Finite-size scaling of helix–coil transitions in poly-alanine studied by multicanonical simulations

Ulrich H. E. Hansmann, and Yuko Okamoto

Citation: J. Chem. Phys. **110**, 1267 (1999); doi: 10.1063/1.478169
View online: https://doi.org/10.1063/1.478169
View Table of Contents: http://aip.scitation.org/toc/jcp/110/2
Published by the American Institute of Physics

Finite-size scaling of helix–coil transitions in poly-alanine studied by multicanonical simulations

Ulrich H. E. Hansmann[a)]
Department of Physics, Michigan Technological University, Houghton, Michigan 49931-1291

Yuko Okamoto[b)]
Department of Theoretical Studies, Institute for Molecular Science, Okazaki, Aichi 444-8585, Japan

(Received 15 July 1998; accepted 6 October 1998)

We report results from multicanonical simulations of poly-alanine. Homopolymers of up to 30 amino acids were considered and various thermodynamic quantities as a function of temperature calculated. We study the nature of the observed helix–coil transition and present estimates for critical exponents. © *1999 American Institute of Physics.* [S0021-9606(99)50702-2]

INTRODUCTION

It is widely believed that the folding of a protein into its native structure involves one or more transitions between distinct phases. However, these transitions are not fully understood. Are they simply crossovers between conformers or is it justified to refer to them (as commonly done) as phase transitions? Strictly speaking, the latter concept is only well defined for infinite systems. It is not clear whether proteins can be regarded as ''almost'' infinitely large systems (in the sense in which the 10^{23} atoms in a crystal form an ''infinite'' system). The properties of proteins depend strongly on the number and composition of amino acids and therefore follow not from the collective behavior of a great many similar unit objects. For the same reason, it is also not possible to study the transition by considering extrapolation of a given protein to an infinitely long chain: The properties of the protein may change completely by adding or substracting an amino acid. Hence, finite-size scaling analyses, a common tool to study phase transitions, is not applicable. The exceptions are homopolymers of amino acids. Here, an extrapolation to infinitely long chains is possible and it should indeed be possible to describe for these molecules configurational changes within the framework of phase transitions. Hence, homopolymers of amino acids are suitable models for investigation of how the observed transitions are affected by the finite size of the molecule.

In an earlier work[1] we studied, for three characteristic amino acids (alanine, valine, and glycine), α-helix formation in short peptide systems and compared our results with those of recent experiments (which are reviewed in Ref. 2). The helix–coil transition was also studied theoretically by various other groups.[3–12] For poly-alanine we observed a sharp transition between disordered coil conformers and an ordered phase where the polymer is in a helical state. It is tempting to describe the poly-alanine chain in the framework of Zimm–Bragg-type theories[13] in which the homopolymers are described by a one-dimensional Ising model with the residues as ''spins'' taking values ''helix'' or ''coil.'' By construction, the length of helical segments cannot grow indefinitely in such models, since long-range order is not possible for a one-dimensional system with solely short-range interactions. Hence, phase transitions are not possible in the Zimm–Bragg model. However, such models may be too crude a description. In the helix–coil transition, the polymer chain changes from a disordered three-dimensional coil to an ordered and only quasi one-dimensional helix. This is certainly different from spin-flips in a one-dimensional chain of Ising spins. In addition, the amino-acid residues in a homopolymer are also subject to long-range (electrostatic) interactions which allow long-range order and phase transitions even for the case of a one-dimensional spin chain. Hence, it is an open question whether the observed helix–coil transition in poly-alanine is a phase transition or whether it is rather a crossover between the two states. In order to find an answer to this question, we can study finite chains and extrapolate the results to the limit of an infinitely long polymer. The limitations of our previous results did not allow such analyses. Here, we report results from a simulation of poly-alanine with both increased statistics and larger chains.

As in the previous article, the use of the multicanonical algorithm[14] was crucial. The various competing interactions within the polymer lead to an energy landscape characterized by a multitude of local minima. Hence, in the low-temperature region, canonical Monte Carlo or molecular dynamics simulations will tend to get trapped in one of these minima and the simulation will not thermalize within the available CPU time. To overcome this problem, we have proposed the application of the multicanonical algorithm[14] and other generalized-ensemble techniques[15] to the protein folding problem.[16] We could demonstrate that these techniques are indeed superior to standard methods[17,18] and a useful tool for investigations of the proteins and peptides,[19–21] since they allow one not only to find the ground-state conformer but also to calculate thermodynamic quantities at various temperatures from one simulation run. In the latter point, our approach differs from global optimization methods which were also used in previous works of protein folding.[22,23]

a)Author to whom correspondence should be addressed; electronic mail: hansmann@mtu.edu
b)Electronic mail: okamotoy@ims.ac.jp

Here we use the multicanonical algorithm to calculate various thermodynamic quantities as a function of temperature for poly-alanine of four different chain lengths. We concentrate on such quantities (like average number of helical residues or specific heat) where we expect to see the strongest signal for the helix–coil transition. The finite-size scaling of these quantities is studied and estimates for critical exponents are calculated. This allows us to investigate the nature of the phase transition in poly-alanine. Finally, we update our results on the Zimm-Bragg parameters s and σ.

METHODS

Peptide preparation and potential energy function

We considered amino-acid homo-oligomers of alanine, which are known from experiments and our previous work to be a strong helix former at low temperatures. The number of residues, N, was taken to be 10, 15, 20, and 30 in order to examine the N dependence. Since the charges at peptide termini are known to reduce helix content,[24,25] we removed them by taking a neutral NH_2 group at the N-terminus and a neutral -COOH group at the C-terminus.

The potential energy function $E_{\rm tot}$ (in kcal/mol) that we used is given by the sum of the electrostatic term E_C, 12-6 Lennard-Jones term $E_{\rm LJ}$, and hydrogen-bond term $E_{\rm HB}$ for all pairs of atoms in the peptide, together with the torsion term $E_{\rm tor}$ for all torsion angles,

$$E_{\rm tot}=E_C+E_{\rm LJ}+E_{\rm HB}+E_{\rm tor}, \tag{1}$$

$$E_C=\sum_{(i,j)} \frac{332 q_i q_j}{\epsilon r_{ij}}, \tag{2}$$

$$E_{\rm LJ}=\sum_{(i,j)} \left(\frac{A_{ij}}{r_{ij}^{12}}-\frac{B_{ij}}{r_{ij}^{6}}\right), \tag{3}$$

$$E_{\rm HB}=\sum_{(i,j)} \left(\frac{C_{ij}}{r_{ij}^{12}}-\frac{D_{ij}}{r_{ij}^{10}}\right), \tag{4}$$

$$E_{\rm tor}=\sum_{l} U_l(1\pm\cos(n_l\chi_l)). \tag{5}$$

Here, r_{ij} (in Å) is the distance between the atoms i and j, and χ_l is the torsion angle for the chemical bond l. The factor 332 in E_C is to give the energy in units of kcal/mol. The parameters (q_i, A_{ij}, B_{ij}, C_{ij}, D_{ij}, U_l, and n_l) for the energy function were adopted from ECEPP/2.[26–28] Note that in ECEPP/2, irrelevant constant terms (that do not depend on conformations of the peptide) are already subtracted from $E_{\rm tot}$. Since one can avoid the complications of electrostatic and hydrogen-bond interactions of side chains with the solvent for nonpolar amino acids, explicit solvent molecules were neglected for simplicity and the dielectric constant ϵ was set equal to 2. The computer code KONF90[29] was used which differs slightly in conventions from the original version of ECEPP/2 (for example, ϕ_1 of KONF90 is equal to $\phi_1 - 180°$ of ECEPP/2, and energies are also different by small irrelevant constant terms). The peptide-bond dihedral angles ω were fixed at the value 180° for simplicity, which leaves ϕ_i, ψ_i, and χ_i ($i=1,...,N$) as independent degrees of freedom. Since alanine has only one χ angle in the side chain, the numbers of independent degrees of freedom are $3N$, where N is the number of residues.

Simulation techniques

The Monte Carlo method that we used is the multicanonical algorithm,[14] which is sometimes also referred to as entropic sampling[30] (in Ref. 31 it was shown that both algorithms are mathematically identical). Unlike in canonical simulations, configurations with energy E are here assigned a weight,

$$w_{mu}(E)\propto\frac{1}{n(E)}=e^{-S(E)}, \tag{6}$$

where

$$S(E)=\log n(E) \tag{7}$$

is the microcanonical entropy. A standard update scheme such as the Metropolis algorithm[32] will realize a Markov chain in that ensemble and yield a uniform distribution of energy

$$P_{mu}(E)\propto n(E)w_{mu}(E)=\text{const.} \tag{8}$$

All energies appear with equal probability and a free random walk in the energy space is enforced. Hence, the simulation can overcome any energy barrier and will not get trapped in one of the many local minima. Since a large range of energies is sampled, the use of the reweighting techniques[33] allows us to calculate the expectation value of any physical quantity $\mathcal{O}$ for a wide range of temperatures T by

$$\langle\mathcal{O}\rangle_T=\frac{\int dE\mathcal{O}(E)P_{mu}(E)w_{mu}^{-1}(E)e^{-\beta E}}{\int dE P_{mu}(E)w_{mu}^{-1}(E)e^{-\beta E}}. \tag{9}$$

To further improve sampling, we used from time to time an improved update, the cut-and-paste move, which is inspired by the Lin–Kernighan update[34] for the traveling salesman problem. Here, a new polymer conformation is proposed by randomly choosing a segment of $m\leqslant N/2$ consecutive residues, ''removing'' it and ''replacing'' it between another randomly chosen part of the polymer chain. The energy of the new conformation is calculated, and this conformation is accepted or rejected according to the usual Metropolis criterion so that detailed balance is satisfied. The advantage of the new update is that it changes not only locally a single dihedral angle, but globally the whole polymer chain, and in this way enhances thermalization. We found that while the CPU time increased moderately by about $\approx$10%, the gain in efficiency was much larger and increased with chain length. It is obvious that the new update can be used only for simulation of homopolymers where the number and type of dihedral angles are the same for each residue.

Computational details

One Monte Carlo (MC) sweep updates every dihedral angle (in both the backbone and the side chain) of the poly-alanine chain once. After each MC sweep we tried the global ''cut-and-paste''-update once.

TABLE I. Tunneling times τ_{tun} measured in MC sweeps as obtained from multicanonical simulations of poly-alanine with length $N=10$, 15, 20, and 30.

N	τ_{tun} [MC sweeps]
10	16 440 (3498)
15	22 964 (3157)
20	92 872 (31710)
30	30 8104 (24254)

The multicanonical weight factors were determined by the iterative procedure described in Refs. 1 and 17. We needed between 40 000 sweeps (for $N=10$) and 500 000 sweeps (for $N=30$) for their calculation. All thermodynamic quantities were then calculated from one production run of Nsw Monte Carlo sweeps following additional 10 000 sweeps for equilibrization. We chose $N_{sw}=200\,000$ for $N=10$, $N_{sw}=250\,000$ for $N=15$, $N_{sw}=500\,000$ for $N=20$, and $N_{sw}=1\,000\,000$ for $N=30$. In all cases, each simulation started from a completely random initial conformation.

After each fourth sweep, we stored the total energy E_{tot}, the component energies E_C, E_{LJ}, E_{HB}, E_{tor}, the number of helical residues n_H, and the number of helical segments n_S of the configuration for further analyses. The latter two quantities were calculated in the following way: We consider that a residue is in a (right-handed) α-helix configuration when the dihedral angles (ϕ,ψ) fall in the range $(-70\pm 20°, -37\pm 20°)$. The length ℓ of a helical segment is then defined by the number of successive residues which are in the α-helix configuration. We define a conformation as helical if it has a helical segment with $\ell \geqslant 3$, since a helix has to have more than three consecutive residues in a helical state to form the characteristic hydrogen bond. The conformation is regarded as completely helical if $\ell \geqslant N-2$, since the two end residues are flexible. Similarly, a conformation is considered to be in a coil state if there exists no helical segment with $\ell \geqslant 3$. The number n_H of helical residues in a conformation is defined by the sum of ℓ over all n_S helical segments in the conformation. The same definition was used in our previous work.[1]

The total number of MC sweeps N_{sw} were chosen so that the number of ''tunneling events'' $N_{\text{tun}} \geqslant 2$. A tunneling event is defined as a series of MC sweeps in which the poly-alanine molecule changes from a completely helical state (lowest-energy state) to a coil state (high-energy state) and *back*. Hence, the value of N_{tun} gives a lower bound on the number of independent low-energy states found in the simulation. To further quantify how the numerical effort of the multicanonical simulation of our system increases with the number of residues, we define as ''tunneling time'' τ_{tun} the average time (in MC sweeps) needed for a tunneling event and display its values in Table I. The tunneling time τ_{tun}, and therefore the numerical effort, increases with the size of the molecule as $\tau_{\text{tun}} \propto N^x$ where $x=3.1(2)$ when the time is measured in MC sweeps and $x=4.1(2)$ when measured in updates (since each MC sweep consists of N updates). While this exponent is large, the numerical performance is still much better than with a canonical algorithm where one would expect a supercritical slowing down, i.e., $\tau_{\text{tun}} \propto e^{aN}$ (a is an unknown constant). We remark that the above exponent is close to the optimal value for the multicanonical algorithm, since the energy grows as $E \approx N^2$ with the number of amino acids. Hence, the numerical effort in our simulations grows like $\approx E^2$, which is what one expects for a random walk in energy. Similar results were found for spin systems. For these systems it is also known that the numerical effort for calculating the multicanonical weights scales like $\approx E^{2.5}$.[35] While we did not systematically investigate this question, we expect that the same relation holds also in our case.

RESULTS AND DISCUSSION

In order to elucidate the efficiency of sampling, we first show in Fig. 1 four conformations obtained during the simulation of poly-alanine with length $N=30$. The four structures were chosen from a span of half of a tunneling event that moved between the 100 000th MC sweeps and the 156 000th MC sweep from a very high-energy state [Fig. 1(a)] to a lowest-energy conformer [Fig. 1(d)]. The conformation changes from a random coil to a completely helical conformation. One sees that our simulation indeed covers a large portion of the conformational space.

In order to study helix–coil transitions in our homopolymers, it is necessary to first define an order parameter which allows us to distinguish between the two phases. Such a quantity can be easily constructed from the number n_H of helical residues as follows:

$$q=\frac{\tilde{n}_H}{N-2}, \tag{10}$$

where $\tilde{n}_H$ is the number of helical residues in a conformation, however, without counting the first and last residues. We chose this definition in order to have $q=1$ for a completely helical conformation. Since the residues at the end of the polymer chain can move freely, they will not be part of a helical segment and therefore should not be counted. On the other hand, $q=0$ indicates, with our definition, a coil conformation without any helical segments. We display $\langle q \rangle_T$ for various chain length in Fig. 2. The transition between a high-temperature phase characterized by coil states ($\langle q \rangle \approx 0$) and a low-temperature phase where helical conformations are dominant ($q \approx 1$) is obvious, and becomes sharper with increasing length of the poly-alanine chain.

The helix–coil transition is correlated with a change in the average potential energy. This can be seen from Figs. 3(a)–3(d) where we show for all four homopolymers the average total energy density $\langle E_{\text{tot}} \rangle_T / N$ and the average contributions from the electrostatic interaction ($\langle E_C \rangle_T / N$), Lennard-Jones term ($\langle E_{\text{LJ}} \rangle_T / N$), hydrogen-bond energy ($\langle E_{\text{HB}} \rangle_T / N$), and torsion energy ($\langle E_{\text{tor}} \rangle_T / N$). We display energy densities instead of the energies in order to have non-extensive quantities. Again we see in $\langle E_{\text{tot}} \rangle_T / N$ a clear signal for a transition which again becomes more pronounced with increasing chain length. We observe that in the high-temperature region the quantity $\langle E_{\text{tot}} \rangle / N$ is independent of the length of the chain. On the other hand, for low temperatures, $\langle E_{\text{tot}} \rangle_T / N$ is a decreasing function of chain length, indicating

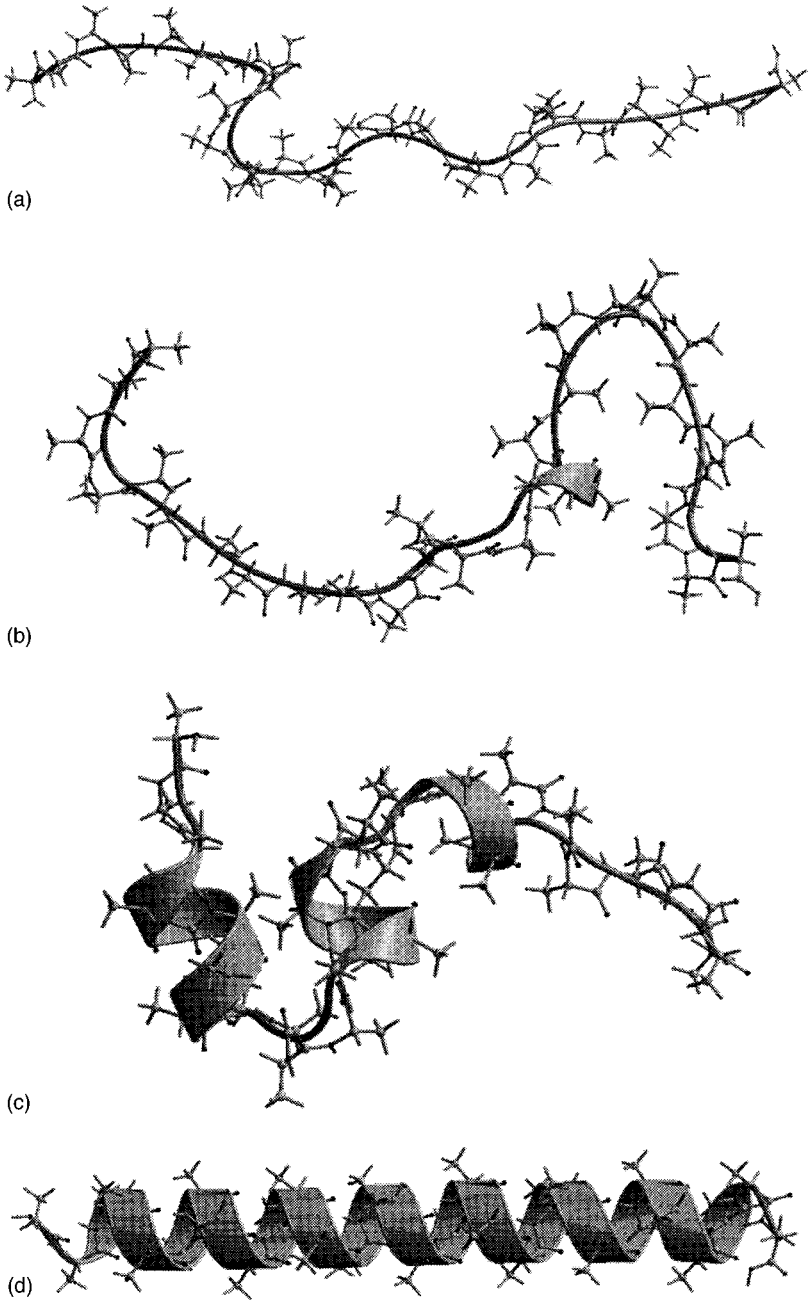

FIG. 1. Four example configurations chosen from a part of the multicanonical simulation where the random walk in energy moved from a high-energy state (a) to one of the lowest-energy states (d). The figure was created with MOLSCRIPT (Ref. 36) and RASTER3D (Ref. 37).

that here the energy is not solely a sum of local contributions from each residue. Instead, the total average energy rather behaves like $\langle E_{\text{tot}}\rangle_T/N \propto N^a$ with $a \approx 1.4$ (data not shown). Examining the results for each component of the potential energy, we learn that at small chain length the change in energy seems to be dominated by the change in the Lennard-Jones term. This energy term depends strongly on the overall size of the molecule, and the change in this quantity indicates a transition between extended and compact structures. However, with increasing number of residues, the contribution from the electrostatic interaction becomes more important, indicating the importance of nonlocal interactions for the observed transition, although even for $N=30$ the Lennard-Jones term is still the most important.

To study the observed transition between helix and coil conformers in more detail, we first have to determine the

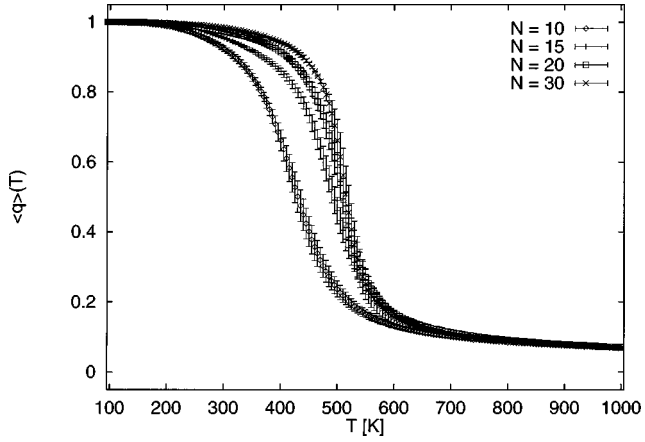

FIG. 2. Order parameter $\langle q\rangle_T$ as a function of temperature T for polyalanine molecules of chain length $N=10$, 15, 20, and 30.

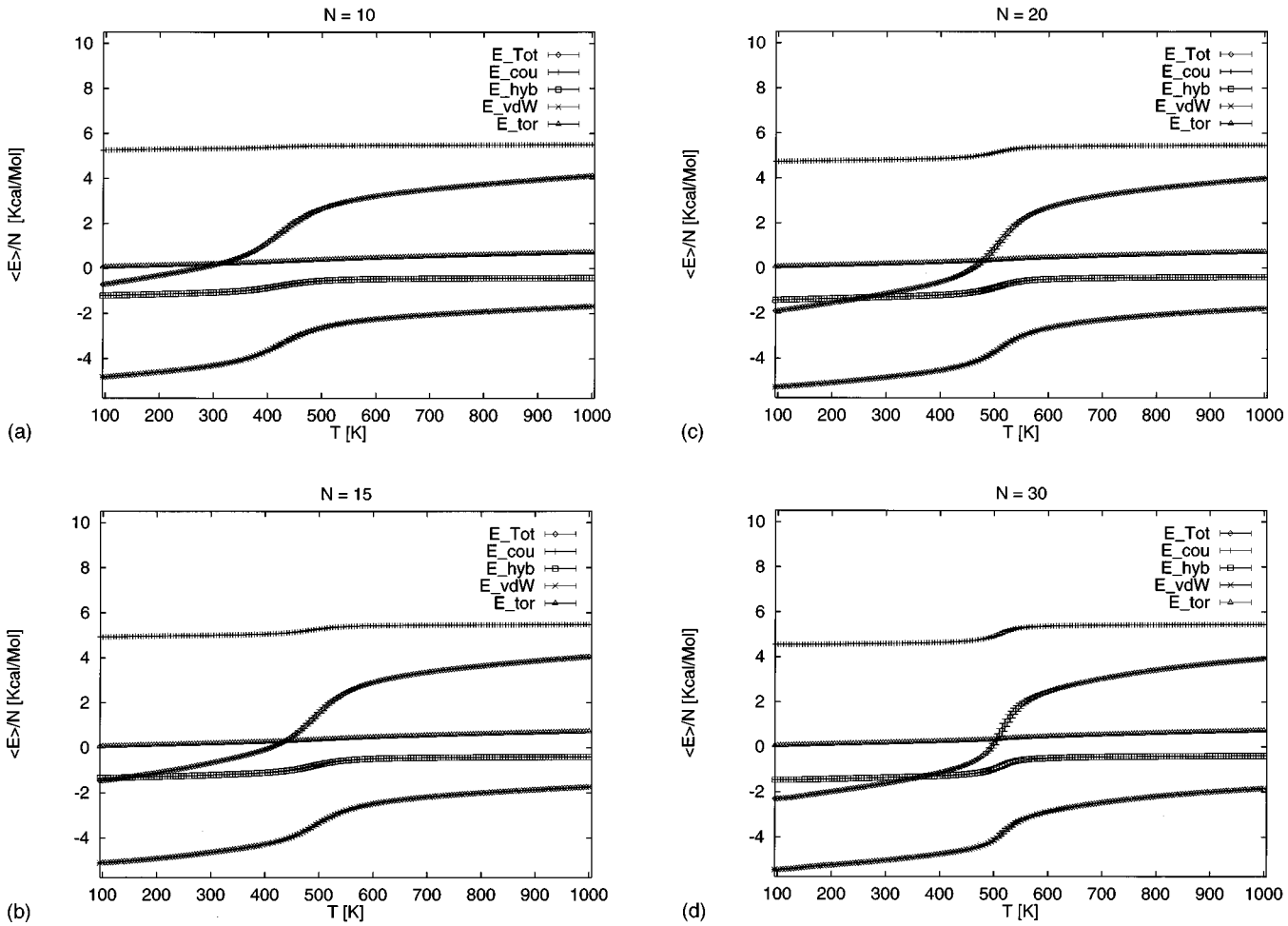

FIG. 3. Total energy density E_{tot}, electrostatic interaction E_C, Lennard-Jones term E_{LJ}, hydrogen-bond energy density E_{HB}, and torsion energy density E_{tor} as a function of temperature for (a) $N=10$, (b) $N=15$, (c) $N=20$, and (d) $N=30$.

transition temperature T_c. One possibility is to define T_c by the condition that $\langle q\rangle_{T=T_c}=0.5$. Another possibility is to define T_c as the temperature where the change in the order parameter is maximal, i.e., where $d/dT\langle q\rangle_T$ has an extremum. Other sensitive quantities are the susceptibility

$$\chi_N(T)=\frac{1}{N-2}\left(\langle q^2\rangle_T-\langle q\rangle_T{}^2\right), \tag{11}$$

which is plotted as a function of temperature in Fig. 4, and the specific heat, which we show in Fig. 5. The latter quantity is defined by

$$C_N(T)\equiv\frac{1}{Nk_B}\frac{d(\langle E_{\mathrm{tot}}\rangle_T)}{dT}=\beta^2\frac{\langle E_{\mathrm{tot}}^2\rangle_T-\langle E_{\mathrm{tot}}\rangle_T^2}{N}. \tag{12}$$

The temperatures where these two quantities have a peak can again be used to determine the transition temperature. In Table II we summarize the estimates for T_c as obtained by the different methods. For our analyses we divided the N_{Sw} MC sweeps in four bins and determined T_c in each bin separately. The shown values are the averages over all bins and the quoted errors were obtained from the variation of the estimates in the single bins. It is reassuring that the obtained estimates for the critical temperature do not depend on the way they were derived. Our final estimate is the average over all estimates as obtained by the different methods and is also displayed in Table II.

We are interested in the transition temperature for an infinitely long poly-alanine chain. Our data for finite chain

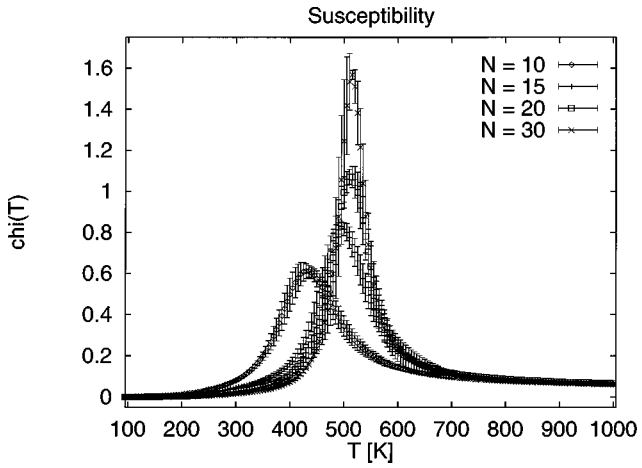

FIG. 4. Susceptibility χ as a function of temperature T for poly-alanine molecules of chain length $N=10$, 15, 20, and 30.

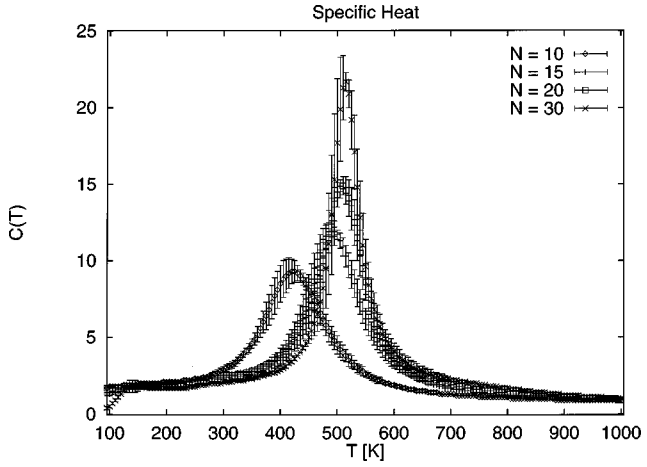

FIG. 5. Specific heat $C(T)$ as a function of temperature T for poly-alamine molecules of chain length $N=10$, 15, 20, and 30.

length show that T_c increases with growing chain length. By trial and error we found that the functional dependence of T_c as a function of chain length N could be best fitted by $T_c(L)=T_c(\infty)-a\cdot e^{-bN}$, which yields our estimate $T_c(\infty)=514$ K.

We remark that we do not find in Fig. 5 any indications for another peak in the specific heat at lower temperature $T<T_c$, which, if existed, could be interpreted as a transition between two helix states. Such a solid–solid transition was observed in a recent study on wormlike polymer chains.[11] Similarly, we do not see a shoulder in the specific heat for $T>T_c$. Hence, we conjecture that no other transitions but a helix–coil one exist for poly-alanine. Our results differ here again from the study of the minimal polymer model in Ref. 11. Our data for the *C*-peptide of ribonuclease *A* indicate that there may be separate transition temperatures for heteropolymers.[38]

The sharp increase in the peak of the susceptibility χ_N (Fig. 4) and the specific heat C_N (Fig. 5) as a function of chain length suggests that the observed transition between helix and coil states is a second-order phase transition. This assumption is supported by Fig. 6(a), where we show the free energy $G(q)=-k_BT\log P(q)$ as a function of order parameter q for $T=T_c$ and chain length $N=30$. Here, $P(q)$ is the probability to find a conformation which has the order parameter value q. We chose a normalization where $G=0$

TABLE II. Estimates for the transition temperature T_c as a function of the number of residues in the poly-alanine chain. Shown are the estimates as obtained from the temperature where the order parameter $\langle q\rangle(T)=0.5$, or where the derivative with respect to T of $\langle q\rangle$ has its maximum, or where the susceptibility $\chi(T)$ is maximal, or where the specific heat $C(T)$ has an extremum. The last column lists our final estimate for the transition temperature.

N	$\langle q\rangle=0.5$	Max $\left(\frac{d\langle q\rangle_T}{dT}\right)$	Max (ξ)	Max $[C(T)]$	T_c
10	432(10)	422(7)	430(5)	423 (7)	427 (7)
15	492 (7)	489 (5)	495(5)	490 (5)	492 (5)
20	512 (5)	509 (5)	513(5)	510 (5)	511 (5)
30	517(10)	513(10)	513(10)	513(10)	513 (10)

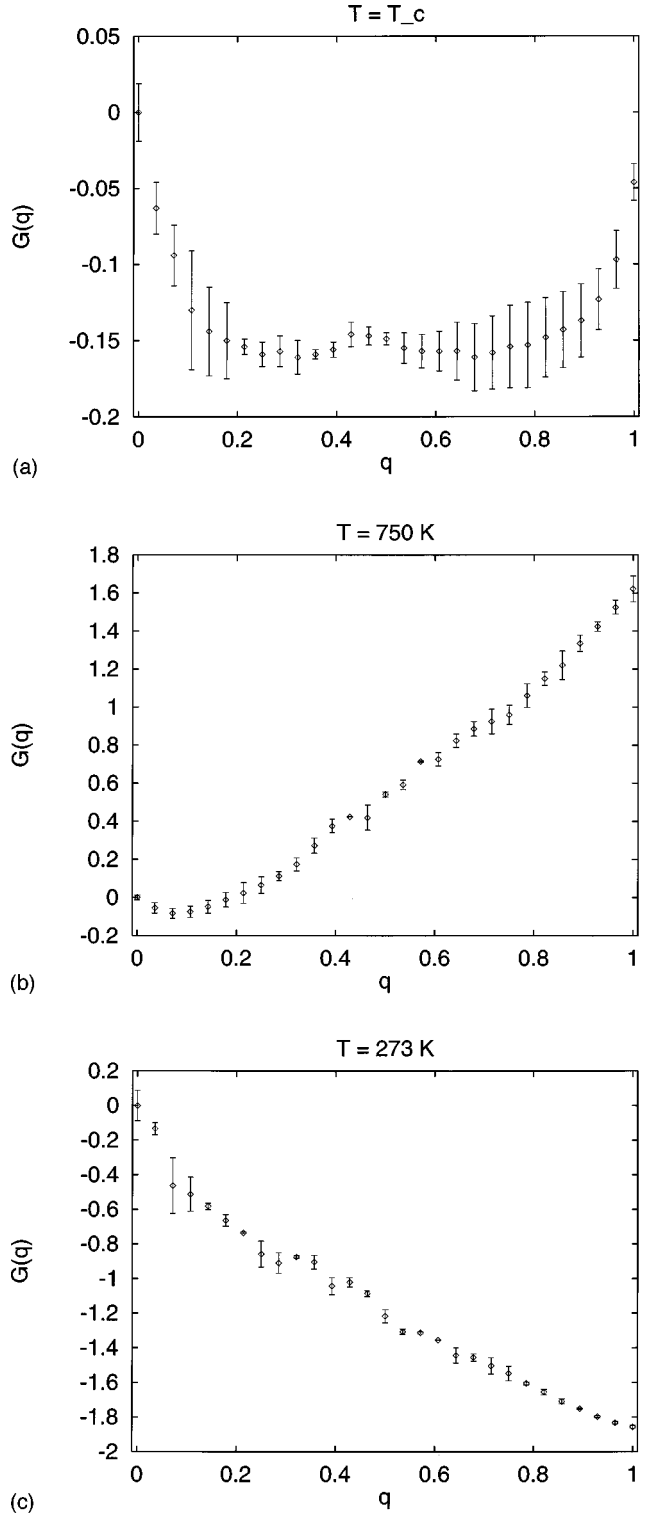

FIG. 6. Free energy $G(q)$ as a function of the order parameter q for (a) $T=T_c$, (b) $T=700$ K, and (c) $T=273$ K. We chose the normalization $G(0)=0$. All results are for $(Ala)_{30}$.

for coil state ($n_H=0$). We see that at the transition temperature, the free energy as a function of order parameter is almost constant for a wide range of values of the order parameter. A similar picture holds for the distribution of potential energy at the transition temperature: states with a wide range of energies appear at T_c with similar probability (data not shown). This is in support for a second-order phase transi-

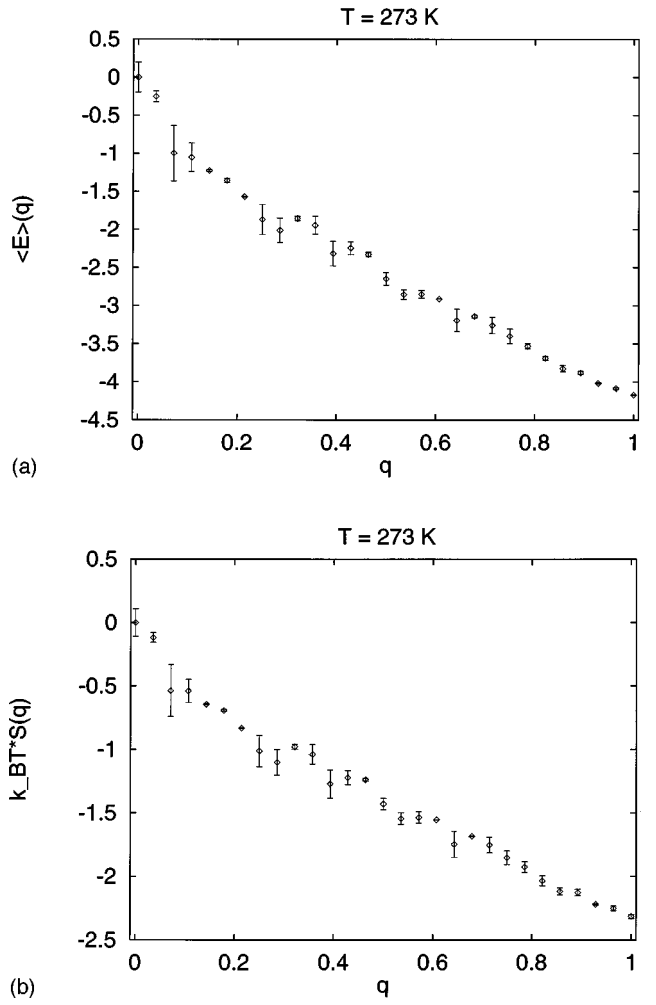

FIG. 7. Average potential energy $\langle E\rangle(q)$ (a) and entropy $\tilde{S}(q)=k_BTS(q)$ (b) as a function of order parameter q at temperature $T=273$ K. We chose the normalization $\langle E\rangle(q=0)=0$ and $\tilde{S}(q=0)=0$. All results are for $(Ala)_{30}$.

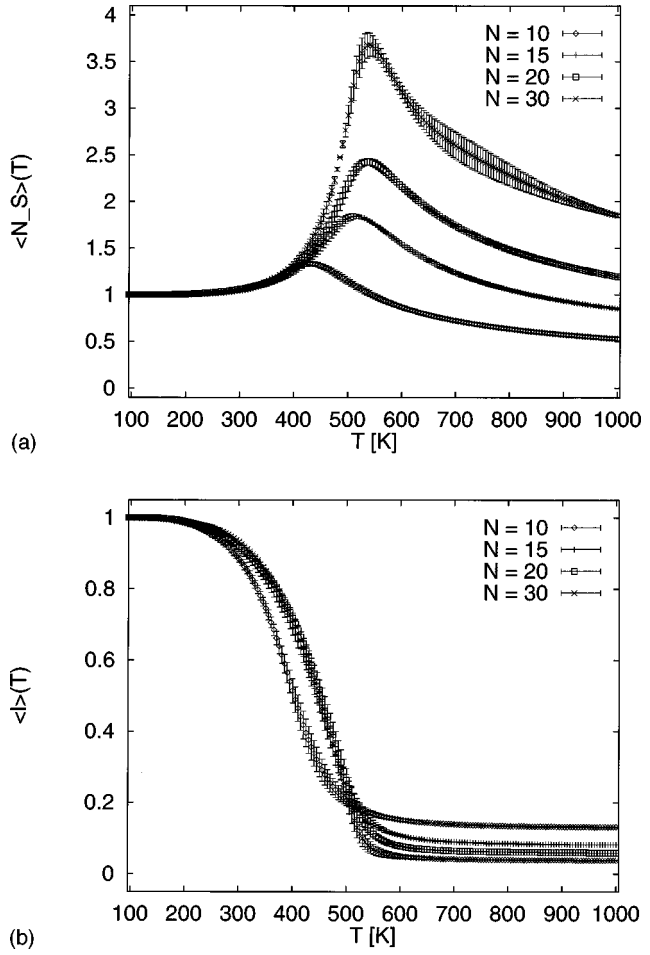

FIG. 8. Average number of helical segments n_s (a) and average length of helical segments $l/(N-2)$ (b) as a function of temperature T for poly-alanine molecules of chain length $N=10$, 15, 20, and 30.

tion, while for a first-order phase transition one would have a double-peak distribution. For comparison, we also show in Fig. 6(b) $G(q)$ for $T=700$ K, which is well within the high-temperature region, and in Fig. 6(c) $G(q)$ for $T=273$ K, an example for the behavior of the free energy in the low-temperature phase. The data are again for our largest chain, $N=30$. At high temperatures, coil states are clearly favored, whereas at low temperature one observes a strong bias toward the total helical state.

In Figs. 7(a) and 7(b) we display the average potential energy $\langle E(q)\rangle$ and entropy $\tilde{S}=k_BTS(q)$, respectively, as a function of q at $T=273$ K ($N=30$). Again we chose a normalization where $\langle E(q=0)\rangle=0$ and $\tilde{S}(q=0)=0$. We observe for both quantities a linear decrease with increasing order parameter. These results indicate again the existence of long-range order, since we find no indications that helical segments become unstable once they reach a critical length. Instead, we see in Fig. 8(a), where we display the average number of helical segments as a function of temperature and chain length, that for each chain the low-temperature region is dominated by a single helix. Around the critical temperature T_c, the number of helical segments is maximal and its average number increases with the size of the chain. This is again consistent with a second-order phase transition where one would also expect fluctuations on all length scales at T_c. The complementary picture is shown in Fig. 8(b), where we display the average length of helical segments $\overline{l}=\langle l\rangle/(N-2)$ as a function of temperature. Our normalization is chosen so that the largest possible segments at each chain length will lead to $\overline{l}=1$. We observe that, for low temperatures, this quantity is independent of the number of residues, which again demonstrates that there is no critical length of helical segments in our system above which helical segments become unstable and break up (or this critical length is larger than 30). For high temperatures we find that $\tilde{l}=\langle l\rangle/(N-2)\rightarrow 1/(N-2)$. This is because $\langle l\rangle\equiv\langle n_H\rangle/\langle n_s\rangle$ is the ratio of the number of helical residues, divided by the number of helical segments. We remark that again neither Fig. 8(a) nor Fig. 8(b) gives any indication for a helix–helix transition as was observed in Ref. 11. It seems that the existence of such a transition depends on the chosen model.

According to the Zimm–Bragg model,[13] the average number of helical residues $\langle n\rangle$ and the average length $\langle l\rangle$ of a helical segment are given for large N by

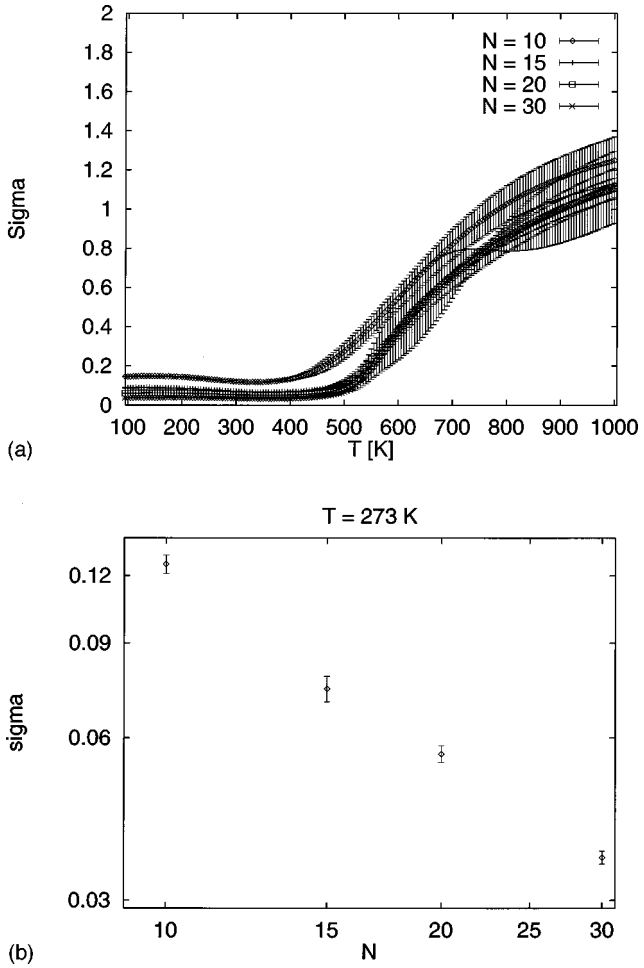

FIG. 9. (a) Nucleation parameter σ as a function of temperature T for poly-alanine molecules of chain length $N=10$, 15, 20, and 30. (b) Log–log plot of the nucleation parameter σ as a function of the number of residues at $T=273$ K.

$$\frac{\langle n\rangle}{N}=\frac{1}{2}-\frac{1-s}{2\sqrt{(1-s)^2+4s\sigma}}, \tag{13}$$

$$\langle l\rangle=1+\frac{2s}{1-s+\sqrt{(1-s)^2+4s\sigma}}, \tag{14}$$

where N is the number of residues, and s and σ are the helix propagation parameter and the nucleation parameter, respectively. Note that $s\geqslant 1$ implies that $\langle n\rangle/N\geqslant 1/2$ (more than 50% helicity). From these equations, with the values of $\langle n\rangle/N$ and $\langle l\rangle$ calculated from the multicanonical production runs, one can obtain estimates of s and σ parameters. We are especially interested here in the nucleation parameter σ which characterizes the probability for a helix–coil junction, and hence is related to the probability of a helical segment breaking apart into two pieces. We show in Fig. 9(a) the average value $\langle\sigma\rangle_T$ as a function of temperature. In accordance with our previous results in Ref. 1, we notice that $\langle\sigma\rangle_T$ is constant below the critical temperature T_c and that its value in the low-temperature region decreases with increasing chain length. From the log–log plot (for $T=273$ K, deep in the low-temperature region) in Fig. 9(b) it follows that $\langle\sigma\rangle$ as a function of chain length follows a power-law behavior,

$$\langle\sigma\rangle(N)=\sigma_0 N^{-c}, \tag{15}$$

[with $\sigma_0=0.48(11)$ and $c=1.13(5)$]. It follows that $\langle\sigma(\infty)\rangle=0$: The probability for a helical segment to break into pieces approaches zero for infinite chain length. This again supports our claim that poly-alanine exhibits a long-range order in the low-temperature phase. How can this result be understood in the framework of the Zimm–Bragg model? The only possibility to have long-range order in a one-dimensional Ising model with short interactions is if the boundary tension is infinitely large. Infinite boundary tension corresponds to $\sigma=0$ in the Zimm–Bragg model and describes here a model where only helical conformers exist (since the probability for helix–coil junctions is per definition zero in this case). But while in the Zimm–Bragg model σ is an input parameter and describes the interaction between coil and helix segments, σ is calculated in our case a posteriori and as a function of temperature. Hence, the poly-alanine chains are described by different Zimm–Bragg models at different temperatures. None of them can exhibit the phase transition observed for the homopolymer. The existence of this phase transition corresponds to the change in the models. Finally, in Table III we summarize the values of σ and, in addition, those of the helix propagation parameter s at $T=273$ K. The latter quantity is an increasing function of chain length N and can be described by $s(N)=s_\infty-ae^{-bN}$, which gives $s_\infty=1.87(3)$.

TABLE III. Nucleation parameter σ and helix propagation parameter s as a function of chain length N for $T=273$ K as determined from multicanonical simulations. Shown are also our extrapolated values for an infinitely long chain.

N	σ	s
10	0.126(4)	1.561(28)
15	0.074(2)	1.679(10)
20	0.056(1)	1.780(13)
30	0.036(1)	1.845(25)
∞	0.0	1.87(3)

After having established the existence of a second-order phase transition for poly-alanine, we further characterize this transition by determining the critical exponents. Conventional arguments for finite-size scaling for a second-order transition are based on the assumption that the singular part of the free energy depends only on the system size N and the correlation length ξ.[39] The critical exponents can be extracted from the finite-size scaling of the heights and width of the peaks in specific heat and susceptibility. With the critical temperature $T_c(N)$ as the position where the peak in the specific heat has its maximum, and $T_1(N)$ and $T_2(N)$ [with $T_1(N)<T_c(N)<T_2(N)$] chosen such that $C(T_1)=1/2C(T_c)=C(T_2)$, we have

$$\Gamma_C(N)=T_2(N)-T_1(N)\propto N^{-1/\nu}, \tag{16}$$

and

$$C_N(T_c)\propto N^{\alpha/\nu}. \tag{17}$$

TABLE IV. Maximum of specific heat $C_{\max}$ and susceptibility $\chi_{\max}$ together with width of peak in specific heat Γ_C and width in peak of susceptibility Γ_χ for various chain lengths.

N	T_c	$C_{\max}$	Γ_C	$\chi_{\max}$	Γ_χ
10	427(7)	9.3(8)	146(7)	0.61(4)	136(7)
15	492(5)	12.0(4)	122(5)	0.82(3)	110(5)
20	511(5)	14.9(6)	95(5)	1.07(4)	90(5)
30	513(10)	21.7(1.5)	82(3)	1.57(8)	78(3)

Similarly, we find from the heights of the peak in the susceptibility

$$\chi_N(T_c) \propto N^{\gamma/\nu}, \tag{18}$$

and from the temperatures where $\chi(T)=1/2\chi(T_c)$, we get a second, independent estimate for the critical exponent ν by

$$\Gamma_\chi(N) = T_2(N) - T_1(N) \propto N^{-1/\nu}. \tag{19}$$

The various quantities are summarized in Table IV. From these values we obtain, using the above equations, the following estimates for the critical exponents: Eq. (16) yields an estimate for $1/\nu = 0.54(5)$ [with a goodness of $q=0.3$ (see Ref. 40 for the definition of q) for the fit] which is comparable to $1/\nu = 0.51(6)$, the value we obtained with a goodness of $q=0.6$ from the fitting of Eq. (19). Combining both values, we have as our final estimate for the correlation length exponent for the helix–coil transition in poly-alanine,

$$\nu = 1.9(2). \tag{20}$$

With a value of $\alpha/\nu = 0.79(9)$, obtained with goodness $q=0.6$ by fiting Eq. (17), we find the following specific heat exponent:

$$\alpha = 1.5(2). \tag{21}$$

Similarly, from Eq. (18) we obtain, with a goodness of $q=0.7$, a value $\gamma/\nu = 0.88(7)$, from which we get our estimate for the susceptibility exponent,

$$\gamma = 1.7(1). \tag{22}$$

The nontrivial values we obtained for these critical exponents give further evidence for the second-order phase transition.

CONCLUSIONS

We have performed multicanonical simulations of poly-alanine for chains of up to 30 residues long with high statistics. Our data allowed us to extrapolate our results to the limit of infinitely long chains and showed that the change between helical and coil conformers can indeed be understood as a temperature-driven, second-order phase transition. We were able to determine some critical exponents for this transition.

ACKNOWLEDGMENTS

Our simulations were performed on computers of the Institute for Molecular Science (IMS), Okazaki, Japan, and the Department of Physics, Michigan Technological University, Houghton, Michigan. This work was supported, in part, by a Research Excellence Fund (E27448) of the state of Michigan and by a Grant-in-Aid for Scientific Research from the Japanese Ministry of Education, Science, Sports and Culture.

[1] Y. Okamoto, U. H. E. Hansmann, and T. Nakazawa, Chem. Lett. **1995**, 391; Y. Okamoto and U. H. E. Hansmann, J. Phys. Chem. **99**, 11276 (1995).
[2] A. Chakrabartty and R. L. Baldwin, in *Protein Folding: In Vivo and In Vitro*, edited by J. Cleland and J. King (ACS, Washington, D.C., 1993), pp. 166–177.
[3] T. Ooi and M. Oobatake, Proc. Natl. Acad. Sci. USA **88**, 2859 (1991).
[4] J. Tirado-Rives and W. L. Jorgensen, Biochemistry **30**, 3864 (1991).
[5] D. J. Tobias and C. L. Brooks, III, Biochemistry **30**, 6059 (1991).
[6] V. Daggett, P. A. Kollman, and I. D. Kuntz, Biopolymers **31**, 1115 (1991).
[7] J. Vila, R. L. Williams, J. A. Grant, J. Wójcik, and H. A. Scheraga, Proc. Natl. Acad. Sci. USA **89**, 7821 (1992).
[8] T. P. Creamer and G. D. Rose, Proc. Natl. Acad. Sci. USA **89**, 5937 (1992).
[9] Y. Okamoto, Proteins: Struct., Funct., Genet. **19**, 14 (1994); Biopolymers **34**, 529 (1994).
[10] Y. Zhou, M. Karplus, J. M. Wichert, and C. K. Hall, J. Chem. Phys. **107**, 10691 (1997).
[11] J. P. Kemp and Z. Y. Chen, Formation of Helical States in Wormlike Polymer Chains, preprint cond-mat/9804178.
[12] A. Yu. Grosberg and A. R. Khokhlov, *Statistical Physics of Macromolecules* (AIP, New York, 1994).
[13] B. H. Zimm and J. K. Bragg, J. Chem. Phys. **31**, 526 (1959).
[14] B. A. Berg and T. Neuhaus, Phys. Lett. B **267**, 249 (1991); Phys. Rev. Lett. **68**, 9 (1992); B. A. Berg, Int. J. Mod. Phys. C **3**, 1083 (1992).
[15] U. H. E. Hansmann and Y. Okamoto, in *Annual Reviews of Computational Physics*, edited by D. Stauffer (World Scientific, Singapore, in press), Vol. VI.
[16] U. H. E. Hansmann and Y. Okamoto, J. Comput. Chem. **14**, 1333 (1993).
[17] U. H. E. Hansmann and Y. Okamoto, Physica A **212**, 415 (1994).
[18] U. H. E. Hansmann and Y. Okamoto, J. Comput. Chem. **18**, 920 (1997).
[19] F. Eisenmenger and U. H. E. Hansmann, J. Phys. Chem. B **101**, 3304 (1997).
[20] U. H. E. Hansmann, M. Masuya, and Y. Okamoto, Proc. Natl. Acad. Sci. USA **94**, 10,652 (1997).
[21] U. H. E. Hansmann and Y. Okamoto, J. Phys. Chem. **102**, 653 (1998).
[22] S. R. Wilson and W. Cui, in *The Protein Folding Problem and Tertiary Structure Prediction*, edited by K. M. Merz, Jr. and S. M. Le Grand (Birkhäuser, Boston, 1994), pp. 43–70.
[23] C. A. Floudas, J. L. Klepeis, and P. M. Pardalos, DIMACS Series in Discrete Mathematics and Theoretical Computer Science, edited by F. Roberts (American Mathematical Society, Providence, in press).
[24] S. Ihara, T. Ooi, and S. Takahashi, Biopolymers **21**, 131 (1982).
[25] K. R. Shoemaker, P. S. Kim, E. J. York, and R. L. Baldwin, Nature (London) **326**, 563 (1987).
[26] F. A. Momany, R. F. McGuire, A. W. Burgess, and H. A. Scheraga, J. Phys. Chem. **79**, 2361 (1975).
[27] G. Némethy, M. S. Pottle, and H. A. Scheraga, J. Phys. Chem. **87**, 1883 (1983).
[28] M. J. Sippl, G. Némethy, and H. A. Scheraga, J. Phys. Chem. **88**, 6231 (1984).
[29] H. Kawai, Y. Okamoto, M. Fukugita, T. Nakazawa, and T. Kikuchi, Chem. Lett. **1991**, 213 (1991); Y. Okamoto, M. Fukugita, T. Nakazawa, and H. Kawai, Protein Eng. **4**, 639 (1991).
[30] J. Lee, Phys. Rev. Lett. **71**, 211 (1993); **71**, 2353(E) (1993).
[31] B. A. Berg, U. H. E. Hansmann, and Y. Okamoto, J. Phys. Chem. **99**, 2236 (1995).
[32] M. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).
[33] A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **61**, 2635 (1988); **63**, 1658(E) (1989), and references given in the erratum.
[34] S. Lin and B. W. Kernighan, Oper. Res. **21**, 498 (1973).
[35] B. A. Berg, in *Multiscale Phenomena and Their Simulation*, edited by F. Karsch, B. Monien, and H. Satz (World Scientific, Singapore, 1997), p. 137.

[36] P. J. Kraulis, J. Appl. Crystallogr. **24**, 946 (1991).
[37] D. Bacon and W. F. Anderson, J. Mol. Graphics **6**, 219 (1988); E. A. Merritt and M. E. P. Murphy, Acta Crystallogr. D **50**, 869 (1994).
[38] U. H. E. Hansmann and Y. Okamoto, Effects of side-chain charges on α-helix stability in *C*-peptide of ribonuclease *A* studied by multicanonical algorithm, MTU-PHY-HA98/4 (submitted).
[39] M. E. Fisher and M. N. Barber, Phys. Rev. Lett. **28**, 1516 (1972).
[40] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes*, 2nd ed. (Cambridge University Press, New York, 1992), p. 657.