On the Pseudo Phase Diagram of Single Semi-Flexible Polymer Chains: A Flat-Histogram Monte Carlo Study

Benno Werlich [1], Mark P. Taylor [2], Timur Shakirov [1] and Wolfgang Paul [1],*

[1]   Institut für Physik, Martin-Luther-Universität, 06099 Halle, Germany;
      Benno.Werlich@physik.uni-halle.de (B.W.); Timur.Shakirov@physik.uni-halle.de (T.S.)
[2]   Department of Physics, Hiram College, Hiram, OH 44234, USA; taylormp@hiram.edu
*   Correspondence: Wolfgang.Paul@physik.uni-halle.de; Tel.: +49-345-552-5440

Abstract: Local stiffness of polymer chains is instrumental in all structure formation processes of polymers, from crystallization of synthetic polymers to protein folding and DNA compactification. We present Stochastic Approximation Monte Carlo simulations—a type of flat-histogram Monte Carlo method—determining the density of states of a model class of single semi-flexible polymer chains, and, from this, their complete thermodynamic behavior. The chains possess a rich pseudo phase diagram as a function of stiffness and temperature, displaying non-trivial ground-state morphologies. This pseudo phase diagram also depends on chain length. Differences to existing pseudo phase diagrams of semi-flexible chains in the literature emphasize the fact that the mechanism of stiffness creation matters.

1. Introduction

Local stiffness or semi-flexibility of polymer chains is probably by far the most important design concept for the creation of ordered polymeric structures. This is obviously true for (semi-)crystalline synthetic polymers which order into lamellar structures [1], displaying not only positional order of the monomers but also orientational order of the chains. Similarly, nature makes use of stiffness of bio-polymers in processes such as the encapsulation of DNA [2] and the formation of secondary structure (α-helices, β-sheets) in proteins [3], as well as in the formation of pathogenic amyloid [4].

Theoretical and computational modeling of semi-flexible chains started very early on, e.g., by Flory's approach to model the isotropic-nematic transition in semi-flexible polymers using a simple lattice model [5]. The most important theoretical model is certainly the worm-like chain [6] with its prediction of an exponential decay of bond-orientation correlation along the chain on the scale of its persistence length. The property of this model, that the Kuhn-length of a chain is twice its persistence length, has become such common folklore that it is often taken as a general relation, although it is only valid within this model of a persistent type of semi-flexibility. The fact that not all stiffnesses are created equal has been stressed and clearly worked out in the theoretical analysis of the isotropic-nematic transition of melts of semi-flexible chains [7]. In addition, in recent years, it has been established by simulations that predictions of the worm-like chain model are actually never realized in nature due to the neglect of excluded volume interactions in the model. Orientational correlations within a chain asymptotically always decay algebraically [8–12] and not exponentially.

In a computer simulation of semi-flexible chains, the persistent mechanism of flexibility is most easily realized by a bond angle potential that favors zero degrees in its ground state. Such a modeling approach has been used to explain the occurrence of toroidal structures in the collapsed

state of stiff DNA chains [13,14]. The phase diagram of this type of model as a function of stiffness and non-bonded attraction was first studied by Grassberger [15] for the simple cubic self-avoiding walk, where it was established that the collapse transition of a polymer chain becomes a first order transition beyond a certain stiffness. The pseudo phase diagram at finite chain length was studied for the bond-fluctuation model (which has 87 different bond angles compared to 2 in the simple cubic lattice model) using extended ensemble Monte Carlo simulations [16–18] and was shown to exhibit regions of stability for lamellar structures, toroidal structures and tennis racket like structures (a combination of the two former ones). While these studies were performed for chains of lengths $N = 40 - 256$, a Wang–Landau Monte Carlo [19,20] study of such semi-flexible bead-spring chains in the continuum, performed for chains of length $N = 13 - 55$ [21] later confirmed the occurrence of these morphologies. Recently a similar model was studied by multi-canonical simulations [22,23], and it was shown that ground states at different stiffness can also be characterized by their knottedness (here, knots are defined by joining the two ends of the chain by an additional long bond).

A different mechanism of introducing stiffness was suggested for the modeling of proteins. In one variant, chains described as thick tubes were suggested [24,25], which, upon collapse, are able to form helical ground states. In a different variant [26,27], the chains were built out of overlapping hard spheres, which, upon collapse, exhibit helical ground states for short enough chains. This model is promising for a systematic study of the effect of such steric stiffness on the single chain phase diagram [28], and we will report in the following on such a study.

2. Materials and Methods

In our model, semiflexible behavior is introduced by using steric hindrance. The polymer is modeled as a chain of tangent or fused hard spheres, as shown in Figure 1. The larger the overlap of the spheres, i.e., the smaller the ratio of bond length, *L*, to sphere diameter, σ, the more the bond angles are restricted to large values ($\vartheta_{\min} = 2 \sin^{-1}(1/2L)$). As a consequence, the chains get stiffer when the bond length is reduced, so we can employ the bond length as a parameter characterizing the stiffness of our chains. We choose $\sigma = 1$ as our unit of length and vary the bond length in the range $0.51 \leq L \leq 1$.
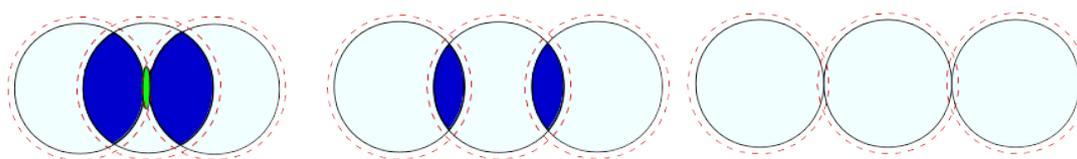

Figure 1. Stiff chains (**left**), semi-flexible chains (**middle**) and flexible chains (**right**) differ by the amount of overlap between adjacent spheres. The dashed red line indicates the range of the square-well attraction, the blue ellipses the inaccessible surface and the green ellipse the overlap of square wells of next-nearest neighbors along the chain.

The case $L = 1$ (shown on the right of Figure 1) is the limit of flexible chains, for which the phase diagram has been studied in detail before [29]. When we introduce an effective attraction between the monomers in the form of a square well potential,

$$U(r_{ij}) = \begin{cases} \infty & r_{ij} < 1 \\ -\varepsilon & 1 < r_{ij} < \lambda \\ 0 & r_{ij} > \lambda \end{cases} , \tag{1}$$

the chain will collapse upon temperature reduction. We choose $\varepsilon = 1$ as our unit of energy and temperatures are measured in units of ε/k_B. Although not the focus of the present work, the square-well potential can be used to study electrostatic interactions. For example, in [30], the phase behavior

of charged colloids in solution is studied by mapping the screened Coulomb potential onto an effective square well attraction. Similarly, the square well chain model has been shown to provide a qualitatively valid representation for the thermodynamics of peptides in aqueous solution [31]. For $\lambda < 1.18$ [32], the collapse is a first order transition in the thermodynamic limit, whereas for $\lambda \geq 1.18$, it is a second order transition followed by a first order crystallization transition at lower temperatures. With decreasing chain length, this crossover shifts to smaller λ, and for the chain lengths we study, the value of $\lambda = 1.1$ that we employed in the simulation means that we have a second order collapse transition for the flexible case. With increase in the overlap (going from right to left in Figure 1), the chains get stiffer due to steric hindrance, and, at the same time, the accessible surface for non-bonded contacts gets smaller (semi-flexible regime with $0.55 < L < 1$). For very short bonds, the square well attractions of next-nearest neighbors overlap leading to a permanent energy contribution (stiff regime with $0.51 \leq L \leq 0.55$).

We have performed Stochastic Approximation Monte Carlo simulations (SAMC) [33,34], a version of flat-histogram Monte Carlo simulations [35], which we adapted to the simulation of polymer models [36–38]. These simulations determine a numerically exact (if the simulation is converged sufficiently) result for the density of states, $g(E)$, of the model under study, or, more precisely, its logarithm, which is the micro-canonical entropy, $S(E) = \ln[g(E)]$, where we set $k_B = 1$. In SAMC, as in Wang–Landau MC, the density of states is determined iteratively starting from an estimate $g_0(E) = 1$ for the complete energy range, by performing an unbiased Monte Carlo simulation in configuration space and accepting new configurations with the acceptance probability

$$A(E) = \min\left(1, \frac{g(E_{\text{old}})}{g(E_{\text{new}})}\right) .$$

The SAMC update of the current estimate for the logarithm of the density of states is [36]

$$\ln[g(E)] = \ln[g(E)] + \gamma_t \left(\delta_{E,E'} - 1/M\right) , \tag{2}$$

where M is the number of energy values or energy intervals for which one wants to estimate $g(E)$. In each update, M is subtracted for all E, and only for the new energy E' (equal to either E_{old} or E_{new}) accepted in the Monte Carlo procedure, the modification factor γ_t is added to $\ln[g(E')]$. The following two necessary conditions [33,34] exist for this method to converge:

$$\sum_{t=1}^{\infty} \gamma_t = \infty, \tag{3}$$

$$\sum_{t=1}^{\infty} \gamma_t^{\nu} < \infty \quad \text{for some } \nu \in (1, 2) . \tag{4}$$

From the density of state, one obtains, for instance, the canonical partition function as

$$Z(T) = \sum_x \exp\{-\beta U(x)\} = \sum_E g(E) \exp\{-\beta E\} , \tag{5}$$

where x is a point in configuration space, $U(x)$ is the potential energy of the configuration and $\beta = 1/k_B T$. Both $g(E)$ and $Z(T)$ contain the complete thermodynamic information on the model, and a determination of its thermodynamic properties is most efficiently performed in a combination of a micro-canonical (starting from $g(E)$) and canonical (starting from $Z(T)$) analysis [35]. We will analyze mainly the behavior of the derivative of the inverse micro-canonical temperature, $\gamma(E) = dT^{-1}(E)/dE = d^2 S(E)/dE^2$ and of the canonical specific heat

$$C_N(T) = \frac{2TZ'Z + T^2 ZZ'' - T^2 Z'^2}{Z^2}, \tag{6}$$

where the prime indicates a derivative with respect to T. The derivative of the micro-canonical temperature has peaks with a value larger than zero at first order transitions and peaks with a value smaller than zero at second order ones [39]. The canonical specific heat shows transitions as peaks and sometimes as shoulders, when transitions can not be well separated. A structural quantity that monitors the collapse transition is the radius of gyration of the chains. We obtain this—and all equilibrium observables—first as an arithmetic average, $\overline{R_g^2}(E)$, over visited configurations at a given energy in a simulation using the converged density of states, which can then be transformed into the canonical ensemble

$$\langle R_g^2 \rangle(T) = \frac{1}{Z(T)} \sum_E \overline{R_g^2}(E) g(E) \exp\{-\beta E\} . \tag{7}$$

3. Results

In the following, we will discuss pseudo phase diagrams for chains of length $N = 20$ and $N = 40$ as a function of bond length (i.e., stiffness) and temperature. As these are pseudo phase diagrams, they depend quantitatively (location of transitions) and qualitatively (order of the transition, number of transitions, resulting morphologies) on chain length. Such a study is therefore in spirit closer to the analysis of peptide pseudo phase behavior than to a study of the properties in the thermodynamic limit. The limiting behavior shows true phase transitions, but it loses pseudo phases of interesting morphologies. After all, chain stiffness is also a determining factor for the native state morphology of peptides and proteins.

3.1. Chain Length $N = 20$

For this chain length, we start by presenting the basic result from our simulations, i.e., the micro-canonical entropy $S(E) = \ln[g(E)]$ in Figure 2.
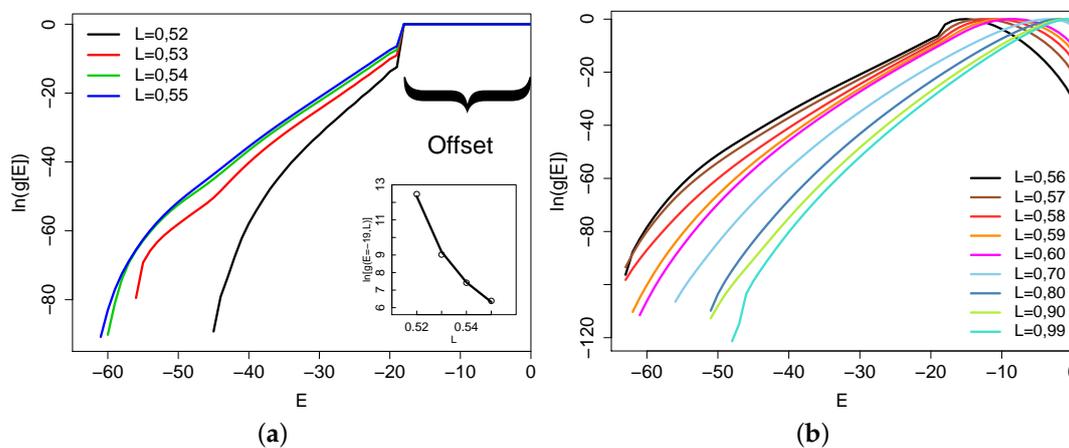

Figure 2. (**a**) shows the micro-canonical entropy for the stiff chains (note that, for $N = 20$, we only studied bond lengths larger than 0.51); (**b**) for the semi-flexible chains. Note the shift in the energy spectrum in the stiff case due to the permanent overlap of interaction zones of next-nearest neighbors along the chain. The inset shows the jump at the step in $\ln[g(E)]$ and a fit with the law $(L - 0.5)^{-3/4}$.

For the stiff chains ($0.52 \leq L \leq 0.55$), there is a jump in the density of states at $E = -18$, which, for these chains, is the energy of permanently interacting next-nearest neighbors along the chain. The jump of the entropy seems to diverge for $L = 0.5$, where the only possible bond angle is $\vartheta = \pi$ and the chain is a rigid rod (see inset on the left). For the semi-flexible chains ($0.56 \leq L < 1$), a smeared out step is still there for $L = 0.56$ but is smoothed out for larger L, while, at the same time, a maximum in the entropy occurs at finite energy. For energies larger than this value, the inverse

micro-canonical temperature becomes negative, a well known behavior for the Boltzmann entropy of finite systems, which can be avoided by using the Gibbs definition of entropy [40,41].

　　Although the entropy curves in Figure 2 contain the complete thermodynamics, information about pseudo phase transitions of the chain are buried in subtle curvature changes of these data. They become more clear when one calculates, e.g., the specific heat from them as is done for Figure 3a or determines the radius of gyration of the chains, for which its temperature derivative is shown in Figure 3b. Peaks in either of these quantities can be associated with structural transitions. The large peaks in the main panel in Figure 3b clearly locate the collapse transition of the chains, as one can also glean from the upper right inset of this figure, which shows the temperature dependence of the radius of gyration itself. As is well known for the collapse transition of polymers, the peak positions in different quantities signaling the transition are shifted with respect to each other. However, the main peaks agree well between the two quantities. The difference in peak positions, combined with information from other quantities, defines our error estimate for the determination of the transition temperature. The order of the transition can be obtained by looking at the energy histogram at the transition temperature, which is shown in the upper right inset in Figure 3a. For $L = 0.53$ and $L = 0.54$, we see a clear double peak structure indicating a first order pseudo phase transition. For $L = 0.55$, the peaks are not well separated and the order of the transition is not well determined from this canonical analysis. A micro-canonical analysis of the peaks in the temperature derivative (as shown for $N = 40$ in the next subsection) tells us that the $L = 0.55$ collapse transition is of first order. The same is obtained for $L = 0.52$, for which we have a highly asymmetric energy distribution due to the small available energy range for this bond length (see Figure 2).
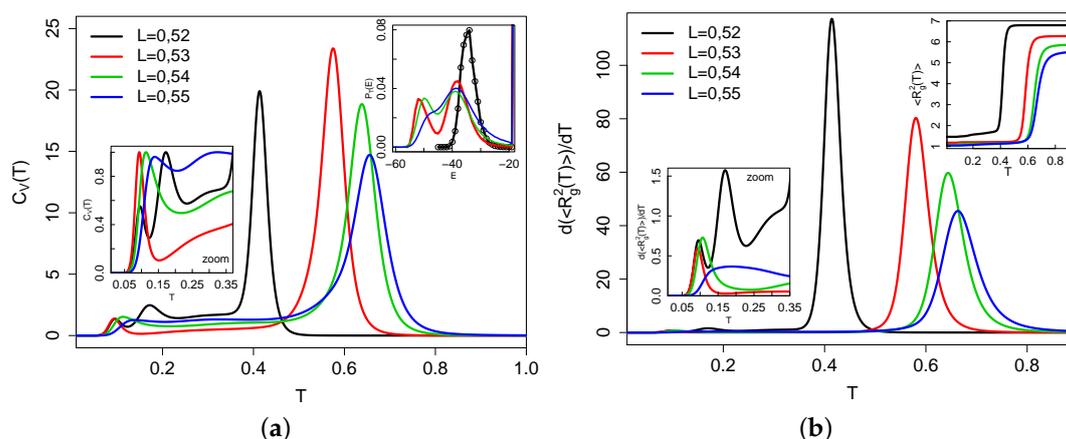

Figure 3. Part (**a**) shows the canonical specific heat for the stiff chains. The inset in the lower left corner shows the low temperature region in more detail. The inset in the upper right corner shows the probability distribution of the energy for the temperatures where the specific heat has a maximum. Part (**b**) shows the temperature derivative of the gyration radius for the same chains. The inset in the lower left corner highlights the low temperature behavior, the inset in the upper right corner shows the temperature dependence of the radius of gyration itself.

　　The determination of the bond length dependence of the transition in the semi-flexible range is illustrated in Figure 4 using the canonical specific heat. For $L = 0.99$ (panel d), a broad high temperature peak indicates the continuous collapse transition of this chain, while a shoulder and a sharp peak at lower temperature indicate further structural transitions. The low temperature crystallization of the flexible chain at $L = 1$ is not well resolvable here because the chain is so short that the globule basically consists of the surface and can not undergo a further volume phase transition. For $L = 0.9$, two peaks are visible with the lower temperature one being the main peak. Micro-canonical analysis indicates that this is a first order transition. In the range between $L = 0.7$ and $L = 0.85$, a broad plateau

region develops between $T \simeq 0.3$ and $T \simeq 0.6$ (panel c), with the relative height between the high temperature and the low temperature part varying as a function of stiffness. For $L < 0.70$, the peaks in the specific heat become again more pronounced (panel b) with the peak at higher temperature becoming the main transition for $L \leq 0.61$. For $L = 0.56$ (shown in panel a), we cross over to the stiff chain behavior with a sharp and clear high temperature peak as in Figure 3a.
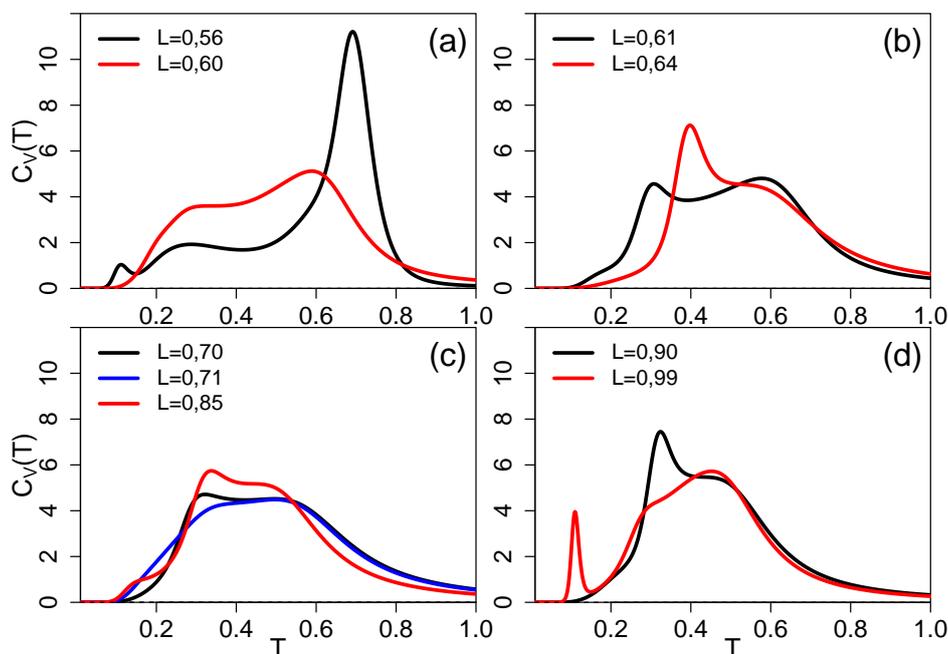

Figure 4. Canonical specific heat as a function of temperature for the bond length range of semi-flexible chains. (**a**–**d**) show different parts of this range.

All the above analysis cumulates in the pseudo phase diagram of Figure 5. As expected for this chain length, the collapse transition is of second order for flexible and moderately semi-flexible chains, but it becomes a first order transition for $L \simeq 0.6$. For all stiffnesses, we observe a regime of further structural transitions at lower temperature, which are of first order for two regimes of bond lengths. Further isolated points indicate local structural transitions which—for a short chain of length $N = 20$—lead to observable features in the thermodynamic quantities, but which are not suspected to survive in the thermodynamic limit. For $N \to \infty$, the line of collapse transitions can be expected to become of first order for all stiffnesses, as it is already of first order for this choice of interaction range for the flexible chains of $L = 1$, and stiffness tends to strengthen the first order character of the transition [15,16]. However, this conclusion is only valid if entropic stiffness (steric exclusion) like we introduce here has the same consequences as energetic stiffness (bending energy) studied in references [15,16]. The configuration snapshots included in Figure 5 show the transitions from random coil (h) to crystalline globule (j) on the flexible end. For semi-flexible chains (configurations g, i, k and l) the collapsed and ground states get elongated and some indications of lamellar packing (i and l) are observable. Stiff chain morphologies for $L = 0.52$ above the "collapse" transition (d) and below it (a) are very similar and only differ by some winding of the chain ends around each other. For $0.53 \leq L \leq 0.55$ (configurations b, c, e, f, and m), this winding gets easier as the chains are more flexible and tighter structures are formed. Which of the low temperature transitions and configurations survive for longer chains and what they signify is a different question. We will address this for a chain of length $N = 40$, which still is in the "peptide" range in the next subsection.
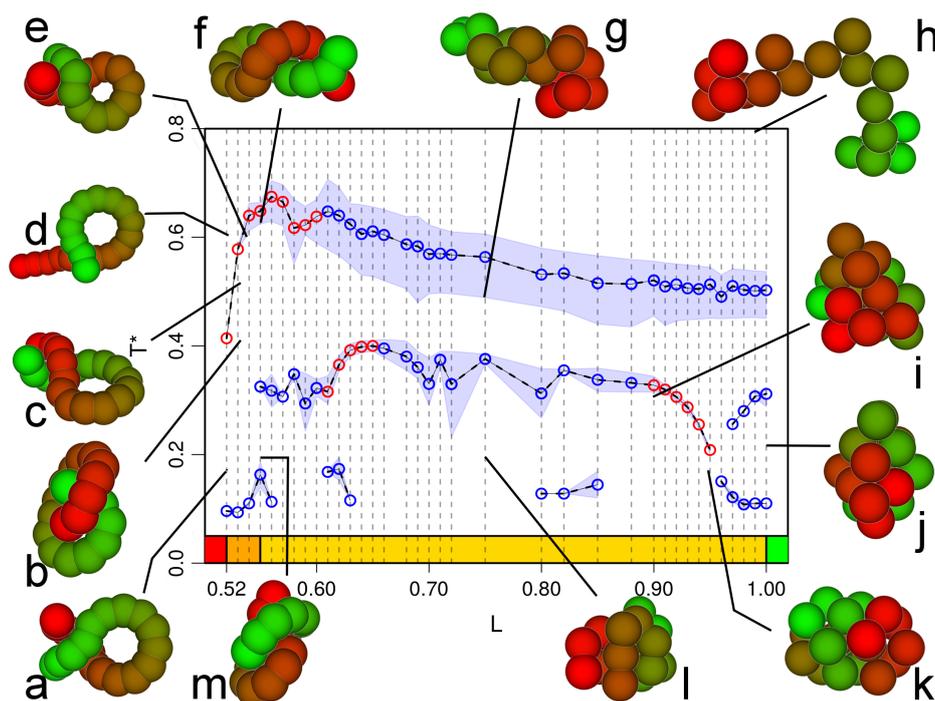
Figure 5. Pseudo phase diagram for chain length $N = 20$. Blue symbols indicate second order transitions, red symbols indicate first order ones. The shaded areas are error estimates based on the different quantities analyzed. Isolated points are included when they occur in more than one observable and again error estimates are given. The configurations (**a**) to (**m**) indicate most probable configurations at a given bond length and temperature pair.

3.2. Chain Length N = 40

For this chain length, we studied the same quantities as discussed in the last subsection for chain length $N = 20$. Here, we illustrate the micro-canonical version of the analysis which is based on the second derivative of the entropy function, as shown in Figure 6 for chains in the semi-flexible range.

As a function of energy, we see (in panel d) a broad plateau region with a high energy (very broad) maximum for $L = 0.95$ and $L = 0.85$ with a value below zero, which indicates the second order collapse transition. At the low energy end of the plateau, both bond lengths exhibit a peak with a maximum below zero, indicating an additional second order transition. For $L = 0.75$ and $L = 0.7$ (panel c), this low energy feature splits in small oscillations of the curve, while, at $L = 0.65$, it is visible again (panel b). For $L = 0.6$ (panel b), we see a step at high energies, a remnant of the step in the entropy between $E = -19$ and $E = -18$, which gives rise to the high energy maximum seen in (panel a) for both $L = 0.58$ and $L = 0.56$. Both of these bond lengths also give a high energy peak which is larger than zero, indicating a first order transition. For $L = 0.56$, another first order transition occurs at low energy, and two intermediate maxima very close to zero are barely visible in the intermediate energy regime.

Combining the analyses of different micro-canonical and canonical observables, we obtain the pseudo phasediagram shown in Figure 7. Here, the regime of second order collapse transitions extends to smaller bond lengths (down to $L = 0.57$) than for $N = 20$. If the hypothesis that the entropic stiffness we are studying behaves similarly to the behavior found in the literature for energetic stiffness is valid, this would indicate that the approach to the thermodynamic limit of this pseudo phasediagram is not monotonous. However, as stated before, this hypothesis has not been proven yet. Below $L = 0.57$, a further intermediate line of transitions occurs, which is first order for $L \leq 0.54$. For this chain, it is instructive to look at the equilibrium morphologies in the different parts of the phase diagram in more detail. Flexible chains at high energy (and temperature) are of course random coils (configuration *h*),

which then collapse into liquid globules (configurations *i* and *j*), and, for $L = 1$, collapse further into an fcc crystal (configuration *k*). For $L = 0.9$, the low temperature configuration (*h*) shows no crystalline order but begins to show toroidal segments or helical structures also visible in configuration *m*, which is an intermediate temperature configuration at $L = 0.6$. For $L \leq 0.6$, we observe the occurrence of knotted structures like those shown in configurations *a*, *b*, *c* and *g*, next to hairpin structures like configuration *f* and tennis racket like structures like configuration *d*, which occur in this bond length range at high temperatures at or above the collapse transition.
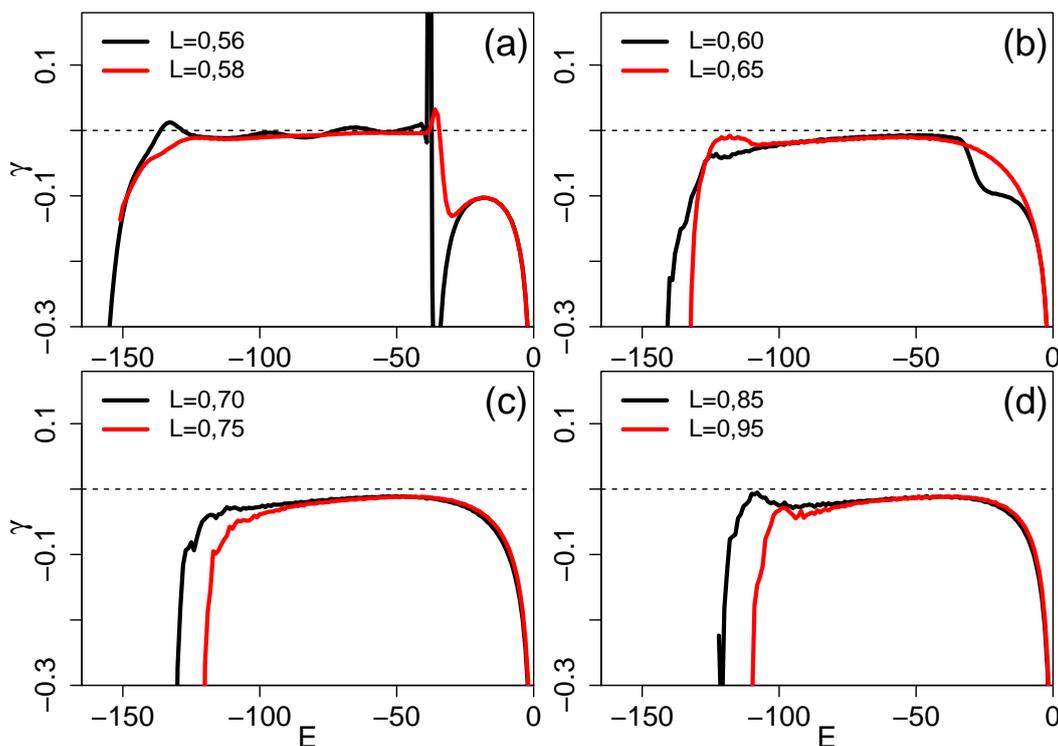

Figure 6. Second derivative of the entropy with respect to energy, $\gamma(E) = d^2S/dE^2$, for different bond lengths in the semi-flexible regime, going from the most flexible chains in (**d**) via increasing stiffness in (**c**) and (**b**) to the border of the stiff regime in (**a**). Peaks with heights above zero indicate a first order phase transition, the others indicate second order transitions.

Configuration *b* is of special interest because it is a true trefoil knot. In this context, configuration *d* is also of interest because it is the first step on the path to the trefoil knot, which is shown in Figure 8. Starting from the tennis racket structure (*a* in Figure 8), the two ends slide over each other (*b2*) and start to twist around each other *c*. Finally, the knot tightens (*d*) and, for very low temperatures, the two ends come into contact (*e1* and *e2* show two views of a configuration slightly before that) and a closed trefoil knot is formed. Configurations *b1* and *b3* are degenerate in energy with *b2*. Configuration *y* shows the ground state for $L = 0.52$ and $N = 20$, which is close to configuration *b1*; the chain length $N = 20$ is too short for the development of a knot at this stiffness, which is formed in the ground state for $L = 0.52$ and $N = 40$ shown in configuration *z*. The slightly larger flexibility for this bond length allows the knot to tighten and the ends to slip over each other.
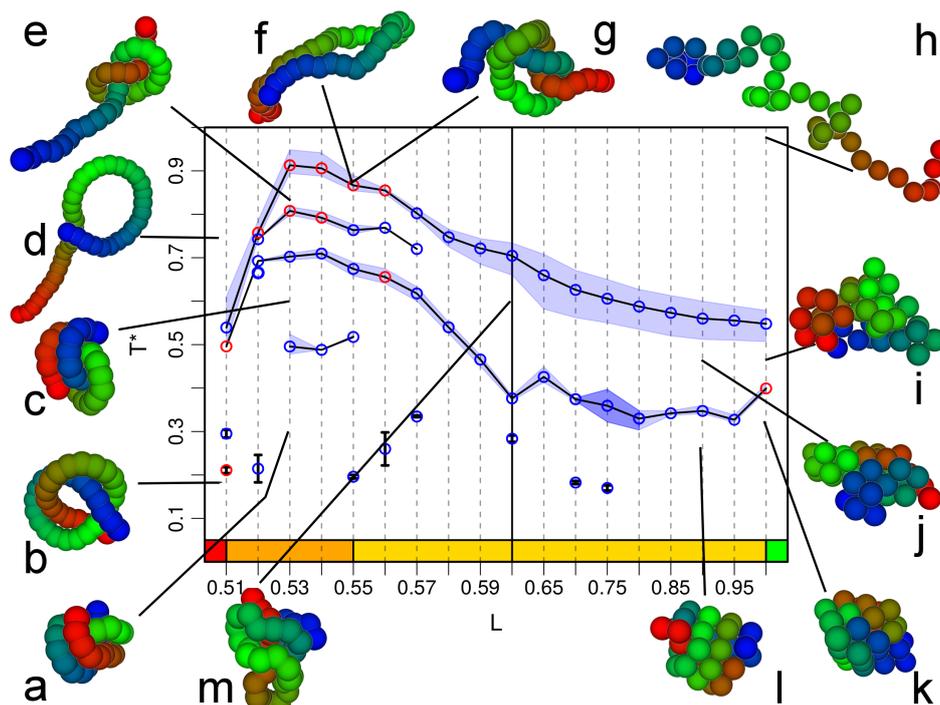
Figure 7. Pseudo phase diagram for chain length $N = 40$ including examples for configurations (**a**–**m**) which are the most probable in the different regions. Blue symbols are for second order transitions, and red symbols are for first order transitions. The shaded regimes indicate error estimates coming from different observables. Only points resulting from a signal in more than one observable are included.

To characterize such a transition into a non-trivially ordered ground state like the trefoil knot more quantitatively, we make use of the contact matrix, a typical tool for the analysis of protein ground states. In the folded state, i.e., the ordered low temperature configuration, one has a fixed neighborhood for each monomer. The same is true for the trefoil knot. We have already shown that one can determine the temperature dependence of contact matrices from SAMC simulations in studies of a Gō model [42] and get an idea about the formation process of the native state upon decreasing the temperature. In Figure 9, we show a sequence of contact matrices for the chain of length $N = 40$ and temperatures going from $T = 0.8$ in a, to $T = 0.62$ in b, $T = 0.57$ in c and $T = 0.2$ in d. For all panels, due to the short bond length of $L = 0.51$, there is always an off-diagonal next-nearest neighbor contact $(n, n2)$, which is the diagonal of red dots. At the highest temperature, one has a high probability for an end contact forming (the snapshot shows one of the less probable contacts from the yellow regime). For smaller temperatures, the region of highest contact probability expands and moves inwards from the ends (panel b). The front boundary of the contact region, however, stays stationary as a function of temperature. For still lower temperatures (panel c), the region of high contact probability has extended into a strip parallel to the diagonal meaning that parallel chain segments at the two ends get into contact as shown in the snapshot. Finally, for the ground state, a well defined strip of neighbor contacts forms, indicating the unique native state (the trefoil knot) of this chain.
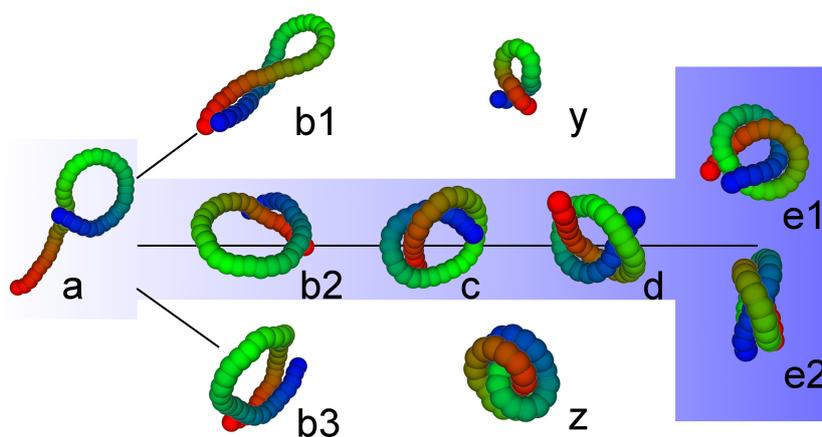
Figure 8. Path to the trefoil knot for $N = 40$ and $L = 0.51$ (**e1** and **e2** show two different views) starting from a tennis racket like configuration (**a**) when one reduces temperature (or energy) going from left to right through **b2**, **c** and **d**. Configurations **b1** and **b3** coexist with **b2**, configuration **y** is the ground state for $N = 20$ and $L = 0.52$, while **z** is the ground state for $N = 40$ and $L = 0.52$.
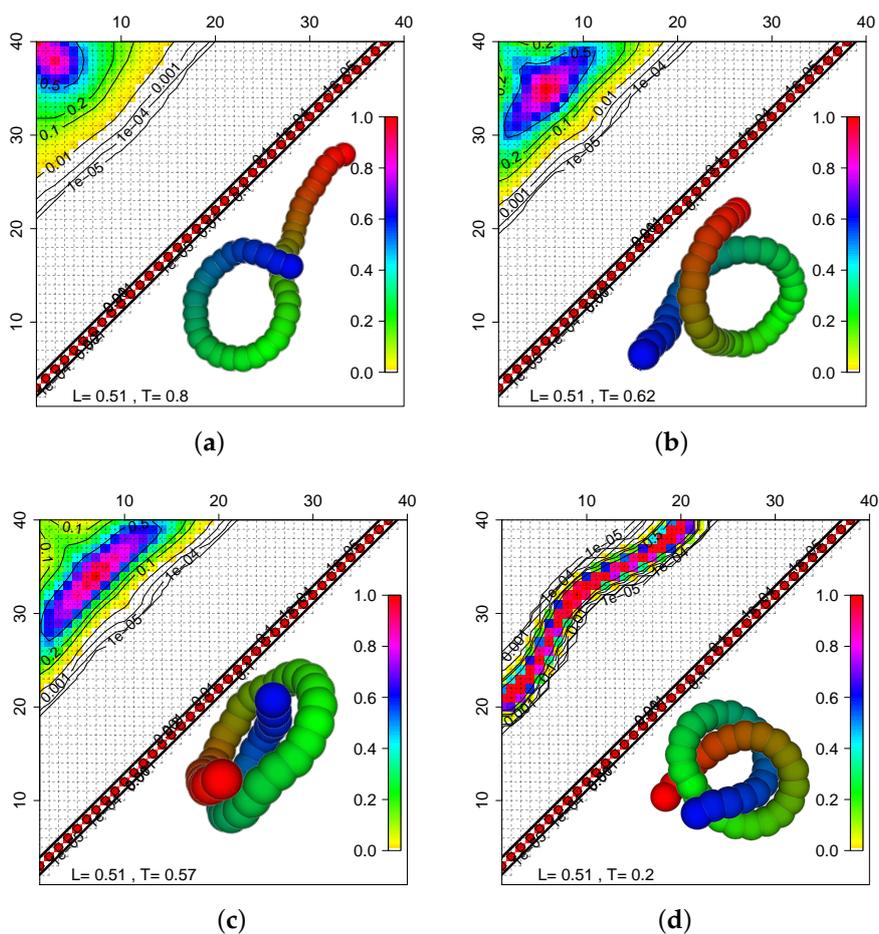

Figure 9. Contact matrices for the chain of length $N = 40$, bond length $L = 0.51$ and for the temperatures $T = 0.8$ in (**a**); $T = 0.62$ in (**b**); $T = 0.57$ in (**c**) and $T = 0.2$ in (**d**). The snapshots visualize the configurations giving rise to the contact matrices.

4. Discussion

In simulations of stiff chains with the persistent mechanism of flexibility [16–18,21–23], one determines phase behavior as a function of two energetic parameters, the temperature and the stiffness energy κ, taking the strength of the non-bonded attraction as the energy scale of the model. This can be done either at constant stiffness, i.e., constant κ/T as in [16–18] or at fixed stiffness energy κ like in [21–23]. The folded, toroidal and hairpin like structures (or tennis rackets) found in [21] were very similar to what was found at constant stiffness in [16–18], except for the explicit differentiation of multiply folded structures in [21]. Keeping the bond length constrained [22,23] but the rest of the model as in [21], one loses the toroidal structure but creates knotted configurations at intermediate stiffnesses for about the same chain length.

In contrast to these studies, stiffness in our model is introduced by steric effects, i.e., it is of entropic not energetic nature. Our phase diagram has only one energy parameter, while the second parameter (the bond length) is kept constant. Our very stiff chains can not fold back, and consequently the lamellar, folded structures are not found in the phase diagram. Hairpins and toroids exist at high stiffness, but they are no longer the ground state, which is a knotted structure, ideally the trefoil knot as for $L = 0.51$ and $N = 40$. The knotted structures in general have moved to the high stiffness region (they occur for $L \leq 0.6$) and are not found in the intermediate stiffness region, like in [22,23], where the high stiffness region is dominated by hairpins and folded chains. This comparison thus confirms our claim from the introduction: not all stiffnesses are created equal [7], and this is true even for models exhibiting the persistent mechanism of flexibility. There are common features in the phase diagram of stiff chains, but they are sufficiently different that one has to decide carefully which model really maps to a given stiff synthetic or bio-polymer. Finally, as we have also shown here, the pseudo phase diagram, of course, depends on chain length, such that structures which are dominant for one chain length might not be possible or be suppressed for another one.

5. Conclusions

We have presented flat histogram Monte Carlo simulations using the SAMC approach to determine the pseudo phase diagram of single semi-flexible chains as a function of stiffness and temperature. Our model consists of overlapping hard spheres interacting with a square well attraction. The degree of overlap determines the stiffness of the chains. In this aspect, the model is close to the behavior of real polymers, where typically the van der Waals radius of the atoms is almost twice as large as the bond length. We presented pseudo phase diagrams for chains of length $N = 20$ and $N = 40$, which exhibit different stable morphologies for low temperatures (or energies). The phase diagrams were determined from a combined canonical and micro-canonical analysis based on the micro-canonical entropy determined in the SAMC simulations. To extend the chain length range that can be studied, it might be useful to resort to the flat-histogram chain growth methods [43,44].

For our model, in which stiffness is induced by steric effects, i.e., it is of entropic origin, the ground state structures for stiff chains are knotted (similar to what was found by Marenz et al. [22] for intermediate stiffness in a model where stiffness is of energetic origin). We showed that, for the stiffest chain of length $N = 40$, the ground state is given by an exact trefoil knot, where the two chain ends are in contact. We have established the route to the trefoil knot upon reducing temperature using contact matrices, a tool known from the analysis of protein configurations. Although the general types of collapsed and ground state morphologies observed in different models for semi-flexible polymers tend to be similar, the locations of different morphologies in the pseudo phase diagrams is model dependent. Thus, comparison with experiment must guide the appropriateness of using different mechanisms (e.g., energetic vs entropic) for chain stiffness. Noting that the pseudo phase diagrams are also chain length dependent, it is apparent that there exists a huge variety of possible collapse and folding behaviors to non-trivial morphologies resulting from the competition between chain stiffness and monomer attraction. Going beyond the homopolymer case [37,38] introduces further complexity

on the path to protein-like polymers, for which the native state in addition is stabilized by specific interactions like, e.g., hydrogen bonds.

Author Contributions: Benno Werlich performed the simulations and analysis with the help of Timur Shakirov; Wolfgang Paul devised the project together with Mark P. Taylor who performed additional Wang–Landau simulations, and both guided the analysis and interpretation, Wolfgang Paul wrote the first draft of the paper and all authors completed it together.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Jiang, X.; Reiter, G.; Hu, W. How Chain-folding crystal growth determines thermodynamic stability of polymer crystals. *J. Phys. Chem. B* **2016**, *120*, 566–571.
2. Maurstad, G.; Stokke, B.T. Toroids of stiff polyelectrolytes. *Curr. Opin. Colloid Interface Int. Sci.* **2005**, *10*, 16–21.
3. Finkelstein, A.V.; Ptitsyn, O. *Protein Physics*; Academic Press: London, UK, 2016.
4. Chiti, F.; Dobson, C.M. Amyloid formation by globular proteins under native conditions. *Nat. Chem. Biol.* **2009**, *5*, 15–22.
5. Flory, P.J. Statistical Thermodynamics of Semi-Flexible Chain Molecules. *Proc. R. Soc. A* **1956**, *234*, 60–73.
6. Kratky, O.; Porod, G. Röntgenuntersuchung gelöster Fadenmoleküle. *J. R. Neth. Chem. Soc.* **1949**, *68*, 1106–1122.
7. Khokhlov, A.R.; Semenov, A.N. Theory of nematic ordering in the melts of macromolecules with different flexibility mechanisms. *Macromolecules* **1986**, *19*, 373–378.
8. Schäfer, L.; Ostendorf, A.; Hager, J. Scaling of the correlations among segment directions of a self-repelling polymer chain. *J. Phys. A Math. Gen.* **1999**, *32*, 7875–7899.
9. Wittmer, J.P.; Beckrich, P.; Meyer, H.; Cavallo, A.; Johner, A.; Baschnagel, J. Intramolecular long-range correlations in polymer melts: The segmental size distribution and its moments. *Phys. Rev. E* **2007**, *76*, 011803.
10. Hsu, H.-P.; Paul, W.; Binder, K. Standard Definitions of Persistence Length do not Describe the Local "Intrinsic" Stiffness of Real Polymer Chains. *Macromolecules* **2010**, *43*, 3094–3102.
11. Hsu, H.-P.; Paul, W.; Binder, K. Polymer chain stiffness vs. excluded volume: A Monte Carlo study of the crossover towards the worm-like chain model. *Eurphys. Lett.* **2010**, *92*, 28003.
12. Hsu, H.-P.; Paul, W.; Binder, K. Breakdown of the Kratky-Porod wormlike chain model for semiflexible polymers in two dimensions. *Europhys. Lett.* **2011**, *95*, 68004.
13. Noguchi, H.; Yoshikawa, K. First-order phase transition in a stiff polymer chain. *Chem. Phys. Lett.* **1997**, *278*, 184–188.
14. Noguchi, H.; Yoshikawa, K. Morphological variation in a collapsed single homopolymer chain. *J. Chem. Phys.* **1998**, *109*, 5070–5077.
15. Bastolla, U.; Grassberger, P. Phase transitions of single semistiff polymer chains. *J. Stat. Phys.* **1997**, *89*, 1061–1078.
16. Ivanov, V.A.; Paul, W.; Binder, K. Finite chain length effects on the coil-globule transition of stiff-chain macromolecules: A Monte Carlo simulation. *J. Chem. Phys.* **1998**, *109*, 5659–5669.
17. Stukan, M.R.; Ivanov, V.A.; Grosberg, A.Y.; Paul, W.; Binder, K. Chain length dependence of the state diagram of a single stiff-chain macromolecule: Theory and Monte Carlo simulation. *J. Chem. Phys.* **2003**, *118*, 3392–3400.
18. Martemyanova, J.A.; Stukan, M.R.; Ivanov, V.A.; Müller, M.; Paul, W.; Binder, K. Dense orientationally ordered states of a single semiflexible macromolecule: An expanded ensemble Monte Carlo simulation. *J. Chem. Phys.* **2005**, *122*, 174907:1–174907:10.
19. Wang, F.; Landau, D.P. Efficient, multiple-range random walk algorithm to calculate the density of states. *Phys. Rev. Lett.* **2001**, *86*, 2050–2053.
20. Wang, F.; Landau, D.P. Determining the density of states for classical statistical models: A random walk algorithm to produce a flat histogram. *Phys. Rev. E* **2001**, *64*, 056101:1–056101:16.

21. Seaton, D.T.; Schnabel, S.; Landau, D.P.; Bachmann, M. From Flexible to Stiff: Systematic Analysis of Structural Phases for Single Semiflexible Polymers. *Phys. Rev. Lett.* **2013**, *110*, 028103:1–028103:5.

22. Marenz, M.; Janke, W. Knots as a Topological Order Parameter for Semiflexible Polymers. *Phys. Rev. Lett.* **2016**, *116*, 128301.

23. Zierenberg, J.; Marenz, M.; Janke, W. Dilute semiflexible polymers with attraction: Collapse, folding and aggregation. *Polymers* **2016**, *8*, 333.

24. Banavar, J.R.; Maritan, A. Colloquium: Geometrical approach to protein folding: A tube picture. *Rev. Mod. Phys.* **2003**, *75*, 23–34.

25. Vogel, T.; Neuhaus, T.; Bachmann, M.; Janke, W. Thickness-dependent secondary structure formation of tubelike polymers. *Europhys. Lett.* **2009**, *85*, 10003.

26. Magee, J.E.; Vasquez, V.R.; Lue, L. Helical structures from an isotropic homopolymer model. *Phys. Rev. Lett.* **2006**, *96*, 207802:1–207802:4.

27. Bannerman, M.N.; Magee, J.E.; Lue, L. Structure and stability of helices in square-well homopolymers. *Phys. Rev E* **2009**, *80*, 021801.

28. Taylor, M.P.; Paul, W.; Binder, K. Applications of the Wang-Landau algorithm to phase transitions of a single polymer chain. *Polym. Sci. Ser. C* **2013**, *55*, 23–38.

29. Taylor, M.P.; Paul, W.; Binder, K. Phase transitions of a single polymer chain: A Wang-Landau simulation study. *J. Chem. Phys.* **2009**, *131*, 114907.

30. Noro, M.G.; Frenkel, D. Extended corresponding-states behavior for particles with variable range attractions. *J. Chem. Phys.* **2000**, *113*, 2941–2944.

31. Taylor, M.P.; Paul, W.; Binder, K. On the polymer physics origins of protein folding thermodynamics. *J. Chem. Phys.* **2016**, *145*, 174903.

32. Taylor, M.P.; Aung, P.P.; Paul, W. Partition function zeros and phase transitions for a square-well polymer chain. *Phys. Rev. E* **2013**, *88*, 012604.

33. Liang, F. A generalized Wang-Landau algorithm for Monte Carlo computation. *J. Am. Stat. Assoc.* **2005**, *100*, 1311–1327.

34. Liang, F.; Liu, C.; Carroll, R.J. Stochastic approximation in Monte Carlo computation. *J. Am. Stat. Assoc.* **2007**, *102*, 305–320.

35. Janke, W.; Paul, W. Thermodynamics and structure of macromolecules from flat-histogram Monte Carlo simulations. *Soft Matter* **2016**, *12*, 642–657.

36. Werlich, B.; Shakirov, T.; Taylor, M.P.; Paul, W. Stochastic approximation Monte Carlo and Wang-Landau Monte Carlo applied to a continuum polymer model. *Comput. Phys. Commun.* **2015**, *186*, 65–70

37. Zablotskiy, S.V.; Martemyanova, J.A.; Ivanov, V.A.; Paul, W. Diagram of states and morphologies of flexible-semiflexible copolymer chains: A Monte Carlo Simulation. *J. Chem. Phys.* **2016**, *144*, 244903.

38. Zablotskiy, S.V.; Martemyanova, J.A.; Ivanov, V.A.; Paul, W. Stochastic Approximation Monte Carlo Algorithm for Calculation of Diagram of States of a Single Flexible-Semiflexible Copolymer Chain. *Polym. Sci. Ser. A* **2016**, *58*, 899–915.

39. 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.

40. Hänggi, P.; Hilbert, S.; Dunkel, J. Meaning of temperature in different thermostatistical ensembles. *Philos. Trans. R. Soc. A* **2016**, *374*, 20150039.

41. Shakirov, T.; Zablotskiy, S.; Böker, A.; Ivanov, V.A.; Paul, W. Comparison of Boltzmann and Gibbs entropies for the analysis of single-chain phase transitions. *Eur. Phys. J. E. Spec. Top.* under review.

42. Böker, A.; Paul, W. Wang-Landau simulations of Gō model molecules. *Eur. Phys. J. E* **2016**, *39*, 5.

43. Owczarek, A.L.; Prellberg, T. Monte carlo investigation of lattice models of polymer collapse in five dimensions. *Int. J. Mod. Phys. C* **2003**, *14*, 621–633.

44. Bachmann, M.; Janke, W. Thermodynamics of lattice heteropolymers. *J. Chem. Phys.* **2004**, *120*, 6779–6791.