

Article

# Microcanonical Analysis of Helical Homopolymers: Exploring the Density of States and Structural Characteristics

Matthew J. Williams 

Institute of Engineering, Murray State University, Murray, KY 42071, USA; mwilliams72@murraystate.edu

**Abstract:** This study investigates the density of states and structural characteristics of helical homopolymers. Comprising repeating identical units, the model enables the exploration of complex behaviors arising from a simple, yet generalized, set of potentials. Utilizing microcanonical analysis, transitions between helical structures are identified and categorized. Through a systematic comparison of results under varying conditions, we develop a nuanced understanding of the system's general behavior. A two-dimensional plot illustrates the relative distribution of different structural types, effectively showcasing their prevalence. The findings of this study substantially advance our understanding of the density of states and structural transformations of helical homopolymers across a range of conditions. Additionally, the prevalence plot offers valuable insights into the occurrence of suppressed intermediate states, particularly in models featuring stiff helix segments. This research significantly enhances our understanding of the complex interactions governing helix bundling phenomena within the context of helical homopolymers.

**Keywords:** semiflexible polymers; conformational phases; phase transitions; helical polymers; Monte Carlo simulations; microcanonical analysis



**Citation:** Williams, M.J. Microcanonical Analysis of Helical Homopolymers: Exploring the Density of States and Structural Characteristics. *Polymers* **2023**, *15*, 3870. <https://doi.org/10.3390/polym15193870>

Academic Editor: Arash Nikoubashman

Received: 27 August 2023

Revised: 17 September 2023

Accepted: 21 September 2023

Published: 24 September 2023



**Copyright:** © 2023 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

This study explores the use of microcanonical inflection point analysis [1,2] to understand structural transitions in the helical homopolymer model [3]. This model is a coarse-grained representation of polymer chains that exhibit helical secondary structures and helix-bundled tertiary structures. Helical structures are common in various biopolymer systems [4–7]. Finite-size effects are an essential aspect of the behavior of these systems. Although the mechanisms responsible for helix formation vary [8], we aim to study the general thermodynamic behavior in a context-independent manner.

The model used in this study consists of a set of effective potentials that coordinate the interaction of the repeating polymer units. These potentials include representations of bonded interactions, non-bonded interactions, bending angle sensitivity, and torsion angle sensitivity. Polymer structures are generated using Markov chain Monte Carlo simulation with the Metropolis–Hastings algorithm [9,10]. The efficiency of this simulation approach can be improved by including multiple parallel replicas that are permitted to exchange replicas at regular intervals [11–13]. In this study, the array of simulation threads is chosen to span a two-dimensional array of conditions defined by temperature and a carefully chosen model parameter that influences structure formation [14,15]. Hamiltonian exchanges are performed between threads with different model parameters [16–18]. Throughout the simulation, each thread produces a canonical ensemble of structures and tracks canonical averages for parameters of interest.

Although the canonical data have previously been analyzed to provide a basic understanding of structural transitions and stability [19], this paper extends that work by using histograms of canonical ensemble quantities to generate microcanonical results pertaining to each possible structure's energy. Microcanonical analysis is used to gain deeper insight into the nature of the observed structural transitions. With the microcanonical ensemble,

inflection point analysis is employed to identify and classify structural transitions [20]. Furthermore, the nature of these transitions can be better understood by considering a two-dimensional structural prevalence plot, which shows the relative prevalence of various structure types within each microcanonical ensemble [21–24].

## 2. Materials and Methods

### 2.1. Model

This study makes use of a helical homopolymer model that generates polymer structures exhibiting both secondary and tertiary structures. This model demonstrates phase transitions between random-coil, globular, and various helical structure phases. The model incorporates four potentials: the finitely extensible nonlinear elastic (FENE) potential [25], the Lennard-Jones (LJ) potential [26], the bending potential, and the torsion potential [27].

Bonded monomers interact according to the FENE potential, which is described by Equation (1). The strength of this potential depends on the distance between monomers,  $r$ , where a minimum potential is achieved for  $r = r_0 \equiv 1$ . The bond length is not allowed to deviate from this equilibrium value by more than  $R \equiv 3/7$ . Any move that separates two bonded monomers by more than  $r + R$  or brings them closer than  $r - R$  is immediately rejected.

Non-bonded monomers interact according to the LJ potential if they are separated by a distance of less than  $r_c \equiv 2.5\sigma$ . The LJ potential is described by Equation (2), where an energetic minimum is found for  $r = r_0$ . This is achieved with  $\sigma \equiv 2^{-1/6}r_0$ . To avoid a discontinuity at  $r = r_c$ , we add an energy shift of  $v_c = 4[(\sigma/r_c)^{12} - (\sigma/r_c)^6]$ .

The polymer chain is additionally subject to torsion and bending potentials, which precipitate the formation of helical secondary structures. These potentials are described by Equations (3) and (4), respectively. In the case of the torsion potential, each series of three bonds has a dihedral angle  $\tau$ . An energy penalty is assigned based on each dihedral angle's variance from  $\tau_0 \equiv 0.873$ . Similarly, a series of two bonds can be used to calculate their bending angle,  $\theta$ . The bending energy is calculated from a bending angle's deviation from  $\theta_0 \equiv 1.742$ .

$$v_{\text{FENE}}(r) = \log\{1 - [(r - r_0)/R]^2\}. \quad (1)$$

$$v_{\text{LJ}}(r) = 4[(\sigma/r)^{12} - (\sigma/r)^6] - v_c. \quad (2)$$

$$v_{\text{bend}}(\theta) = 1 - \cos(\theta - \theta_0). \quad (3)$$

$$v_{\text{tor}}(\tau) = 1 - \cos(\tau - \tau_0). \quad (4)$$

The total energy of a polymer with configuration  $\mathbf{X}$  can be calculated from a summation over all of the potentials as shown in Equation (5). When included in the Hamiltonian, each potential is scaled by an energy scale. These energy scales are denoted as  $S_{\text{FENE}}$ ,  $S_{\text{LJ}}$ ,  $S_{\tau}$ , and  $S_{\theta}$ .

$$H(\mathbf{X}) = S_{\text{FENE}} \sum_i v_{\text{FENE}}(r_{i,i+1}) + S_{\text{LJ}} \sum_{i>j} v_{\text{LJ}}(r_{ij}) \\ + S_{\tau} \sum_l v_{\text{tor}}(\tau_l) + S_{\theta} \sum_k v_{\text{bend}}(\theta_k). \quad (5)$$

The energy scale used for each potential is chosen from commonly used values in the literature [21,28]. These values are:  $S_{\text{FENE}} \equiv -(98/5)r_0^2R^2/2 = -1.8$ ,  $S_{\text{LJ}} \equiv 1$ , and  $S_{\theta} = 200$ . The value of  $S_{\tau}$  varies between simulation threads and influences the structure types produced by a particular simulation.

### 2.2. Two-Dimensional Parallel Tempering

The polymer model is simulated using a parallel tempering Monte Carlo approach that employs a two-dimensional array of threads with varied temperatures and Hamiltonians. This approach is used to simulate polymers comprising 30 and 40 monomers. The 30-monomer simulation consists of 160 threads, with 16 temperature values ranging from

0.2 to 1.6. These values are chosen such that the lowest temperature will settle into a distinct low-energy structure type and the highest temperature threads will produce random coil structures that easily free a replica from any local free energy minimum. Intermediate temperatures are spaced exponentially so that successive temperatures differ by the same factor. There are 10 threads with each temperature value; these threads each simulate a unique Hamiltonian with  $S_\tau$  values between 5 and 14. The presumptive ground state structure varies as a function of  $S_\tau$ , so a single simulation is able to span several different system types. The range of values for  $S_\tau$  is chosen so that the simulation would include single-helix systems, two-helix bundles, and intermediate systems between them. A second set of simulations is run for 40-monomer chains with a similar computational setup. In the 40-monomer simulations,  $S_\tau$  varies between 5 and 25. Exchanges between threads  $i$  and  $j$  are accepted with a probability given in Equation (6). In this equation,  $\beta_i$  is the inverse temperature in thread  $i$ ,  $X_i$  is the polymer configuration coming from thread  $i$ , and  $H_i$  is the Hamiltonian for thread  $i$ .

$$P_{\text{exch}} = \min\left(1, \frac{e^{\beta_i H_i(X_i)} e^{\beta_j H_j(X_j)}}{e^{\beta_i H_i(X_j)} e^{\beta_j H_j(X_i)}}\right). \quad (6)$$

Each simulation thread uses a standard Metropolis algorithm with displacement updates between exchanges. Updates displace a single random monomer by a random amount. Updates are accepted with a probability  $P_{\text{acc}} = \min(1, e^{-\beta\Delta E})$ , where  $\beta = 1/(k_B T)$ . We use units in which  $k_b \equiv 1$ . To improve simulation efficiency, three different global updates are included. For a global displacement, a monomer is chosen at random and every monomer beyond the chosen one is displaced by the same amount. To perform a bending update, a single monomer is randomly selected and all subsequent monomers rotate around an axis perpendicular to the chosen monomers neighboring bonds. Torsion update chooses a single monomer and rotates all monomers beyond the chosen monomer around an axis formed by the previous bond. The use of global updates and the 2D parallel tempering approach leads to efficient sampling of the entire range of energies for systems with a variety of tertiary structure formations.

As data are collected, each simulation thread tracks canonical ensemble average values for several measurable order parameters, most notably, energy and an order parameter,  $q \equiv \sum_{|i-j|>6} (v_{\text{LJ}}(r_{ij})) / \sum_{|i-j|\leq 6} (v_{\text{LJ}}(r_{ij}))$ . This order parameter measures the ratio of the total Lennard-Jones interaction between monomers separated by more than six bonds to the total Lennard-Jones interaction between monomers separated by six or fewer bonds. This parameter increases when the chain is folded to include interactions between monomers separated by many bonds. It is useful for distinguishing different helix bundling phases.

### 2.3. Multiple Histogram Reweighting

This paper focuses on a microcanonical analysis of helical polymer systems. We characterize phase transitions by analyzing the microcanonical entropy and its derivatives as functions of energy. The microcanonical entropy, defined as  $S(E) = k_B \log(g(E))$ , is calculated by reweighting the histograms produced by individual simulation threads. A simulation thread,  $i$ , with temperature,  $T$ , produces a histogram,  $h_i(E)$ . This thread's histogram estimates the density of states using  $\bar{g}_i(E) = h_i(E)e^{\beta_i E}$ . The accuracy of this estimate at a given value of  $E$  naturally depends on the number of counts in that energy bin.

If histograms from threads at different temperatures overlap sufficiently, they can be weighted to produce a density of states that spans all energies generated by all threads,  $g(E)$ . To achieve this, we start with an initial guess for the partition function of  $Z = 1$  for all temperatures. Next, an estimate of  $\hat{g}(E)$  is computed using Equation (7), where  $M_i$  represents the total number of measurements in thread  $i$ . The partition function estimate can be improved using Equation (8). These equations are applied iteratively until results converge.

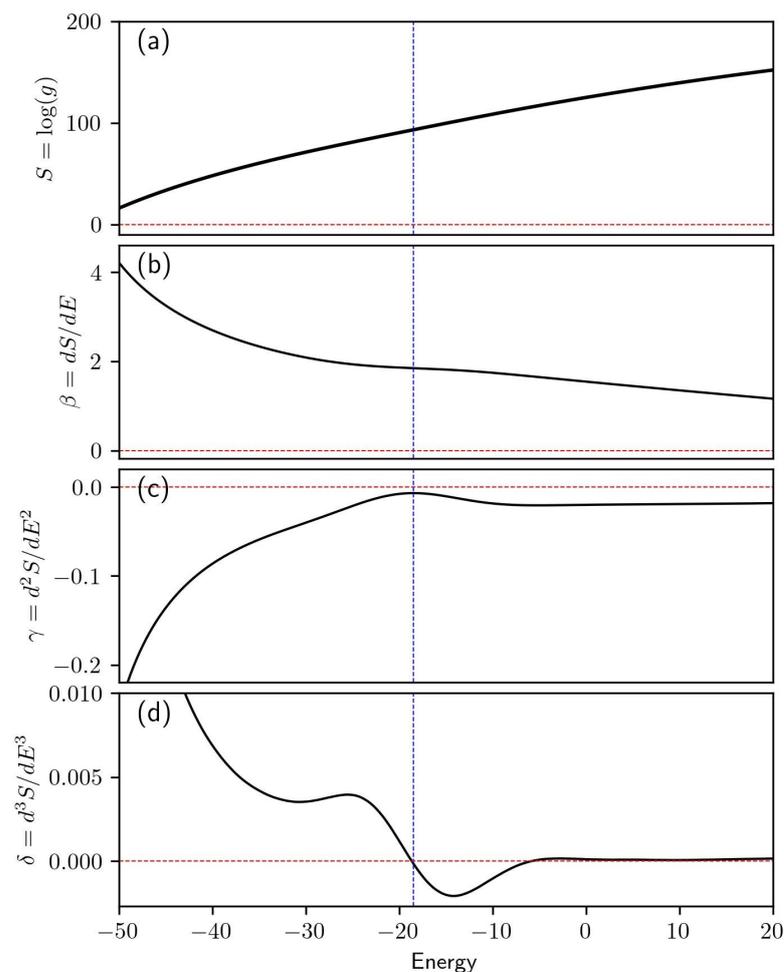
$$\hat{g}(E) = \frac{\sum_i h_i(E)}{\sum_i M_i Z_i^{-1} e^{-\beta_i E}}. \quad (7)$$

$$Z_i = \sum_E \hat{g}(E) e^{-\beta_i E}. \quad (8)$$

In practice, the values of  $g(E)$  are too large to work with directly, so the logarithms of  $g(E)$  and  $Z_i$  are used instead.

From the density of states, the microcanonical entropy can be recalculated. We then employ the Savitzky–Golay technique [29] to remove noise from  $S$  and calculate its first three derivatives while preserving the underlying trends. The derivatives of  $S$ , denoted as  $\beta(E) \equiv dS/dE$ ,  $\gamma(E) \equiv d^2S/dE^2$ , and  $\delta(E) \equiv d^3S/dE^3$  are used to analyze phase transitions.

Microcanonical inflection point analysis can be used to identify phase transitions and determine their order. In the absence of phase transitions,  $S(E)$  and its derivatives will be monotonic functions that alternate between increasing and decreasing, with all derivatives asymptotically approaching zero. An  $n$ -th order phase transition in a finite system will be exhibited as a region of minimal sensitivity within the  $(n - 1)$ -th derivative of the microcanonical entropy. This behavior is illustrated in Figure 1, where sample data are used to illustrate the inflection point or region of least sensitivity seen in the first derivative of the  $S(E)$ . This inflection point corresponds to a second-order phase transition and is visible in panel (b). The second-order phase transition can also be identified by analyzing the higher-order derivatives. In panel (c), the transition manifests as a peak approaching zero and it becomes a zero-crossing in panel (d).



**Figure 1.** Illustration of example data for microcanonical inflection point analysis. The vertical blue line marks the transition energy for a second-order phase transition. The entropy  $S$  is provided in panel (a). Panels (b–d) give the first, second, and third derivatives of the entropy, respectively.

We can further use the density of states to calculate canonical results with higher resolution than initially sampled. The histogram of energies for a canonical ensemble with inverse temperature,  $\beta$ , can be calculated from the density of states using  $h(E) = g(E)e^{-\beta E}$ . From this histogram, the average energy can be computed  $\bar{E} = \sum h(E)E / \sum h(E)$ . By performing this calculation for an array of temperatures, we can determine the specific heat from the temperature derivative of the ensemble average energy,  $C_V = d\bar{E}/dT$ .

#### 2.4. Two-Dimensional Density of States

In addition to recording the energy histogram for each thread, a two-dimensional histogram along  $E$  and a carefully selected order parameter are also recorded for each temperature thread. In this study, the order parameter  $q$  is chosen. This parameter measures the Lennard-Jones interaction between monomers separated by more than six monomers relative to the interaction with monomers less than or equal to six bonds. For the helical homopolymer system, this order parameter serves as a measure of the helix segment bundling. A two-dimensional analog of the microcanonical entropy can be calculated from these histograms using Equation (9).

$$g(E, q) = \frac{\sum_i h_i(E, q)}{\sum_i M_i Z_i^{-1} e^{-\beta_i E}} \quad (9)$$

Cross-sections of the density of states can be extracted from  $g(E, q)$  and scaled by  $g(E)$  to determine the relative prevalence of various structure types as a function of  $E$ . The structural prevalence for a particular value of  $q$  is calculated from  $p(E, q) = g(E, q)/g(E)$ .

### 3. Results

In this section, key findings from this study are presented. Phase transitions are identified and classified by plotting the microcanonical entropy and its derivatives. For several example systems, the transition signals observed in the microcanonical analysis are compared with those in the canonical analysis. These phase transitions can be further understood by examining multiple cross-sections of a two-dimensional representation of the density of states across energy and  $q$ . We begin by presenting data for an array of simulations in which all polymers have a length of  $N = 30$  and will discuss polymers with  $N = 40$  in Section 3.2.

#### 3.1. Microcanonical Results

In Figure 2, data are provided for three distinct system types all with length,  $N = 30$ . The first column, Figure 2a,d,g,j,m, presents data pertaining to a simulation in which  $S_\tau = 5$ . Here, random coil structures are observed at high energy and two-helix bundles appear at low-energy. The single-helix phase is not dominant for any energy range. The rightmost column, Figure 2c,f,i,l,o, pertains to simulations with  $S_\tau = 13$ . This system generates distinctly single-helix low-energy structures. The center column, Figure 2b,e,h,k,n, represents systems with  $S_\tau = 8$ . This system is intermediate to the other two, producing both single-helix and two-helix structures at low energies. Ultimately, the lowest-energy structures for this system are two-helix bundles. Further discussion on the specific makeup of the structural phases for these systems will be provided in Section 3.1.1.

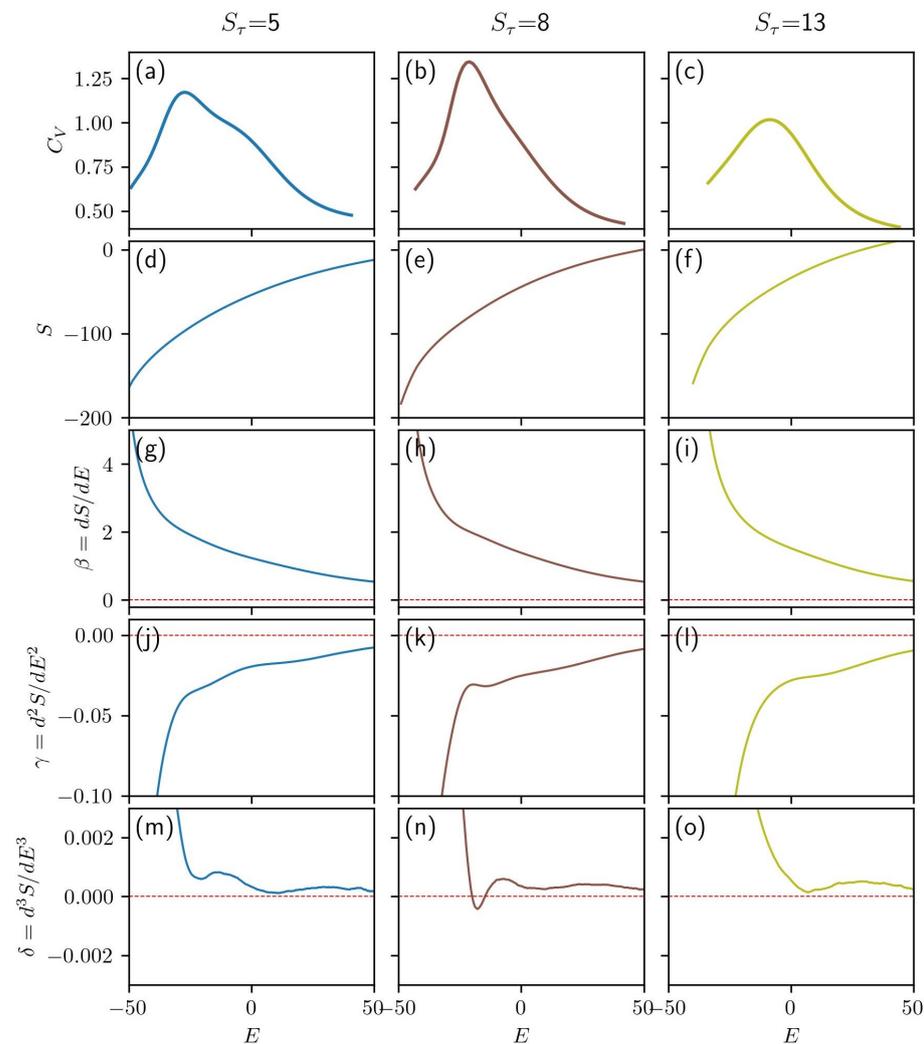
In the top row of Figure 2, panels (a)–(c), the canonical ensemble specific heat is given as a function of that ensemble's average energy,  $\bar{E}$ . Phase transitions typically correspond to peaks in the specific heat. A single strong peak is evident in each system shown here. Each peak represents the transition to the lowest-energy structure type. In panel (a), an additional phase transition is visible as a shoulder at a higher energy than the prominent peak. Panel (b) also exhibits a nearly hidden shoulder at a higher energy than the primary peak. These specific heat plots serve as a reference against which inflection points in the microcanonical quantities can be compared.

The microcanonical entropy,  $S$ , is given for three different values of  $S_\tau$  in Figure 2d–f. The absence of inflection points in the microcanonical entropy suggests that there are no

first-order phase transitions in these systems. The derivatives of the microcanonical entropy are presented in the bottom three rows, panels (g)–(o).

Figure 2g–i gives  $\beta$ . A region of decreased sensitivity is evident in panel (h). This region, apparent in the second derivative of  $S$ , corresponds to a second-order phase transition. This transition continues to be seen in the higher-order derivatives as peaks approaching zero. Additionally, this transition is observable in panel (b) as a peak in the specific heat.

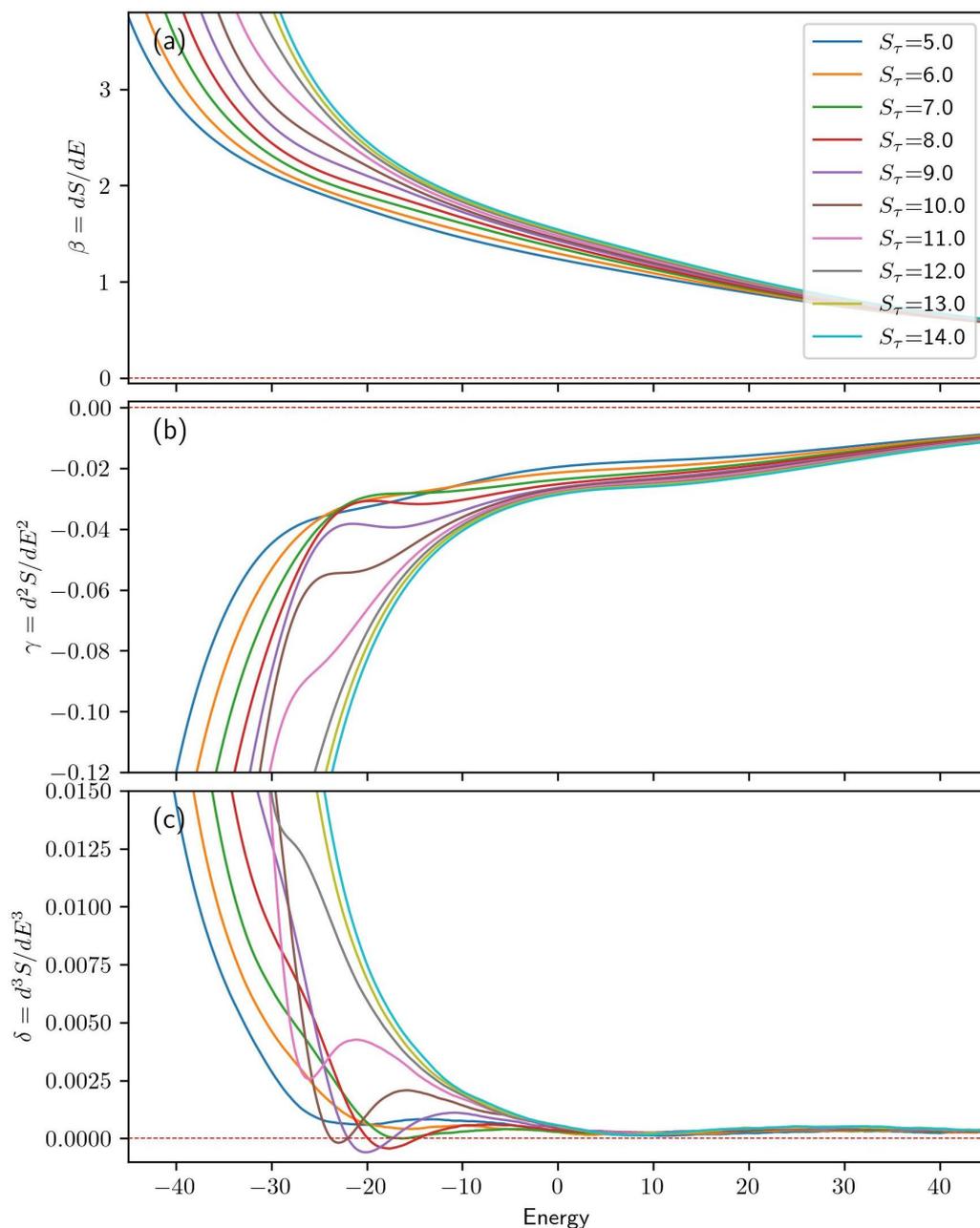
All three plots of  $\gamma$  in Figure 2j–l feature inflection points that correspond to third-order phase transitions. In panel (j), two distinct inflection points align well with the peak and shoulder observed in the specific heat plot in panel (a). In panel (k), we see the peak associated with the second-order transition as well as an inflection point at higher energy. Again, the transitions in the microcanonical entropy plots align closely with the peak and shoulder in the specific heat. In panel (l), a single inflection point is present signifying a lone third-order transition. This transition again aligns with a transition present in the specific heat. Each of the third-order transitions, represented as regions of decreased sensitivity in the  $\gamma$  plots, can be clearly identified as inverted peaks in the  $\delta$  plots from Figure 2m,n.



**Figure 2.** (a–c) For three different values of  $S_\tau$ , the canonical ensemble specific heat is given as a function of each ensemble’s average energy. The microcanonical entropy is given in (d–f). The first (g–i), second (j–l), and third (m–o) derivatives of microcanonical entropy are given as a function of energy. All panels in this figure represent polymers of length  $N = 30$ .

Not only are the transitions more easily identified using microcanonical analysis when compared to the canonical representation in the top row, but the transition order can be inferred from the order of derivative that first shows a region of minimal sensitivity.

A broader range of systems is explored in Figure 3. Panel (a) presents  $\beta$  for systems with  $S_\tau$  values ranging from 5 to 14. Although subtle, inflection points corresponding to second-order transitions are discernible for  $S_\tau$  of 8, 9, and 10. These transition energies are more easily identified by considering Figure 3b, which shows the second derivative of the microcanonical entropy,  $\gamma$ . In this figure, the three second-order transitions occur at transition energies of  $-19.6$ ,  $-21.6$ , and  $-23.1$  for the  $S_\tau = 8, 9$ , and  $10$  curves, respectively.



**Figure 3.** (a) For each value of  $S_\tau$  in the  $N = 30$  simulation, the first derivative of microcanonical entropy is given as a function of energy. Panels (b,c) present the second and third derivatives of the microcanonical entropy, respectively.

All systems display inflection points in  $\gamma$  corresponding to third-order transitions with approximate transition energies of  $7.0$ . Additionally, lower-energy third-order transitions

are observed for  $S_\tau$  values of 5, 6, 7, and 11, with transition energies of  $-20.8$ ,  $-15.5$ ,  $-16.5$ , and  $-26.2$ , respectively. These third-order transition energies are most easily identified by examining the inverted peaks in  $\delta$ , as shown in panel (c). For  $S_\tau = 12$ , a region of minimal sensitivity at  $E = -27.9$  represents a fourth-order transition.

### 3.1.1. Two-Dimensional Density of States

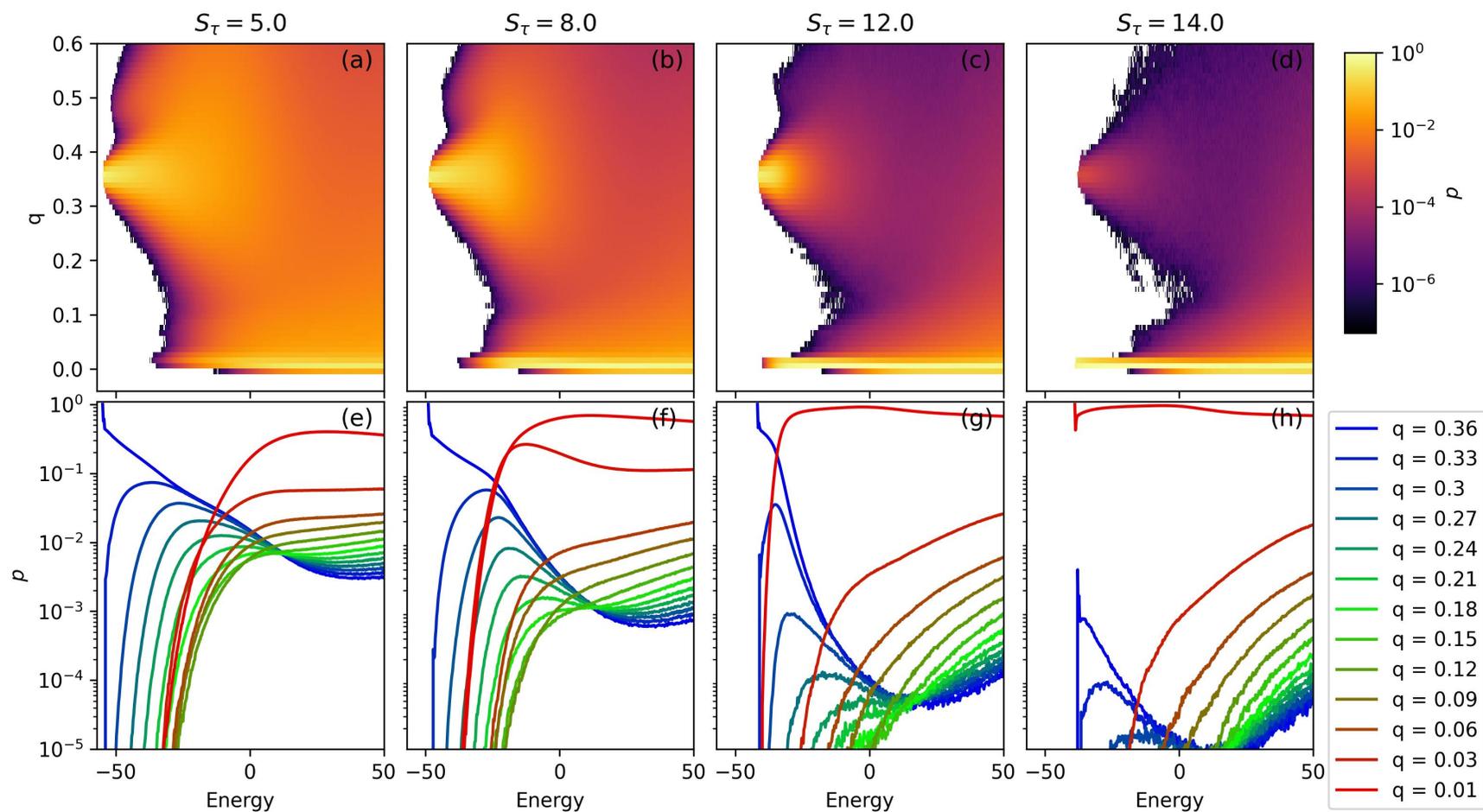
Structural prevalence is a useful way to visualize the full behavior of a system. In Figure 4, we analyze structural prevalence as a function of both  $E$  and order parameter  $q$ . Single-helix bundles have little to no LJ interaction between monomers separated by more than six bonds and therefore have a  $q$  value of approximately 0.01. In contrast, for two-helix bundles, each helix segment has monomers that interact with monomers from the other segment, more than six bonds away. Two-helix bundles tend to have  $q$  of approximately 0.35.

In Figure 4a–d, each panel gives the relative prevalence of structures across a space of  $E$  and  $q$ . A vertical slice through a single value of energy can be thought of as a normalized histogram of the structures found at each  $q$ . Bright yellow regions represent values of  $q$  which are highly represented in a single microcanonical ensemble with energy  $E$ . Darker regions represent more sparsely populated regions of the  $q$  vs.  $E$  space. No structures were observed within the bins that are colored white. Because the color scale is logarithmic, small changes in shade can represent rather large differences in structural prevalence. For this reason, panels (a)–(d) should be seen as more qualitative results. In order to read quantitative information from this chart, Figure 4e–h give horizontal cross-sections from each two-dimensional prevalence plot.

In all four  $S_\tau$  examples given, the lowest-energy structures are made up entirely of structures within a single  $q$  bin. Figure 4a–c represents systems with  $S_\tau$  values between 5 and 12. In all of these cases, the lowest energy structures have a value of  $q$  of 0.36 corresponding to two-helix bundles. As  $S_\tau$  increases, the energetic advantage for folding decreases, and the energy of the ground state two-helix bundle approaches that of the lowest energy single-helix structures. At  $S_\tau = 14$ , single-helix structures become energetically preferable to two-helix bundles. For this reason, the lowest energy populated bin in panel (d) occurs at  $q \approx 0$ .

As energy increases, structural variability increases as well. For the two-helix bundles, this effect is very apparent; as  $E$  increases the variability in  $q$  also increases. We see that when structures are constrained with a stronger torsion potential ( $S_\tau$ ), this structure variability increase is less dramatic than in the more weakly constrained cases. For  $S_\tau = 5$ , there is a secondary peak at relatively low energy and  $q \approx 0.5$ . This secondary peak corresponds to two-helix bundles in which the two ends of the polymer begin to wrap around each other instead of remaining constrained to their own helix segment. Within the range of  $S_\tau$  values explored within this study, these structures are never the dominant structure type. The increase in structural variability with increasing energy is not captured as well in the single-helix case ( $S_\tau = 14$ ), because variability does not as quickly affect the value of  $q$ .

Figure 4e–h show horizontal cross-sections from the above graphs. Each curve gives the prevalence of structures of a particular value of  $q$  as a function of energies. In each case, the highest value of  $q$  given corresponds to two-helix bundles and the lowest corresponds to single-helix structures. In panels (a)–(c) we see the  $q = 0.36$  curve trend toward 1 at low energy, as all structures are represented within this single  $q$  bin. As energy increases, varied two-helix bundle states begin to increase in prevalence and at some point, single-helix structures also begin to increase in prevalence. The energy at which the single-helix structures increase in prevalence corresponds well to the low-energy phase transitions seen in Figure 3. Single-helix structures dominate at low energies in (d), but as energy increases we do see the presence of some two-helix bundles.

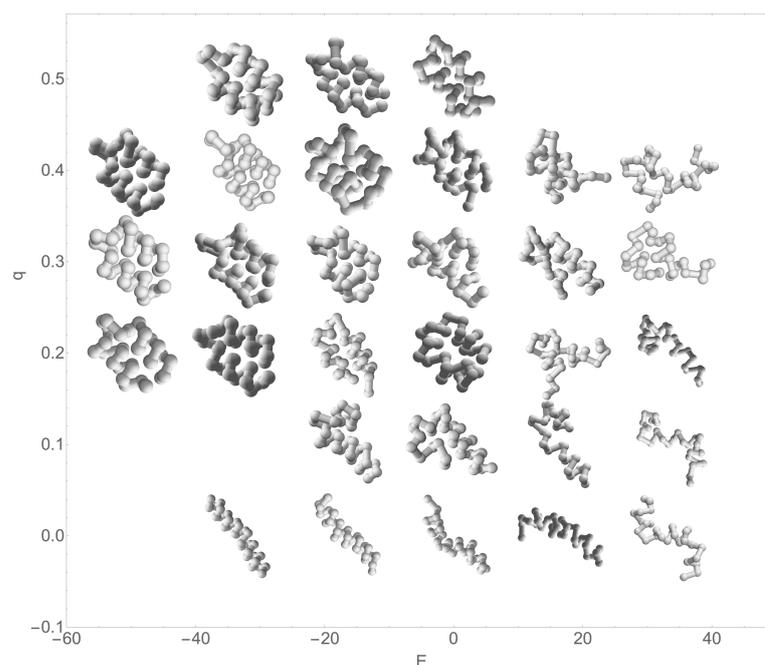


**Figure 4.** (a–d) A two-dimensional structural prevalence plot is given for several interesting values of  $S_\tau$  in the  $N = 30$  simulation. In these figures, the structural prevalence is given as a function of structure energy and order parameter,  $q$ . Low-energy two-helix bundles have  $q \approx 0.35$  and single-helix structures have  $q \approx 0.01$ . Below, in (e–h), cross-sections of the two-dimensional structural prevalence plot are given. Values of  $q$  are chosen such that  $p(E)$  is shown for two-helix bundles, single-helix structures, and an array of intermediate structures.

Of particular interest is the suppression of intermediate states between single-helix structures and two-helix bundles. These intermediate states are depicted by green curves in panels (e)–(h). Figure 4e makes it clear that for the third-order phase transition previously noted in the  $S_{tau} = 5$  system, intermediate states between the single-helix phase and the two-helix phase have significant populations. At larger values of  $S_{tau}$ , the intermediate states are increasingly suppressed within the energetic window surrounding the transition from single-helix to two-helix bundle structures.

The  $S_{tau} = 12$  system has a fourth-order phase transition from the single-helix phase to the two-helix bundle at  $E = -25$ . In Figure 4g, the intermediate structures are completely unrepresented as the representation of the single-helix phases decreases and that of the two-helix phase increases. At  $S_{tau} = 14$ , the presumptive ground state structure is a single-helix segment. The microcanonical analysis shows that there is no low-energy phase transition present in this system. However, we can observe interesting behavior with the presence of some two-helix bundles at energies slightly above those of lowest energy single-helix structures. Although they are never the dominant structure type, their presence should not be ignored. In this system, we also observe significant suppression of the intermediate states.

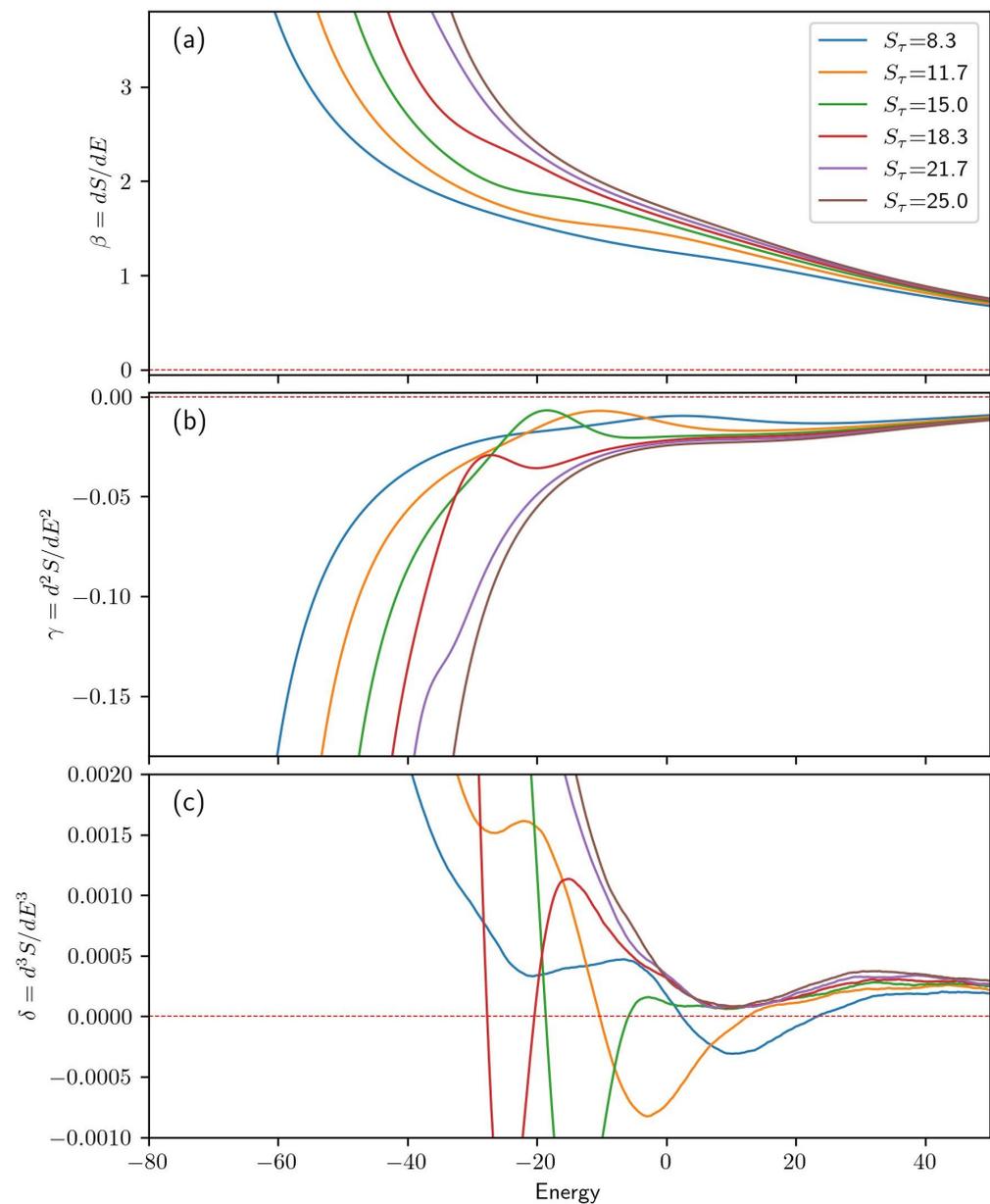
Figure 5 provides some example structures from the simulation with an  $S_{tau}$  of 8. Although the structures and energies differ in simulations with other values of  $S_{tau}$ , they are qualitatively arranged in the same manner. Organized single-helix structures occur at low  $E$  and low  $q$ , whereas organized two-helix bundles occur at low  $E$  and  $q \approx 0.35$ . As energy increases, some structural variability is evident in the structures presented above. This variability manifests as inconsistent spacing of monomers in successive helix turns and variable spacing between helical segments. At higher energies, variability in the joint between helix segments is also found. As  $E \approx -18$ , one helix segment of a two-helix bundle can be longer than the other, and at  $E \approx 20$ , random coil structures become dominant and structures no longer have distinct helix segments. The  $S_{tau} = 8$  system was chosen because there is a significant representation of the intermediate structures falling in between the single-helix phase and the two-helix bundle phase. As  $S_{tau}$  increases, this intermediate region becomes so sparse that a simple time series of structures from all canonical ensembles fails to capture a representative sample.



**Figure 5.** Example structures from the  $S_{tau} = 8$  simulation. Structures are plotted in the approximate location of their  $E$  and  $q$ .

### 3.2. Alternate System Size

In addition to 30-monomer systems, the same approach has been applied to a system with 40 monomers. Figure 6 presents the one-dimensional microcanonical quantities for the 40-monomer helical homopolymer model. The range of  $S_\tau$  values is adjusted to stabilize the same tertiary structures observed in the 30-monomer results. At  $S_\tau = 8.3$ , two-helix bundles are the sole low-energy structure type produced and at  $S_\tau = 25$ , single-helix structures dominate at low energies. Intermediate values of  $S_\tau$  are also presented.



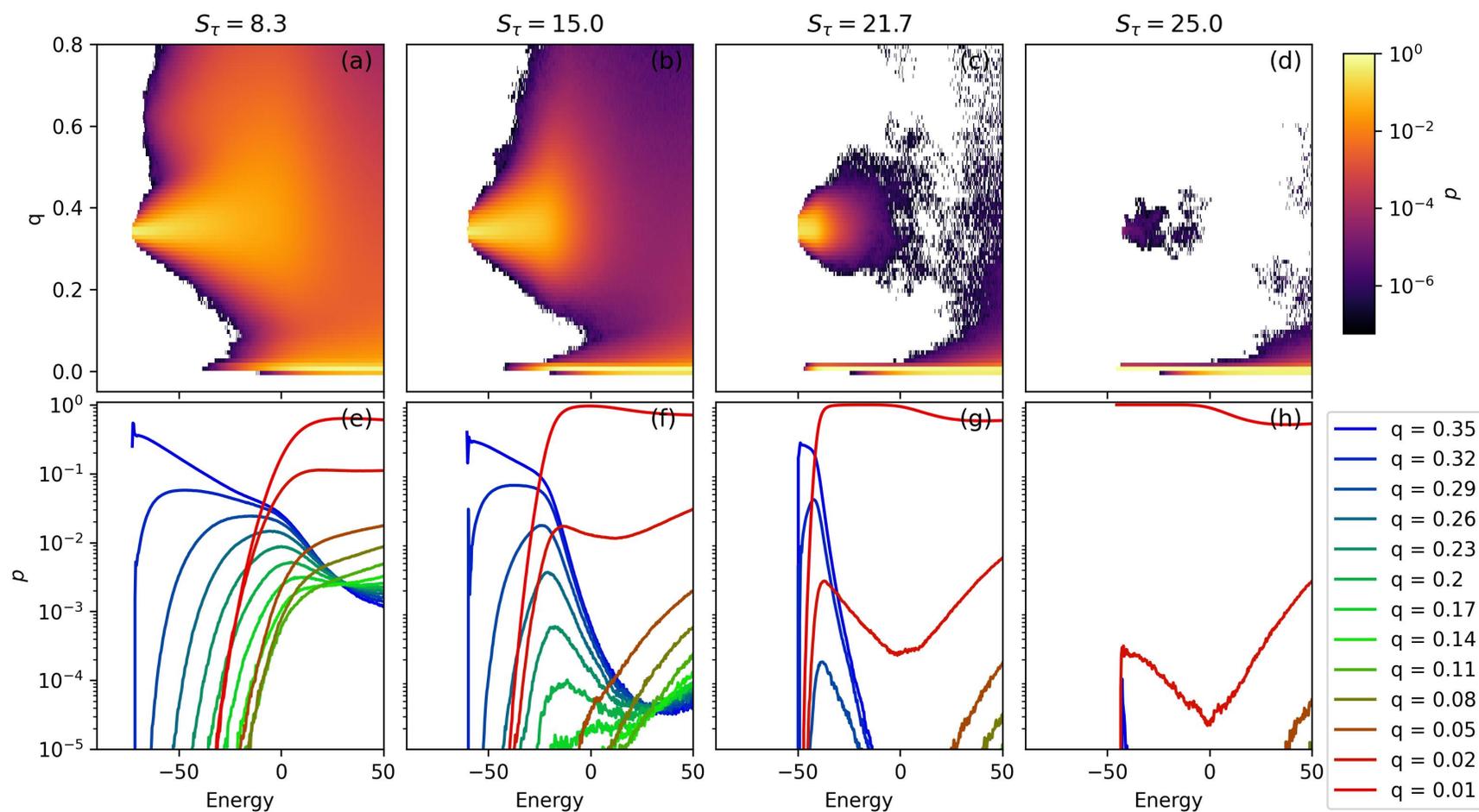
**Figure 6.** (a) For each value of  $S_\tau$  in the  $N = 40$  simulation, the first derivative of microcanonical entropy is given as a function of energy. (b,c) give the second and third derivatives of the microcanonical entropy, respectively.

We can see in Figure 6a that systems with  $S_\tau$  of 8.3, 11.7, 15, and 18.3 all exhibit second-order phase transitions at energies of 3.0,  $-10.4$ ,  $-18.8$ , and  $-27.5$ , respectively. These correspond to the transition between low-energy two-helix bundles and a single-helix phase, which dominates at an intermediate energy range. For  $S_\tau = 21.7$ , the transition between two-helix bundles and single helices is a third-order transition occurring at an energy of  $-34.9$ .

Systems with  $S_\tau$  values of 18.3, 21.7, and 25.0 clearly exhibit a third-order transition at  $E \approx 10$ . This corresponds to a transition between random-coil structures and more organized single-helix structures. Although this transition is present in all of the systems studied here, the signal is obscured by the peak corresponding to lower-order transitions seen in the  $S_\tau = 8.3, 11.7$ , and  $15.0$  cases. Evidence of this transition is most easily observed as an inverted peak in the  $\delta$  plot from panel (c).

Figure 7 presents a two-dimensional prevalence representation for the  $N = 40$  systems. Qualitatively, the behavior of the 40-monomer system closely resembles that of the  $N = 30$ . At  $S_\tau = 8.3$ , there is a strong low-energy representation of two-helix bundle structures at  $q \approx 0.35$ . With a slight increase in energy, some structures emerge with a  $q$  value of approximately 0.6; this peak corresponds to three-helix bundles. These structures are poorly stabilized and form with several different orientations. Although there are single-helix structures formed for  $S_\tau = 8.3$ , they are only represented in the relatively higher energy range of  $E > -30$ .

For systems with  $S_\tau$  greater than 8.3, several changes become apparent. At  $S_\tau = 15$  the presence of three-helix bundles completely disappears. The lowest energy for the single-helix phase decreases and eventually approaches that of the two-helix bundles in the  $S_\tau = 21.7$  case. As the increased  $S_\tau$  brings the energies of the two-helix and single-helix phases closer together, intermediate structures become increasingly entropically suppressed. This effect is evident at  $S_\tau = 15$  when considering the green curves in panel (f). The  $S_\tau = 21.7$  system demonstrates this effect even more strongly, with intermediate phases nearly entirely unrepresented in our simulation. The low-energy dominance of the single-helix structures is apparent in the  $S_\tau = 25$  system. In panels (d) and (h), very few two-helix bundle structures are observed.



**Figure 7.** (a–d) A two-dimensional structural prevalence plot is given for several interesting values of  $S_\tau au$  in the  $N = 40$  simulation. In these figures, the structural prevalence is given as a function of structure energy and order parameter,  $q$  and  $E$ . Low-energy two-helix bundles have  $q \approx 0.35$  and single-helix structures have  $q \approx 0.01$ . Below, in (e–h), cross-sections of the two-dimensional structural prevalence plot are given. Values of  $q$  are chosen such that  $p(E)$  is shown for two-helix bundles, single-helix structures, and an array of intermediate structures.

#### 4. Discussion

Microcanonical inflection point analysis is gaining wider acceptance in the realm of computational statistical physics, particularly for its ability to provide a nuanced understanding of phase transitions [30–33]. In our study, this method has proven useful for identifying, classifying, and understanding structural transitions in helical homopolymers. The flexibility of the approach means that it can be applied wherever the density of states is known, suggesting its utility across a broad range of disciplines including condensed matter physics and biological systems. Additionally, we find that incorporating an order parameter to differentiate phases allows us to produce a two-dimensional prevalence plot, which offers additional insights into the physical behavior of the system under study.

In this paper, we expand on earlier research on the helical homopolymer model, where phase transitions were identified and a hyperphase diagram was constructed through analysis of canonical quantities [28]. We employ microcanonical inflection point analysis to gain a deeper understanding of phase transitions in polymers with lengths of  $N = 30$  and  $N = 40$ . Specifically, we are able to more precisely measure phase transition energies and determine the order of each transition. The energy scale for torsion potential,  $S_\tau$ , is tuned to achieve varying helical segment stiffness [3]. Across all systems investigated in this paper, we find a third-order phase transition between single-helix structures and random coil structures. In cases where there is an energetic preference for two-helix bundles, an additional transition from single-helix to two-helix bundles occurs at lower energy. The order of this transition is influenced by the stiffness of the helix: stiffer systems show second-order transitions, whereas more weakly confined systems exhibit third-order transitions. Despite the variations in transition energy and helix stiffness, the order of phase transitions remains consistent across both system sizes examined in this study.

Utilizing two-dimensional prevalence plots, we observe varying degrees of entropic suppression for intermediate states between the two-helix and single-helix phases. Generally, we find that systems with stiffer helical segments exhibit greater entropic suppression of these intermediate structures. This trend holds true for both the 30-monomer and the 40-monomer systems. The entropic suppression of these intermediate states stems from their significantly increased energy levels in models with stiff helical segments. From a practical standpoint, this entropic suppression complicates accurate sampling when using traditional one-dimensional replica exchange schemes. In our case, incorporating a two-dimensional replica exchange scheme greatly improved simulation efficiency.

#### 5. Conclusions

This study underscores the utility of microcanonical inflection point analysis for studying complex systems. We used this method to analyze phase transitions in the helical homopolymer model. We found consistent third-order transitions between single-helix and random coil structures across different system sizes. Additionally, systems with an energetic preference for two-helix bundles exhibited an extra transition, the order of which varied based on helix stiffness. The methodology proved to be a robust tool for understanding phase transitions, thereby broadening its applicability across statistical systems.

There are several intriguing avenues for future work related to this publication; although the utility of microcanonical inflection point analysis is evident, this technique could benefit from further formalization and refinement to promote its broader adoption across various disciplines. With respect to the helical homopolymer model, conducting a microcanonical analysis of longer chains would permit a more comprehensive investigation of transitions in systems featuring stable three-helix bundle.

**Funding:** This research received no external funding.

**Data Availability Statement:** The data that support the findings of this study are available from the corresponding author upon reasonable request.

**Conflicts of Interest:** The author declares no conflict of interest.

## References

1. Junghans, C.; Bachmann, M.; Janke, W. Microcanonical analyses of peptide aggregation processes. *Phys. Rev. Lett.* **2006**, *97*, 218103. [[CrossRef](#)]
2. Schnabel, S.; Seaton, D.T.; Landau, D.P.; Bachmann, M. Microcanonical entropy inflection points: Key to systematic understanding of transitions in finite systems. *Phys. Rev. E* **2011**, *84*, 011127. [[CrossRef](#)] [[PubMed](#)]
3. Williams, M.J.; Bachmann, M. Significance of bending restraints for the stability of helical polymer conformations. *Phys. Rev. E* **2016**, *93*, 062501. [[CrossRef](#)]
4. Sluysmans, D.; Willet, N.; Thevenot, J.; Lecommandoux, S.; Duwez, A. Single-molecule mechanical unfolding experiments reveal a critical length for the formation of alpha-helices in peptides. *Nanoscale Horiz.* **2020**, *5*, 671–678. [[CrossRef](#)]
5. Barlow, D.; Thornton, J. Helix geometry in proteins. *J. Mol. Biol.* **1988**, *201*, 601–619. [[CrossRef](#)]
6. Hao, B.; Zhou, W.; Theg, S.M. The polar amino acid in the TatA transmembrane helix is not strictly necessary for protein function. *J. Biol. Chem.* **2023**, *229*, 102998. [[CrossRef](#)] [[PubMed](#)]
7. Song, Z.; Tan, Z.; Zheng, X.; Fu, Z.; Ponnusamy, E.; Cheng, J. Manipulating the helix-coil transition profile of synthetic polypeptides by leveraging side-chain molecular interactions. *Polym. Chem.* **2020**, *11*, 1445–1449. [[CrossRef](#)]
8. Phase diagram of flexible polymers with quenched disordered charged monomers. *Phys. A: Stat. Mech. Appl.* **2022**, *604*, 127787. [[CrossRef](#)]
9. Metropolis, N.; Rosenbluth, A.W.; Rosenbluth, M.N.; Teller, A.H.; Teller, E. Equation of State Calculations by Fast Computing Machines. *J. Chem. Phys.* **2004**, *21*, 1087–1092. [[CrossRef](#)]
10. Hastings, W.K. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **1970**, *57*, 97–109. [[CrossRef](#)]
11. Swendsen, R.H.; Wang, J.S. Replica Monte Carlo Simulation of Spin-Glasses. *Phys. Rev. Lett.* **1986**, *57*, 2607–2609. [[CrossRef](#)] [[PubMed](#)]
12. Katzgraber, H.; Trebst, S.; Huse, D.; Troyer, M. Feedback-optimized parallel tempering Monte Carlo. *J. Stat. Mech. Theory Exp.* **2006**, *2006*, P03018. [[CrossRef](#)]
13. Gront, D.; Kolinski, A. Efficient scheme for optimization of parallel tempering Monte Carlo method. *J. Phys. Condens.* **2007**, *19*, 036225. [[CrossRef](#)]
14. Aierken, D.; Bachmann, M. Impact of bending stiffness on ground-state conformations for semiflexible polymers. *J. Chem. Phys.* **2023**, *158*, 214905. [[CrossRef](#)] [[PubMed](#)]
15. Majumder, S.; Marenz, M.; Paul, S.; Janke, W. Knots are Generic Stable Phases in Semiflexible Polymers. *Macromolecules* **2021**, *54*, 5321–5334. [[CrossRef](#)]
16. Fukunishi, H.; Watanabe, O.; Takada, S. On the Hamiltonian replica exchange method for efficient sampling of biomolecular systems: Application to protein structure prediction. *J. Chem. Phys.* **2002**, *116*, 9058–9067. [[CrossRef](#)]
17. Liao, Q. Enhanced sampling and free energy calculations for protein simulations. In *Computational Approaches for Understanding Dynamical Systems: Protein Folding and Assembly*; Elsevier: Hoboken, NJ, USA, 2020; Volume 170, pp. 177–213.
18. Sugita, Y.; Kitao, A.; Okamoto, Y. Multidimensional replica-exchange method for free-energy calculations. *J. Chem. Phys.* **2000**, *113*, 6042–6051. [[CrossRef](#)]
19. Williams, M.J.; Bachmann, M. System-Size Dependence of Helix-Bundle Formation for Generic Semiflexible Polymers. *Polymers* **2016**, *8*, 245. [[CrossRef](#)] [[PubMed](#)]
20. Qi, K.; Bachmann, M. Classification of Phase Transitions by Microcanonical Inflection-Point Analysis. *Phys. Rev. Lett.* **2018**, *120*, 180601. [[CrossRef](#)] [[PubMed](#)]
21. Qi, K.; Liewehr, B.; Koci, T.; Pattanasiri, B.; Williams, M.J.; Bachmann, M. Influence of bonded interactions on structural phases of flexible polymers. *J. Chem. Phys.* **2019**, *150*, 054904. [[CrossRef](#)]
22. Aierken, D.; Bachmann, M. Stable intermediate phase of secondary structures for semiflexible polymers. *Phys. Rev. E* **2023**, *107*, L032501. [[CrossRef](#)]
23. Sitarachu, K.; Bachmann, M. Evidence for additional third-order transitions in the two-dimensional Ising model. *Phys. Rev. E* **2022**, *106*, 014134. [[CrossRef](#)]
24. Sitarachu, K.; Zia, R.K.P.; Bachmann, M. Exact microcanonical statistical analysis of transition behavior in Ising chains and strips. *J. Stat. Mech. Theory Exp.* **2020**, *2020*, 073204. [[CrossRef](#)]
25. Kremer, K.; Grest, G.S. Dynamics of entangled linear polymer melts: A molecular-dynamics simulation. *J. Chem. Phys.* **1990**, *92*, 5057–5086. [[CrossRef](#)]
26. Lennard-Jones, J.E. Cohesion. *Proc. Phys. Soc.* **1931**, *43*, 461. [[CrossRef](#)]
27. Rapaport, D.C. Molecular dynamics simulation of polymer helix formation using rigid-link methods. *Phys. Rev. E* **2002**, *66*, 011906. [[CrossRef](#)] [[PubMed](#)]
28. Williams, M.J.; Bachmann, M. Stabilization of Helical Macromolecular Phases by Confined Bending. *Phys. Rev. Lett.* **2015**, *115*, 048301. [[CrossRef](#)] [[PubMed](#)]
29. Savitzky, A.; Golay, M.J.E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Anal. Chem.* **1964**, *36*, 1627–1639. [[CrossRef](#)]
30. Chen, T.; Lin, X.; Liu, Y.; Liang, H. Microcanonical analysis of association of hydrophobic segments in a heteropolymer. *Phys. Rev. E* **2007**, *76*, 046110. [[CrossRef](#)] [[PubMed](#)]

31. Lauer, C.; Paul, W. Dimerization of Polyglutamine within the PRIME20 Model using Stochastic Approximation Monte Carlo. *Macromol. Theory Simul.* **2023**, *32*, 2200075. [[CrossRef](#)]
32. Aierken, D.; Bachmann, M. Comparison of Conformational Phase Behavior for Flexible and Semiflexible Polymers. *Polymers* **2020**, *12*, 3013. [[CrossRef](#)] [[PubMed](#)]
33. Trugilho, L.F.; Rizzi, L.G. Microcanonical Characterization of First-Order Phase Transitions in a Generalized Model for Aggregation. *J. Stat. Phys.* **2022**, *186*, 40. [[CrossRef](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.