

LU TP 98-27  
Revised version  
May 11, 1999

# Monte Carlo Study of the Phase Structure of Compact Polymer Chains

Anders Irbäck<sup>1</sup> and Erik Sandelin<sup>2</sup>

Complex Systems Group, Department of Theoretical Physics  
University of Lund, Sölvegatan 14A, S-223 62 Lund, Sweden  
<http://www.thep.lu.se/tf2/complex/>

Submitted to *Journal of Chemical Physics*

Abstract:

We study the phase behavior of single homopolymers in a simple hydrophobic/hydrophilic off-lattice model with sequence independent local interactions. The specific heat is, not unexpectedly, found to exhibit a pronounced peak well below the collapse temperature, signalling a possible low-temperature phase transition. The system size dependence at this maximum is investigated both with and without the local interactions, using chains with up to 50 monomers. The size dependence is found to be weak. The specific heat itself seems not to diverge. The homopolymer results are compared with those for two non-uniform sequences. Our calculations are performed using the methods of simulated and parallel tempering. The performances of these algorithms are discussed, based on careful tests for a small system.

---

<sup>1</sup>irback@thep.lu.se

<sup>2</sup>erik@thep.lu.se

# 1 Introduction

The thermodynamic behavior of isolated homopolymers is known in quite some detail at and above their collapse temperature  $T = T_\theta$ , from analytical work [1, 2, 3] and numerical simulations of very long chains [4, 5]. Much less is known about the behavior at low temperatures. Consequently, it is of utmost interest to examine the low- $T$  phase behavior and its model dependence, which, in particular, may shed light on the mechanism of folding for heteropolymers.

The possibility of an interesting low- $T$  phase structure for homopolymers was predicted several years ago [6]. Recent studies [7, 8, 9, 10] of lattice chains with stiffness [11] have shown that such chains, in addition to the coil-globule transition at  $T_\theta$ , exhibit a phase transition from the globule phase to a frozen crystalline phase. Also, Zhou *et al.* [12, 13] have found evidence for two or more phase transitions in simple off-lattice models without stiffness terms.

In this paper, our starting point is a simple off-lattice model [14] for protein folding with two types of monomers, A (hydrophobic) and B (hydrophilic). In addition to “hydrophobicity” forces, the model also contains sequence independent local interactions. It has previously been demonstrated [14] that it is possible to find AB sequences with compact, stable native structures in this model. Furthermore, it was shown that this property is crucially dependent on the local interactions. In the present paper we perform a detailed study of A homopolymers in this model, both with and without the local interactions. The most striking feature of the A homopolymers, at least for small chain lengths, is a pronounced peak in the specific heat, located well below  $T_\theta$ . This could signal that these chains, like lattice chains with stiffness, undergo a low- $T$  phase transition. It is therefore interesting to look into the system size dependence at these temperatures, which is a computationally demanding task.

In this paper we study the behavior at the low- $T$  maximum of the specific heat for chains with up to  $N = 50$  monomers, by using the methods of simulated [15, 16, 17] and parallel [18, 19, 20, 21] tempering, which are found to be much more efficient than standard methods (see below). In this range of  $N$ , we find that the size dependence is very weak. There is no sign that the specific heat, or any other quantity, develops a singularity with increasing  $N$ . This holds both with and without the local interactions, although the crossover is somewhat more abrupt in the presence of these interactions.

We also compare homopolymer results to those for two non-uniform AB sequences; three  $N = 20$  sequences, the A homopolymer and two AB heteropolymers, are studied with and without local interactions. We find that the specific heat and entropy display a relatively weak sequence dependence over wide ranges of low  $T$ . The dependence of these thermodynamic quantities upon the local interactions is, by contrast, strong.

Our ability to simulate the low- $T$  behavior of these chains in a controlled way relies heavily on the efficiency of simulated and parallel tempering. In order to assess the efficiencies of these algorithms, we have performed careful tests for a chain with twelve monomers. We find that the two methods are very similar in efficiency, and that they are much better than a standard Monte Carlo.

The paper is organized as follows. Section 2 contains a brief description of the model and the two Monte Carlo algorithms, and a discussion of the algorithm tests. In Sec. 3 we present the results of our simulations. A summary is given in Sec. 4.

	Sequence					$T_f$
S	AAAA	BBAA	AABA	ABAA	ABBA	0.23
U	BAAA	AAAB	AAAA	BAAB	AABB	< 0.15

Table 1: Two AB sequences and their folding temperatures  $T_f$  for  $(\kappa_1, \kappa_2) = (-1, 0.5)$  [14].

## 2 Methods

### 2.1 The Model

The model studied [14] has two types of monomers, A (hydrophobic) and B (hydrophilic). The monomers are linked by rigid bonds of unit length to form linear chains. The positions of the monomers will be denoted by  $\mathbf{r}_i$  ( $i = 1, \dots, N$ ), and the (unit) bond vectors by  $\mathbf{b}_i = \mathbf{r}_{i+1} - \mathbf{r}_i$  ( $i = 1, \dots, N - 1$ ). The energy function is given by

$$E = \kappa_1 E_1 + \kappa_2 E_2 + E_{LJ}, \quad (1)$$

where

$$E_1 = - \sum_{i=1}^{N-2} \mathbf{b}_i \cdot \mathbf{b}_{i+1}, \quad (2)$$

$$E_2 = - \sum_{i=1}^{N-3} \mathbf{b}_i \cdot \mathbf{b}_{i+2}, \quad (3)$$

$$E_{LJ} = 4 \sum_{i=1}^{N-2} \sum_{j=i+2}^N \epsilon_{ij} \left( \frac{1}{r_{ij}^{12}} - \frac{1}{r_{ij}^6} \right). \quad (4)$$

The model has three parameters,  $\kappa_1$ ,  $\kappa_2$  and  $\epsilon_{ij}$ . The parameter  $\epsilon_{ij}$  sets the depth of the minimum of the nonlocal Lennard-Jones interactions. It is 1 for an AA pair and 1/2 for AB and BB pairs. The parameters  $\kappa_1$  and  $\kappa_2$  determine the strengths of the sequence independent local interactions. We will study the two choices  $(\kappa_1, \kappa_2) = (0, 0)$  and  $(\kappa_1, \kappa_2) = (-1, 0.5)$ . It has previously been shown [14] that there are AB sequences with stable native structures at relatively high temperatures for  $(\kappa_1, \kappa_2) = (-1, 0.5)$ , whereas there are no, or at least extremely few, such sequences for  $(\kappa_1, \kappa_2) = (0, 0)$ .

In this paper the main focus is on A homopolymers. In addition, we study two AB heteropolymers, denoted by S and U, which can be found in Table 1. The folding temperature, defined as the temperature where the dominance of the ground state sets in, is high for sequence S, and more typical for sequence U, for  $(\kappa_1, \kappa_2) = (-1, 0.5)$ .

Our  $(\kappa_1, \kappa_2) = (0, 0)$  A homopolymers are almost identical to the chains studied by Baumgärtner [22, 23] several years ago, except that the minimum of our Lennard-Jones potential is at  $r_{ij} = 2^{1/6}$  rather than 1. The model studied by Zhou *et al.* [12, 13], for fixed and flexible bond lengths, has a square-well potential in place of the Lennard-Jones potential.

## 2.2 Monte Carlo Methods

Our simulations are performed using simulated [15, 16, 17] and parallel [18, 19, 20, 21] tempering. The basic idea of both these methods is to create an enlarged configuration space. This means that a number of different temperatures are studied simultaneously, which in particular can be used as a tool for speeding up the evolution of the system at low temperatures.

In simulated tempering one simulates the joint probability distribution

$$P(r, k) \propto \exp[-g_k - E(r)/T_k], \quad (5)$$

where  $r = \{\mathbf{r}_1, \dots, \mathbf{r}_N\}$  denotes the chain conformation and  $k$  is a temperature index. This distribution contains two sets of parameters; the temperatures that the system is allowed to visit,  $\{T_k\}$ , and a set of tunable simulation parameters,  $\{g_k\}$ . The parameters  $g_k$  are free in the sense that averages corresponding to the desired Boltzmann distribution can be obtained directly, without reweighting, independently of these. However, the  $g_k$  parameters govern the marginal distribution in  $k$ , which can be written as

$$P(k) \propto \exp(-g_k - F_k/T_k), \quad (6)$$

where  $F_k$  is the free energy at  $T = T_k$ . Hence, to have good mobility in  $k$ , it is essential to make a careful choice of  $\{g_k\}$ . This is typically achieved by means of trial runs. The actual simulation of the distribution  $P(r, k)$  can be done by using ordinary separate updates of  $r$  and  $k$ .

Parallel tempering uses an even larger configuration space; the simulated distribution is

$$P(r_1, \dots, r_K) \propto \prod_{k=1}^K \exp[-E(r_k)/T_k], \quad (7)$$

where each  $r_k$  represents a complete chain conformation. The simulation is carried out by using two types of updates: ordinary, parallel updates of the different  $r_k$  and accept/reject controlled swaps  $r_k \leftrightarrow r_{k+1}$ . The latter update is the analog of the  $k$  update in simulated tempering. Compared to simulated tempering, this algorithm has two advantages. First, it is very easy to parallelize; and second, there are no  $g_k$  parameters to be tuned. These parameters are not needed since the  $k$  distribution is uniform by construction.

Using these two methods, we have studied chains with 12–400 monomers. Our simulations for  $N \leq 32$  were carried out on workstations using simulated tempering, whereas those for larger  $N$  were done on a parallel computer using parallel tempering. The temperatures  $T_k$  were chosen according to [24]

$$T_k = T_{\min} \left( \frac{T_{\max}}{T_{\min}} \right)^{(k-1)/(K-1)}, \quad (8)$$

where  $T_{\min} = T_1$  and  $T_{\max} = T_K$  denote the lowest and highest allowed temperatures, respectively. The parameters  $T_{\min}$ ,  $T_{\max}$  and  $K$  can be found in Table 2.

## 2.3 The Dynamics of the Algorithms

In order to compare the performances of simulated and parallel tempering, we carried out calculations with both methods for the  $N = 12$  A homopolymer with  $(\kappa_1, \kappa_2) = (-1, 0.5)$ . In these

$N$	$(\kappa_1, \kappa_2) = (0, 0)$			$(\kappa_1, \kappa_2) = (-1, 0.5)$		
	$T_{\min}$	$T_{\max}$	$K$	$T_{\min}$	$T_{\max}$	$K$
12	0.15	4	20	0.15	4	20
20	0.15	4	25	0.15	4	25
32	0.25	4	25	0.25	4	25
50	0.25	4	35	0.35	4	30
100	1.0	5	40			
200	1.5	4.5	35			
400	2.0	5	30			

Table 2: The parameters  $T_{\min}$ ,  $T_{\max}$  and  $K$  [see Eq. (8)].

calculations we used the same set of temperatures (see Table 2), and the same type of conformation update. We also performed a standard fixed- $T$  simulation at  $T = T_{\min} = 0.15$ , using the same conformation update.

In Fig. 1 we show run-time histories for the sum of all torsional angles along the chain,  $\alpha$ , from these three simulations. The quantity  $\alpha$  is convenient to study since it is known from the symmetry of the model that its probability distribution has to be symmetric,  $P(\alpha) = P(-\alpha)$ . When comparing the different runs in Fig. 1, the horizontal axis can be thought of as CPU time, since the cost of the additional updates used in simulated and parallel tempering is negligible. Therefore, it is evident from this figure that both simulated and parallel tempering are dramatic improvements on the standard method. Note that the system is trapped at negative  $\alpha$  during the whole fixed- $T$  simulation, whereas the true distribution  $P(\alpha)$  is symmetric.

Deciding which is best of simulated and parallel tempering is a more delicate task, which requires a careful analysis of the statistical errors. To get accurate estimates of the errors, we performed two additional runs, a factor of 10 longer than those in Fig. 1. The data from these long runs were analyzed by using the jackknife method [25], which is chosen because the number of visits of a given temperature is a random variable in simulated tempering. As a result, one does not have an unbiased estimator of, say,  $\alpha$  at temperature  $T = T_k$ , but rather one estimates the ratio

$$\langle \alpha \rangle_{T_k} = \frac{\langle \chi_k \cdot \alpha \rangle}{\langle \chi_k \rangle}, \quad (9)$$

where  $\langle \cdot \rangle$  without subscript denotes an average with respect to the joint distribution [see Eq. (5)], and  $\chi_k$  is 1 if  $T = T_k$  and 0 otherwise. For parallel tempering the jackknife procedure is equivalent to a standard block analysis.

The jackknife analysis was carried out for different bin sizes. For simulated tempering the dependence on bin size was weak. For parallel tempering, it was, by contrast, necessary to use relatively large bin sizes in order to get stable error estimates at low  $T$ . For this method, we also computed the autocorrelation function for  $\alpha$ . Consistent with the jackknife analysis, the exponential autocorrelation time was estimated to be approximately  $7 \times 10^4$  sweeps (same unit as in Fig. 1) at  $T = T_{\min}$ .

The final statistical errors for the two methods differed by less than 10%. Thus, the algorithms are very similar in efficiency for this system. This is in line with the findings of Hansmann and

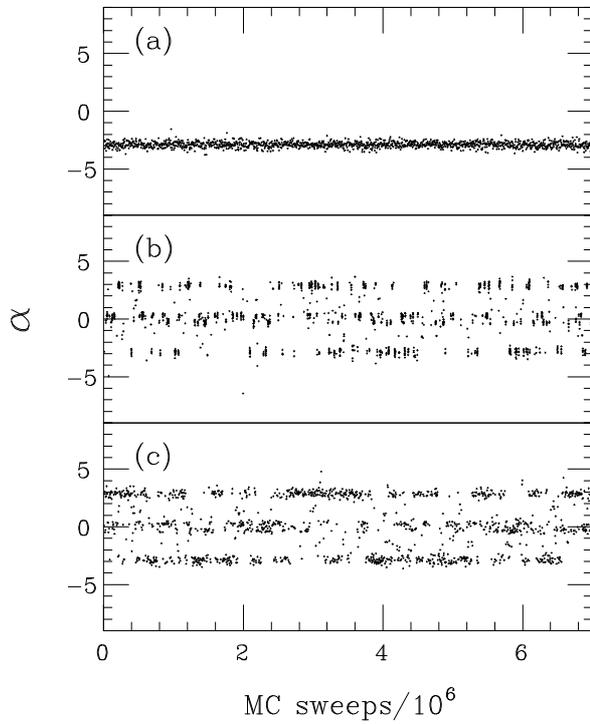


Figure 1: Evolution of the sum of all torsional angles,  $\alpha$ , in three different simulations of the  $N = 12$  A homopolymer for  $(\kappa_1, \kappa_2) = (-1, 0.5)$ . (a) Standard Monte Carlo, (b) simulated tempering, and (c) parallel tempering. All data shown correspond to  $T = T_{\min} = 0.15$ .

Okamoto [24], who measured the time needed to find the ground state in an all-atom model for a small peptide, with similar results for simulated and parallel tempering.

## 2.4 Measurements

Two important observables in our study of the low- $T$  behavior of these chains are the specific heat

$$C_v = \frac{\partial \langle E \rangle}{\partial T} = \frac{1}{T^2} (\langle E^2 \rangle - \langle E \rangle^2) \quad (10)$$

and the bond-bond correlation function

$$C_b(d) = \frac{1}{N-d} \sum_{i=1}^{N-d} \mathbf{b}_i \cdot \mathbf{b}_{i+d}, \quad (11)$$

which for  $d = 1$  and  $2$  can be written as  $C_b(d) = -E_d/(N-d)$  [see Eqs. (2) and (3)].

We estimate the collapse temperature  $T_\theta$  by using the radius of gyration  $r_g$ , defined by

$$r_g^2 = \frac{1}{N} \sum_{i=1}^N \left( \mathbf{r}_i - \frac{1}{N} \sum_{j=1}^N \mathbf{r}_j \right)^2. \quad (12)$$

The ratio  $\langle r_g^2 \rangle / N$  tends to zero for fixed  $T < T_\theta$ , and to infinity for fixed  $T > T_\theta$ , as  $N \rightarrow \infty$ . As is well-known, this means that  $T_\theta$  can be estimated from the intersections of curves corresponding to different  $N$ , in a plot of  $\langle r_g^2 \rangle / N$  vs  $T$ .

In our comparison of different  $N = 20$  chains, we calculate their entropies  $S$ . The simulated-tempering algorithm is well suited for this purpose, since free-energy differences can be directly obtained from the marginal probabilities  $P(k)$  [15]. Using Eq. (6), one finds that

$$S_k - S_l = g_k - g_l + \ln \frac{P(k)}{P(l)} + \frac{\langle E \rangle_k}{T_k} - \frac{\langle E \rangle_l}{T_l}, \quad (13)$$

where  $S_{k,l}$  and  $\langle E \rangle_{k,l}$  denote the entropy and average energy, respectively, at  $T = T_{k,l}$ . This expression is handy since the parameters  $g_k$  are constants and the average energies are relatively easy to measure. The remaining term,  $\ln P(k)/P(l)$ , is just a small correction if the parameters  $g_k$  are well chosen.

Equation (13) can be used to obtain the temperature dependence for a given chain. In order to compare the entropies of two different chains [different sequences and/or different  $(\kappa_1, \kappa_2)$ , but the same  $N$ ], we use the umbrellalike [26] formula

$$S' - S = \frac{\langle E' \rangle' - \langle E \rangle}{T} + \ln \left\langle \exp \left( -\frac{E' - E}{T} \right) \right\rangle, \quad (14)$$

where the prime is used to distinguish the two systems. In Eq. (14) one could, of course, replace  $\langle E' \rangle'$  by expectation values referring to the system without prime.

All statistical errors quoted below are  $1\sigma$  errors, obtained by the jackknife method [25].

### 3 Results

We have performed simulations of the A homopolymers of the model defined in Sec. 2.1 for  $(\kappa_1, \kappa_2) = (0, 0)$  and  $(-1, 0.5)$ . In Fig. 2 we show the temperature dependence of the specific heat for different chain lengths. For both choices of  $(\kappa_1, \kappa_2)$ , it can be seen that the specific heat exhibits a pronounced peak at low temperatures, which could signal a phase transition. The height of the peak depends, however, only weakly on  $N$ . The  $N$  dependence is stronger at higher temperatures where another peak seems to develop. This shoulder or peak can most probably be associated with the collapse transition at  $T_\theta$ . In Sec. 3.1 we estimate  $T_\theta$  by using data for the radius of gyration. We then return to the  $N$  dependence at low temperatures in Sec. 3.2. Finally, in Sec. 3.3, we compare the behaviors of three different  $N = 20$  sequences, the A homopolymer and the two AB heteropolymers in Table 1.

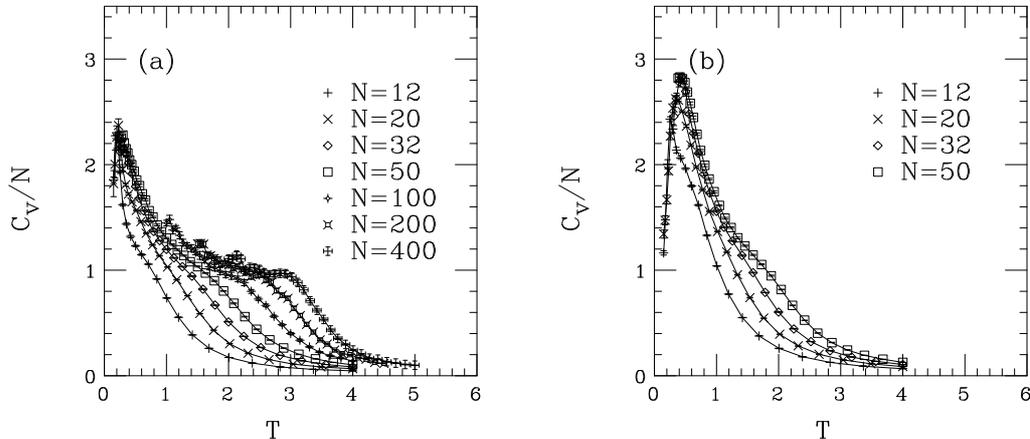


Figure 2: Temperature dependence of the specific heat for A homopolymers for (a)  $(\kappa_1, \kappa_2) = (0, 0)$  and (b)  $(\kappa_1, \kappa_2) = (-1, 0.5)$ .

### 3.1 Estimation of $T_\theta$

In order to estimate  $T_\theta$ , it is easier to use the radius of gyration than the specific heat. In Fig. 3a we show the ratio  $\langle r_g^2 \rangle / N$  against  $T$  for different  $N$  for  $(\kappa_1, \kappa_2) = (0, 0)$ . In the temperature range relevant for the  $T_\theta$  estimation, we studied chains with up to 400 monomers (simulations are much easier here than at low  $T$ ). The temperature at which the curves corresponding to the sizes  $N$  and  $2N$  intersect,  $T_N$ , was estimated to be  $3.76 \pm 0.11$ ,  $4.05 \pm 0.06$  and  $4.33 \pm 0.08$  for  $N = 50$ , 100 and 200, respectively. A least-square fit of these data points to the mean-field form

$$T_N = T_\theta + \frac{\text{const.}}{\sqrt{N}} \quad (15)$$

yields  $T_\theta = 4.88 \pm 0.18$ . The systematic uncertainty in this extrapolation is difficult to estimate without data for larger  $N$ . Nevertheless, this analysis convincingly shows that  $T_\theta$  is indeed located well above the low- $T$  maximum of the specific heat.

In Fig. 3b we show the same plot for  $(\kappa_1, \kappa_2) = (-1, 0.5)$ . All data are for  $N \leq 50$ . They are very similar to those obtained for  $(\kappa_1, \kappa_2) = (0, 0)$  at these  $N$ . In particular, it is clear that the collapse transition occurs well above the low- $T$  maximum of the specific heat in this case too.

### 3.2 The Low- $T$ Behavior

We now turn to the behavior around the low- $T$  maximum of the specific heat. Figure 4, which is an enlargement of Fig. 2, shows the specific heat itself in the vicinity of the maximum. A linear scaling with  $N$  of the height of the peak in  $C_v/N$  would have implied a first-order phase transition. The data show a weak  $N$  dependence. The peak grows very slowly with  $N$  for  $(\kappa_1, \kappa_2) = (-1, 0.5)$ , and not at all for  $(\kappa_1, \kappa_2) = (0, 0)$ . In neither case is there any sign that  $C_v/N$  diverges. We also studied the energy distributions at the maxima. Consistent with the specific heat analysis, these were found to have a single peak, without any trace of bimodality.

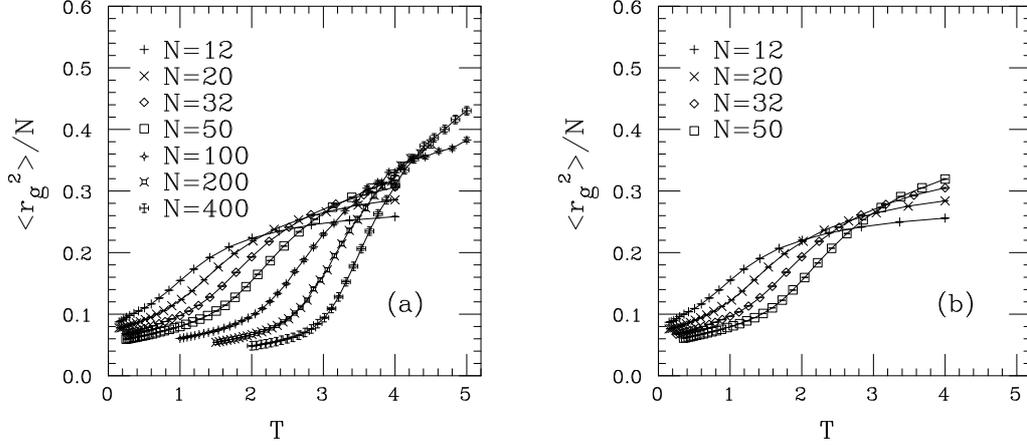


Figure 3: Temperature dependence of  $\langle r_g^2 \rangle / N$  for A homopolymers for (a)  $(\kappa_1, \kappa_2) = (0, 0)$  and (b)  $(\kappa_1, \kappa_2) = (-1, 0.5)$ .

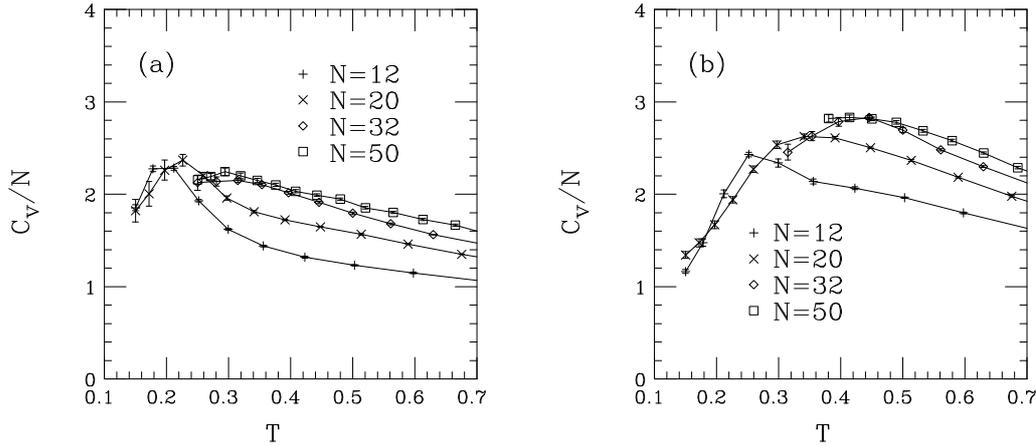


Figure 4: Low- $T$  behavior of the specific heat for A homopolymers for (a)  $(\kappa_1, \kappa_2) = (0, 0)$  and (b)  $(\kappa_1, \kappa_2) = (-1, 0.5)$ .

The size of the chains, as measured by the radius of gyration, changes very little around the low- $T$  maximum of the specific heat (see Fig. 3). The size is similar for the two choices of  $(\kappa_1, \kappa_2)$ , and very small at these temperatures. The bond-bond correlation  $C_b(d)$  [see Eq. (11)] provides information about the local structure of the chains. Unlike the radius of gyration, this quantity is strongly  $(\kappa_1, \kappa_2)$  dependent. This can be seen from Fig. 5, which shows  $C_b(d)$  against  $T$  for  $d = 2$ . The behavior of  $C_b(2)$  for  $(\kappa_1, \kappa_2) = (-1, 0.5)$  shows that strong regularities develop in the local structure at low  $T$ . The change with decreasing  $T$  is, however, gradual, and does not become more abrupt with increasing  $N$ . In the absence of local interactions,  $(\kappa_1, \kappa_2) = (0, 0)$ , the correlations  $C_b(d)$  remain weak at low  $T$ . For small  $N$  there are some irregularities at very low  $T$ , but these disappear with increasing  $N$ .

Let us compare these results with findings for lattice chains. As mentioned in the introduction, it

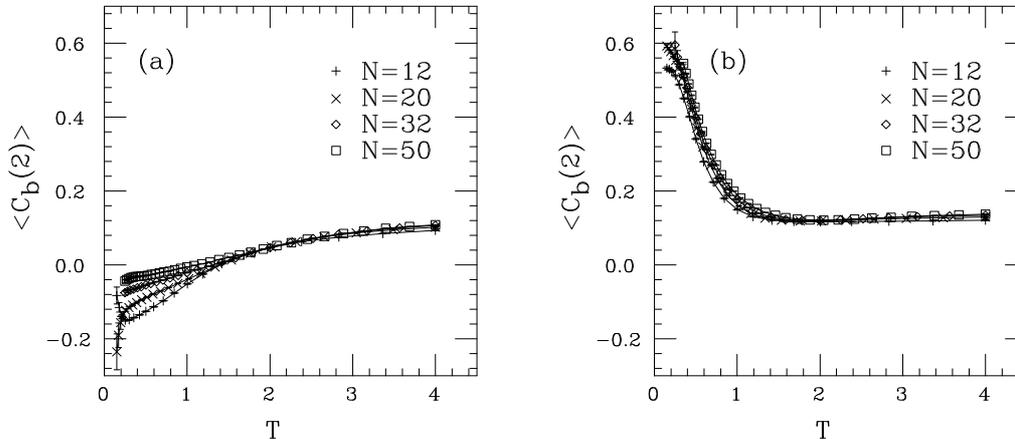


Figure 5: Temperature dependence of  $C_b(2)$  for A homopolymers for (a)  $(\kappa_1, \kappa_2) = (0, 0)$  and (b)  $(\kappa_1, \kappa_2) = (-1, 0.5)$ .

has been found [7, 8, 9, 10] that lattice chains with stiffness exhibit a transition to an ordered low- $T$  phase. Our  $(\kappa_1, \kappa_2) = (-1, 0.5)$  chains, which have a certain degree of stiffness, do show strong regularities in the local structure at low  $T$ , as can be seen from the behavior of the correlation  $C_b(d)$ . Also, the peak in the specific heat is more pronounced for these chains than for the  $(\kappa_1, \kappa_2) = (0, 0)$  chains. However, we do not find any sign that a phase transition takes place. Although the  $N$  range probed is comparatively small, this indicates a possible difference between lattice and off-lattice chains. It would be interesting to systematically study how the off-lattice results depend on the degree of stiffness of the chains, but this is beyond the scope of the present paper.

The low- $T$  behavior of off-lattice models similar to ours has been studied before. As mentioned earlier, Baumgärtner [22, 23] studied a model very closely related to our  $(\kappa_1, \kappa_2) = (0, 0)$  model. Here the low- $T$  peak in the specific heat was associated with a probably first-order phase transition. As we have seen, our results provide no support for the existence of such a transition. Zhou *et al.* [12, 13] studied a model with square-well potentials, both for fixed and flexible bond lengths. For flexible bond lengths, evidence in terms of bimodal energy distributions was found for a first-order low- $T$  phase transition. This contrasts sharply with the behavior of our chains, which, as mentioned above, have unimodal energy distributions.

### 3.3 Heteropolymers

In this section we compare the thermodynamic behaviors of the  $N = 20$  A homopolymer and the two  $N = 20$  heteropolymers S and U in Table 1. Each of the three sequences is studied both for  $(\kappa_1, \kappa_2) = (0, 0)$  and  $(-1, 0.5)$ , which gives us a total of six different chains. The temperatures studied range between 0.15 and 4. One of the chains, sequence S with  $(\kappa_1, \kappa_2) = (-1, 0.5)$ , has a folding temperature  $T_f \approx 0.23 > 0.15$  [14] and is referred to as the stable one, whereas  $T_f < 0.15$  for the others.

In Fig. 6a we show the specific heat, which has a peak at low  $T$  for all the six chains. Note that the

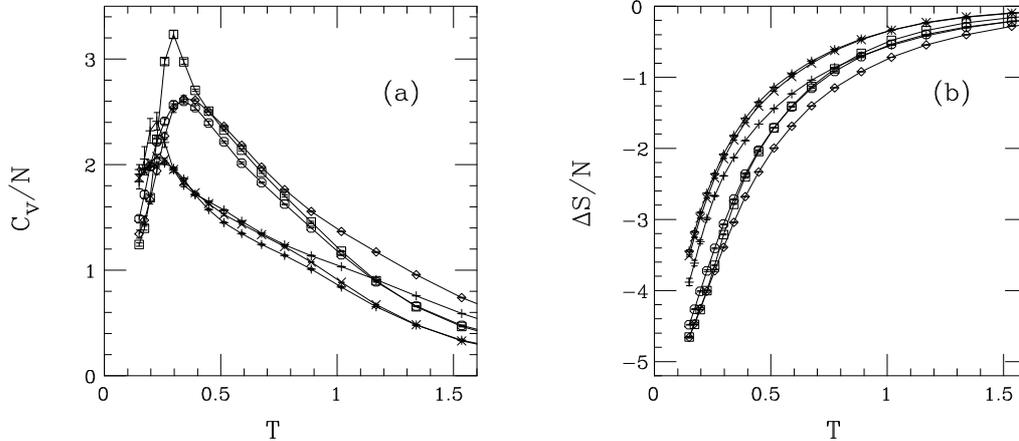


Figure 6: (a) Specific heat and (b) entropy against temperature for the following six  $N = 20$  chains: A homopolymer,  $(\kappa_1, \kappa_2) = (0, 0)$  (+); sequence S,  $(\kappa_1, \kappa_2) = (0, 0)$  ( $\times$ ); sequence U,  $(\kappa_1, \kappa_2) = (0, 0)$  (fancy +); A homopolymer,  $(\kappa_1, \kappa_2) = (-1, 0.5)$  ( $\diamond$ ); sequence S,  $(\kappa_1, \kappa_2) = (-1, 0.5)$  ( $\square$ ); sequence U,  $(\kappa_1, \kappa_2) = (-1, 0.5)$  ( $\circ$ ).

peak is highest for the stable chain, and that the maximum is at  $T \approx 0.30$  for this chain, which is fairly close to its folding temperature. The perhaps most interesting part of this figure is, however, between  $T \approx 0.4$  and  $T \approx 0.9$ . Here the data for the different chains approximately collapse onto two curves, each corresponding to a fixed  $(\kappa_1, \kappa_2)$ . Thus, the dependence on sequence is weak over this fairly wide range of temperature.

We also calculated the entropies of the chains, using the method described in Sec. 2.4, which is convenient since it does not require numerical integration of any observable. In Fig. 6b we show  $\Delta S = S - S_0$ , where  $S_0$  is chosen as the entropy of the  $(\kappa_1, \kappa_2) = (-1, 0.5)$  A homopolymer at  $T = 4$ . Thus, the constant  $S_0$  is the same for all the chains. From the figure it can be seen that the chains with  $(\kappa_1, \kappa_2) = (-1, 0.5)$  have lower entropies than those with  $(\kappa_1, \kappa_2) = (0, 0)$  at low  $T$ . This is not unexpected since the  $(\kappa_1, \kappa_2) = (-1, 0.5)$  chains have stiffness. What still may seem somewhat surprising is that the stable chain, below its folding temperature, does not have lower entropy than the A homopolymer with the same  $(\kappa_1, \kappa_2)$ . This example underlines that the entropy is largely determined by local features at low  $T$ . The entropy does not tell whether the structure is globally stable or not (which it does in lattice models).

## 4 Summary

We have studied isolated off-lattice homopolymers with  $[(\kappa_1, \kappa_2) = (-1, 0.5)]$  and without  $[(\kappa_1, \kappa_2) = (0, 0)]$  local interactions. In both cases the specific heat exhibits a peak at low  $T$ . The peak is most pronounced for  $(\kappa_1, \kappa_2) = (-1, 0.5)$ . Also, the local structure of the chains develops strong regularities at low  $T$  for this choice of  $(\kappa_1, \kappa_2)$ . The dependence on system size is, however, weak at these temperatures; our results, obtained for  $N \leq 50$ , show no sign that a singularity develops with increasing  $N$ .

We have also compared homopolymer results to those for two non-uniform sequences, which were deliberately chosen to represent different types of behavior. Nevertheless, the specific heat and entropy show a relatively weak sequence dependence over wide ranges of low  $T$ . The dependence of these thermodynamic quantities upon the local interactions is, by contrast, strong.

Our calculations have been performed by using simulated and parallel tempering. Our tests show that these two methods are indeed much more efficient than standard methods in the difficult low- $T$  regime. The difference in efficiency between simulated and parallel tempering was found to be small.

*Note added in proof.* After this manuscript was completed, we became aware of two studies [27, 28] of  $(\kappa_1, \kappa_2) = (0, 0)$  A homopolymers. Grassberger and Hegger [28] estimated  $0.2 \leq 1/T_\theta \leq 0.23$ , which is in perfect agreement with our estimate.

## Acknowledgements

This work was supported by the Swedish Foundation for Strategic Research, by the Swedish Council for Planning and Coordination of Research (FRN), and by Paralleldatorcentrum (PDC), Royal Institute of Technology.

## References

- [1] P.G. de Gennes, *Scaling Concepts in Polymer Physics* (Cornell University Press, Ithaca, 1988).
- [2] P.G. de Gennes, “Collapse of a Polymer Chain in Poor Solvents”, *J. Phys. (France) Lett.* **36**, L55 (1975).
- [3] A.Yu. Grosberg and D.V. Kuznetsov, “Quantitative Theory of the Globule-to-Coil Transition. 1. Link Density Distribution in a Globule and Its Radius of Gyration”, *Macromolecules* **25**, 1970 (1992).
- [4] P. Grassberger and R. Hegger, “Simulations of three-dimensional  $\theta$  polymers”, *J. Chem. Phys.* **102**, 6881 (1995).
- [5] P. Grassberger, “Pruned-Enriched Rosenbluth Method: Simulations of  $\theta$  Polymers of Chain Length up to 1 000 000”, *Phys. Rev. E* **56**, 3682 (1997).
- [6] I.M. Lifshitz, A.Yu. Grosberg and A.R. Khokhlov, “Some Problems of the Statistical Physics of Polymer Chains with Volume Interaction”, *Rev. Mod. Phys.* **50**, 683 (1978).
- [7] A. Kolinski, J. Skolnick and R. Yaris, “Monte Carlo Simulations on an Equilibrium Globular Protein Folding Model”, *Proc. Natl. Acad. Sci. USA* **83**, 7267 (1986).
- [8] S. Doniach, T. Garel and H. Orland, “Phase Diagram of a Semiflexible Polymer Chain in a  $\theta$  Solvent: Application to Protein Folding”, *J. Chem. Phys.* **105**, 1601 (1996).
- [9] U. Bastolla and P. Grassberger, “Phase Transitions of Single Semi-Stiff Polymer Chains”, *J. Stat. Phys.* **89**, 1061 (1997).
- [10] J.P.K. Doye, R.P. Sear and D. Frenkel, “The Effect of Chain Stiffness on the Phase Behaviour of Isolated Homopolymers”, *J. Chem. Phys.* **108**, 2134 (1998).
- [11] P.J. Flory, “Statistical Thermodynamics of Semi-Flexible Chain Molecules”, *Proc. Roy. Soc. A* **234**, 60 (1956).
- [12] Y. Zhou, C.K. Hall and M. Karplus, “First-Order Disorder-to-Order Transition in an Isolated Homopolymer Model”, *Phys. Rev. Lett.* **77**, 2822 (1996).
- [13] Y. Zhou, M. Karplus, J.M. Wichert and C.K. Hall, “Equilibrium Thermodynamics of Homopolymers and Clusters: Molecular Dynamics and Monte Carlo Simulations of Systems with Square-Well Interactions”, *J. Chem. Phys.* **107**, 10691 (1997).
- [14] A. Irbäck, C. Peterson, F. Potthast and O. Sommelius, “Local Interactions and Protein Folding: A Three-Dimensional Off-Lattice Approach”, *J. Chem. Phys.* **107**, 273 (1997).
- [15] A.P. Lyubartsev, A.A. Martsinovski, S.V. Shevkunov and P.N. Vorontsov-Velyaminov, “New Approach to Monte Carlo Calculation of the Free Energy: Method of Expanded Ensembles”, *J. Chem. Phys.* **96**, 1776 (1992).
- [16] E. Marinari and G. Parisi, “Simulated Tempering: A New Monte Carlo Scheme”, *Europhys. Lett.* **19**, 451 (1992).
- [17] A. Irbäck and F. Potthast, “Studies of an Off-Lattice Model for Protein Folding: Sequence Dependence and Improved Sampling at Finite Temperature”, *J. Chem. Phys.* **103**, 10298 (1995).

- [18] C.J. Geyer and E.A. Thompson, “Annealing Markov Chain Monte Carlo with Applications to Pedigree Analysis”, University of Minnesota preprint (1993).
- [19] M.C. Tesi, E.J. Janse van Rensburg, E. Orlandini and S.G. Whittington, “Monte Carlo Study of the Interacting Self-Avoiding Walk Model in Three Dimensions”, *J. Stat. Phys.* **82**, 155 (1996).
- [20] K. Hukushima and K. Nemoto, “Exchange Monte Carlo Method and Application to Spin Glass Simulations”, *J. Phys. Soc. (Jap)* **65**, 1604 (1996).
- [21] U.H.E. Hansmann, “Parallel Tempering Algorithm for Conformational Studies of Biological Molecules”, *Chem. Phys. Lett.* **281**, 140 (1997).
- [22] A. Baumgärtner, “Statics and Dynamics of the Freely Jointed Polymer Chain with Lennard-Jones Interaction”, *J. Chem. Phys.* **72**, 871 (1980).
- [23] A. Baumgärtner, “Glass Transition of a Polymer Chain”, *J. Chem. Phys.* **73**, 2489 (1980).
- [24] U.H.E. Hansmann and Y. Okamoto, “Numerical Comparisons of Three Recently Proposed Algorithms in the Protein Folding Problem”, *J. Comput. Chem.* **18**, 920 (1997).
- [25] R.G. Miller, “The Jackknife — A Review”, *Biometrika* **61**, 1 (1974).
- [26] G.M. Torrie and J.P. Valleau, “Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling”, *J. Comput. Phys.* **23**, 187 (1977).
- [27] D. Frenkel, G.C.A.M. Mooij and B. Smit, “Novel Scheme to Study Structural and Thermal Properties of Continuously Deformable Molecules”, *J. Phys.: Condens. Matter* **4**, 3053 (1992).
- [28] P. Grassberger and R. Hegger, “Monte Carlo Simulations of Off-Lattice Polymers”, *J. Phys.: Condens. Matter* **7**, 3089 (1995).