inset in Fig. 4.3 shows only the low-energy clusters corresponding to $S_A = 0$ along with example structures for each discernible cluster. While the lower right hand cluster is made up of three-helix bundles in which all helices are parallel and of approximately the same length, the upper left and cluster contains structures in which two long helical segments are wrapped around each other and a short helical segment connects their ends.

The presence of multiple low-temperature cluster for both $S_A = 0$ and 0.5 is due to local free-energy minima near in phase space to the ground state. Ensembles exhibiting this behavior are inherently unstable due to the varied structure type accessible at a single low-temperature. Additionally they are highly sensitive to changes to their environment, as seen by the drastic change in structure type with the introduction of even a weak adsorption surface. Not shown here is their high sensitivity to other changes such as a small change in S_τ.

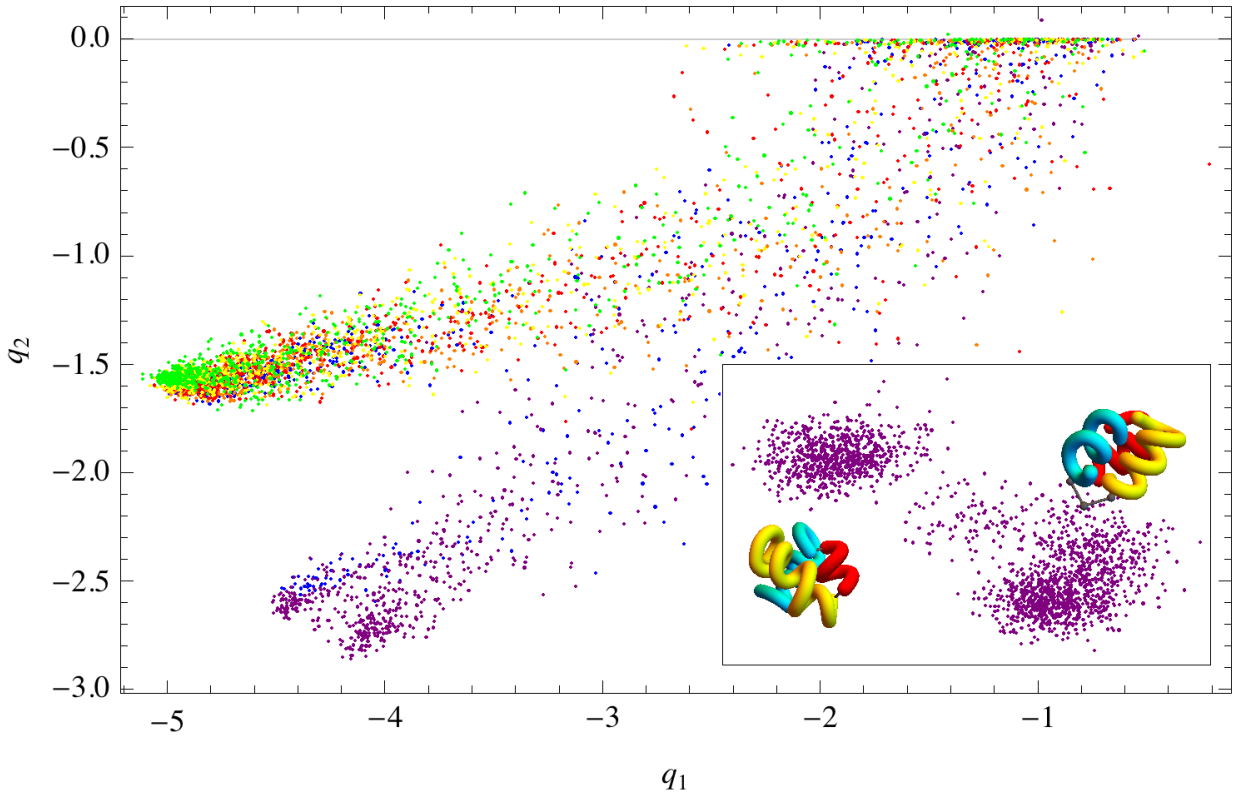

Figure 4.3: Shown above is a plot of q_1 vs q_2 for a collection of example structures produced across all temperatures simulated. The color of each point corresponds to the torsion energy scale at which the structure it represents was formed. Structures in q_1, q_2 space cluster with similar structure types. Low-temperature structures for $S_A = 0$ are shown independently in the lower right inset along with example structures for each cluster. Colors agree with the legend shown in Fig. 4.2.

As we increase S_A and transition from three-helix to two-helix structures we find that the system sacrifices Lennard-Jones energy from globular collapse (q_2) in favor of lower energy in the torsion potential, larger contact with the adsorption surface, and lower Lennard-Jones energy of local contacts associated with q_1. This is demonstrated by a decrease in q_1 and increase in q_2. The inherent instability and high sensitivity shown in the ensembles with $S_A = 0$ and 0.5 is greatly reduced by the inclusion of the adsorption surface as demonstrated by the consistent and single-peaked clusters corresponding to two-helix bundles formed for

all values of S_A between 0.75 and 2.0. At these higher values of adsorption strength the low-temperature configuration are predictable, consistent, and highly resilient to changes in their environment.

We notice also that the primary effect of a varied adsorption strength is the temperature at which the polymer adsorbs to the substrate. The particular strength chosen does not however have a strong effect on the geometry of adsorbed structures in within a wide range of values of S_A between 0.75 and 2. This means that within this adsorption regime the influences on structure formation are a binary effect where once a structure is adsorbed it no longer matters the strength of the adsorption surface. This will remain true until the adsorption surface is strong enough to compete with the torsion potential.

4.2 Adsorbed and free structures for an array S_τ values

The particular case of $S_\tau = 6$ shown above exhibits a dramatic qualitative shift when adsorption is added to the standard helix bundle model. We will further explore the effect of the adsorption surface on a wide array of S_τ values between 0 and 30. To simplify this exploration we will consider only the free case ($S_A = 0$) and a single adsorbed case ($S_A = 2$).

4.2.1 Low temperature free and adsorbed structures

In Fig. 4.4 we see examples of structures formed at low temperature with and without the adsorption surface. We see that for the free case (top row), amorphous solids, 4-helix bundles, 3-helix bundles, 2-helix bundles, and single helixes are formed. In the adsorbed case (bottom row), the types of structures formed are much more restricted. Above $T \approx 2$, we consistently find stable formation of two-helix bundles. This can be observed quantitatively in Fig 4.5, where, for low-temperature (a), in the region $1 \leq S_\tau < 8$, the free case clearly exhibits 3- and 4-helix bundles while the adsorbed case forms predominantly 2-helix bundles.

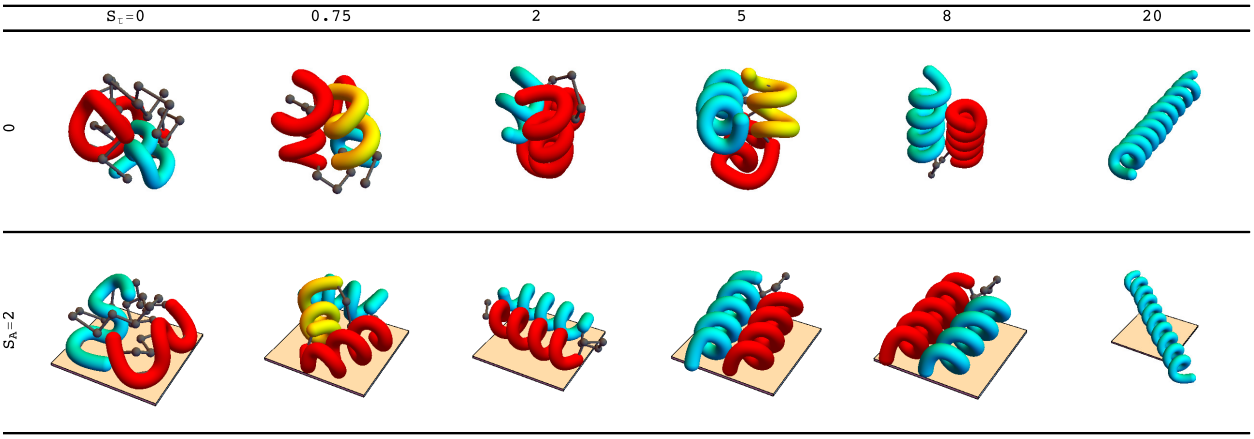

Figure 4.4: Example structures from a collection at low temperature in an array of simulations. This array of structures highlights the differences in structure formation in helical polymers with and without an adsorption surface.

The parameter space over which the 2-helix bundle dominates is expanded to lower values of S_τ due to the adsorption surface. Similarly, the single-helix phase dominates over a larger region of phase space in the adsorbed case because the addition of an adsorption surface entropically suppresses the two-helix phase more than the single-helix phase.

As the temperature is increased structural variability within each canonical ensemble increases also. As this variability increases we see less consistency of structures in a single phase across an array of S_τ values. In panel (b) of Fig. 4.5 this is reflected in more gradual transitions in S_τ space and a steeper slope within the stable regions of each phase for both the free and adsorbed cases. The smearing of the transitions indicates the presence of intermediate structures between the two structure types as well as mixed phase ensembles. The increased slope within phases is a direct result of the increased variability of structures in that phase, with each S_τ value producing slightly different structures. Both of these effects are seen in both the free and adsorbed models. At $T = 0.45$, we also see the transition between two-helix bundles and single helices occurring at a lower value of of S_τ. All three

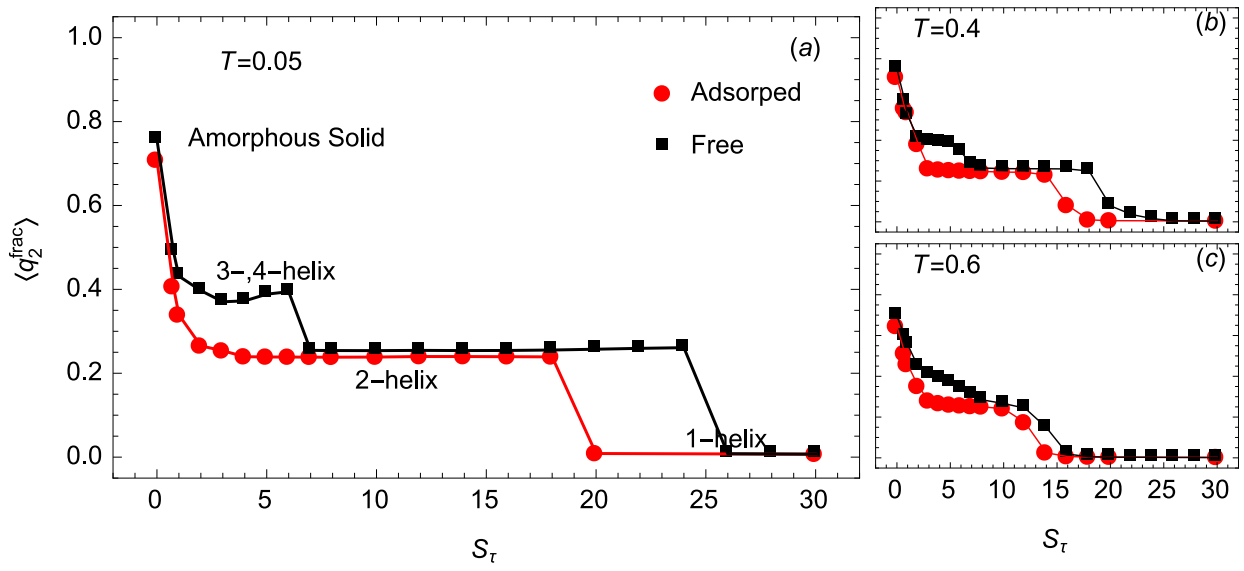

Figure 4.5: All three panels show structural parameter $\langle q_2^{\text{frac}} \rangle$ plotted for a single temperature ensemble from each value S_τ. We can see the different structural phases formed and transitions between them. (a) For $T = 0.05$ the two systems depart from each other strongly in $1 < S_\tau < 8$ where in the free case 3- and 4-helix bundles dominate, whereas the adsorbed polymer directly forms a 2-helix bundle. Similar behavior is observed for (b) and (c) but with the behavior exhibiting progressively less dramatic transitions as temperature is increased.

of these effects are more pronounced in the free case. For ensembles produced at $T = 0.7$ in Fig. 4.5 (c) we see the same effects described for $T = 0.45$ but to a greater extent. Here the free case has almost no stratification between single-, 2-, 3-, 4-helix phases. In contrast the 2-helix phase in the adsorbed case remains far more discernible. This phase stability into higher temperatures is an important feature due to the presence of the substrate.

The ensembles from which the canonical ensembles in Fig. 4.5 come, can be better understood by considering the histograms which are averaged to obtained the canonical mean. In Fig. 4.6 we see the histograms produced by q_2^{frac} values for all structures in each canonical ensemble produced at $T = 0.05$. In the free case we see multi-peaked histograms for many of the 3-helix and 4-helix ensembles. We also see stark variation between the details

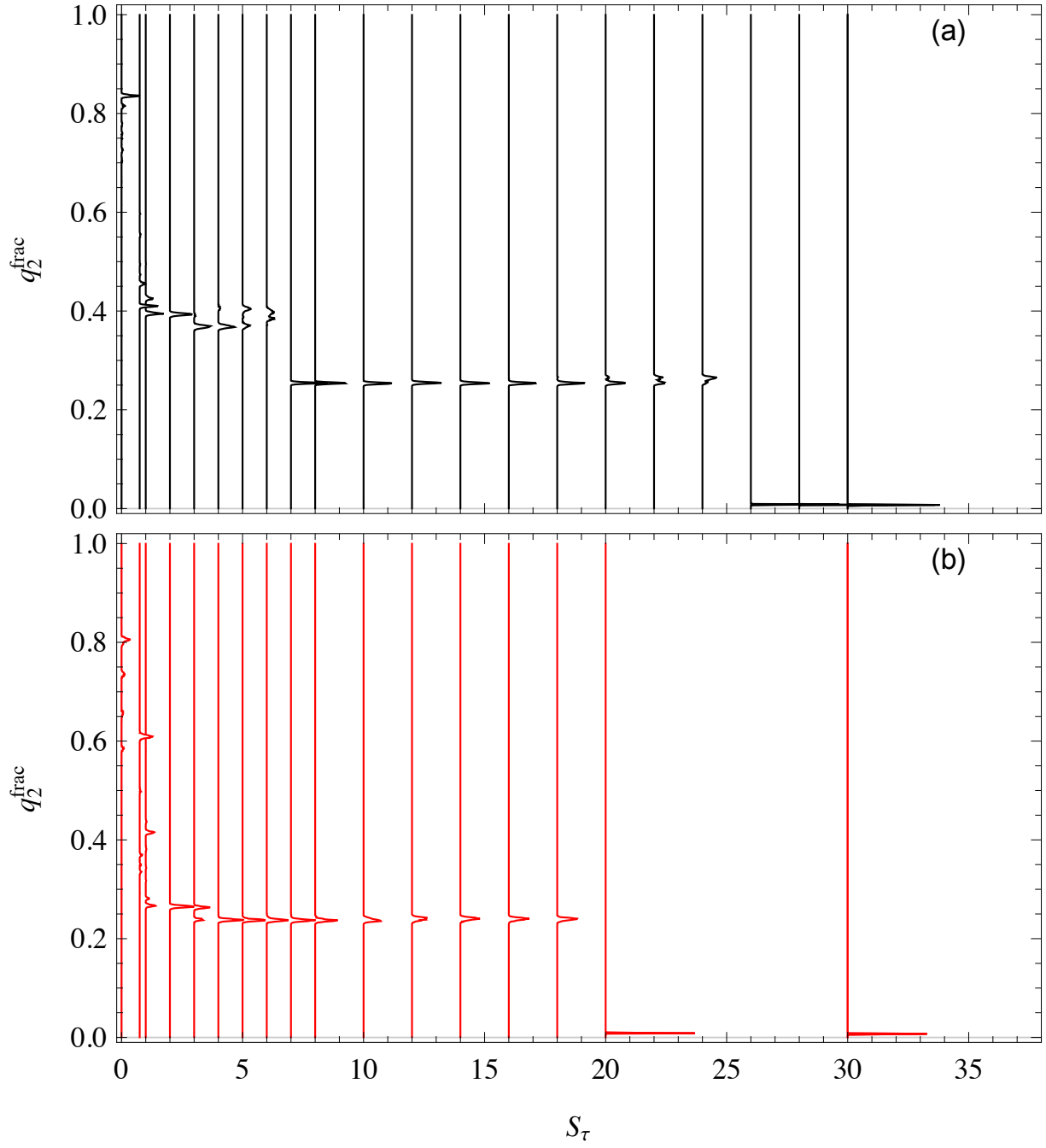

Figure 4.6: q_2^{frac} histograms for canonical ensembles at $T = 0.05$. The base line for each histogram aligns the value of S_τ at which it was generated. Panel (a) shows the histograms generated in the case of a semiflexible free polymer, and panel (b) gives the histograms generated for semiflexible adsorbed polymers.

of neighboring histograms. This variation occurs because as we move from $S_\tau = 1$ to 6 we transition between 3- and 4-helix bundle as well as different helix segment orientations. For the adsorbed ensembles in the same region of S_τ space, we observe much more consistent formation of helix bundles which in this case are 2-helix bundles. For $S_\tau = 20$ to 25 in the free model double peaked histograms are caused by a change in the geometry of the joint between helix segments. In the semiflexible case this type of instability in the two-helix bundle geometry is completely eliminated for all $S_\tau \geq 4$.

4.2.2 Structures in $q_1 - q_2$ space

To further draw distinction between the free and adsorbed cases, we consider the region in $q_1 - q_2$ space where structures are formed for a selection of simulations. Each panel in Fig. 4.7 represents all of the structures generated in a single parallel tempering simulation. The structures collected in the simulation across all temperatures are projected onto a two dimensional space of q_1 and q_2 which is divided into bins. Any bin which reaches a threshold number of structures is colored black. Black region where structures are generated in each simulation is shown on a background of gray representing all of the structures formed across all values of S_τ. Both the black and gray regions represent structures formed access all temperatures which are included in the simulation. As S_τ increases, q_1 increases, sacrificing local order in favor of increased global Lennard-Jones connections and thus decreased q_2.

Structures formed at lower temperature tend to have lower values of both q_1 and q_2 clustering in the lower left of the particular branch along which structures are formed. Regions with structures dominant at $T < 0.075$ are colored in red. Splitting of low-temperature peaks in the $q_1 - q_2$ space indicates local minima surrounding the global minimum. Local minima can decrease the stability of structures as well as increase their variability. For $S_\tau = 0$, we see in both the adsorbed and free case that there are many low-temperature clusters and these represent a very rough free-energy landscape. As S_τ increases, the additional constraint of

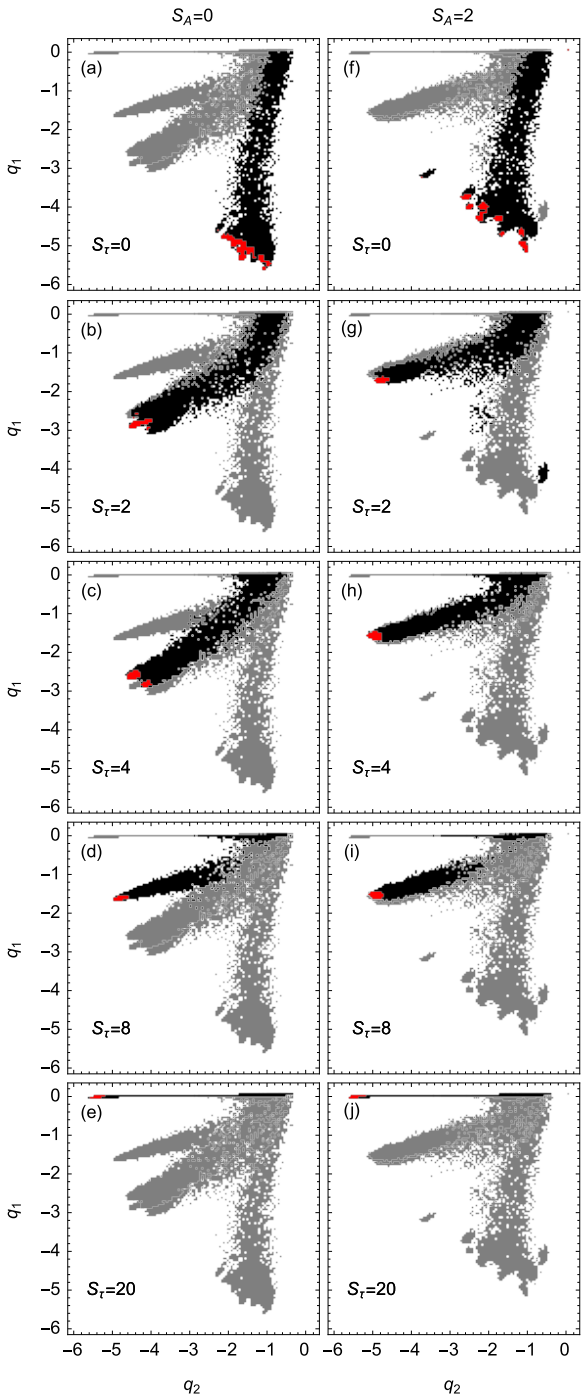

Figure 4.7: For each panel the gray background area represents the region in which structures are formed for all values of S_τ and T for either the free (left column (a)-(e)) or adsorbed (right (f)-(j)) simulations. In the foreground of each panel the area shows where structures are formed for all temperatures at the particular model parameters corresponding to the panel. The red area highlights the region where low-temperature ($T < 0.075$) structures are formed.

the torsion potential works to eliminate the local minima surrounding the global minimum. At $S_\tau = 2$ and 4 the low-temperature structures appear to be more stable in the adsorbed case and still exhibit local minima at $S_\tau = 0$. For $S_\tau \geq 8$, there does not appear to be any significant variation in the results obtained in free and adsorbed simulations.

4.2.3 Hyper-phase diagrams of adsorbed and free helical polymers

For each simulation thread with a unique S_τ and T it is possible for us to assign a phase at which the ensemble exists by studying the locations of structures generated within the q_1 q_2 space. Considering the branch along which the majority of structures occurs can help to distinguish between single helices, two-helix bundles, three-helix bundle, four-helix bundles, and amorphous solids. For polymers of length 40, which are mainly discussed here, the branches formed by three- and four-helix bundles are overlapping and difficult to distinguish in all but the lowest temperature cases.

For each simulation thread a phase is chosen based on transitions observed in the canonical quantities as well as the branch in which structures lie in the $q_1 - q_2$ plot. These phases are shown in Fig. 4.8, in which the color of each region represents the state which is dominant under the conditions of the simulation corresponding to the area covered. The lower left corner of both plots are difficult to classify and unpredictable because of the inherent instability of low-temperature structures with models with low S_τ. The primary contrast between the free and adsorbed cases is the expanded two-helix (green) region, which replaces the blue three-helix region in the free case.

4.3 Summary

In this chapter, we performed simulations to determine the influence of the presence of an adsorption surface on helical polymers. To begin, we performed simulations for a single tor-

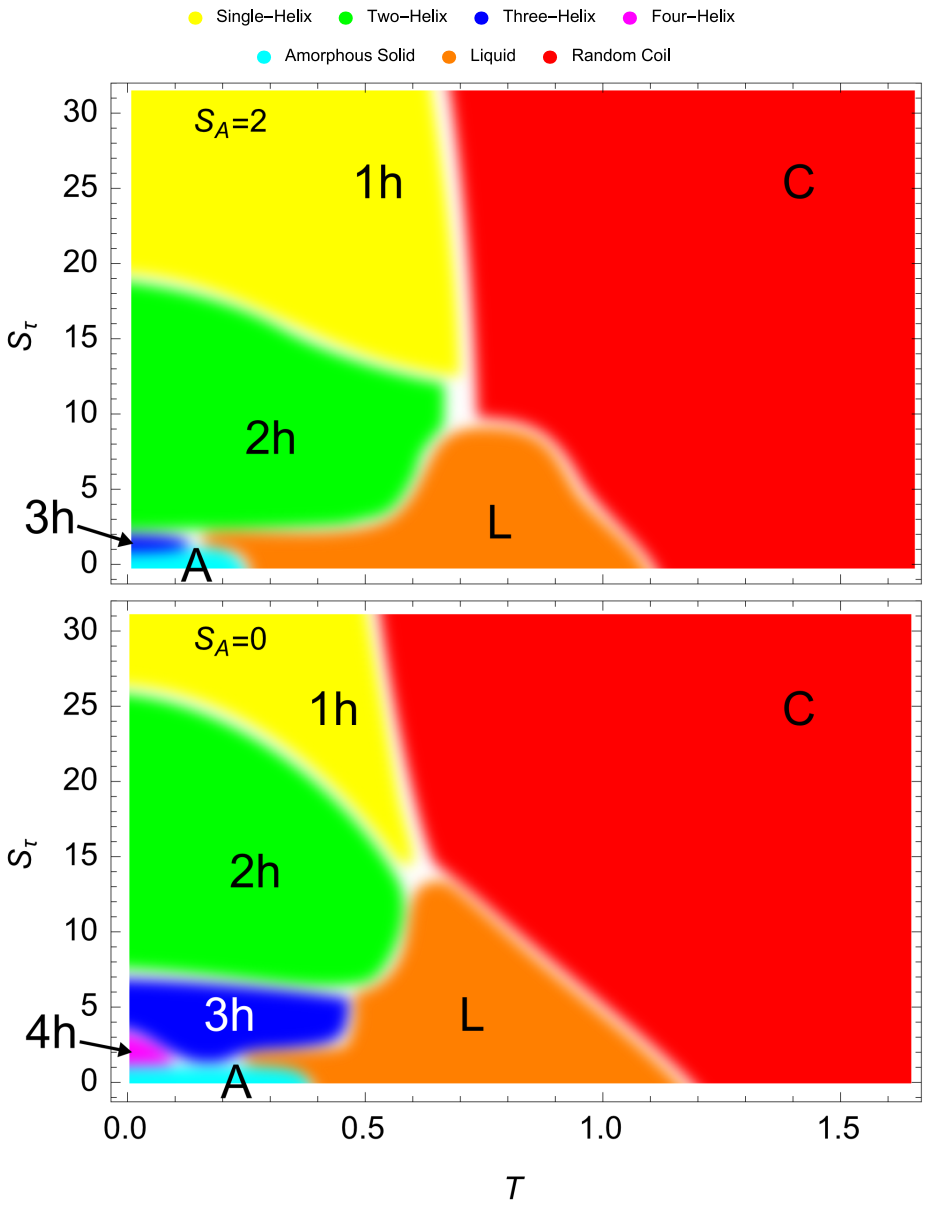

Figure 4.8: Hyper-phase diagrams representing the dominant structural phase present at each S_τ, T combination for both the adsorbed case ($S_A = 2$) and the free case ($S_A = 0$) . Black lines are approximate locations of phase transitions.

sion strength across an array of adsorption strengths. The presence of an adsorption surface can influence the structures formed as well as the stability of those structures. In the free case, we see that for the particular adsorption strength chosen ($S_\tau = 6$), three-helix bundles are formed with two distinct organizations. Multiple low-temperature clusters indicate a local free-energy minimum near the ground state in parameter space. The presence of local minima near the global minimum causes structural instability as well as high sensitivity to slight changes in the environment.

The free-energy local minimum present in the free case disappears for adsorbed polymers. We also see a single well-defined low-temperature cluster representing two-helix bundles. The lack of local minima indicates more stable structures with less variability and more resiliency to changes in the environment.

As S_A increases we note that the adsorption transition occurs at higher and higher temperatures and that the liquid phase exists over a smaller temperature region. The Θ-transition and the freezing transition eventually merge into a single structural phase transition from gas to two-helix bundle.

Simulation of free and adsorbed polymers across an array of torsion strength values gives a broader understanding of the effect of an adsorption surface on many different types of helical structures. With $S_\tau = 0$, the adsorbed and free cases exhibit many low-temperature free-energy minima, indicating high instability as well as high sensitivity. For values of S_τ between 1 and 6, we see high instability and sensitivity in the 3- and 4- helix bundles formed in the free case but reliable 2-helix bundles in the adsorbed case. While structures formed at torsion strength $S_\tau < 8$ prove very different in the adsorbed and free cases, structures formed at $S_\tau \geq 8$ show relatively little influence from the adsorption surface.

Chapter 5

Conclusion

This dissertation has focussed on the structural behavior of helical polymers. These polymers are represented by a homopolymer model based on effective potentials that represent the FENE interaction between neighboring monomers, the Lennard-Jones interaction between nonbonded monomers, a torsion potential associated with the dihedral angles, a bending potential associated with bond angles, and interaction with a substrate via an adsorption potential. Each of these potentials has a tunable energy scale that allows for the relative influence of each interaction to be varied. Using these sets of potentials, we explore structural geometry, transitions between structures, structural stability, and hyper-phase diagrams comparing structural domains under various conditions in order to gain insight into folding pathways, structure transitions, ground state structures, and stability.

Simulations of helical polymers present an obvious comparison to biological polymers such as polypeptides, which often form helical structures and helix bundle structures. It can be observed that bending angles between amino acids are fixed in polypeptide chains. We examine the effect of a bending restraint on coarse-grained polymers with a propensity for local helical order via their torsion angles.

With the inclusion of a torsion potential, both flexible and semiflexible polymers exhibit local helical order. The stabilization provided to the helix segments by a bending restraint present in the semiflexible case leads to formation of helix bundles. The number of helical segments in the bundle is dependent on the helix segment stiffness and the length of the polymer. Since the helical segments are divided by a joint which is formed by the disturbance of several torsion angles, the segment stiffness is dictated by the strength of the torsion potential. In the case of the flexible helical polymer, we find that low-temperature ensembles are highly unstable in both structural consistency and with respect to variation of system parameters.

Predictability and stability are essential to consistent functionality in biological systems. We hypothesize that the bending restraint observed in nature is necessary for the stability of protein structures.

Further stabilization of helical structures can be achieved by the introduction of an adsorption surface. For semiflexible helical polymers at a single torsion strength, structure formation and stability are compared in the presence of adsorbing surfaces with an array of adsorption strengths. The torsion strength is chosen such that the free polymer configurations are dominated by three-helix bundles at low temperatures. When adsorbed to the surface, the low-temperature configuration of the polymer changes from 3-helix bundle to 2-helix bundle. The adsorbed 2-helix phase is stable to a higher temperature and more structurally stable than in the free case. As adsorption strength increases the temperature at which the polymer adsorbs to the surface increases as well. Interestingly, beyond controlling the temperature at which the polymer adsorbs, the adsorption strength has relatively little influence on the particular structure of the adsorbed polymer. The low-temperature 2-helix structures across different adsorption strengths are very similar. Of course if the adsorption strength were increased further, to the point where it overcomes the torsion strength, the helical order would be sacrificed to maximize surface contacts.

We extend this study of adsorbed helical polymers further by considering the entire spectrum of torsion strengths for both the free case and the adsorbed case. For weak torsion potentials, unstable 3- and 4-helix structures dominant for the free case are replaced by the far more stable 2- helix structure phases in the adsorbed case. The transition between the single-helix phase and 2-helix bundle phase also occurs at a weaker value of the torsion potential.

Although we have no means to test the influence on evolution and abiogenesis, we do find increased stability of helical structures when attached to an adsorbing substrate. These results support the findings of Lundqvist *et al.* [78] who showed that the presence of silica nanoparticles would contribute a surface onto which organic macromolecules could attach and be stabilized.

Bibliography

[1] M. J. Williams and M. Bachmann, Phys. Rev. Lett. **115**, 048301 (2015).

[2] M. J. Williams and C. Fertig, Phys. Rev. A **91**, 023432 (2015).

[3] K. Binder, Zeitschrift für Phys. B Condens. Matter **43**, 119 (1981).

[4] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D. Hwang, Phys. Rep. **424**, 175 (2006).

[5] S. Cole, C. G. Lacey, C. M. Baugh, and C. S. Frenk, Mon. Not. R. Astron. Soc. **319**, 168 (2002).

[6] P. J. Flory, *Principles of Polymer Chemistry* (Cornell University Press, 1953).

[7] P.-G. de Gennes, *Scaling Concepts in Polymer Physics* (Cornell University Press, 1979).

[8] D. C. Rapaport, J. E. Johnson, and J. Skolnick, Comput. Phys. Commun. **121-122**, 231 (1999).

[9] F. Chiti and C. M. Dobson, Annu. Rev. Biochem. **75**, 333 (2006).

[10] A. V. Badasyan, A. Giacometti, Y. S. Mamasakhlisov, V. F. Morozov, and A. S. Benight, Phys. Rev. E **81**, 021921 (2010).

[11] D. Poland and H. A. Scheraga, *Theory of Helix-Coil Transitions in Biopolymers: Statistical Mechanical Theory of Order-Disorder Transitions in Biological Macromolecules* (Academic Press, 1970).

[12] M. Bachmann, *Thermodynamics and Statistical Mechanics of Macromolecular Systems* (Cambridge University Press, 2014).

[13] C. Levinthal, Extr. Du J. Chim. Phys. **65**, (1968).

[14] K. A. Dill, Biochem. **24**, 1501 (1985).

[15] J. D. Bryngelson, J. N. Onuchic, N. D. Socci, and P. G. Wolynes, Proteins **21**, 167 (1995).

[16] M. Karplus, Fold. Des. **2**, S69 (1997).

[17] M. Bachmann and W. Janke, Phys. Rev. E **73**, 1 (2006).

[18] G. C. Rollins and K. A. Dill, J. Am. Chem. Soc. **136**, 11420 (2014).

[19] T. Vogel, J. Gross, and M. Bachmann, J. Chem. Phys. **142**, 104901 (2015).

[20] T. Geisinger, M. Müller, and K. Binder, J. Chem. Phys. **111**, 5241 (1999).

[21] M. Watzlawek, C. N. Likos, and H. Löwen, Phys. Rev. Lett. **82**, 5289 (1999).

[22] M. Möddel, M. Bachmann, and W. Janke, J. Phys. Chem. B **113**, 3313 (2009).

[23] S. Schnabel, M. Bachmann, and W. Janke, Phys. Rev. Lett. **98**, 048103 (2007).

[24] D. H. E. Gross, *Microcanonical Thermodynamics: Phase Transitions in "Small" Systems Systems* (World Scientific, 2001).

[25] S. Schnabel, D. T. Seaton, D. P. Landau, and M. Bachmann, Phys. Rev. E **84**, 1 (2011).

[26] G. Cornilescu, F. Delaglio, and A. Bax, J. Biomol. NMR **13**, 289 (1999).

[27] A. MacKerell, D. Bashford, M. Bellott, R. Dunbrack, J. Evanseck, M. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. Lau, C. Mattos, S. Michnick, T. Ngo, D. Nguyen, B. Prodhom, W. Reiher, B. Roux, M. Schlenkrich, J. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus, J. Phys. Chem. B **102**, 3586 (1998).

[28] S. Förster, E. Kohl, M. Ivanov, J. Gross, W. Widdra, and W. Janke, J. Chem. Phys. **141**, 164701 (2014).

[29] J. N. Onuchic, Z. Luthey-Schulten, and P. G. Wolynes, Annu. Rev. Phys. Chem. **48**, 545 (1997).

[30] J. Gross, T. Neuhaus, T. Vogel, and M. Bachmann, J. Chem. Phys. **138**, 074905 (2013).

[31] M. Bachmann and W. Janke, J. Chem. Phys. **120**, 6779 (2004).

[32] S. Schöbl, J. Zierenberg, and W. Janke, Phys. Rev. E **84**, 051805 (2011).

[33] S. Schnabel, M. Bachmann, and W. Janke, J. Chem. Phys. **131**, 124904 (2009).

[34] M. P. Taylor, W. Paul, and K. Binder, J. Chem. Phys. **131**, 114907 (2009).

[35] J. A. Martemyanova, M. R. Stukan, V. A. Ivanov, M. Müller, W. Paul, and K. Binder, J. Chem. Phys. **122**, 174907 (2005).

[36] P. A. Wiggins and P. C. Nelson, Phys. Rev. E **73**, 031906 (2006).

[37] D. T. Seaton, S. Schnabel, M. Bachmann, and D. P. Landau, Int. J. Mod. Phys. C **23**, 1240004 (2012).

[38] J. Zierenberg and W. Janke, Europhys. Lett. **109**, 28002 (2015).

[39] C. Junghans, M. Bachmann, and W. Janke, Phys. Rev. Lett. **97**, 218103 (2006).

[40] R. Hegger and P. Grassberger, J. Phys. A. Math. Gen. **27**, 4069 (1994).

[41] R. Rajesh, D. Dhar, D. Giri, S. Kumar, and Y. Singh, Phys. Rev. E **65**, 056124 (2002).

[42] M. Möddel, W. Janke, and M. Bachmann, Phys. Chem. Chem. Phys. **12**, 11548 (2010).

[43] M. Möddel, W. Janke, and M. Bachmann, Comput. Phys. Commun. **182**, 1961 (2011).

[44] T. Vogel and M. Bachmann, Comput. Phys. Commun. **182**, 1928 (2011).

[45] M. J. Williams and M. Bachmann, *Effect of Bending Restraint on Helical Polymer Structure Formation and Stability*, To Be Published (2016).

[46] D. C. Rapaport, Phys. Rev. E **66**, 1 (2002).

[47] S. Sabeur, Cent. Eur. J. Phys. **12**, 421 (2014).

[48] T. Vogel, T. Neuhaus, M. Bachmann, and W. Janke, Europhys. Lett. **85**, 10003 (2009).

[49] J. A. Scott, J. Gaspar, M. Stuchell, S. Alam, J. Skalicky, and W. I. Sundquist, To Be Publ. (n.d.).

[50] *http://www.rcsb.org/pdb/explore.do?structureId=1YXR*

[51] D. J. Barlow and J. M. Thornton, J. Mol. Biol. **201**, 601 (1988).

[52] A. Maritan, C. Micheletti, A. Trovato, and J. R. Banavar, Nature **406**, 287 (2000).

[53] T. Vogel, T. Neuhaus, M. Bachmann, and W. Janke, Eur. Phys. J. E **30**, 7 (2009).

[54] J. P. Kemp and Z. Y. Chen, Phys. Rev. Lett. **81**, 3880 (1998).

[55] B. H. Zimm and J. K. Bragg, J. Chem. Phys. **31**, 526 (1959).

[56] B. H. Zimm and J. K. Bragg, J. Chem. Phys. **28**, 1246 (1958).

[57] Z. Qin, A. Fabre, and M. J. Buehler, Eur. Phys. J. E **36**, 53 (2013).

[58] Z. Luthey-Schulten, B. E. Ramirez, and P. G. Wolynes, J. Phys. Chem. **99**, 2177 (1995).

[59] T. Bereau, M. Deserno, and M. Bachmann, Biophys. J. **100**, 2764 (2011).

[60] V. Varshney and G. A. Carri, Phys. Rev. Lett. **95**, 168304 (2005).

[61] J. E. Magee, V. R. Vasquez, and L. Lue, Phys. Rev. Lett. **96**, 207802 (2006).

[62] K. Ghosh and K. A. Dill, J. Am. Chem Soc **131**, 2306 (2010).

[63] T. Bereau, M. Bachmann, and M. Deserno, J. Am. Chem. Soc. **132**, 13129 (2010).

[64] C. Zhang, J. Hou, and S.-H. Kim, Proc. Natl. Acad. Sci. U. S. A. **99**, 3581 (2002).

[65] S. W. Bruun, V. Iešmantavičius, J. Danielsson, and F. M. Poulsen, Proc. Natl. Acad. Sci. U.S.A **107**, 13306 (2010).

[66] Z. Guo, C. L. Brooks, and E. M. Boczko, Proc. Natl. Acad. Sci. U.S.A **94**, 10161 (1997).

[67] N. L. Harris, S. R. Presnell, and F. E. Cohen, J. Mol. Biol. **236**, 1356 (1994).

[68] A. Irbäck, F. Sjunnesson, and S. Wallin, Proc. Natl. Acad. Sci. U. S. A. **97**, 13614 (2000).

[69] S. R. Presnell and F. E. Cohen, Proc. Natl. Acad. Sci. U. S. A. **86**, 6592 (1989).

[70] P. Palenčár and T. Bleha, Comput. Theor. Chem. **1006**, 62 (2013).

[71] M. J. Williams and M. Bachmann, in Phys. Procedia (Elsevier B.V., 2015), pp. 130-134.

[72] M. J. Williams and M. Bachmann, *The Effect of Surface Adsorption on Tertiary Structure Formation in Helical Polymers*, To Be Published (2016).

[73] E. Eisenriegler, K. Kremer, and K. Binder, J. Chem. Phys. **77**, 6296 (1982).

[74] N. Källrot and P. Linse, Macromolecules **40**, 4669 (2007).

[75] M. Bachmann and W. Janke, Phys. Rev. Lett. **95**, 058102 (2005).

[76] W. Norde and J. P. Favier, Colloids Surf. **64**, 87 (1992).

[77] G. A. Carri and M. Muthukumar, Phys. Rev. Lett. **82**, 5405 (1999).

[78] M. Lundqvist, P. Nygren, B. H. Jonsson, and K. Broo, Angew. Chemie - Int. Ed. **45**, 8169 (2006).

[79] A. Y. Grosberg and A. R. Khokhlov, *Statistical Physics of Macromolecules* (AIP Press, 1994).

[80] K. A. Dill and H. S. Chan, Nat. Struct. Biol. **4**, 10 (1997).

[81] R. B. Bird, C. F. Curtiss, R. C. Armstrong, and O. Hassager, *Dynamics of Polymeric Liquids, 2nd Ed.* (Wiley, 1987).

[82] K. Kremer and G. S. Grest, J. Chem. Phys. **92**, 5057 (1990).

[83] A. Milchev, A. Bhattacharya, and K. Binder, Macromolecules **34**, 1881 (2001).

[84] J. E. Lennard-Jones, Proc. Phys. Soc. **43**, 461 (1931).

[85] D. T. Seaton, S. Schnabel, D. P. Landau, and M. Bachmann, Phys. Rev. Lett. **110**, 028103 (2013).

[86] D. J. Earl and M. W. Deem, Phys. Chem. Chem. Phys. **7**, 3910 (2005).

[87] K. Hukushima and K. Nemoto, J. Phys. Soc. Japan **65**, 1604 (1996).

[88] K. Hukushima, H. Takayama, and K. Nemoto, Int. J. Mod. Phys. C **07**, 337 (1996).

[89] R. H. Swendsen and J.-S. Wang, Phys. Rev. Lett. **57**, 2607 (1986).

[90] C. J. Geyer, in Comput. Sci. Stat. Proc. 23rd Symp. Interface (Interface Foundation, Fairfax, VA, 1991), pp. 156-163.

[91] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).

[92] K. Qi and M. Bachmann, J. Chem. Phys. **141**, 074101 (2014).