Theoretical studies of transition states by the multioverlap molecular dynamics methods

Satoru G. Itoh[a] and Yuko Okamoto[b]
Department of Physics, School of Science, Nagoya University, Furo-cho, Chikusa-ku Nagoya, Aichi 464-8602, Japan

(Received 17 November 2005; accepted 5 January 2006; published online 9 March 2006)

The multioverlap molecular dynamics method gives a flat probability distribution in the multidimensional dihedral-angle-distance space, where the dihedral-angle distance of a configuration with respect to a reference state gives a measure for structural similarity. Hence, this method realizes a random walk among specific configurations in the multidimensional dihedral-angle-distance space at a constant temperature and explores widely in the configurational space. We applied the multioverlap molecular dynamics method to a pentapeptide, Met-enkephalin, in gas phase as a test system. Comparing the results of this method with those of the conventional canonical and multicanonical algorithms, we demonstrate its effectiveness. Furthermore, from the detailed free-energy landscape obtained from the results of the multioverlap molecular dynamics simulation, we obtain the transition state between two specific reference configurations of Met-enkephalin. We also deduce the transition pathway between the two specific reference configurations. © *2006 American Institute of Physics*. [DOI: 10.1063/1.2171189]

I. INTRODUCTION

In order to understand the protein folding problem, it is essential that the detailed free-energy landscape of the protein system is obtained. By analyzing the free-energy landscape, we can deduce the folding pathways and the stability of any structures of the protein. Furthermore, the transition state between two specific stable states can also be discovered. Exploring the transition state, we can gain information about state transitions. From a point of view of molecular modeling or drug design, moreover, it is also very important that the transition state is found. Accordingly, many efforts are devoted to obtain the detailed free-energy landscape by computer simulations.

Canonical-ensemble simulations[1–6] are widely employed in computer simulations. In the canonical ensemble at a fixed temperature, the probability distribution of the potential energy is given by the product of the density of states and the Boltzmann weight factor, and we have a bell-shaped probability distribution of the potential energy. However, this simulation method is not suitable for applying to complex systems such as proteins. Because such complex systems have many local-minimum free-energy states, canonical-ensemble simulations tend to get trapped at the local-minimum states. At low temperatures, in particular, the usual canonical-ensemble simulations cannot realize efficient sampling in the configurational space. This is because in canonical simulations energy fluctuations are small at a low temperature and energy barriers cannot be overcome. Therefore, if we employ the conventional canonical-ensemble method in complex systems, we may estimate inaccurately various thermodynamic quantities at low temperatures. To overcome this difficulty, the generalized-ensemble algorithms have been proposed (for a review, see Ref. 7).

The multicanonical algorithm[8–11] is perhaps one of the most well-known methods among the generalized-ensemble algorithms. In the multicanonical ensemble, the probability distribution of the potential energy is expressed by the product of the density of states and a non-Boltzmann weight factor, which we refer to as the multicanonical weight factor, and we have a flat probability distribution of the potential energy. Therefore, multicanonical-ensemble simulations realize a free random walk in the potential energy space and overcome energy barriers. Because the multicanonical simulations do not get trapped in local-minimum states, we need much less simulation time to get an accurate free-energy landscape than conventional canonical simulations. Therefore, the application of the multicanonical algorithm to the protein folding was proposed.[12] Since then there have been many works based on this method and its variants in protein and related systems (for a review, see Ref. 13).

The multicanonical algorithm aims at achieving a wide range sampling in the configurational space. However, because of the very nature of this algorithm, it is difficult to focus on specific configurations. Consequently, the free-energy landscape around or among these configurations of interest may be incorrectly estimated in multicanonical-ensemble simulations. To understand protein folding, we must investigate the stability of certain configurations and the transition states among these configurations. Accordingly, the detailed free-energy landscape in the neighborhood of specific configurations is necessary. Recently, a new algorithm, which is a generalization of the multicanonical algorithm and is referred to as the multioverlap algorithm,[14] was proposed to overcome this difficulty: This method focuses on

[a]Electronic mail: itoh@tb.phys.nagoya-u.ac.jp
[b]Electronic mail: okamoto@phys.nagoya-u.ac.jp

specific configurations but can still overcome potential energy barriers. Here, an overlap of a configuration is a measure of structural similarity to a reference configuration. In the multioverlap ensemble, the probability distribution is expressed by the product of the density of states and a non-Boltzmann weight factor, which we refer to as the multioverlap weight factor, and we have a flat probability distribution in the overlap space. Consequently, the multioverlap ensemble realizes a random walk in the overlap space and efficiently samples the conformational space, and we can obtain the detailed free-energy landscape in the neighborhood of specific configurations.

A Monte Carlo (MC) version of this algorithm was proposed in Ref. 14. In general, for chain molecules such as proteins, MC simulations are mostly based on the updates of dihedral angles, not Cartesian coordinates, in order to maintain the covalent geometry of such chain molecules. A small update of a single dihedral angle can then result in a large motion of the molecule, and the trial MC step will be almost always rejected. Therefore, in many particle systems such as proteins in solution, MC algorithm would sample inefficiently the conformational space, and it is difficult to estimate correctly the free-energy landscape.

To avoid such problems in the MC simulations, the molecular dynamics (MD) algorithm is often employed. For instance, the MD version of multicanonical algorithm was developed in Refs. 10 and 11. We also proposed a MD version of the multioverlap method in Ref. 15. In this article, we present detailed comparisons of the multioverlap MD algorithm with the conventional canonical MD method and the multicanonical MD method, taking Met-enkephalin in gas phase as a benchmark system. Moreover, from the detailed free-energy landscape obtained from the results of the multioverlap MD simulation, we predict a transition pathway between two specific configurations of Met-enkephalin.

In Sec. II we summarize the formulation of the multioverlap MD algorithm. The details of the condition of various simulation methods are given in Sec. III. We present the results of the application of these methods to Met-enkephalin in Sec. IV. Section V is devoted to conclusions.

II. FORMULATION OF THE MULTIOVERLAP MOLECULAR DYNAMICS ALGORITHM

In this section we explain the formulation of the multioverlap MD algorithm.[15] The dihedral-angle distance[14] is defined as a reaction coordinate in Sec. II A. In Sec. II B we introduce a non-Boltzmann weight factor, which we refer to as the multioverlap weight factor. The multioverlap weight factor realizes a constant probability distribution in a multidimensional dihedral-angle-distance space. In Sec. II C we present the equations of motion in the multioverlap ensemble. We present details of the updating procedure of the multioverlap weight factor in Sec. II D. In Sec. II E we explain the reweighting techniques.[16,17] Utilizing the reweighting techniques, we can calculate appropriate physical quantities and obtain the free-energy landscape at any temperature.

A. Definition of dihedral-angle distance

In order to explore transition states among any reference configurations, we would like to perform a simulation which focuses on the reference configurations and does not get trapped in local-minimum states. While a free random walk in the potential energy space is realized in the multicanonical MD method, we would like to perform a random walk in some reaction coordinate so that the reference configuration can be efficiently sampled. In the multioverlap algorithm,[14] the overlap is introduced as this reaction coordinate. The overlap O with respect to a reference configuration is defined as follows:[14,18]

$$O = 1 - d, \tag{1}$$

where d is the dihedral-angle distance given by

$$d = \frac{1}{n\pi}\sum_i d_a(v_i, v_i^0). \tag{2}$$

Here, n is the total number of dihedral angles, v_i is the dihedral angle i, and v_i^0 is the dihedral angle i of the reference configuration. The distance $d_a(v_i, v_i^0)$ between two dihedral angles is defined by

$$d_a(v_i, v_i^0) = \min(|v_i - v_i^0|, 2\pi - |v_i - v_i^0|). \tag{3}$$

The dihedral-angle distance d in Eq. (2) takes on a value in the range $0 \leqslant d \leqslant 1$. From Eq. (1), correspondingly, $0 \leqslant O \leqslant 1$. In particular, if we consider a system at infinite temperature ($T_0 = \infty$), the average values of the dihedral-angle distance d and the overlap O are $\frac{1}{2}$. This is because the distance $d_a(v_i, v_i^0)$ in Eq. (3) will have a uniform distribution in the range between 0 and π at $T_0 = \infty$. Furthermore, if $d = 0$ ($O = 1$), all dihedral angles are coincident with those of the reference configuration. The dihedral-angle distance (the overlap) is thus an indicator of how similar the conformation is to the reference conformation. As one can see in Eq. (1), the dihedral-angle distance d is equivalent to the overlap O. Hereafter, we employ the dihedral-angle distance d as the reaction coordinate in the multioverlap algorithm.

B. Constant probability distribution in dihedral-angle-distance space

We want the simulation to realize a random walk in a multidimensional dihedral-angle-distance space. In other words, the simulation needs to have a constant probability distribution with the dihedral-angle-distance reaction coordinates. In the case of canonical ensemble at a constant temperature T_0, the probability distribution P_c of potential energy E is represented by the product of the density of states $n(E)$ and the Boltzmann weight factor W_c,

$$P_c(E; T_0) = n(E) W_c(E; T_0) = n(E) e^{-\beta_0 E}, \tag{4}$$

where β_0 is given by $\beta_0 = 1/k_B T_0$ (k_B is the Boltzmann constant). From Eq. (4) the probability distribution P_c is not constant and takes a much smaller value in the high-energy, making the probability for the escape from local-minimum

states small. Consequently, canonical-ensemble simulations would get trapped in the local-minimum states at a low temperature. In the multioverlap ensemble at a constant temperature T_0, on the other hand, the probability distribution is determined by the following non-Boltzmann weight factor, which we refer to as the multioverlap weight factor:

$$W_{\mathrm{muov}}(d;E) = e^{-\beta_0 E_{\mathrm{muov}}(d;E)}, \tag{5}$$

where $E_{\mathrm{muov}}(d;E)$ is the "multioverlap potential energy" defined by

$$E_{\mathrm{muov}}(d;E) = E - k_{\mathrm{B}}T_0 f(d). \tag{6}$$

The function $f(d)$ is the dimensionless free energy at dihedral-angle distance d.

The generalization to the multidimensional dihedral-angle-distance space is straightforward and the multioverlap weight factor is given by

$$W_{\mathrm{muov}}(d_1, \ldots, d_N;E) = e^{-\beta_0 E_{\mathrm{muov}}(d_1,\ldots,d_N;E)} \tag{7}$$

and

$$E_{\mathrm{muov}}(d_1, \ldots, d_N;E) = E - k_{\mathrm{B}}T_0 f(d_1, \ldots, d_N), \tag{8}$$

where N is the number of the reference configurations and d_i is the dihedral-angle distance with respect to reference configuration $i(i=1,\ldots,N)$. The function $f(d_1,\ldots,d_N)$ is the dimensionless free energy with the fixed values of dihedral-angle distances $d_1,\ldots,d_N$. The dimensionless free energy $f(d_1,\ldots,d_N)$ is defined so that the probability distribution P_{muov} is flat:

$$\begin{aligned} P_{\mathrm{muov}}(d_1, \ldots, d_N) &= \int dE n(d_1, \ldots, d_N;E) \\ &\quad \times W_{\mathrm{muov}}(d_1, \ldots, d_N;E) \\ &= \int dE n(d_1, \ldots, d_N;E) e^{-\beta_0 E + f(d_1,\ldots,d_N)} \\ &\equiv \text{const}, \end{aligned} \tag{9}$$

where $n(d_1,\ldots,d_N;E)$ is the density of states. Thus, we are able to perform simulations, which realize a random walk in the multidimensional dihedral-angle-distance space.

In this paper we use only the two-dimensional version of these methods. Namely, $N=2$ in Eqs. (7)–(9). We can then perform a simulation which is focused on two specific reference configurations. Accordingly, we can explore a transition state between the two reference configurations. We will only deal with the two-dimensional version of these methods hereafter.

C. Equations of motion in multioverlap MD simulations

The canonical MD simulations are performed by solving numerically the following equations of motion with a Gaussian thermostat,[2,3] which we refer to as the Gaussian constraint method,

$$\begin{aligned} \dot{\boldsymbol{q}}_i &= \frac{d\boldsymbol{q}_i}{dt} = \frac{\boldsymbol{p}_i}{m_i}, \\ \dot{\boldsymbol{p}}_i &= \boldsymbol{F}_i - \zeta_c \boldsymbol{p}_i, \end{aligned} \tag{10}$$

where m_i, $\boldsymbol{q}_i$, and $\boldsymbol{p}_i$ are the mass, coordinate vector, and momentum vector of atom i. The force $\boldsymbol{F}_i$ acting on atom i is given by

$$\boldsymbol{F}_i = -\frac{\partial E}{\partial \boldsymbol{q}_i}. \tag{11}$$

The coefficient ζ_c is chosen so as to guarantee that the total kinetic energy is constant,

$$\zeta_c = \frac{\sum_i \boldsymbol{F}_i \cdot \dot{\boldsymbol{q}}_i}{2\sum_i \boldsymbol{p}_i^2/2m_i}. \tag{12}$$

Correspondingly, the molecular dynamics algorithm in the multioverlap ensemble naturally follows from Eq. (7) (see Refs. 10 and 11 for the case of multicanonical MD). The multioverlap MD simulation is carried out by solving the following modified equations of motion with a Gaussian thermostat:

$$\begin{aligned} \dot{\boldsymbol{q}}_i &= \frac{d\boldsymbol{q}_i}{dt} = \frac{\boldsymbol{p}_i}{m_i}, \\ \dot{\boldsymbol{p}}_i &= \boldsymbol{F}_i^{\mathrm{muov}} - \zeta_{\mathrm{muov}} \boldsymbol{p}_i. \end{aligned} \tag{13}$$

The "force" $\boldsymbol{F}_i^{\mathrm{muov}}$ acting on atom i is calculated from [see Eq. (8)]

$$\boldsymbol{F}_i^{\mathrm{muov}} = -\frac{\partial E_{\mathrm{muov}}}{\partial \boldsymbol{q}_i} = \boldsymbol{F}_i + k_{\mathrm{B}}T_0 \frac{\partial f(d_1,d_2)}{\partial \boldsymbol{q}_i}. \tag{14}$$

The coefficient ζ_{muov} is defined by

$$\zeta_{\mathrm{muov}} = \frac{\sum_i \boldsymbol{F}_i^{\mathrm{muov}} \cdot \dot{\boldsymbol{q}}_i}{2\sum_i \mathbf{p}_i^2/2m_{\mathbf{i}}}. \tag{15}$$

D. Determination of the multioverlap potential energy

The multioverlap potential energy E_{muov}, or the dimensionless free energy $f(d_1,d_2)$, in Eq. (8) is not *a priori* known and we must obtain its estimate by iterations of short simulations. Several methods[19–24] to determine the dimensionless free energy $f(d_1,d_2)$ exist and we determine it by the following process.[21] We update the dimensionless free energy $f(d_1,d_2)$ at each MD step of a short multioverlap MD simulation, and we iterate this procedure. Suppose that we have $f=f^{(l)}(d_1,d_2:k-1)$ at the $(k-1)$th MD step of the lth iteration of the short multioverlap MD simulation, and that the configuration at the kth MD step has the values $d_1=c_1$ and $d_2=c_2$. We then update the dimensionless free energy by

$$f^{(l)}(d_1=c_1,d_2=c_2;k)=f^{(l)}(d_1=c_1,d_2=c_2;k-1)-a^{(l)}, \tag{16}$$

where $a^{(l)}$ is an appropriately chosen positive constant. The lth iteration of the multioverlap MD simulation with the updating procedure of Eq. (16) is continued until the probability distribution $P_{\mathrm{muov}}(d_1,d_2)$ in Eq. (9) becomes reasonably flat with fluctuations of order $a^{(l)}$. For the $(l+1)$th iteration, we make the value of the constant a smaller, i.e., $a^{(l+1)} \leqslant a^{(l)}$, and repeat the updating procedure of Eq. (16) with l replaced by $l+1$. The initial value can be set as follows:

$$f^{(1)}(d_1,d_2;0)=0. \tag{17}$$

The iteration is terminated when the probability distribution $P_{\mathrm{muov}}(d_1,d_2)$ becomes satisfactorily flat. After the dimensionless free energy $f(d_1,d_2)$ is determined, we make a long production multioverlap MD simulation of Eqs. (13) and (14) with this $f(d_1,d_2)$.

E. Reweighting techniques

The results of the multioverlap production run can be analyzed by the reweighting techniques. Suppose that we have determined the dimensionless free energy $f(d_1,d_2)$ at a constant temperature T_0 and that we have made a production run at this temperature. The expectation value of a physical quantity A at any temperature T is calculated from

$$\langle A\rangle_T=\frac{\sum_{d_1,d_2,E}A(d_1,d_2;E)n(d_1,d_2;E)e^{-\beta E}}{\sum_{d_1,d_2,E}n(d_1,d_2;E)e^{-\beta E}}, \tag{18}$$

where the best estimate of the density of states is given by the single-histogram reweighting techniques,[16,17]

$$n(d_1,d_2;E)=\frac{N_{\mathrm{muov}}(d_1,d_2;E)}{W_{\mathrm{muov}}(d_1,d_2;E)}, \tag{19}$$

and $N_{\mathrm{muov}}(d_1,d_2;E)$ is the histogram of the probability distribution that was obtained by the multioverlap production run. By substituting Eqs. (7), (8), and (19) into Eq. (18), we have

$$\langle A\rangle_T=\frac{\sum_{d_1,d_2,E}A(d_1,d_2;E)N_{\mathrm{muov}}(d_1,d_2;E)e^{\beta_0E-f(d_1,d_2)-\beta E}}{\sum_{d_1,d_2,E}N_{\mathrm{muov}}(d_1,d_2;E)e^{\beta_0E-f(d_1,d_2)-\beta E}}. \tag{20}$$

We can also calculate the free energy (or the potential of mean force) with appropriate reaction coordinates. For example, the free energy $F(\xi_1,\xi_2;T)$ with reaction coordinates ξ_1,ξ_2 at temperature T is defined by

$$F(\xi_1,\xi_2;T)=-k_{\mathrm{B}}T\ln P_c(\xi_1,\xi_2;T), \tag{21}$$

where $P_c(\xi_1,\xi_2;T)$ is the reweighted canonical probability distribution of ξ_1 and ξ_2 and given by [see Eq. (20)]

$$P_c(\xi_1,\xi_2;T)=\frac{\sum_{d_1,d_2,E}N_{\mathrm{muov}}(\xi_1,\xi_2;d_1,d_2;E)e^{\beta_0E-f(d_1,d_2)-\beta E}}{\sum_{\xi_1,\xi_2,d_1,d_2,E}N_{\mathrm{muov}}(\xi_1,\xi_2;d_1,d_2;E)e^{\beta_0E-f(d_1,d_2)-\beta E}}, \tag{22}$$

and $N_{\mathrm{muov}}(\xi_1,\xi_2;d_1,d_2;E)$ is the histogram of the probability distribution that was obtained from the multioverlap production run.

III. COMPUTATIONAL DETAILS

Met-enkephalin is one of the simplest peptides and has the amino-acid sequence Tyr-Gly-Gly-Phe-Met. This peptide is often adopted as a test system in biomolecular simulations. Therefore, we also adopted Met-enkephalin in vacuum as a test system of the multioverlap MD method. In our simulations the N terminus and the C terminus were blocked with the acetyl group and the N-methyl group, respectively. This is because we wanted the total charge of the Met-enkephalin system to be neutral. Accordingly, the total number of atoms of Met-enkephalin in our simulations is 84. The force field that we adopted is the CHARMM param 22 parameter set.[25] Our multioverlap MD simulations were performed by implementing the method in the CHARMM macromolecular mechanics program.[26] The main part of implementation is shown as follows. We introduced the Gaussian constraint method (Gaussian thermostat)[2,3] to the CHARMM macromolecular mechanics program. The corresponding equations of motion were implemented. Namely, we used Eq. (10) for the canonical MD simulations and Eq. (13) for the multioverlap MD simulations. We also implemented the method in the CHARMM macromolecular mechanics program to perform multicanonical MD simulations. The time step was taken to be 0.5 fs and the leap-frog algorithm[27] was employed for the numerical integration.

We consider two local-minimum-energy states of Met-enkephalin as reference configurations. These configurations were obtained by the simulated annealing MD method.[28] During the simulated annealing run, the temperature was decreased linearly from 1000 to 100 K with an increment of 50 K, and the canonical MD simulations were performed for 500 ps at each temperature (9.5 ns in total). This simulated annealing MD run was repeated ten times with different initial random numbers. The obtained final conformations were further minimized by the conjugate gradient method, and two conformations were identified as the reference configurations from the backbone hydrogen bond patterns. In Fig. 1 we show these reference configurations of Met-enkephalin. Reference configuration 1 (RC1) has a β-turn structure with two backbone hydrogen bonds between Gly-2 and Met-5, and reference configuration 2 (RC2) has a γ-turn structure with two backbone hydrogen bonds between Gly-2 and Phe-4. Reference configuration 1 also has a hydrogen bond between hydrogen bond acceptor CO of Gly-2 and hydrogen bond donor NH of Phe-4. We remark that with ECEPP/2 energy function[29–31] RC1 corresponds to the global-minimum state and RC2 corresponds to a local-minimum state.[32]

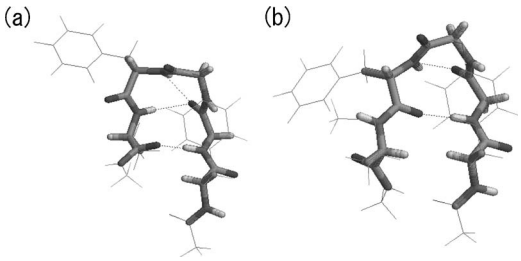

FIG. 1. (a) Reference configuration 1 and (b) reference configuration 2. The dotted lines denote the hydrogen bonds. The N terminus and the C terminus are on the right-hand side and on the left-hand side, respectively. The figures were created with RASMOL (Ref. 36).

The backbone dihedral angles are of three types: the rotation angle about the $N-C^{\alpha}$ bond of the backbone (ϕ), that about the $C^{\alpha}-C$ bond (ψ), and that about the peptide bond C–N (ω). Our multioverlap MD simulation was performed using the all-atom model, but we used only ϕ and ψ angles in the definition of the dihedral-angle distances in Eq. (2). This is because the dihedral angles of the backbone ω have almost the fixed value of 180° for the peptide bond C–N. Furthermore, by using only the backbone dihedral angles (and not side-chain dihedral angles) as the elements of the dihedral-angle distances, we focused on the backbone structures of Met-enkephalin. In Eq. (2), consequently, the number n of the elements of the dihedral-angle distances is ten because Met-enkephalin has five pairs of ϕ and ψ. In Table I we list the dihedral angles ϕ and ψ of the two reference configurations in Fig. 1.

Our multioverlap MD simulation was carried out at $T_0=300$ K. We first have to determine the multioverlap weight factor $W_{\mathrm{muov}}(d_1,d_2;E)$ in Eq. (7), or the dimensionless free energy $f(d_1,d_2)$ in Eq. (8), to get a flat probability distribution in the two-dimensional dihedral-angle-distance space (d_1,d_2). For that purpose we used the procedure in Sec. II D. We first set $f^{(1)}(d_1,d_2)=0$ according to Eq. (17). We then performed the multioverlap MD simulation of Eq. (13) for 14 ns. The dimensionless free energy $f^{(1)}(d_1,d_2)$ was updated by Eq. (16) at each MD step with $a^{(1)}=0.0001$. For this calculation, the dihedral-angle distances (d_1,d_2) were discretized with a bin size of 0.01. This 14 ns MD simulation was sufficient to obtain an optimal multioverlap weight factor, and we did not further iterate the process. Finally, the multioverlap MD production run was then performed with this weight factor for 24 ns after equilibration of 1 ns. Because the multioverlap MD simulations perform a random walk in the configurational space, the results will not depend on the initial conformation. For the initial conformation of the multioverlap MD simulation production run, we thus simply adopted one of the final conformations obtained by the above simulated annealing runs. In Fig. 2 we show this initial conformation and list their backbone dihedral angles in Table II. For the purpose of comparisons, we also performed a usual canonical MD simulation and a multicanonical MD simulation for 24 ns at $T_0=300$ K. The initial conformation for both the canonical production run and the multicanonical production run was the same as that for the multioverlap production run.

TABLE I. Backbone dihedral angles ϕ and ψ for reference configurations 1 and 2.

Reference configuration 1			Reference configuration 2		
Residue	Type	Angle (°)	Residue	Type	Angle (°)
1	ϕ_1	−100.1	1	ϕ_1	−136.0
1	ψ_1	136.2	1	ψ_1	139.3
2	ϕ_2	−149.2	2	ϕ_2	−163.8
2	ψ_2	56.6	2	ψ_2	68.8
3	ϕ_3	76.4	3	ϕ_3	88.7
3	ψ_3	−78.2	3	ψ_3	−61.0
4	ϕ_4	−87.9	4	ϕ_4	−108.3
4	ψ_4	−37.5	4	ψ_4	−179.7
5	ϕ_5	−79.8	5	ϕ_5	−92.2
5	ψ_5	138.9	5	ψ_5	146.1

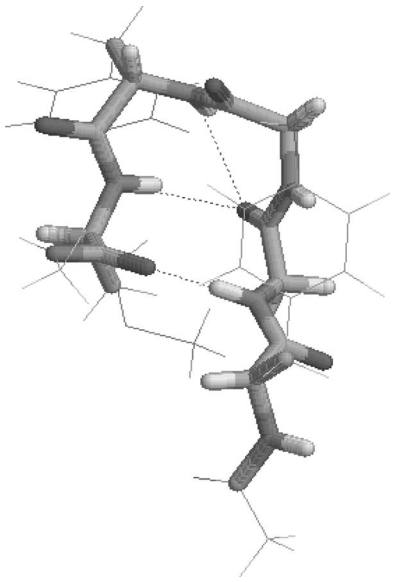

FIG. 2. The common initial conformation of the usual canonical, multicanonical, and multioverlap MD simulations. See also the caption of Fig. 1.

IV. RESULTS AND DISCUSSION

In this section we present the results of the multioverlap MD simulation of Met-enkephalin in vacuum. Furthermore, we compare with the results of the usual canonical, multica-

TABLE II. Backbone dihedral angles ϕ and ψ for the initial conformation.

Initial conformation		
Residue	Type	Angle (°)
1	ϕ_1	−107.3
1	ψ_1	149.5
2	ϕ_2	−156.7
2	ψ_2	64.6
3	ϕ_3	68.3
3	ψ_3	−89.2
4	ϕ_4	−89.3
4	ψ_4	−13.4
5	ϕ_5	−77.7
5	ψ_5	110.3

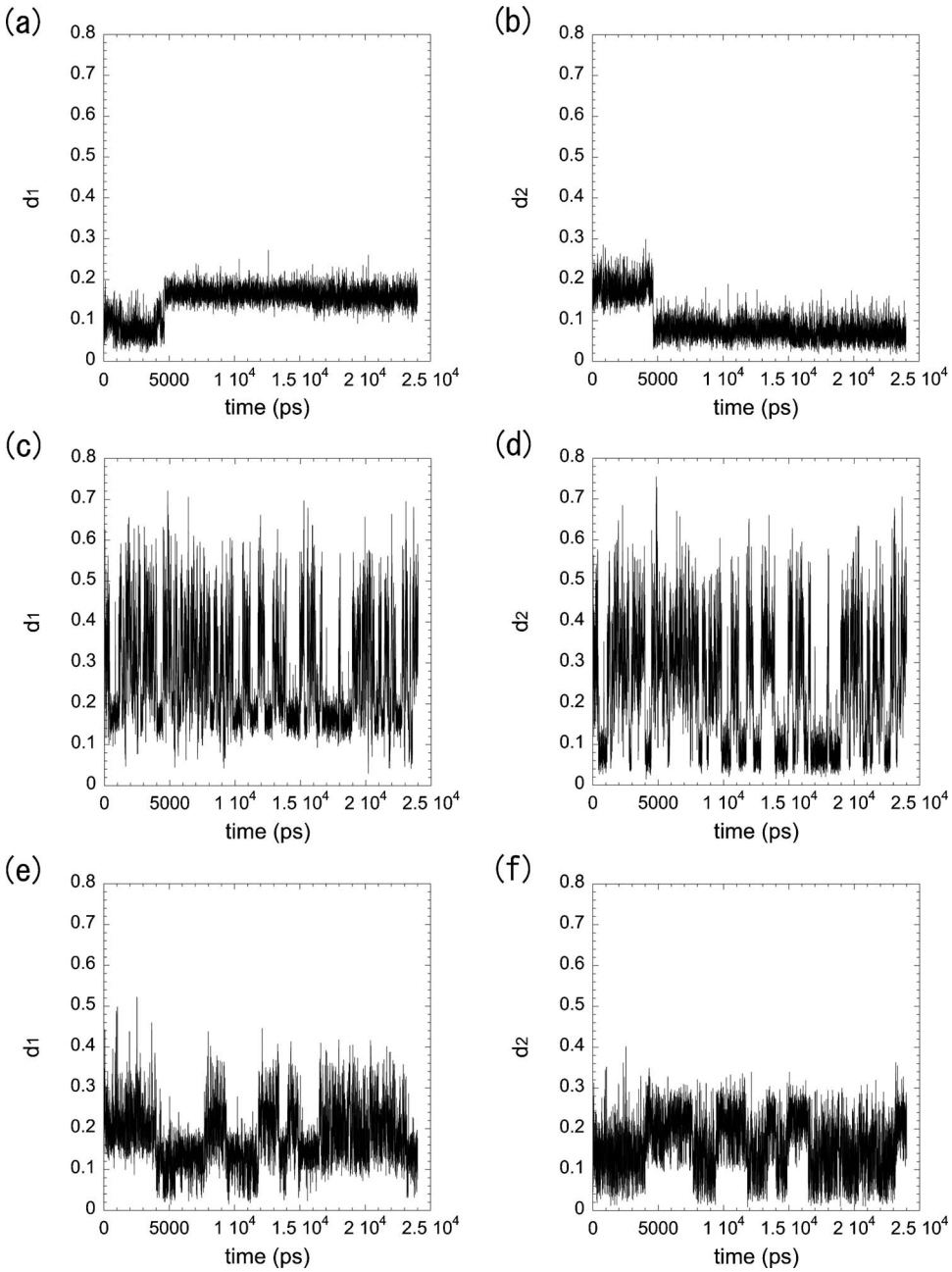

FIG. 3. The time series of the dihedral-angle distances d_1 and d_2. (a) and (b) are the results from the usual canonical MD simulation, (c) and (d) are from the multicanonical MD simulation, and (e) and (f) are from the multioverlap MD simulation at T_0 =300 K.

nonical, and multioverlap MD simulations. The various time series are given in Sec. IV A. In Sec. IV B we show the raw data of the probability distributions and discuss the effectiveness of the multioverlap MD method. The physical quantities can be calculated by the reweighting techniques.[16,17] In Sec. IV C the physical quantities, which were obtained from the usual canonical and multioverlap MD simulations, are compared with those from the multicanonical MD simulation. In the last section we describe the detailed free-energy landscape calculated from the multioverlap MD simulation and identify the conformations in the transition state between RC1 and RC2.

A. Time series of simulations

We first examine the time series of various quantities from the usual canonical, multicanonical, and multioverlap MD simulations. Figure 3 shows the time series of the dihedral-angle distances with respect to each of the two reference configurations. When $d_1=0$, the values of the backbone dihedral angles are completely coincident with those of reference configuration 1 and it turned out that $d_2=0.159$. Conversely, when $d_2=0$, we have $d_1=0.159$. In the usual canonical MD simulation at $T_0=300$ K [see Figs. 3(a) and 3(b)], the configuration transited from a RC1-like state to a RC2-like state near 5 ns and did not transit back from the RC2-like state to the RC1-like state. In other words, the canonical MD simulation got trapped in the RC2-like local-minimum state. Thus, the usual canonical MD simulation does not sample efficiently the conformational space and we cannot calculate an accurate free-energy landscape. On the one hand, the multicanonical MD simulation did not get trapped in the local-minimum states, as we can see in Figs. 3(c) and 3(d). For both d_1 and d_2 we observe random walks both in d_1 space and in d_2 space; both dihedral-angle dis-

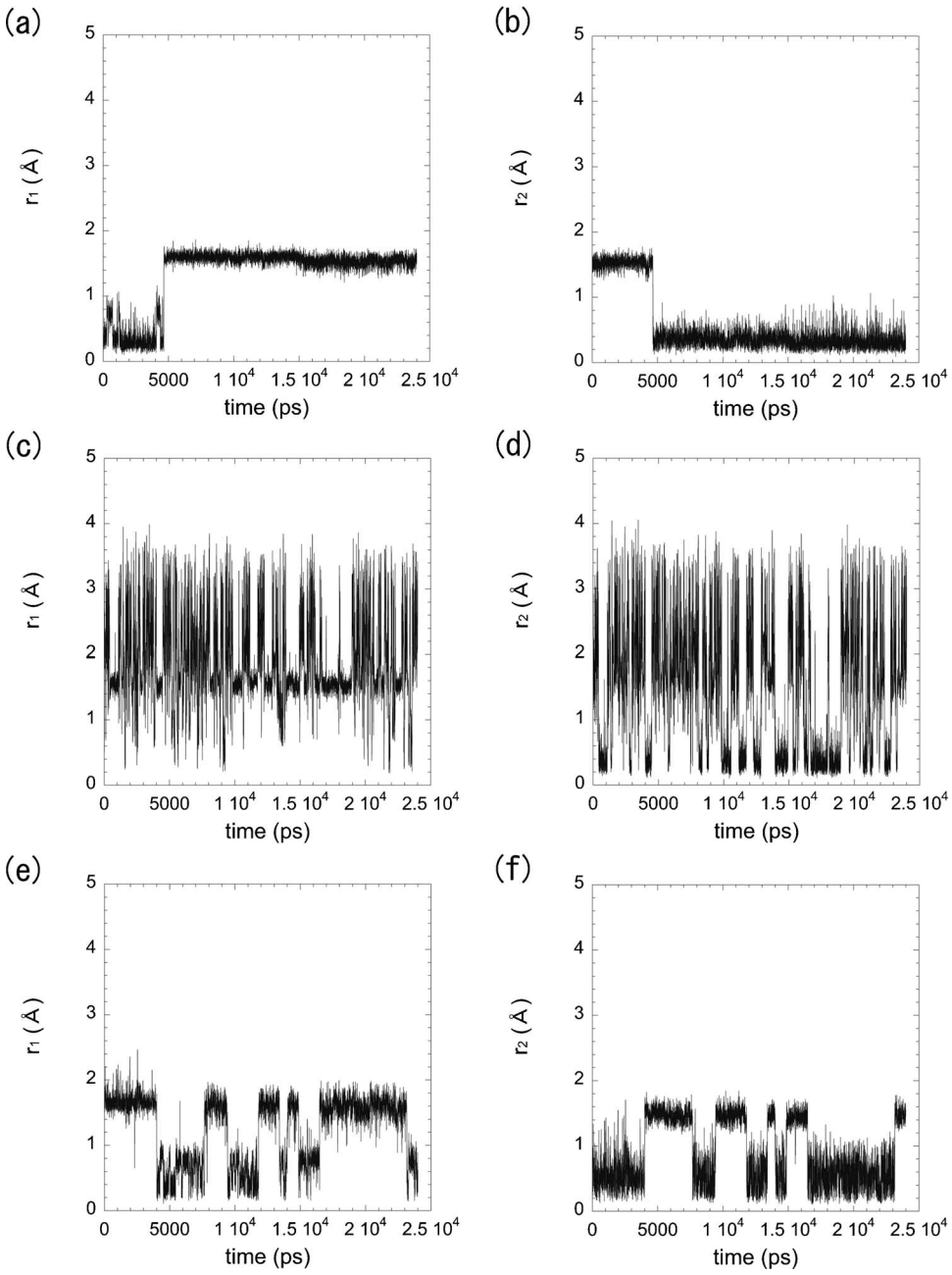

FIG. 4. The time series of the RMSD r_1 and r_2. (a) and (b) are the results from the usual canonical MD simulation, (c) and (d) are from the multicanonical MD simulation, and (e) and (f) are from the multioverlap MD simulation at $T_0=300$ K.

tances often visited small values as well as large values beyond 0.5, which is the average value at $T_0=\infty$. Therefore, we had an efficient sampling in the conformational space in the multicanonical MD simulation. When we look into Fig. 3(c) more carefully, however, we find that the multicanonical MD simulation did not sample around the RC1-like state very much (d_1 values did not take very small values). Accordingly, we may not obtain an accurate free-energy landscape near RC1 from the results of the multicanonical MD simulation. Finally, as one can see in Figs. 3(e) and 3(f), the multioverlap MD simulation did not get trapped in the local-minimum states, either. Although the ranges of the dihedral-angle distances that were covered are less in the multioverlap MD simulation than in the multicanonical MD simulation (reflecting the fact that the latter explores a wide conformational space than the former), the multioverlap simulation indeed visited both the RC1 state and RC2 state. We observe transitions between the RC1 state and RC2 state several times in the figure. Thus, the multioverlap MD simulation can realize a random walk in the two-dimensional dihedral-angle-distance space and yet focus on the two reference configurations RC1 and RC2.

In Fig. 4 we show the time series of the root-mean-square distance (RMSD) of the backbone of Met-enkephalin with respect to each of the two reference configurations from the canonical, multicanonical, and multioverlap MD simulations. The RMSD r_i with respect to reference configuration i is defined by

$$r_i = \min\left[\sqrt{\frac{1}{N}\sum_j (\boldsymbol{q}_j - \boldsymbol{q}_j^{(i)})^2}\right], \tag{23}$$

where N is the number of atoms, $\{\boldsymbol{q}_j^{(i)}\}$ are the coordinates of reference configuration i, and the minimization is over the

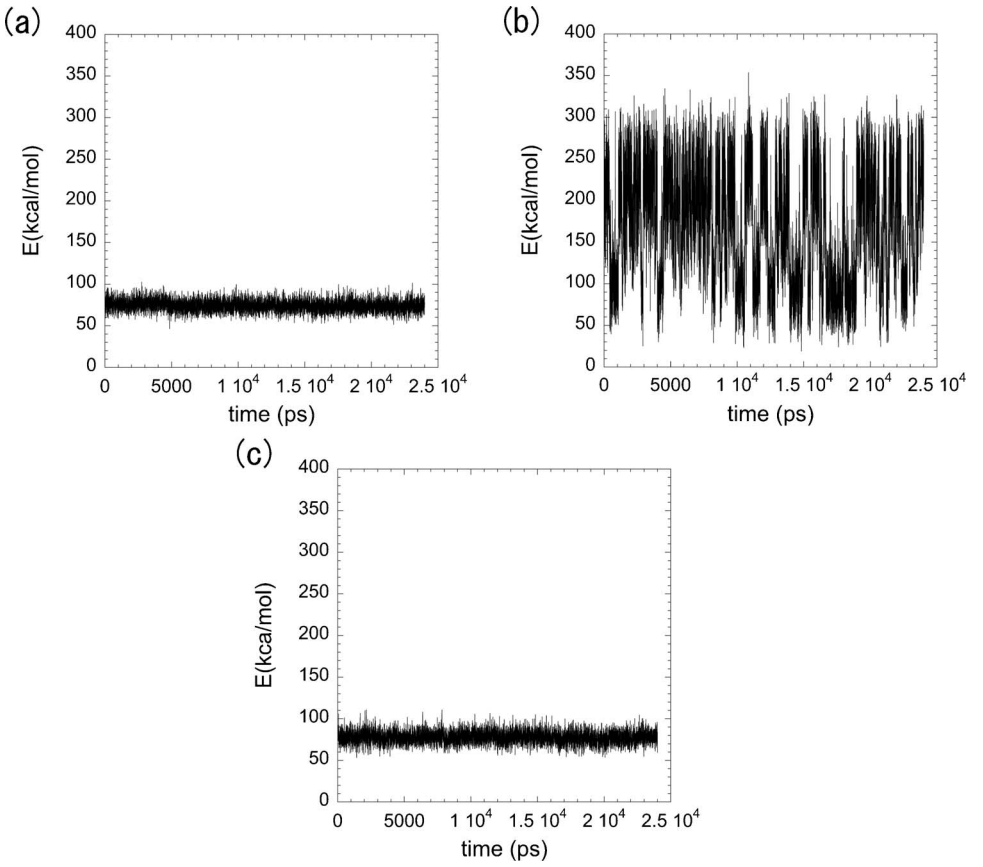

FIG. 5. The time series of the potential energy E. (a) is the results from the usual canonical MD simulation, (b) is from the multicanonical MD simulation, and (c) is from the multioverlap MD simulation at $T_0=300$ K.

rigid translations and rigid rotations of the coordinates of the configuration $\{\boldsymbol{q}_j\}$ with respect to the center of geometry. The behavior of the three simulations in Fig. 4 is the same as in Fig. 3; there are strong correlations between the dihedral-angle distance d_1 (d_2) and the RMSD r_1 (r_2). By employing the RMSD as the reaction coordinates, however, the boundary between the RC1-like state and RC2-like state is more clarified. Incidentally, when $r_1=0$ ($r_2=0$), we have $r_2=1.52$ ($r_1=1.52$).

Figure 5 shows the time series of the potential energy of the three simulations. The multicanonical MD simulation covers widely the potential energy space, as we can see in Figs. 5(b). The time series of the potential energy of the multioverlap MD simulation, however, is not much different from that of the canonical MD simulation [compare Figs. 5(a) and 5(c)]. This is because the multioverlap algorithm is based on the Boltzmann weight factor at temperature T_0 as far as energy dependence is concerned [see Eqs. (7) and (8)], while the multicanonical algorithm is independent of temperature. The multioverlap MD method aims at a random walk in the dihedral-angle-distance space, not in the potential energy space.

B. Probability distributions of simulations

We discuss the probability distributions of configuration from the three simulations, the usual canonical, multicanonical, and multioverlap MD simulations. In Fig. 6 we show the raw data of the histograms with respect to the two dihedral-angle-distance coordinates. The bin size of the two-dimensional histograms is 0.01×0.01. These histograms represent the probability distributions in the two-dimensional dihedral-angle-distance space for the three ensembles. From Figs. 6(a) and 6(b), it is obvious that the probability distributions of the usual canonical and multicanonical MD simulations are biased towards RC2; there are pronounced peaks near $(d_1,d_2)=(0.159,0.0)$. In other words, as previously stated, the usual canonical and multicanonical MD simulations did not sample efficiently the RC1-like states [near $(d_1,d_2)=(0.0,0.159)$]. In Fig. 6(c), on the other hand, we confirm that the multioverlap MD simulation has a rather flat probability distribution in the two-dimensional dihedral-angle-distance space containing both the RC1 state and the RC2 state [see Eq. (9)].

In Fig. 7 we show the raw data of the histograms with respect to the two RMSD coordinates. The bin size of the two-dimensional histograms is 0.1×0.1 Å^2. In this case, the histograms were taken every 100 MD steps (50 fs). Therefore, these histograms are rugged in comparison with those with the dihedral-angle-distance coordinates in Fig. 6, where the data were taken every MD step (0.5 fs). The two peaks that correspond to the RC1 and RC2 states are disconnected in the case of the canonical MD simulation [see Fig. 7(a)], and they are connected in both the multicanonical MD simulation and the multioverlap MD simulation [see Figs. 7(b) and 7(c)]. However, while the multicanonical MD simulation has to visit a region with large r_1 and r_2 (high-energy region) in order to have transitions between RC1 and RC2, the multioverlap MD simulation can connect both states within a region with small r_1 and r_2. The characteristics of the probability distributions in Fig. 7 are essentially the same as in

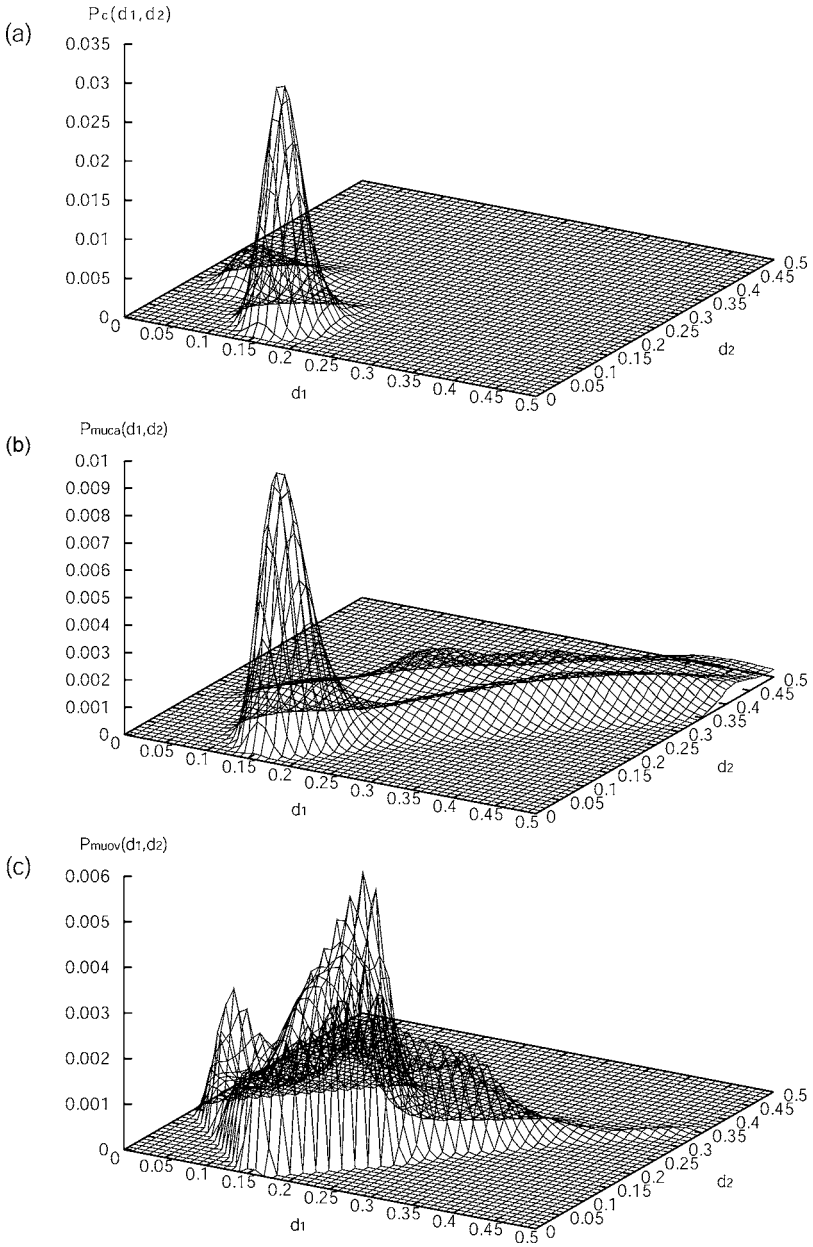

FIG. 6. The raw data of the probability distribution with respect to the dihedral-angle-distance axes d_1 and d_2. (a) is the results from the usual canonical MD simulation, (b) is from the multicanonical MD simulation, and (c) is from the multioverlap MD simulation at T_0 =300 K.

Fig. 6. Namely, in Figs. 7(a) and 7(b) the probability distributions are biased distributions towards RC2, and that from the multioverlap MD simulation in Fig. 7(c) has finite contributions in both the RC1 state and the RC2 state. In the multioverlap ensemble, however, the probability distribution is not needed to become flat in the RMSD space [compare Figs. 6(c) and 7(c)]. This is because the multioverlap ensemble is devised to obtain a flat probability distribution in the dihedral-angle-distance space and not in the RMSD space. Nevertheless, the multioverlap MD simulation realized an efficient sampling in the RMSD space between RC1 and RC2. Thus, the multioverlap MD simulation is suitable to sample between the reference configurations in comparison with the other methods.

Figure 8 shows the raw data of the probability distributions of the potential energy. The bin size of the histograms is 1.0 kcal/mol. The probability distribution of the potential energy in the multicanonical MD simulation, as a matter of course, is flat. Thus, the multicanonical methods are suitable to sample the potential energy space, not the conformational space between the specific reference configurations. The probability distribution of the potential energy in the multioverlap MD simulation is almost the same as that in the usual canonical MD simulation. The probability distribution in the multioverlap MD simulation is, however, a little wider than in the usual canonical MD simulation. This is because the multioverlap MD simulation has to sample a little higher-energy region in order to overcome the potential energy barrier between the RC1-like state and RC2-like state.

C. Physical quantities calculated by the reweighting techniques

We now examine the physical quantities calculated from the results of the three simulations, the usual canonical, multicanonical, and multioverlap MD simulations, by the reweighting techniques. The reweighting techniques for the

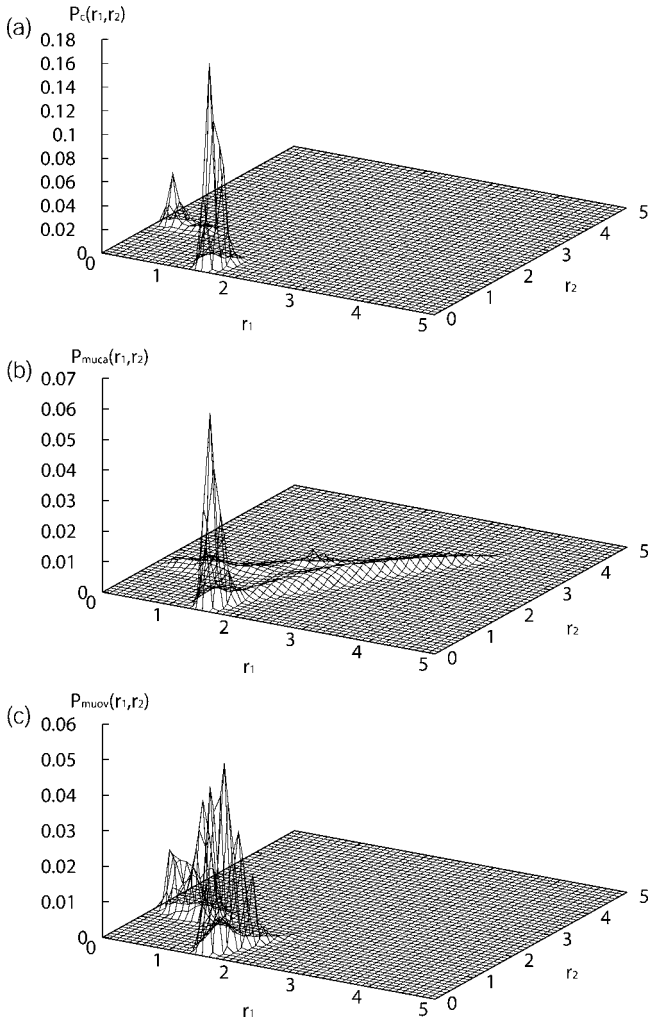

FIG. 7. The raw data of the probability distribution with respect to the RMSD axes r_1 and r_2. (a) is the results from the usual canonical MD simulation, (b) is from the multicanonical MD simulation, and (c) is from the multioverlap MD simulation at T_0=300 K.

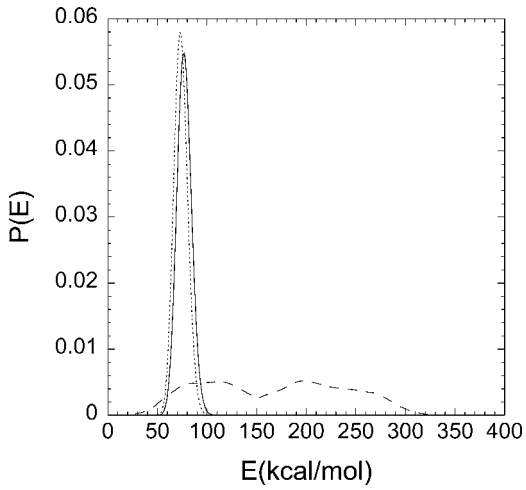

FIG. 8. The raw data of the probability distribution of the potential energy E. The dotted line, the dashed line, and the solid line show the results from the usual canonical MD simulation at T_0=300 K, the results from the multicanonical MD simulation at T_0=300 K, and the results from the multioverlap MD simulation at T_0=300 K, respectively.

multioverlap MD method were explained in Sec. II E. The reweighting techniques for the other MD methods are accounted in Ref. 7.

In Fig. 9 we show the probability distributions and physical quantities calculated by the reweighting techniques. Here, the error bars were calculated by the jackknife method.[33–35] The number of bins was taken to be 8. The results from the multicanonical MD simulation are shown as a reference in the figure, because the multicanonical algorithm is well known for giving accurate expectation values for a wide range of temperature.[7] As we can see in Fig. 9(a), the probability distributions of the potential energy at T =300 K calculated from the results of the usual canonical and multioverlap MD simulations are in good agreement with those of the multicanonical MD simulation. Furthermore, the average potential energy as a function of temperature is also in agreement with that from the multicanonical MD simulation, although we see slight deviations beyond error bars below $T \approx 250$ K and above $T \approx 350$ K in the case of the canonical MD simulation. In Fig. 9(c), however, we see that the specific heat as a function of temperature calculated from the results of the canonical MD simulation does not coincide with those of the multicanonical MD simulation in the entire temperature range (the error bars do not overlap). This is because the usual canonical MD simulation got trapped in the local-minimum states and did not have enough sampling in the conformational space. The specific heat here is defined by

$$C_v = \frac{1}{k_B}\frac{d\langle E\rangle_T}{dT} = \beta^2(\langle E^2\rangle_T - \langle E\rangle_T^2). \tag{24}$$

The specific heat is the derivative of the average potential energy and it is more difficult to obtain accurate results than the average potential energy itself. In the case of the multioverlap MD simulation, the results well coincide with those from the multicanonical MD simulation between about 250 and 350 K. In the region under 250 K and above 350 K, however, we see deviations between the results of the two simulations. This sets a reliable range of temperature where accurate thermodynamic quantities can be calculated by the multioverlap MD simulation. The reason for the deviations is that the multioverlap algorithm samples conformations in the dihedral-angle-distance space but not in the energy space. Accordingly, the multioverlap simulation is difficult to give an accurate estimate of the density of states in Eq. (19) over a wide potential energy range. Thus, in the multioverlap MD method, the expectation values calculated by the reweighting techniques in Eq. (20) are correct only in the neighborhood of the temperature at which simulations were performed.

D. Transition state estimated from free-energy landscape

We now study the transition between the two states, RC1 and RC2. The free-energy landscape was calculated from Eq. (21) with appropriate reaction coordinates by the reweighting techniques. In Fig. 10 we show the free-energy landscape at T=300 K obtained from the three simulations with respect to the reaction coordinates of the two dihedral-angle distances.

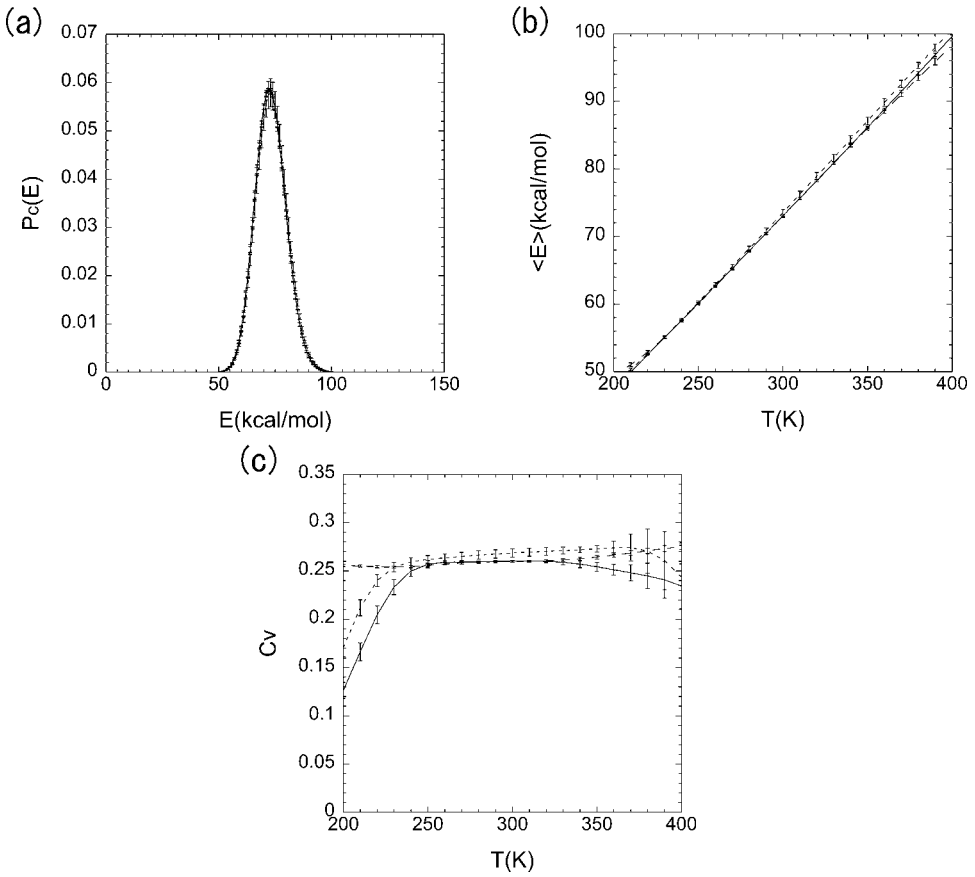

FIG. 9. (a) Probability distribution of the potential energy at $T=300$ K, (b) average potential energy as a function of temperature, and (c) specific heat as a function of temperature. These results were calculated from the usual canonical MD simulation (dotted line), the multicanonical MD simulation (dashed line), and the multioverlap MD simulation (solid line) by the reweighting techniques.

The free-energy landscape of the usual canonical MD simulation is inaccurate due to insufficient sampling in the conformational space, as previously mentioned. The results from the multicanonical MD simulation have a rugged surface but cover a wide region in the two-dimensional dihedral-angle-distance space in comparison with those of the multioverlap MD simulation [compare Figs. 10(b) and 10(c)]. This is because the multioverlap method samples efficiently and selectively the conformational space between the two reference configurations. On the other hand, the multicanonical MD simulation makes wide sampling in the conformational space but does not focus on specific reference configurations. Thus,

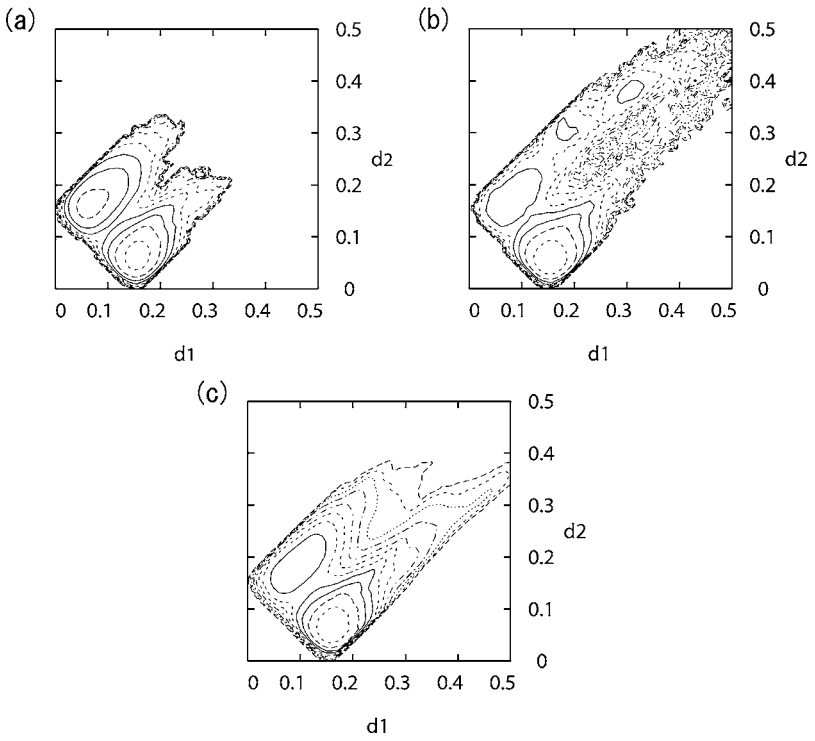

FIG. 10. The free-energy landscape obtained from (a) the usual canonical MD simulation, (b) the multicanonical MD simulation, and (c) the multioverlap MD simulation at $T=300$ K with respect to the dihedral-angle-distance axes d_1 and d_2. The contour lines are drawn every 1 kcal/mol.

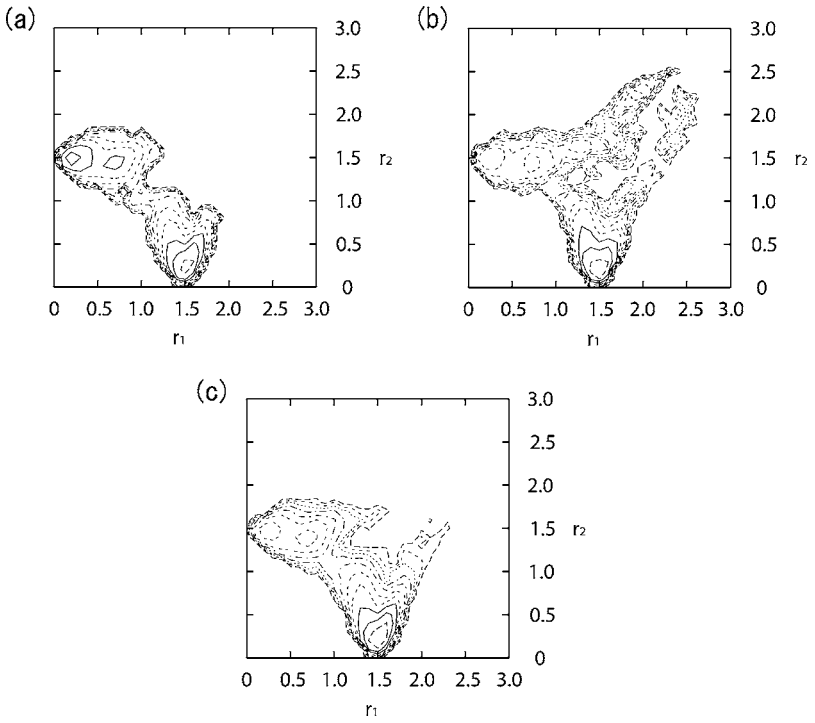

FIG. 11. The free-energy landscape obtained from (a) the usual canonical MD simulation, (b) the multicanonical MD simulation, and (c) the multioverlap MD simulation at $T=300$ K with respect to the RMSD axes r_1 and r_2. The contour lines are drawn every 1 kcal/mol.

the multioverlap method is better in the sense that a detailed free-energy landscape in the neighborhood and between the two specific reference configurations can be obtained.

In Fig. 11 we show the free-energy landscape at $T=300$ K calculated from the three simulations with respect to the two RMSD axes. Although the characteristics of these figures are essentially the same as those in Fig. 10, the saddle point between the two local-minimum states (RC1 and RC2 states) can be clearly identified. Figure 12 shows the free-energy landscape obtained from the multioverlap MD simulation. This figure is the same as Fig. 11(c). We labeled the local-minimum states (A_1, A_2, and B) and the transition state (C). In Fig. 13 we show representative conformations in the local-minimum states A_1, A_2, and B. The conformations in the local-minimum states A_1 and A_2 have the same backbone hydrogen bonds as in RC1. The local-minimum state B, which has the same backbone hydrogen bonds as in RC2, corresponds to the global-minimum free-energy state at $T=300$ K. The free-energy difference between the global-minimum state (B) and the local-minimum state (A_1) [or (A_2)] is about 3 kcal/mol.

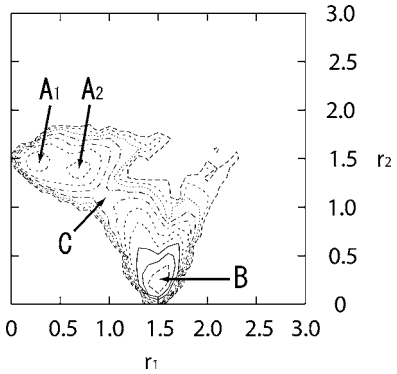

FIG. 12. The free-energy landscape obtained from the multioverlap MD simulation at $T=300$ K with respect to the RMSD axes r_1 and r_2. The contour lines are drawn every 1 kcal/mol. The labels A_1, A_2, and B locate the local-minimum states. The label C stands for the saddle point between A_1 (or A_2) and B.

The saddle point C in Fig. 12 corresponds to the transition state between the global-minimum state (B) and the local-minimum state (A_1) [or (A_2)]. The free-energy difference between B and C is about 6 kcal/mol and that between A_1 (or A_2) and C is about 3 kcal/mol. Because $k_BT \approx 0.6$ kcal/mol at $T=300$ K, these barrier heights are rather high. This is why the usual canonical MD simulation got trapped in the vicinity of the global-minimum state B (RC2-like state). In Table III we list the free-energy difference among the states. Two representative conformations in C are shown in Fig. 14. These structures have a backbone hydrogen bond between CO of Gly-2 and NH of Phe-4. This hydrogen bond in C is common to both RC1 and RC2. The hydrogen bonds between NH of Gly-2 and CO of Met-5 which exists in RC1 and that between NH of Gly-2 and CO of Phe-4 which exists in RC2 are missing in C. These structures are thus more extended than reference configurations 1

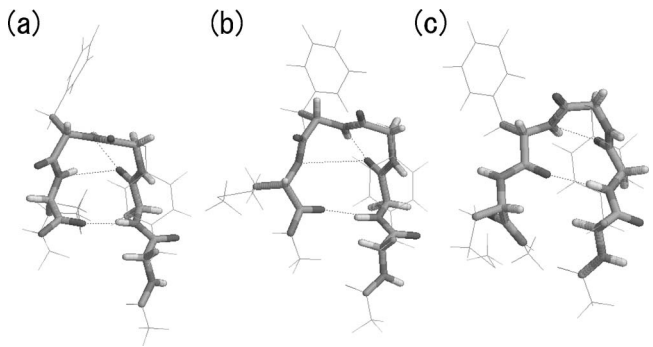

FIG. 13. (a) The structure in A_1, (b) A_2, and (c) B in Fig. 12. See also the caption of Fig. 1.

TABLE III. Free-energy difference (kcal/mol) among the states.

	A_1	A_2	B	C
A_1	0	−0.1	3.2	−2.7
A_2	0.1	0	3.3	−2.6
B	−3.2	−3.3	0	−5.9
C	2.7	2.6	5.9	0

and 2. Accordingly, the conformations in C are very reasonable as intermediate structures between RC1 and RC2.

In Table IV we list the backbone dihedral angles ϕ and ψ of the conformations in Figs. 13 and 14. From Tables I and IV and Figs. 13 and 14, we can deduce the transition pathways from RC1 to RC2. Note that the major difference between RC1 and RC2 in Table I is the value of ψ_4. The two hydrogen bonds (between Gly-2 and Met-5) in RC1 will be simultaneously broken by a large rotation of ψ_4, but this direct pathway is impossible because of high-energy barriers. In the following we focus on the relation between the changes of the backbone dihedral angle and the formation/breakage of the backbone hydrogen bonds in order to elucidate a possible transition pathway from RC1 to RC2. The dihedral angle ϕ_5 first rotates while keeping the hydrogen bonds. This process corresponds to the conformational change from Fig. 13(a) to Fig. 13(b). The dihedral angles ϕ_2 and ϕ_5 then rotate and the hydrogen bond between NH of Gly-2 and CO of Met-5 is broken [transition from Fig. 13(b) to Fig. 14(a)]. From Fig. 14(a) to Fig. 14(b), we also see that the hydrogen bond between CO of Gly-2 and NH of Met-5 is brink of collapse. Finally, the dihedral angle ψ_4 rotates again and the hydrogen bond between NH of Gly-2 and CO of Phe-5 is formed [transition from Fig. 14(b) to Fig. 13(c)]. In summary, we have the following transition pathway: A_1 [Fig. 13(a)] $\rightarrow A_2$ [Fig. 13(b)] $\rightarrow C$ [Fig. 14(a)] $\rightarrow C$ [Fig. 14(b)] $\rightarrow B$ [Fig. 13(c)].

V. CONCLUSIONS

In this article we presented detailed comparisons of the canonical, multicanonical, and multioverlap MD methods, taking a pentapeptide system of Met-enkephalin in vacuum as a benchmark system. The multioverlap MD simulation was performed so that it will realize a random walk between two reference configurations. The conventional canonical MD simulation got trapped in the vicinity of one of the two reference configurations. The multicanonical MD simulation, on the other hand, did not get trapped in the states of energy local minima, but it sampled widely only around one of the two reference configurations. Finally, the multioverlap MD simulation did sample the configurational space around both reference configurations. Therefore, we have shown the effectiveness of the multioverlap MD method over the canonical and multicanonical MD methods. We could obtain the detailed free-energy landscape between the two reference configurations from the results of the multioverlap MD simulation. From the free-energy landscape we identified the transition state and deduced the transition pathway between the two local-minimum states. Thus, the multioverlap MD method is a very powerful tool for studying the free-energy landscape and transition state between two specific configurations.

Some of the possible future applications of the present method are as follows. Firstly, we used only the dihedral angles of the backbone as the elements of the dihedral-angle distances. If dihedral angles of side chains are also included, we will be able to investigate the effects of the side-chain conformations on protein folding. Secondly, we studied a peptide in vacuum. We can easily apply the method to a protein in solution, which is a more realistic system. Thirdly, we presented the case with two reference configurations. Be-

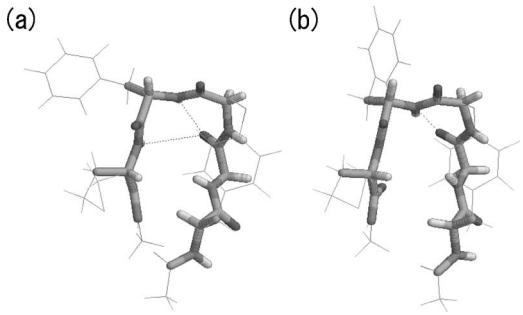

FIG. 14. Two conformations in the saddle point C in Fig. 12. See also the caption of Fig. 1.

TABLE IV. Backbone dihedral angles ϕ and ψ for the structures in Figs. 13 and 14.

Conformation in Fig. 13(a)			Conformation in Fig. 13(b)			Conformation in Fig. 13(c)			Conformation in Fig. 14(a)			Conformation in Fig. 14(b)		
Residue	Type	Angle (°)	Residue	Type	Angle (°)	Residue	Type	Angle (°)	Residue	Type	Angle (°)	Residue	Type	Angle (°)
1	ϕ_1	−157.3	1	ϕ_1	−145.9	1	ϕ_1	−131.2	1	ϕ_1	−148.5	1	ϕ_1	−147.5
1	ψ_1	122.7	1	ψ_1	122.0	1	ψ_1	142.3	1	ψ_1	136.6	1	ψ_1	163.7
2	ϕ_2	−130.4	2	ϕ_2	−126.7	2	ϕ_2	−171.6	2	ϕ_2	−166.6	2	ϕ_2	−174.9
2	ψ_2	47.2	2	ψ_2	59.1	2	ψ_2	64.6	2	ψ_2	66.1	2	ψ_2	64.7
3	ϕ_3	88.1	3	ϕ_3	74.3	3	ϕ_3	90.7	3	ϕ_3	86.1	3	ϕ_3	73.3
3	ψ_3	−91.1	3	ψ_3	−64.7	3	ψ_3	−59.4	3	ψ_3	−66.1	3	ψ_3	−62.5
4	ϕ_4	−95.0	4	ϕ_4	−83.1	4	ϕ_4	−118.3	4	ϕ_4	−86.9	4	ϕ_4	−86.0
4	ψ_4	−34.5	4	ψ_4	−69.9	4	ψ_4	−167.9	4	ψ_4	−68.2	4	ψ_4	−72.9
5	ϕ_5	−74.1	5	ϕ_5	−136.2	5	ϕ_5	−82.9	5	ϕ_5	−170.1	5	ϕ_5	−164.9
5	ψ_5	135.9	5	ψ_5	−169.3	5	ψ_5	139.1	5	ψ_5	151.8	5	ψ_5	177.2

cause we generalized this method to a multidimensional version, it is straightforward to deal with more than two reference configurations.

ACKNOWLEDGMENTS

The authors thank Dr. Bernd A. Berg and Dr. Hiroshi Noguchi for useful discussions. The computations were performed on the computers at the Research Center for Computational Science, Okazaki National Research Institutes. This work was supported, in part, by the Grants-in-Aid for the NAREGI Nanoscience Project and for Scientific Research in Priority Areas, "Water and Biomolecules," from the Ministry of Education, Culture, Sports, Science and Technology, Japan.

[1] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).
[2] W. G. Hoover, A. J. C. Ladd, and B. Moran, Phys. Rev. Lett. **48**, 1818 (1982).
[3] D. J. Evans, J. Chem. Phys. **78**, 3297 (1983).
[4] S. Nosé, Mol. Phys. **52**, 255 (1984).
[5] S. Nosé, J. Chem. Phys. **81**, 511 (1984).
[6] W. G. Hoover, Phys. Rev. A **31**, 1695 (1985).
[7] A. Mitsutake, Y. Sugita, and Y. Okamoto, Biopolymers **60**, 96 (2001).
[8] B. A. Berg and T. Neuhaus, Phys. Lett. B **267**, 249 (1991).
[9] B. A. Berg and T. Neuhaus, Phys. Rev. Lett. **68**, 9 (1992).
[10] U. H. E. Hansmann, Y. Okamoto, and F. Eisenmenger, Chem. Phys. Lett. **259**, 321 (1996).
[11] N. Nakajima, H. Nakamura, and A. Kidera, J. Phys. Chem. B **101**, 817 (1997).
[12] U. H. E. Hansmann and Y. Okamoto, J. Comput. Chem. **14**, 1333 (1993).
[13] U. H. E. Hansmann and Y. Okamoto, in *Annual Reviews of Computational Physics VI*, edited by D. Stauffer (World Scientific, Singapore, 1999), pp. 129–157.
[14] B. A. Berg, H. Noguchi, and Y. Okamoto, Phys. Rev. E **68**, 036126 (2003).
[15] S. G. Itoh and Y. Okamoto, Chem. Phys. Lett. **400**, 308 (2004).
[16] A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **61**, 2635 (1988).
[17] A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **63**, 1658 (1989).
[18] U. H. E. Hansmann, M. Masuya, and Y. Okamoto, Proc. Natl. Acad. Sci. U.S.A. **94**, 10652 (1997).
[19] B. A. Berg and T. Celik, Phys. Rev. Lett. **69**, 2292 (1992).
[20] Y. Okamoto and U. H. E. Hansmann, J. Phys. Chem. **99**, 11276 (1995).
[21] F. Wang and D. P. Landau, Phys. Rev. Lett. **86**, 2050 (2001).
[22] Y. Sugita and Y. Okamoto, Chem. Phys. Lett. **329**, 261 (2000).
[23] A. Mitsutake, Y. Sugita, and Y. Okamoto, J. Chem. Phys. **118**, 6664 (2003).
[24] A. Mitsutake, Y. Sugita, and Y. Okamoto, J. Chem. Phys. **118**, 6676 (2003).
[25] A. D. MacKerell, Jr., D. Bashford, M. Bellott *et al.*, J. Phys. Chem. B **102**, 3586 (1998).
[26] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, J. Comput. Chem. **4**, 187 (1983).
[27] D. Brown and J. H. R. Clarke, Mol. Phys. **51**, 1243 (1984).
[28] S. Kirkpatrick, C. D. Gelatt, Jr., and M. P. Vecchi, Science **220**, 671 (1983).
[29] F. A. Momany, R. F. McGuire, A. W. Burgess, and H. A. Scheraga, J. Phys. Chem. **79**, 2361 (1975).
[30] G. Némethy, M. S. Pottle, and H. A. Scheraga, J. Phys. Chem. **87**, 1883 (1983).
[31] M. J. Sippl, G. Némethy, and H. A. Scheraga, J. Phys. Chem. **88**, 6231 (1984).
[32] A. Mitsutake, U. H. E. Hansmann, and Y. Okamoto, J. Mol. Graphics Modell. **16**, 226 (1998).
[33] M. H. Quenouille, Biometrika **43**, 353 (1956).
[34] R. G. Miller, Biometrika **61**, 1 (1974).
[35] B. A. Berg, *Markov Chain Monte Carlo Simulations and Their Statistical Analysis* (World Scientific, Singapore, 2004).
[36] R. A. Sayle and E. J. Milner-White, Trends Biochem. Sci. **20**, 374 (1995).