Helix-coil transitions of amino-acid homo-oligomers in aqueous solution studied by multicanonical simulations

Ayori Mitsutake, and Yuko Okamoto

Citation: J. Chem. Phys. **112**, 10638 (2000); doi: 10.1063/1.481697
View online: https://doi.org/10.1063/1.481697
View Table of Contents: http://aip.scitation.org/toc/jcp/112/23
Published by the American Institute of Physics

Helix-coil transitions of amino-acid homo-oligomers in aqueous solution studied by multicanonical simulations

Ayori Mitsutake[a] and Yuko Okamoto[b]
Department of Functional Molecular Science, The Graduate University for Advanced Studies and Department of Theoretical Studies, Institute for Molecular Science, Okazaki, Aichi 444-8585, Japan

(Received 2 February 2000; accepted 24 March 2000)

Helix-coil transitions of homo-oligomers in aqueous solution are studied by multicanonical Monte Carlo simulations. The solvation effects are represented by the sum of the terms that are proportional to the solvent-accessible surface area of the atomic groups. Homo-oligomers of length 10 are considered for three characteristic amino acids, alanine, valine, and glycine, which are helix former, helix indifferent, and helix breaker, respectively. We calculated as a function of temperature the distributions of the backbone dihedral angles, the average values of total energy, and its component terms of the homo-oligomers. It is shown that for homo-alanine, the helix-coil transition exists and that the transition temperature in water is considerably lower than in gas phase, which implies that the effects of solvation tend to reduce helical content. Moreover, the helix propagation parameter s and nucleation parameter σ of the Zimm-Bragg model were calculated. The s values that were obtained from the simulations in aqueous solution are in remarkable agreement with the experimental results. © *2000 American Institute of Physics.* [S0021-9606(00)51723-1]

I. INTRODUCTION

The α-helix formation usually occurs at an earlier stage of protein folding. Thus, the study of helix-coil transitions provides important insight into the protein folding problem. Many experiments[1–11] have studied α-helix forming propensities of each amino acid (for a review, see Ref. 12).

The helix-forming propensities of various amino acids have been measured by using block copolymers,[1,6] synthetic peptide fragments,[2–4,7,10,11] a noncovalent α-helix dimer,[5] or proteins (including helical sequences).[8,9,11] The helix-forming tendencies of the amino acids determined from these studies are quite similar but not completely in agreement with each other. This implies that the helix-forming tendencies depend on the amino acid sequences and the environment. However, it has been clear that the rank order of helix propensity of different amino acids is essentially identical among different experiments and that alanine, valine, and glycine are helix former, helix indifferent, and helix breaker, respectively.

The α-helix forming tendencies and the helix-coil transitions have also been theoretically studied mainly with short peptides systems. Various methods have been employed for these studies; for instance, Monte Carlo simulations,[13,14] molecular dynamics,[15–20] simulated annealing,[21–24] and multicanonical Monte Carlo simulations[25–27] have been applied.

The multicanonical algorithm[28] is particularly useful for studying helix-coil transitions, because a single simulation run can give thermodynamics quantities in a wide range of temperatures covering both helix phase at low temperatures and coil phase at high temperatures. This method and its generalizations have already been used in many applications in protein and related systems.[25–27,29–51] In a previous work,[25] thermodynamics of helix-coil transitions of homo-oligomers in gas phase were studied by multicanonical algorithms. Home-oligomers of length 10 were considered for three characteristic amino acids, alanine (helix former), valine (helix indifferent), and glycine (helix breaker). The helix-coil transition from an ideal helix to a random coil was observed in homo-alanine. Moreover, it was shown that the obtained helix propagation parameters s of the Zimm-Bragg model[52] for the three amino acids were in agreement with the experimental values. It turned out, however, that the helix-coil transition temperature was above 400 K and unrealistically high.[25] This discrepancy is presumably caused by the lack of solvent in the simulations.

In this paper, we study thermodynamics of helix-coil transitions in amino-acid homo-oligomers of length 10 in aqueous solution by multicanonical Monte Carlo simulations and compare the results with those in the gas phase. The preliminary results were reported elsewhere.[27] The solvation effects are represented by the sum of the terms that are proportional to the solvent-accessible surface area of the atomic groups. We calculate the distributions of the backbone dihedral angles, average values of total potential energy, its component terms, specific heat, and helicity as a function of temperature. It is found that overall behaviors and average values of various quantities are similar to those in the gas phase. The helix-coil transition temperature for homo-alanine in aqueous solution, however, is indeed found to be much lower than that in the gas phase, implying that the effects of solvation reduce helical content. Moreover, the helix propagation parameter s and nucleation parameter σ of the Zimm-Bragg model in aqueous solution are calculated and are found to be in remarkable agreement with experimental results.

[a]Electronic mail: ayori@ims.ac.jp
[b]Electronic mail: okamotoy@ims.ac.jp

II. METHODS

A. Peptide preparation and potential energy function

Homo-oligomers of length 10 are considered for three characteristic amino acids, alanine, valine, and glycine, which are helix former, helix indifferent, and helix breaker, respectively. Since the charges at peptide termini are known to reduce helix content,[53,54] we removed them by taking a neutral NH_2- group at the N-terminus and a neutral $-COOH$ group at the C-terminus.

The energy function E_{TOT} that we used is the sum of the conformational energy term of the solute E_P and the solvation free energy term $\mathcal{E}_{SOL}$ for the interaction of the protein with the surrounding solvent,

$$E_{TOT}=E_P+\mathcal{E}_{SOL}. \tag{1}$$

The conformational energy E_P (in kcal/mol) consists of electrostatic term E_C, Lennard-Jones term E_V, hydrogen-bond term E_H, and torsion term E_T,

$$\begin{aligned} E_P&=E_C+E_V+E_H+E_T,\\ E_C&=\sum_{(i,j)}\frac{332q_iq_j}{\epsilon r_{ij}},\\ E_V&=\sum_{(i,j)}\left(\frac{A_{ij}}{r_{ij}^{12}}-\frac{B_{ij}}{r_{ij}^{6}}\right),\\ E_H&=\sum_{(i,j)}\left(\frac{C_{ij}}{r_{ij}^{12}}-\frac{D_{ij}}{r_{ij}^{10}}\right),\\ E_T&=\sum_i U_i(1+\cos(n_i\chi^i+\theta_i)). \end{aligned} \tag{2}$$

Here, r_{ij} is the distance (in Å) between atoms i and j, ϵ is the dielectric constant, and χ^i is the torsion angle for the chemical bond i. Each atom is expressed by a point at its center of mass, and the partial charge q_i (in units of electronic charges) is assumed to be concentrated at that point. The factor 332 in E_C is a constant to express energy in units of kcal/mol. The parameters in the energy function as well as the molecular geometry are based on ECEPP/2.[55–57] The computer code KONF90[21] was modified to accommodate the multicanonical algorithm. The dielectric constant ϵ is set equal to 2 (according to the prescription of ECEPP/2[55–57]). The peptide-bond dihedral angles ω are fixed at the value 180° for simplicity. The remaining dihedral angles (ϕ_i and ψ_i in the main chain of each residue i and χ_i^j in its side chains) constitute the degrees of freedom in the simulations.

The solvation free energy that we used is the sum of terms that are proportional to the solvent-accessible surface area of the atomic groups of the solute,

$$\Delta\mathcal{E}_{SOL}=\sum_i \sigma_iA_i, \tag{3}$$

where the summation extends over all groups i of atoms, and A_i is the corresponding solvent-accessible surface area. The constant of proportionality σ_i represents the contribution to the solvent free energy of group i per unit accessible surface area. In this paper we used the parameters of Ref. 58. The groups containing hydrogen atoms are treated as ''united atoms.'' The accessible surface area A_i is obtained by the surface area of fused spheres centered at each united atom. The radius of the sphere is R_i+R_w, where R_w is the effective radius of the solvent molecule. Here we set R_i to van der Walls radius and R_w to 1.4 Å. The coefficients σ_i were determined by applying the least-squares procedure to the experimental data of 22 small molecules.[58] Then, the term represents the contribution to complete solvation free energy. For the calculation of solvent-accessible surface area, we use the computer code NSOL,[59] which is based on the code NSC.[60]

B. Multicanonical algorithm

While a regular Monte Carlo method generates states according to the canonical distribution, the multicanonical algorithms generate states so that a one-dimensional random walk in energy space is realized. Hence, any energy barrier can be overcome, and one can avoid getting trapped in states of energy local minima. In the multicanonical ensemble the probability distribution of energy, $P_{mu}(E)$, is defined in such a way that a configuration with any energy enters with equal probability,

$$P_{mu}(E)\propto n(E)w_{mu}(E)=\text{const}, \tag{4}$$

where $n(E)$ is the density of states and $w_{mu}(E)$ is the multicanonical weight factor. The multicanonical weight factor then satisfies

$$w_{mu}(E)\propto n^{-1}(E). \tag{5}$$

Since this weight factor is not *a priori* known, one has to determine it for each system by a few iterations of trial Monte Carlo simulations. The detailed procedures for this process are described in detail in Refs. 25, 61, and 62. In this work we used the one in Ref. 62. Once the optimal weight factor $w_{mu}(E)$ is determined, one performs with this weight factor a multicanonical simulation with high statistics. From this simulation run one can not only locate the energy global minimum but also calculate the canonical-ensemble average of any physical quantity A at any temperature $T(=1/R\beta)$ for a wide range of temperatures by the reweighting techniques,[63]

$$\langle A\rangle_T=\frac{\int dEA(E)P_{mu}(E)w_{mu}^{-1}(E)e^{-\beta E}}{\int dEP_{mu}(E)w_{mu}^{-1}(E)e^{-\beta E}}, \tag{6}$$

where $P_{mu}(E)$ is the distribution of energy obtained by the final simulation.

C. Computational details

In this paper, the numbers of degrees of freedom (dihedral angles) are 30, 50, and 20 for $(Ala)_{10}$, $(Val)_{10}$, and $(Gly)_{10}$, respectively. One Monte Carlo (MC) sweep consists of updating all these angles once with Metropolis evaluation[64] for each update. For the calculation of multicanonical weight factors, it required between 110 000 MC sweeps in gas phase and 400 000 MC sweeps in aqueous solution for $(Ala)_{10}$, 900 000 MC sweeps in gas phase and 1 000 000 MC sweeps in aqueous solution for $(Val)_{10}$, and 450 000 MC sweeps in gas phase and 250 000 MC sweeps in

aqueous solution for $(Gly)_{10}$. After the optimal weighting factor $w_{mu}(E)$ was determined, we then made one production run with 1 000 000 MC sweeps and obtained various thermodynamic quantities as a function of temperature by the reweighting techniques.[63] Initial conformations were randomly generated.

III. RESULTS AND DISCUSSION

A. Time series of the total potential energy

We first examine how much of the configuration space the multicanonical simulations explore. As explained in Sec. II, a simulation in the multicanonical ensemble performs a one-dimensional random walk in the potential energy space. It should visit not only the ground-state region but also states with very high energy. This ensures that the simulations avoid getting trapped in configurations with energy local minima. We display in Fig. 1 the ''time series'' of the total potential energy E_{TOT} for $(Ala)_{10}$, $(Val)_{10}$, and $(Gly)_{10}$ in aqueous solution. The results indeed exhibit a random walk in energy space covering a range of 60–100 kcal/mol. We confirmed that random walks in energy space for $(Ala)_{10}$, $(Val)_{10}$, and $(Gly)_{10}$ were also obtained in gas phase (data not shown). Since configurations are sampled over a large range of energies, the reweighting techniques allow one to calculate thermodynamic quantities as functions of temperature for a wide range of temperatures.[63]

B. The lowest-energy conformations

We investigate the lowest-energy conformations of homo-oligomers obtained during the present simulations. In Tables I and II, we list the dihedral angles of the lowest-energy conformations for $(Ala)_{10}$, $(Val)_{10}$, and $(Gly)_{10}$ in gas phase and in aqueous solution, respectively. For $(Ala)_{10}$ in both environments, the residues of the dihedral angles from residue 2 to residue 9 are in α-helix state. Here, the criterion we adopt for the α-helix state of residue is as follows: We consider that a residue is in the α-helix configuration when the dihedral angles (ϕ, ψ) fall in the range $(-70\pm30°, -37\pm30°)$. The length ℓ of a helical segment is then defined by the number of successive residues which are in the helix configuration. The lowest-energy conformations for $(Ala)_{10}$ have the helix length $\ell=8$ and are completely helical conformations (the terminal residues tend to be frayed). The side-chain structures of $(Ala)_{10}$ are also uniquely determined for the lowest-energy conformations in both environments; namely, the values of χ are close to one of 60, -60, and 180°, which are all equivalent angles because of the three-fold rotational symmetry of the alanine side chain. These lowest-energy conformations for $(Ala)_{10}$ in gas phase and in aqueous solution are shown in Figs. 2(a) and 2(b), respectively. These conformations have six intrachain backbone hydrogen bonds that characterize the α-helix and are indeed completely helical.

As shown in Tables I and II, the dihedral angles of $(Val)_{10}$ for the lowest-energy conformations in both environments are again in almost ideal helix state (from residue 2 to residue 9 in gas phase and from residue 2 to residue 8 in aqueous solution). These lowest-energy conformations are shown in Figs. 2(c) and 2(d). We remark that the lowest-energy conformation of $(Val)_{10}$ in aqueous solution discussed here was actually obtained by another multicanonical simulation run that was made separately from the one shown in Fig. 1(b) (the energy difference between the lowest-energy conformations obtained by the two runs was only about 1 kcal/mol). This run, however, got trapped in the ground-state region and did not perform a random walk in energy space (this means that the obtained multicanonical weight factor for this particular run was not optimal). We thus used the results of the run in Fig. 1(b) instead for the calculation of thermodynamic quantities as a function of temperature in a wide temperature range (which is presented below). The thermodynamic quantities were calculated for the tempera-

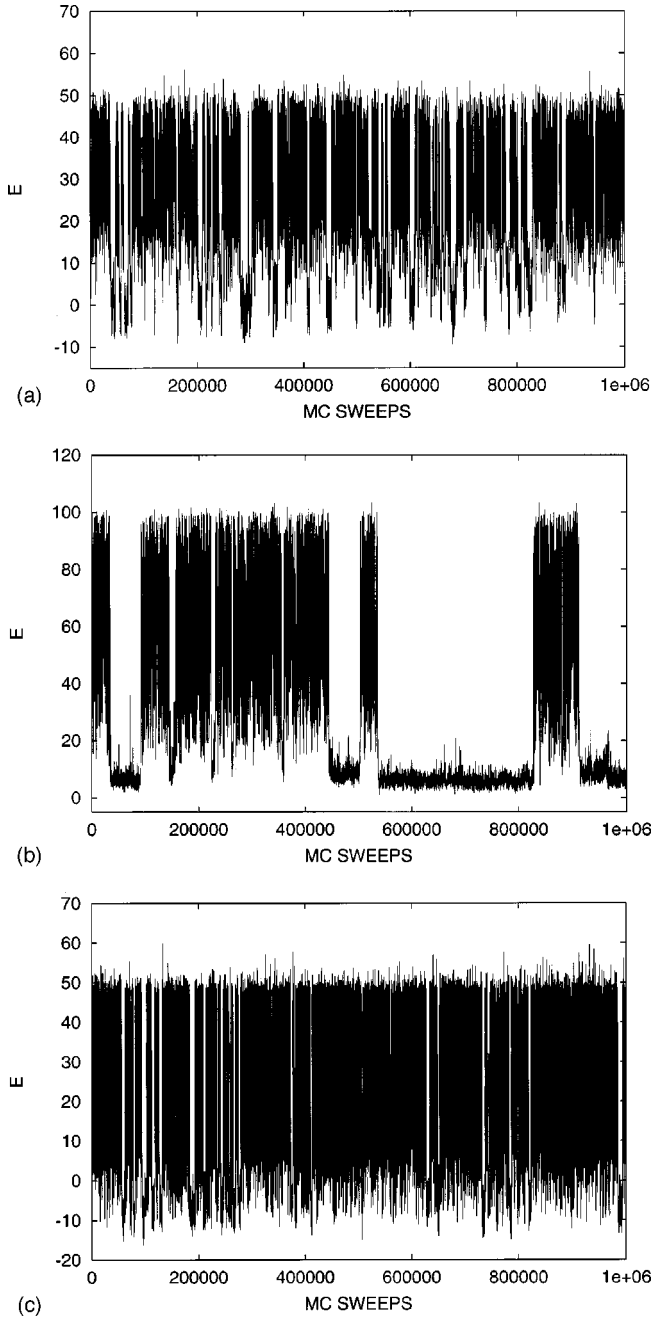

FIG. 1. Time series of the total potential energy E_{TOT} (kcal/mol) for $(Ala)_{10}$ (a), $(Val)_{10}$ (b), and $(Gly)_{10}$ (c) in aqueous solution.

TABLE I. The dihedral angles of the lowest-energy conformations for $(Ala)_{10}$ (a), $(Val)_{10}$ (b), and $(Gly)_{10}$ (c) in gas phase.

	Residue	ϕ	ψ	χ
(a)	1	−24	−42	−68
	2	−66	−38	−67
	3	−68	−36	65
	4	−66	−38	65
	5	−69	−35	−57
	6	−71	−35	63
	7	−73	−30	177
	8	−75	−36	51
	9	−75	−34	−50
	10	−153	97	178

	Residue	ϕ	ψ	χ^1	χ^2	χ^3
(b)	1	3	−20	65	−51	177
	2	−70	−34	173	−67	−59
	3	−64	−31	173	172	−54
	4	−73	−39	174	−68	60
	5	−64	−34	166	47	39
	6	−66	−34	163	50	−71
	7	−73	−36	163	170	49
	8	−63	−34	161	−72	50
	9	−78	−47	166	48	168
	10	−99	91	−176	178	−57

	Residue	ϕ	ψ
(c)	1	122	−50
	2	160	−86
	3	−148	69
	4	61	−109
	5	−56	−52
	6	−79	78
	7	146	−34
	8	84	31
	9	170	174
	10	−51	144

TABLE II. The dihedral angles of the lowest-energy conformations for $(Ala)_{10}$ (a), $(Val)_{10}$ (b), and $(Gly)_{10}$ (c) in aqueous solution.

	Residue	ϕ	ψ	χ
(a)	1	19	149	62
	2	−76	−29	−166
	3	−68	−38	−175
	4	−74	−33	174
	5	−69	−39	−46
	6	−67	−36	63
	7	−68	−40	67
	8	−64	−36	69
	9	−70	−42	63
	10	−154	107	174

	Residue	ϕ	ψ	χ^1	χ^2	χ^3
(b)	1	13	−32	62	−51	55
	2	−62	−33	173	−63	−48
	3	−67	−40	171	172	62
	4	−68	−35	170	48	173
	5	−65	−34	163	−65	−72
	6	−77	−39	165	−72	−72
	7	−60	−40	162	−73	42
	8	−79	−47	170	−72	55
	9	−106	84	−172	−63	71
	10	−79	128	175	−61	57

	Residue	ϕ	ψ
(c)	1	107	−83
	2	79	−89
	3	170	−46
	4	−78	173
	5	61	35
	6	62	35
	7	80	44
	8	−152	29
	9	176	172
	10	−69	−79

ture range between 200 and 700 K, while the ground state corresponds to the temperature 0 K. Thus, it does not matter whether the run did reach the ground state or not, as long as it performs a random walk in the relevant energy range.

For $(Gly)_{10}$ in the gas phase, the dihedral angles of the lowest-energy conformation seem to imply that it is a coil structure (see Table I). The lowest-energy conformation of $(Gly)_{10}$ in aqueous solution, on the other hand, has a left-hand α-helix from residue 5 to residue 7 (see Table II). Both conformations are compared in Figs. 2(e) and 2(f). It is apparent that they are rather compact and round. A close examination of the structures revealed that both conformations have β-sheet-like characteristics. Namely, they are stabilized by the intrachain backbone hydrogen bonds that are found in β-sheet structures. For the lowest-energy conformation in the gas phase [Fig. 2(e)], the carbonyl oxygen (and amide nitrogen) of residue 3 and amide nitrogen (and carbonyl oxygen) of residue 6 are hydrogen bonded. There also exist two such hydrogen bonds that connect residues 4 and 9. For the lowest-energy conformation in aqueous solution [Fig. 2(f)], there exist two such hydrogen bonds between residue 2 and residue 9.

C. Distributions of backbone dihedral angles

The lowest-energy conformations discussed in detail in the last subsection correspond to the conformations at low temperatures. We next calculated the distributions of the backbone dihedral angles of homo-oligomers in aqueous solution as a function of temperature in order to study how these conformations change as the temperature is raised. In Fig. 3 we show the distributions of the backbone dihedral angles of the sixth residue of $(Ala)_{10}$, $(Val)_{10}$, and $(Gly)_{10}$ in aqueous solution at $T=200$ K and 1000 K. For $(Ala)_{10}$ there is a single peak at $T=200$ K and this peak corresponds to the dihedral angle of a right-hand α-helix state. The distributions for other residues have essentially the same peakexcept for the terminal residues. This implies that around $T=200$ K there exists only a completely helical conformation. At $T=1000$ K the distributions are widely spread, implying the large thermal fluctuations. These results suggest the existence of a transition between an ensemble of well-defined compact conformations (ideal α-helix state) and random-coil structures.

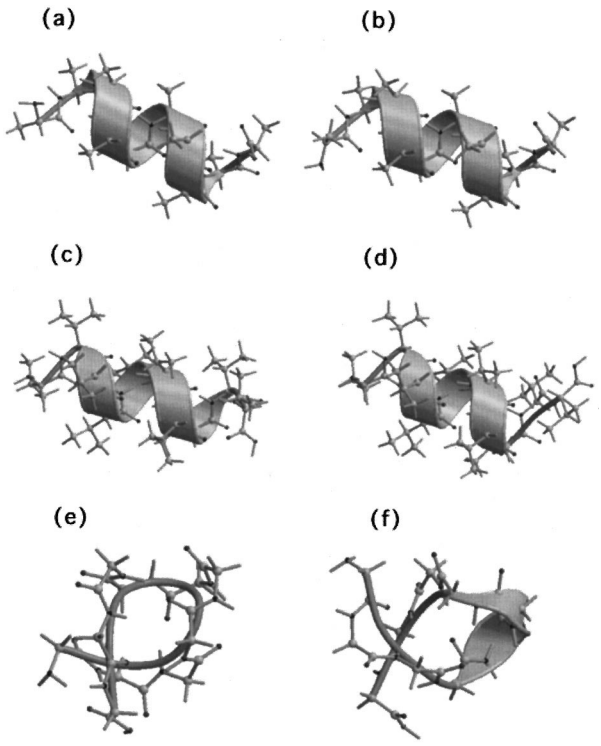

FIG. 2. The lowest-energy conformations of $(\mathrm{Ala})_{10}$ [(a) and (b)], $(\mathrm{Val})_{10}$ [(c) and (d)], and $(\mathrm{Gly})_{10}$ [(e) and (f)] in gas phase and in aqueous solution, respectively.

For $(\mathrm{Val})_{10}$ [see Fig. 3(b)], the results are similar to those for $(\mathrm{Ala})_{10}$ in the sense that at $T=200$ K there is a dominant peak in the distribution that corresponds to the lowest-energy conformation and that at $T=1000$ K the distribution is widely spread (random-coil state).

The situation is slightly different for $(\mathrm{Gly})_{10}$. Since glycine dose not have a side chain, $(\mathrm{Gly})_{10}$ is much more flexible than the other two homo-oligomers. We observe two dominant peaks in the distributions at $T=200$ K [see Fig. 3(c)], which implies that this temperature is not low enough to single out the lowest-energy conformation. The large flexibility of $(\mathrm{Gly})_{10}$ is most clearly seen in the distributions at $T=1000$ K.

D. Average energy and specific heat

We investigate how each energy term varies as a function of temperature. We calculated the average values of total energy and its component terms of the homo-oligomers, $(\mathrm{Ala})_{10}$, $(\mathrm{Val})_{10}$, and $(\mathrm{Gly})_{10}$, as a function of temperature in the gas phase and in aqueous solution. In Fig. 4 the results are shown. For homo-alanine in gas phase, all the conformational energy terms increase monotonically as temperature increases. The changes of each component's terms are very small, except for the Lennard-Jones term, E_V, indicating that E_V plays an important role in the folding of homo-alanine.[25] The Lennard-Jones term, in principle, has contributions from all possible pairs of atoms, while the hydrogen-bond term is only from the donor-acceptor pairs. Therefore, the Lennard-Jones term is responsible for the large conformational

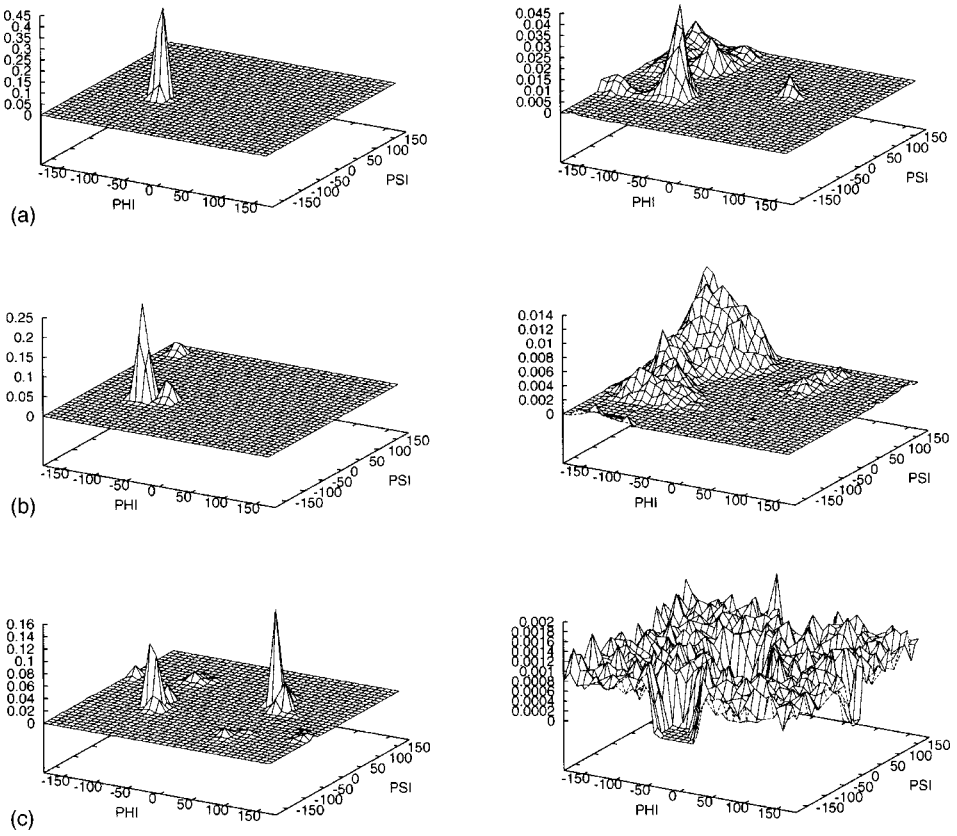

FIG. 3. Distributions of the backbone dihedral angles of $(\mathrm{Ala})_{10}$ (a), $(\mathrm{Val})_{10}$ (b), and $(\mathrm{Gly})_{10}$ (c) in aqueous solution as a function of temperature. The results for the sixth residue from the *N*-terminus at $T=200$K (left-hand side) and 1000 K (right-hand side) are shown. The values for each case were calculated from one multicanonical production run of 1 000 000 MC sweeps.

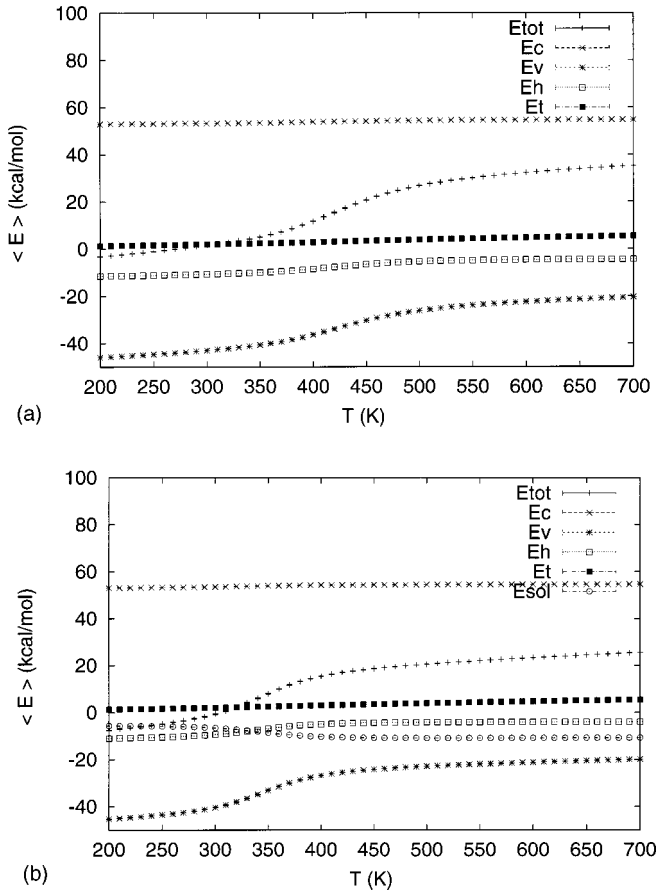

FIG. 4. Average total energy E_{TOT} and averages of its component terms, Coulomb energy E_C, hydrogen-bond energy E_H, Lennard-Jones energy E_V, torsion energy E_T, and solvation free energy E_{SOL} (only for the case in aqueous solution) for homo-alanine as a function of temperature T in gas phase (a) and in aqueous solution (b). The values for each case were calculated from one multicanonical production run of 1 000 000 MC sweeps.

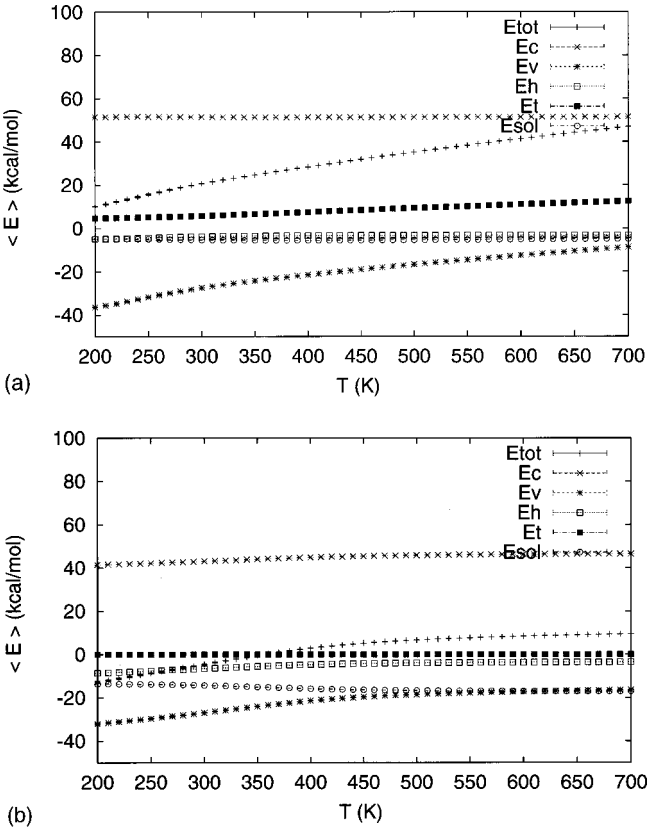

FIG. 5. Average total energy E_{TOT} and averages of its component terms, Coulomb energy E_C, hydrogen-bond energy E_H, Lennard-Jones energy E_V, torsion energy E_T, and solvation free energy E_{SOL} (only for the case in aqueous solution) for homo-valine (a) and homo-glycine (b) as a function of temperature T in aqueous solution. The values for each case were calculated from one multicanonical production run of 1 000 000 MC sweeps.

changes from a random coil to an ideal helix. Once the (near-)helical conformation is obtained, the hydrogen-bond term further stabilizes it.

In aqueous solution the overall behaviors of the conformational energy terms are very similar to those in gas phase. The solvation term, on the other hand, decreases monotonically as temperature increases. These results imply that the solvation term favors random-coil conformations, while the conformational terms favor helical conformations. This is because the solvation free energy of conformations at high temperatures (random coil) is lower than that at low temperatures (α-helix conformations) and because the conformational energies at high temperatures (random coil) are higher than those at low temperatures (α-helix conformations). The rapid changes (decrease for the solvation term and increase for the rest of the terms) of all the average values occur at the same temperature (around at 420 K in gas phase and 340 K in solvent). This suggests the existence of a certain phase transition. As was shown in gas phase in Ref. 25 and is discussed below for the case with solvent, this transition indeed corresponds to a helix-coil transition. It is interesting to note that the helix-coil transition in solvent is the result of two conflicting effects between conformational energy and solvation free energy, which lowers the helix-coil transition temperature compared to the gas-phase value.

In Fig. 5 the average energy values as a function of temperature for $(\mathrm{Val})_{10}$ and $(\mathrm{Gly})_{10}$ in aqueous solution are shown. For homo-valine and homo-glycine, the behaviors of the conformational energy terms in aqueous solution are quite similar to those in gas phase (data not shown) and all the conformational terms increase monotonically as temperature increases. The solvation term, on the other hand, decreases monotonically as a function of temperature. For homo-valine and homo-glycine, the change in total energy is not as conspicuous as in homo-alanine. Hence, the helix-coil transition in homo-valine and homo-glycine is not as clearly observed as in homo-alanine. We calculated the specific heat for the homo-oligomers as a function of temperature. The specific heat here is defined by the following equation:

$$C(T)=\beta^2\frac{\langle E_{\mathrm{TOT}}^2\rangle_T-\langle E_{\mathrm{TOT}}\rangle_T^2}{N}, \tag{7}$$

where $N(=10)$ is the number of residues in the oligomer. In Fig. 6 we show the specific heat as a function of temperature

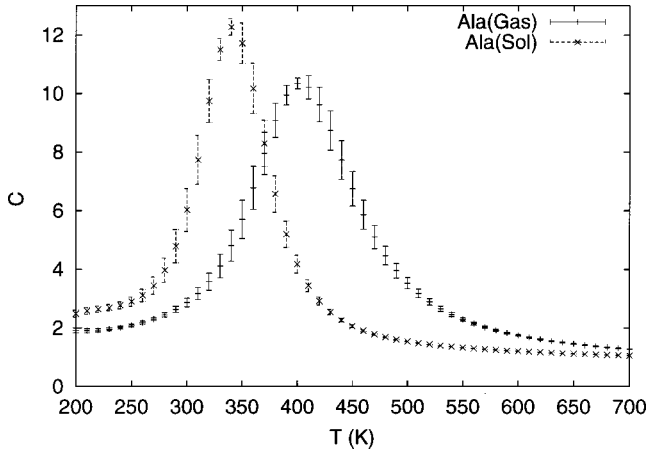

FIG. 6. Specific heat C as a function of temperature T for $(\mathrm{Ala})_{10}$ in gas phase and in aqueous solution. The values for each case were calculated from one multicanonical production run of 1 000 000 MC sweeps.

for $(\mathrm{Ala})_{10}$ in gas phase and in aqueous solution. We observe sharp peaks in the specific heat for both environments, which implies the existence of some phase transition. The temperatures at the peak, transition temperatures, are $T_c \approx 420$ and 340 K in gas phase and in aqueous solution, respectively. The transition temperature T_c for $(\mathrm{Ala})_{10}$ in aqueous solution is thus significantly lower than that in the gas phase and much closer to experimentally relevant temperatures.

E. Helicity of homo-oligomers

We calculated the average number of helical residues $\langle n \rangle_T$ in a conformation as a function of temperature. In Fig. 7 we show the average helicity $\langle n \rangle_T$ as a function of temperature for $(\mathrm{Ala})_{10}$, $(\mathrm{Val})_{10}$, and $(\mathrm{Gly})_{10}$ in gas phase and in aqueous solution. The average helicity tends to decrease monotonically as the temperature increases because of the increased thermal fluctuations.

At $T=200$ K, $\langle n \rangle_T$ for homo-alanine in both environments are 8. If we neglect the terminal residues, in which α-helix tends to be frayed, $n=8$ corresponds to the maximal helicity, and the conformation can be considered completely helical. Then homo-alanine in both environments is in an ideal helical structure at $T=200$ K. Around the room temperature, $(\mathrm{Ala})_{10}$ in gas phase and in aqueous solution is still substantially helical ($\approx 80\%$ and $\approx 70\%$ helicity, respectively). This is consistent with the fact that alanine is a strong helix former. For homo-alanine, at transition temperatures (around 420 K in gas phase and 340 K in aqueous solution), $\langle n \rangle_T$ is 5 (50% helicity). This implies that the phase transitions observed above by the peak in specific heat are indeed helix-coil transitions between an ideal helix and a random coil. Hence, as far as the helix-coil transition is concerned, the solvation effects do not alter the nature of the transition and just shift the transition temperature.

As is shown in Fig. 7, the average helicity of homo-glycine in aqueous solution is similar to that in gas phase and is very low ($<20\%$ helicity). It is apparent that homo-glycine does not favor helix formation over the whole temperature range in both environments. The average helicity of

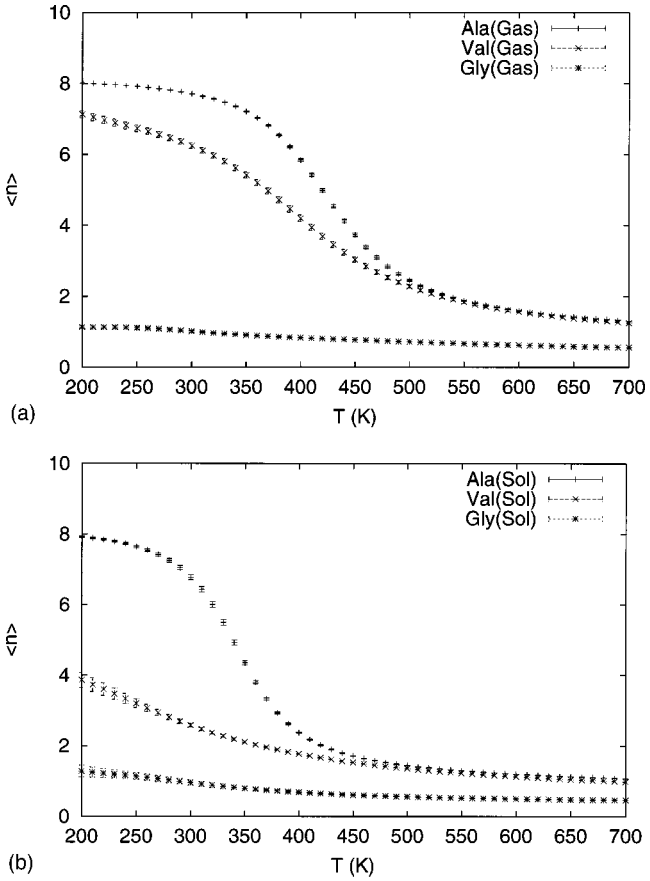

FIG. 7. Average helicity $\langle n \rangle_T$ as a function of temperature T for $(\mathrm{Ala})_{10}$, $(\mathrm{Val})_{10}$, and $(\mathrm{Gly})_{10}$ in gas phase (a) and in aqueous solution (b). The values for each case were calculated from one multicanonical production run of 1 000 000 MC sweeps.

homo-valine in aqueous solution is lower than that in gas phase and is less than 40% helicity for a wide range of temperatures. The percent helicity lies between that of alanine and glycine. All these results are consistent with the fact that alanine is a helix former and glycine is a helix breaker, while valine comes in between the two.

We next calculated the percent helicity as a function of residue number for the three homo-oligomers. The results at $T=200$ and 1000 K are shown in Fig. 8. The percent helicity is in general lower at the terminal residues than in the internal residues (i.e., fraying is observed) for a wide range of temperatures, because the dihedral angles of terminal residues are less constrained than those of the internal residues. For $(\mathrm{Ala})_{10}$ and $(\mathrm{Val})_{10}$, the internal residues are significantly helical at $T=200$ K. For $(\mathrm{Gly})_{10}$, the residues prefer a coil state rather than helix for a wide range of temperatures. We do observe fraying of the termini for all cases. The contrast is most outstanding for $(\mathrm{Ala})_{10}$ because it has high helicity. The increase of fraying as the temperature is raised is also clearly seen for $(\mathrm{Ala})_{10}$.

F. Zimm-Bragg s and σ parameters

Finally, the helix propagation parameter s and nucleation parameter σ of the Zimm-Bragg model[52] were calculated as a function of temperature.

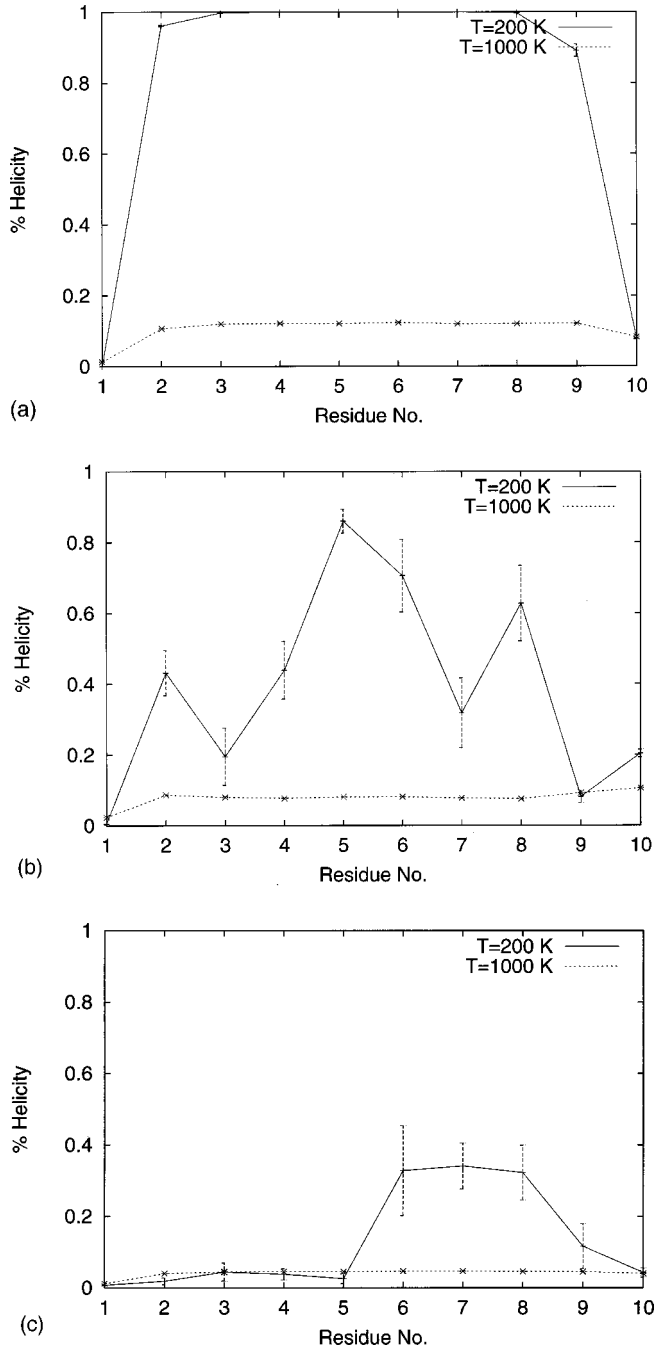

FIG. 8. Percent helicity as a function of residue number for $(\mathrm{Ala})_{10}$ (a), $(\mathrm{Val})_{10}$ (b), and $(\mathrm{Gly})_{10}$ (c) in aqueous solution at 200 and 1000 K. The values for each case were calculated from one multicanonical production run of 1 000 000 MC sweeps.

According to the Zimm-Bragg model, the average number of helical residues $\langle n\rangle_T$ and the average length $\langle \ell\rangle_T$ of a helical segment are given for large N by

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

$$\langle \ell\rangle_T=1+\frac{2s}{1-s+\sqrt{(1-s)^2+4s\sigma}}, \tag{8}$$

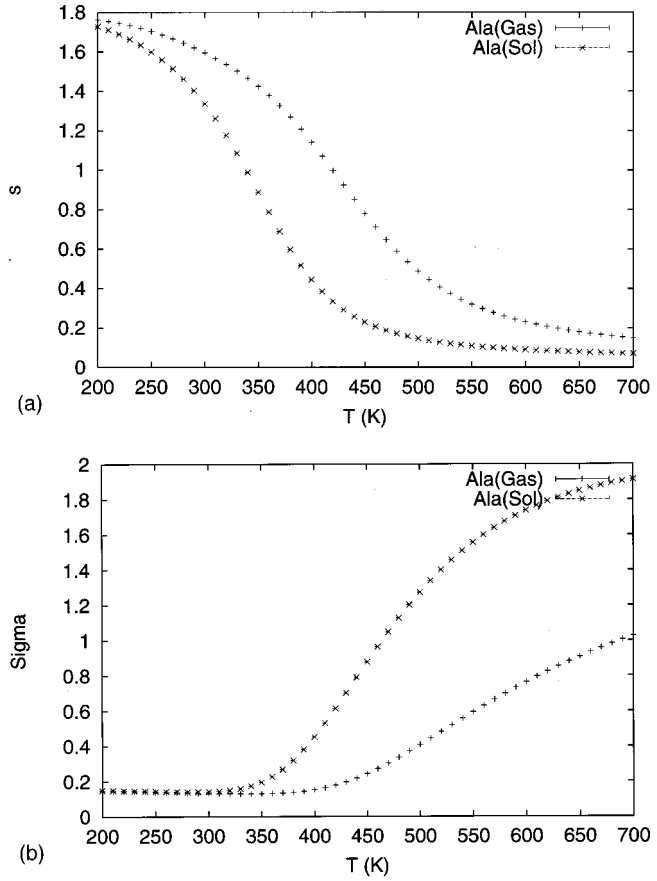

FIG. 9. Helix propagation parameter s (a) and nucleation parameter σ (b) of the Zimm-Bragg model as a function of temperature T for alanine in gas phase and in aqueous solution. The values for each case were calculated from one multicanonical production run of 1 000 000 MC sweeps.

where N is the number of residues. Here, the length ℓ of a helical segment is defined by the number of successive residues which are in the helix configuration. Note that from Eq. (8) the temperature where $s=1$ holds corresponds to 50% helicity, which in turn gives the helix-coil transition temperature. From these equations with the values of $\langle n\rangle/N$ and $\langle \ell\rangle$ calculated from the multicanonical production runs, one can obtain estimates of s and σ parameters.

In Fig. 9 we show the s and σ values for alanine as a function of temperature. The s parameter monotonically decreases as the temperature increases. It is shown that $s(\mathrm{Ala})$ in aqueous solution decreases more rapidly than that in gas phase as the temperature increases. As noted above, the helix-coil transition temperature T_c can also be identified as the temperature where $s=1$ holds (i.e., 50% helicity) in Fig. 9(a). It is 420 K in gas phase and 340 K in aqueous solution, in agreement with the previous estimates by the peaks of specific heat (see Fig. 6).

As is clear from Fig. 9(b), in the helix phase ($T<T_c$) the σ parameter for alanine is small and constant, but in the random-coil phase ($T>T_c$) σ starts to grow as temperature increases. This growth of σ value reflects the increased thermal fluctuations that prevent the formation of a long helix. That is, below T_c cooperativety for helix formation wins over thermal fluctuations, but above T_c thermal fluctuations win and no long helices can be formed.

TABLE III. The helix propagation parameter s and nucleation parameter σ of Zimm-Bragg model for alanine (Ala), valine (Val), and glycine (Gly) at $T=273$ K in gas phase (Gas) and in aqueous solution (Sol), together with the experimental values (Expt.) (Ref. 12).

Amino acid	s (Gas)	s (Sol)	s (Expt.)
Ala	1.67	1.51	1.07∼2.19
Val	1.36	0.31	0.2∼0.93
Gly	0.17	0.21	0.02∼0.57

We next make a comparison between the s values obtained by the present simulations and those by experiments.[12] In Table III we list s for the three amino acids at $T=273$ K in gas phase and in aqueous solution obtained by simulations together with those by experiments. The values in aqueous solution are slightly smaller than those in gas phase but not too much different except for valine. The s value for valine is significantly lower in aqueous solution than in gas phase. One finds that our values in aqueous solution are in remarkable agreement with those determined by experiments.[12]

IV. CONCLUSIONS

In the present work we have performed multicanonical Monte Carlo simulations to study helix-coil transitions of homo-oligomers in aqueous solution. Homo-oligomers of length 10 are considered for three characteristic amino acids, alanine (helix former), valine (helix indifferent) and glycine (helix breaker). Various thermodynamic quantities as a function of temperature were calculated and compared with those obtained in gas phase. It was found that solvation effects reduce helix formation slightly, but that overall, nature of the helix-coil transition is unaltered by the addition of solvent. For instance, for homo-alanine we still observed in solvent a helix-coil transition from a completely helical conformation at low temperatures to random-coil conformations at high temperatures. This agreement can presumably be accounted for by the fact that we dealt with only nonpolar amino acids, with which we can avoid the complications of electrostatic interactions between the homo-oligomer and solvent. It was also shown that the transition temperature for homo-alanine gets significantly lowered in aqueous solution compared with that in gas phase. This indeed rectified the unrealistically high value in gas phase, which was the most serious discrepancy between theory and experiments found in the previous work.

ACKNOWLEDGMENTS

The simulations were performed on the computers at the Computer Center of the Institute for Molecular Science. This work was supported, in part, by grants from Research Fellowships of the Japan Society for the Promotion of Science for Young Scientists and from the Research for the Future Program of the Japan Society for the Promotion of Science (JSPS-RFTF98P01101).

[1] M. Sueki, S. Lee, S. P. Powers, J. B. Denton, Y. Konishi, and H. A. Scheraga, Macromolecules **17**, 148 (1984).
[2] S. Padmanabhan, S. Marqusee, T. Ridgeway, T. M. Laue, and R. L. Baldwin, Nature (London) **344**, 268 (1990).
[3] G. Merutka, W. Lipton, W. Shalango, S.-H. Park, and E. Stellwagen, Biochemistry **29**, 121 (1990).
[4] P. C. Lyu, M. I. Liff, L. A. Marky, and N. R. Kallenbach, Science **250**, 669 (1990).
[5] K. T. O'Neil and W. F. DeGrado, Science **250**, 646 (1990).
[6] J. Wojcik, K. H. Altman, and H. A. Scheraga, Biopolymers **30**, 121 (1990).
[7] E. Stellwagen, D.-H. Park, W. Shalango, and A. Jain, Biopolymers **32**, 1193 (1992).
[8] A. Horovitz, J. M. Matthews, and A. R. Fersht, J. Mol. Biol. **227**, 560 (1992).
[9] M. Blaber, X. Zhang, and B. W. Matthews, Science **260**, 1637 (1993).
[10] A. Chakrabartty, T. Kortemme, and R. L. Baldwin, Protein Sci. **3**, 843 (1994).
[11] J. K. Myers, C. N. Pace, and J. M. Scholtz, Proc. Natl. Acad. Sci. USA **94**, 2833 (1997).
[12] A. Chakrabartty and R.L. Baldwin, in *Protein Folding: In Vivo and In Vitro*, edited by J. Cleland and J. King (American Chemical Society, Washington, D.C., 1993), p. 166.
[13] T. P. Creamer and G. D. Rose, Proc. Natl. Acad. Sci. USA **89**, 5937 (1992).
[14] S.-S. Sung, Biophys. J. **68**, 826 (1995).
[15] J. Tirado-Rives and W. L. Jorgensen, Biochemistry **30**, 3864 (1991).
[16] D. J. Tobias and C. L. Brooks III, Biochemistry **30**, 6059 (1991).
[17] K. V. Soman, A. Karimi, and D. A. Case, Biopolymers **31**, 1351 (1991).
[18] V. Daggett and M. Levitt, J. Mol. Biol. **223**, 1121 (1992).
[19] M. Takano, T. Takahashi, and K. Nagayama, Phys. Rev. Lett. **80**, 5691 (1998).
[20] M. Takano, T. Yamato, J. Higo, A. Suyama, and K. Nagayama, J. Am. Chem. Soc. **121**, 605 (1999).
[21] H. Kawai, Y. Okamoto, M. Fukugita, T. Nakazawa, and T. Kikuchi, Chem. Lett. **213**, (1991); Y. Okamoto, M. Fukugita, T. Nakazawa, and H. Kawai, Protein Eng. **4**, 639 (1991).
[22] M. Fukugita, T. Nakazawa, H. Kawai, and Y. Okamoto, Chem. Lett. **1991**, 1279.
[23] Y. Okamoto, Proteins: Struct., Funct., Genet. **19**, 14 (1994).
[24] Y. Okamoto, Biopolymers **34**, 529 (1994).
[25] Y. Okamoto, U. H. E. Hansmann, and T. Nakazawa, Chem. Lett. **1995**, 391; J. Phys. Chem. **99**, 11276 (1995).
[26] U. H. E. Hansmann and Y. Okamoto, J. Chem. Phys. **110**, 1267 (1999).
[27] A. Mitsutake and Y. Okamoto, Chem. Phys. Lett. **309**, 95 (1999).
[28] B. A. Berg and T. Neuhaus, Phys. Lett. B **267**, 249 (1991); Phys. Rev. Lett. **68**, 9 (1992).
[29] U. H. E. Hansmann and Y. Okamoto, J. Comput. Chem. **14**, 1333 (1993).
[30] U. H. E. Hansmann and Y. Okamoto, J. Phys. Soc. Jpn. **63**, 3945 (1994); Physica A **212**, 415 (1994).
[31] M. H. Hao and H. A. Scheraga, J. Phys. Chem. **98**, 4940 (1994).
[32] M. H. Hao and H. A. Scheraga, J. Phys. Chem. **98**, 9882 (1994).
[33] A. Kidera, Proc. Natl. Acad. Sci. USA **92**, 9886 (1995).
[34] A. Kolinski, W. Galazka, and J. Skolnick, Proteins **26**, 271 (1996).
[35] N. Urakami and M. Takasu, J. Phys. Soc. Jpn. **65**, 2694 (1996).
[36] S. Kumar, P. Payne, and M. Vásquez, J. Comput. Chem. **17**, 1269 (1996).
[37] U. H. E. Hansmann, Y. Okamoto, and F. Eisenmenger, Chem. Phys. Lett. **259**, 321 (1996).
[38] N. Nakajima, H. Nakamura, and A. Kidera, J. Phys. Chem. **101**, 817 (1997).
[39] H. Noguchi and K. Yoshikawa, Chem. Phys. Lett. **278**, 184 (1997).
[40] U. H. E. Hansmann and F. Eisenmenger, J. Phys. Chem. B **101**, 3304 (1997).
[41] J. Higo, N. Nakajima, H. Shirai, A. Kidera, and H. Nakamura, J. Comput. Chem. **18**, 2086 (1997).
[42] A. Kolinski, W. Galazaka, and J. Skolnick, J. Chem. Phys. **108**, 2608 (1998).
[43] C. Bartels and M. Karplus, J. Phys. Chem. B **102**, 865 (1998).
[44] Y. Iba, G. Chikenji, and M. Kikuchi, J. Phys. Soc. Jpn. **67**, 3327 (1998).
[45] N. Nakajima, Chem. Phys. Lett. **288**, 319 (1998).
[46] M. H. Hao and H. A. Scheraga, J. Mol. Biol. **277**, 973 (1998).
[47] H. Shirai, N. Nakajima, J. Higo, A. Kidera, and H. Nakamura, J. Mol. Biol. **278**, 481 (1998).
[48] M. Schaefer, C. Bartels, and M. Karplus, J. Mol. Biol. **284**, 835 (1998).
[49] C. Bartels, R. H. Stote, and M. Karplus, J. Mol. Biol. **284**, 1641 (1998).

[50] U. H. E. Hansmann and Y. Okamoto, J. Phys. Chem. B **102**, 653 (1998); **103**, 1595 (1999).
[51] A. Mitsutake, U. H. E. Hansmann, and Y. Okamoto, J. Mol. Graphics Mod. **16**, 226 (1998).
[52] B. H. Zimm and J. K. Bragg, J. Chem. Phys. **31**, 526 (1959).
[53] S. Ihara, T. Ooi, and S. Takahashi, Biopolymers **21**, 131 (1982).
[54] K. R. Shoemaker, P. S. Kim, E. J. York, J. M. Stewart, and R. L. Baldwin, Nature (London) **326**, 563 (1987).
[55] F. A. Momany, R. F. McGuire, A. W. Burgess, and H. A. Scheraga, J. Phys. Chem. **79**, 2361 (1975).
[56] G. Némethy, M. S. Pottle, and H. A. Scheraga, J. Phys. Chem. **87**, 1883 (1983).
[57] M. J. Sippl, G. Némethy, and H. A. Scheraga, J. Phys. Chem. **88**, 6231 (1984).
[58] T. Ooi, M. Oobatake, G. Némethy, and H. A. Scheraga, Proc. Natl. Acad. Sci. USA **84**, 3086 (1987).
[59] M. Masuya (in preparation).
[60] F. Eisenhaber, P. Lijnzaad, P. Argos, C. Sander, and M. Scharf, J. Comput. Chem. **16**, 273 (1995).
[61] B. A. Berg, Int. J. Mod. Phys. C **3**, 1083 (1992).
[62] B. A. Berg, Nucl. Phys. B (Proc. Suppl.) **63 A–C**, 982 (1998).
[63] A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **61**, 2635 (1988).
[64] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).