Chapter 4
Generalized-Ensemble Algorithms for Simulations of Complex Molecular Systems

Hisashi Okumura, Satoru G. Itoh, and Yuko Okamoto

Abstract In molecular simulations of complex systems with many degrees of freedom, conventional Monte Carlo and molecular dynamics simulations in canonical ensemble or isobaric-isothermal ensemble suffer from a great difficulty, in which simulations tend to get trapped in states of energy local minima. A simulation in generalized ensemble performs a random walk in specified variables and overcomes this difficulty. In this chapter, we review the generalized-ensemble algorithms. Replica-exchange method, multicanonical algorithm, and their extensions are described. Some simulation results based on these generalized-ensemble algorithms are also presented.

H. Okumura • S.G. Itoh
Department of Theoretical and Computational Molecular Science, Institute for Molecular Science, Okazaki, Aichi 444-8585, Japan

Research Center for Computational Science, Okazaki, Aichi 444-8585, Japan

Department of Structural Molecular Science, The Graduate University for Advanced Study, Okazaki, Aichi 444-8585, Japan
e-mail: hokumura@ims.ac.jp; itoh@ims.ac.jp

Y. Okamoto (✉)
Department of Physics, Graduate School of Science, Nagoya University, Nagoya, Aichi 464-8602, Japan

Structural Biology Research Center, Graduate School of Science, Nagoya University, Nagoya, Aichi 464-8602, Japan

Center for Computational Science, Graduate School of Engineering, Nagoya University, Nagoya, Aichi 464-8603, Japan
e-mail: okamoto@phys.nagoya-u.ac.jp

J. Leszczynski and M.K. Shukla (eds.), *Practical Aspects of Computational Chemistry II*, DOI 10.1007/978-94-007-0923-2_4,

4.1 Introduction

In complex molecular systems such as biomolecular systems, conventional Monte Carlo (MC) and molecular dynamics (MD) simulations at low temperatures in the canonical ensemble and those at low temperatures or high pressures in the isobaric-isothermal ensemble tend to get trapped in states of energy local minima, giving results in error. In order to overcome this difficulty, a class of simulation methods, which are referred to as the *generalized-ensemble algorithms*, are often employed (for reviews, see e.g., Refs. [1–5]). In a generalized-ensemble simulation, each state is weighted by a non-Boltzmann probability weight factor so that a random walk in potential energy space may be realized. The random walk allows the simulation to overcome any energy barrier and to sample much wider conformational space than by conventional methods. The generalized-ensemble algorithm was introduced to the molecular simulation field almost 20 years ago [6].

One of the most well-known generalized-ensemble algorithms is perhaps *replica-exchange method* (REM) [7] (see Ref. [8] for the MD version). Multiple replicas of the system in the canonical ensemble at different temperatures are simulated simultaneously, and every few steps, a pair of replicas at neighboring temperatures is exchanged. This causes a random walk in temperature for each replica, and the simulation can avoid getting trapped in states of energy local minima.

REM was extended to multidimensions/multivariables so that not only temperature but also other parameter values of the system are exchanged, and the method is referred to as *multidimensional replica-exchange method* (MREM) [9]. Various special cases of MREM were then proposed [10–15] (MREM is also known as *Hamiltonian replica-exchange method* [10]).

Another widely used generalized-ensemble algorithm is *multicanonical algorithm* (MUCA) [16, 17] (for a textbook, see, e.g., Ref. [18]; see also Refs. [19, 20] for the MD version). The probability weight factor, which is referred to as the multicanonical weight factor, is defined to be inversely proportional to the density of states so that a flat distribution in potential energy may be obtained. The uniform distribution induces a free random walk in the potential energy space, and the multiple-minima problem is overcome.

MUCA was extended so that flat distributions in parameters other than potential energy and/or multidimensional parameter space may be realized [21–28].

We remark that general formulations for multidimensional/multivariable generalized-ensemble algorithms (including REM and MUCA) were recently worked out [29–31].

In this chapter, we describe both REM and MUCA. We then present several of newly developed generalized-ensemble algorithms that are multidimensional/multicomponent extensions of REM and MUCA. The first algorithm is an example of MREM and referred to as the *van der Waals replica-exchange method* (vWREM) [32], where different values of van der Waals radius are exchanged. The second one is the *multioverlap algorithm* (MUOV), which performs a random walk

in the overlap space instead of the potential energy space [33–35]. Further extension of MUOV, which is referred to the *multicanonical-multioverlap algorithm* (MUCA-MUOV) [36–38] and realizes a random walk both in the potential energy space and the overlap space, is then given. The fourth method that we present here is the *multibaric-multithermal algorithm* (MUBATH), which realizes a random walk both in the potential energy space and in the volume space [39–45]. We remark that other generalized-ensemble algorithms for the isobaric-isothermal ensemble have also been developed [46–49]. Finally, examples of some simulation results based on these methods are presented.

4.2 Generalized-Ensemble Algorithms

4.2.1 Replica-Exchange Method

Let us consider a system of N atoms of mass m_k ($k = 1, \ldots, N$) with their coordinate vectors and momentum vectors denoted by $q \equiv \{\boldsymbol{q}_1, \ldots, \boldsymbol{q}_N\}$ and $p \equiv \{\boldsymbol{p}_1, \ldots, \boldsymbol{p}_N\}$, respectively. The Hamiltonian $H(q, p)$ of the system is the sum of the kinetic energy $K(p)$ and the potential energy $E(q)$:

$$H(q, p) = K(p) + E(q), \tag{4.1}$$

where

$$K(p) = \sum_{k=1}^{N} \frac{\boldsymbol{p}_k{}^2}{2m_k}. \tag{4.2}$$

In the canonical ensemble at temperature T, each state $x \equiv (q, p)$ with the Hamiltonian $H(q, p)$ is weighted by the Boltzmann factor:

$$W_{\mathrm{B}}(x; T) = \exp\left(-\beta H(q, p)\right), \tag{4.3}$$

where the inverse temperature β is defined by $\beta = 1/k_{\mathrm{B}}T$ (k_{B} is the Boltzmann constant). The average kinetic energy at temperature T is then given by

$$\langle K(p)\rangle_T = \left\langle \sum_{k=1}^{N} \frac{\boldsymbol{p}_k{}^2}{2m_k} \right\rangle_T = \frac{3}{2} N k_{\mathrm{B}} T. \tag{4.4}$$

Because the coordinates q and momenta p are decoupled in Eq. 4.1, we can suppress the kinetic energy part and can write the Boltzmann factor as

$$W_{\mathrm{B}}(x; T) \propto W_{\mathrm{B}}(E; T) = \exp(-\beta E). \tag{4.5}$$

The canonical probability distribution of potential energy $P_{\mathrm{NVT}}(E;T)$ is then given by the product of the density of states $n(E)$ and the Boltzmann weight factor $W_{\mathrm{B}}(E;T)$:

$$P_{\mathrm{NVT}}(E;T) \propto n(E) W_{\mathrm{B}}(E;T). \tag{4.6}$$

Because $n(E)$ is a rapidly increasing function and the Boltzmann factor decreases exponentially, the canonical ensemble yields a bell-shaped distribution of potential energy which has a maximum around the average potential energy at temperature T. The conventional MC or MD simulations at constant temperature are expected to yield $P_{\mathrm{NVT}}(E;T)$. A MC simulation based on the Metropolis algorithm [50] is performed with the following transition probability from a state x of potential energy E to a state x' of potential energy E':

$$w(x \to x') = \min\left(1, \frac{W_{\mathrm{B}}(E';T)}{W_{\mathrm{B}}(E;T)}\right) = \min\left(1, \exp\left(-\beta \Delta E\right)\right), \tag{4.7}$$

where

$$\Delta E = E' - E. \tag{4.8}$$

A MD simulation, on the other hand, is based on the following Newton equations of motion:

$$\dot{\boldsymbol{q}}_k = \frac{\boldsymbol{p}_k}{m_k}, \tag{4.9}$$

$$\dot{\boldsymbol{p}}_k = -\frac{\partial E}{\partial \boldsymbol{q}_k} = \boldsymbol{F}_k, \tag{4.10}$$

where $\boldsymbol{F}_k$ is the force acting on the kth atom ($k = 1, \cdots, N$). This set of equations actually yield the microcanonical ensemble, however, and we have to add a thermostat in order to obtain the canonical ensemble at temperature T. Here, we just follow Nosé's prescription [51, 52], and we have

$$\dot{\boldsymbol{q}}_k = \frac{\boldsymbol{p}_k}{m_k}, \tag{4.11}$$

$$\dot{\boldsymbol{p}}_k = -\frac{\partial E}{\partial \boldsymbol{q}_k} - \frac{\dot{s}}{s}\boldsymbol{p}_k = \boldsymbol{F}_k - \frac{\dot{s}}{s}\,\boldsymbol{p}_k, \tag{4.12}$$

$$\dot{s} = s\,\frac{P_s}{Q}, \tag{4.13}$$

$$\dot{P}_s = \sum_{k=1}^{N} \frac{{\boldsymbol{p}_k}^2}{m_k} - 3Nk_{\mathrm{B}}T = 3Nk_{\mathrm{B}}\left(T(t) - T\right), \tag{4.14}$$

where s is Nosé's scaling parameter, P_s is its conjugate momentum, Q is its mass, and the "instantaneous temperature" $T(t)$ is defined by

$$T(t) = \frac{1}{3Nk_{\mathrm{B}}} \sum_{k=1}^{N} \frac{\boldsymbol{p}_k(t)^2}{m_k}. \tag{4.15}$$

However, in practice, it is very difficult to obtain accurate canonical distributions of complex systems at low temperatures by conventional MC or MD simulation methods. This is because simulations at low temperatures tend to get trapped in one or a few of local-minimum-energy states. This difficulty is overcome by, for instance, the generalized-ensemble algorithms, which greatly enhance conformational sampling.

The *replica-exchange method* (REM) [7] is one of effective generalized-ensemble algorithms. The system for REM consists of M *noninteracting* copies (or replicas) of the original system in the canonical ensemble at M different temperatures T_m ($m = 1, \ldots, M$). We arrange the replicas so that there is always exactly one replica at each temperature. Then there exists a one-to-one correspondence between replicas and temperatures; the label i ($i = 1, \ldots, M$) for replicas is a permutation of the label m ($m = 1, \ldots, M$) for temperatures, and vice versa:

$$\begin{cases} i = i(m) \equiv f(m), \\ m = m(i) \equiv f^{-1}(i), \end{cases} \tag{4.16}$$

where $f(m)$ is a permutation function of m and $f^{-1}(i)$ is its inverse.

Let $X = \left\{x_1^{[i(1)]}, \ldots, x_M^{[i(M)]}\right\} = \left\{x_{m(1)}^{[1]}, \ldots, x_{m(M)}^{[M]}\right\}$ stand for a "state" in this generalized ensemble. Each "substate" $x_m^{[i]}$ is specified by the coordinates $q^{[i]}$ and momenta $p^{[i]}$ of N atoms in replica i at temperature T_m:

$$x_m^{[i]} \equiv \left(q^{[i]}, p^{[i]}\right)_m. \tag{4.17}$$

Because the replicas are noninteracting, the weight factor for the state X in this generalized ensemble is given by the product of Boltzmann factors for each replica (or at each temperature):

$$W_{\mathrm{REM}}(X) = \prod_{i=1}^{M} \exp\left\{-\beta_{m(i)} H\left(q^{[i]}, p^{[i]}\right)\right\} = \prod_{m=1}^{M} \exp\left\{-\beta_m H\left(q^{[i(m)]}, p^{[i(m)]}\right)\right\}, \tag{4.18}$$

where $i(m)$ and $m(i)$ are the permutation functions in Eq. 4.16.

We now consider exchanging a pair of replicas in this ensemble. Suppose we exchange replicas i and j which are at temperatures T_m and T_n, respectively:

$$X = \left\{\ldots, x_m^{[i]}, \ldots, x_n^{[j]}, \ldots\right\} \longrightarrow X' = \left\{\ldots, x_m^{[j]\prime}, \ldots, x_n^{[i]\prime}, \ldots\right\}. \tag{4.19}$$

Here, i, j, m, and n are related by the permutation functions in Eq. 4.16, and the exchange of replicas introduces a new permutation function f':

$$\begin{cases} i = f(m) \longrightarrow j = f'(m), \\ j = f(n) \longrightarrow i = f'(n). \end{cases} \tag{4.20}$$

The exchange of replicas can be written in more detail as

$$\begin{cases} x_m^{[i]} \equiv \left(q^{[i]}, p^{[i]}\right)_m \longrightarrow x_m^{[j]\prime} \equiv \left(q^{[j]}, p^{[j]\prime}\right)_m, \\ x_n^{[j]} \equiv \left(q^{[j]}, p^{[j]}\right)_n \longrightarrow x_n^{[i]\prime} \equiv \left(q^{[i]}, p^{[i]\prime}\right)_n, \end{cases} \tag{4.21}$$

where the definitions for $p^{[i]\prime}$ and $p^{[j]\prime}$ will be given below.

In the original implementation of the REM [7], Monte Carlo method was used, and only the coordinates q (and the potential energy function $E(q)$) had to be taken into account. In molecular dynamics method, on the other hand, we also have to deal with the momenta p. We proposed the following momentum assignment in Eq. 4.21 [8]:

$$\begin{cases} p^{[i]\prime} \equiv \sqrt{\dfrac{T_n}{T_m}}\, p^{[i]}, \\ p^{[j]\prime} \equiv \sqrt{\dfrac{T_m}{T_n}}\, p^{[j]}, \end{cases} \tag{4.22}$$

which we believe is the simplest and the most natural. This assignment means that we just rescale uniformly the velocities of all the atoms in the replicas by the square root of the ratio of the two temperatures so that the temperature condition in Eq. 4.4 may be satisfied immediately after replica exchange is accepted. We remark that general momentum rescaling formulae were derived for various thermostats in Ref. [53].

The transition probability of this replica-exchange process is given by the usual Metropolis criterion:

$$w(X \to X') \equiv w\left(x_m^{[i]} \,\middle|\, x_n^{[j]}\right) = \min\left(1, \frac{W_{\mathrm{REM}}(X')}{W_{\mathrm{REM}}(X)}\right) = \min\left(1, \exp\left(-\Delta\right)\right), \tag{4.23}$$

where in the second expression (i.e., $w(x_m^{[i]}|x_n^{[j]})$), we explicitly wrote the pair of replicas (and temperatures) to be exchanged. From Eq. 4.22, the kinetic energy terms all cancel out in Eq. 4.23, and Δ becomes

$$\Delta = \beta_m \left(E\left(q^{[j]}\right) - E\left(q^{[i]}\right)\right) - \beta_n \left(E\left(q^{[j]}\right) - E\left(q^{[i]}\right)\right), \tag{4.24}$$

$$= \left(\beta_m - \beta_n\right)\left(E\left(q^{[j]}\right) - E\left(q^{[i]}\right)\right). \tag{4.25}$$

Here, i, j, m, and n are related by the permutation functions in Eq. 4.16 before the replica exchange:

$$\begin{cases} i = f(m), \\ j = f(n). \end{cases} \tag{4.26}$$

Note that after introducing the momentum rescaling in Eq. 4.22, we have the same Metropolis criterion for replica exchanges, i.e., Eqs. 4.23 and 4.25, for both MC and MD versions.

Without loss of generality, we can assume $T_1 < T_2 < \cdots < T_M$. The lowest temperature T_1 should be sufficiently low so that the simulation can explore the global-minimum-energy region, and the highest temperature T_M should be sufficiently high so that no trapping in an energy-local-minimum state occurs. A REM simulation is then realized by alternately performing the following two steps:

1. Each replica in canonical ensemble of the fixed temperature is simulated *simultaneously* and *independently* for a certain MC or MD steps.
2. A pair of replicas at neighboring temperatures, say $x_m^{[i]}$ and $x_{m+1}^{[j]}$, is exchanged with the probability $w\left(x_m^{[i]} \middle| x_{m+1}^{[j]}\right)$ in Eq. 4.23.

A random walk in "temperature space" is realized for each replica, which in turn induces a random walk in potential energy space. This alleviates the problem of getting trapped in states of energy local minima.

After a long production run of a REM simulation, the canonical expectation value of a physical quantity A at temperature T_m ($m = 1, \ldots, M$) can be calculated by the usual arithmetic mean:

$$\langle A \rangle_{T_m} = \frac{1}{n_m} \sum_{k=1}^{n_m} A\left(x_m(k)\right), \tag{4.27}$$

where $x_m(k)$ ($k = 1, \ldots, n_m$) are the configurations obtained at temperature T_m and n_m is the total number of measurements made at $T = T_m$. The expectation value at any intermediate temperature T ($= 1/k_{\mathrm{B}}\beta$) can also be obtained as follows:

$$\langle A \rangle_T = \frac{\sum_E A(E) P_{\mathrm{NVT}}(E;T)}{\sum_E P_{\mathrm{NVT}}(E;T)} = \frac{\sum_E A(E) n(E) \exp(-\beta E)}{\sum_E n(E) \exp(-\beta E)}. \tag{4.28}$$

Here, the explicit form of the physical quantity A should be known as a function of potential energy E. For instance, $A(E) = E$ gives the average potential energy $\langle E \rangle_T$ as a function of temperature, and $A(E) = \beta^2 (E - \langle E \rangle_T)^2$ gives specific heat.

The density of states $n(E)$ in Eq. 4.28 is given by the multiple-histogram reweighting techniques [54, 55] as follows (an extension of the multiple-histogram method is also referred to as the *weighted histogram analysis method* (WHAM) [55]). Let $N_m(E)$ and n_m be respectively the potential energy histogram and the total number of samples obtained at temperature $T_m = 1/k_B\beta_m$ ($m = 1, \cdots, M$). The best estimate of the density of states is then given by [54, 55]

$$n(E) = \frac{\sum_{m=1}^{M} g_m^{-1} N_m(E)}{\sum_{m=1}^{M} g_m^{-1} n_m \exp(f_m - \beta_m E)}, \tag{4.29}$$

where we have for each m ($= 1, \ldots, M$)

$$\exp(-f_m) = \sum_E n(E) \exp(-\beta_m E). \tag{4.30}$$

Here, $g_m = 1 + 2\tau_m$, and τ_m is the integrated autocorrelation time at temperature T_m. For many systems, the quantity g_m can safely be set to be a constant in the reweighting formulae [55], and hereafter, we set $g_m = 1$. Note that Eqs. 4.29 and 4.30 are solved self-consistently by iteration [54, 55] to obtain the density of states $n(E)$ and the dimensionless Helmholtz free energy f_m.

Moreover, the ensemble averages of any physical quantity A (including those that cannot be expressed as functions of potential energy) at any temperature T ($=1/k_B\beta$) can now be obtained from the "trajectory" of configurations of the production run. Namely, we first obtain f_m ($m = 1, \cdots, M$) by solving Eqs. 4.29 and 4.30 self-consistently, and then we have [56]

$$\langle A \rangle_T = \frac{\sum_{m=1}^{M} \sum_{k=1}^{n_m} A(x_m(k)) \dfrac{1}{\sum_{\ell=1}^{M} n_\ell \exp[f_\ell - \beta_\ell E(x_m(k))]} \exp[-\beta E(x_m(k))]}{\sum_{m=1}^{M} \sum_{k=1}^{n_m} \dfrac{1}{\sum_{\ell=1}^{M} n_\ell \exp[f_\ell - \beta_\ell E(x_m(k))]} \exp[-\beta E(x_m(k))]}, \tag{4.31}$$

where $x_m(k)$ ($k = 1, \cdots, n_m$) are the configurations obtained at temperature T_m.

4.2.2 Extensions of the Replica-Exchange Method

4.2.2.1 Multidimensional Replica-Exchange Method

We now describe the *multidimensional replica-exchange method* (MREM) [9]. The crucial observation that led to this algorithm is as follows: As long as we have M *noninteracting* replicas of the original system, the Hamiltonian $H(q, p)$ of the system does not have to be identical among the replicas, and it can depend on a parameter with different parameter values for different replicas.

Let us consider a generalized potential energy function $E_{\boldsymbol{\lambda}}(x)$, which depends on L parameters $\boldsymbol{\lambda} = (\lambda^{(1)}, \ldots, \lambda^{(L)})$, of a system in state x. The system for MREM consists of M noninteracting replicas of the original system in the "canonical ensemble" with $M(= M_0 \times M_1 \times \cdots \times M_L)$ different parameter sets $\boldsymbol{\Lambda}_m$ $(m = 1, \ldots, M)$, where $\boldsymbol{\Lambda}_m \equiv (T_{m_0}, \boldsymbol{\lambda}_m) \equiv (T_{m_0}, \lambda^{(1)}_{m_1}, \ldots, \lambda^{(L)}_{m_L})$ with $m_0 = 1, \ldots, M_0, m_\ell = 1, \ldots, M_\ell$ $(\ell = 1, \ldots, L)$. Because the replicas are noninteracting, the weight factor is given by the product of Boltzmann-like factors for each replica:

$$W_{\mathrm{MREM}} \equiv \prod_{m_0=1}^{M_0} \prod_{m_1=1}^{M_1} \cdots \prod_{m_L=1}^{M_L} \exp\left(-\beta_{m_0} E_{\boldsymbol{\lambda}_m}\right). \tag{4.32}$$

Without loss of generality, we can order the parameters so that $T_1 < T_2 < \cdots < T_{M_0}$ and $\lambda^{(\ell)}_1 < \lambda^{(\ell)}_2 < \cdots < \lambda^{(\ell)}_{M_\ell}$ (for each $\ell = 1, \cdots, L$). A MREM simulation is realized by alternately performing the following two steps:

1. For each replica, a "canonical" MC or MD simulation at the fixed parameter values is carried out simultaneously and independently for a certain steps.
2. We exchange a pair of replicas i and j which are at the parameter sets $\boldsymbol{\Lambda}_m$ and $\boldsymbol{\Lambda}_{m+1}$, respectively. The transition probability for this replica-exchange process is given by

$$w(\boldsymbol{\Lambda}_m \leftrightarrow \boldsymbol{\Lambda}_{m+1}) = \min\left(1, \exp(-\Delta)\right), \tag{4.33}$$

where we have

$$\Delta = (\beta_{m_0} - \beta_{m_0+1}) \left(E_{\boldsymbol{\lambda}_m}\left(q^{[j]}\right) - E_{\boldsymbol{\lambda}_m}\left(q^{[i]}\right)\right), \tag{4.34}$$

for T-exchange, and

$$\Delta = \beta_{m_0} \left[\left(E_{\boldsymbol{\lambda}_{m_\ell}}(q^{[j]}) - E_{\boldsymbol{\lambda}_{m_\ell}}(q^{[i]})\right) - \left(E_{\boldsymbol{\lambda}_{m_\ell+1}}(q^{[j]}) - E_{\boldsymbol{\lambda}_{m_\ell+1}}(q^{[i]})\right)\right], \tag{4.35}$$

for $\lambda^{(\ell)}$-exchange (for one of $\ell = 1, \cdots, L$). Here, $q^{[i]}$ and $q^{[j]}$ stand for configuration variables for replicas i and j, respectively, before the replica exchange.

4.2.2.2 van der Waals Replica-Exchange Method

We now describe a special example of MREM, which we refer to as the *van der Waals Replica-Exchange Method* (vWREM) [32].

We consider a system consisting of solute molecule(s) in explicit solvent. We can write the total potential energy as follows:

$$E_\lambda(q) = E_\mathrm{p}(q_\mathrm{p}) + E_\mathrm{ps}(q_\mathrm{p}, q_\mathrm{s}) + E_\mathrm{s}(q_\mathrm{s}), \tag{4.36}$$

where E_p is the potential energy for the atoms in the solute only, E_ps is the interaction term between solute atoms and solvent atoms, and E_s is the potential energy for the atoms of the solvent molecules only. Here, $q = \{q_\mathrm{p}, q_\mathrm{s}\}$, where q_p and q_s are the coordinate vectors of the solute atoms and the solvent atoms, respectively, and denoted by $q_\mathrm{p} \equiv \left\{\boldsymbol{q}_1, \ldots, \boldsymbol{q}_{N_\mathrm{p}}\right\}$ and $q_\mathrm{s} \equiv \left\{\boldsymbol{q}_{N_\mathrm{p}+1}, \ldots, \boldsymbol{q}_N\right\}$. ($N_\mathrm{p}$ is the total number of atoms in the solute.)

We are more concerned with effective sampling of the conformational space of the solute itself than that of the solvent molecules. The steric hindrance of the solute conformations are governed by the van der Waals radii of each atom in the solute. Namely, when the van der Waals radii are large, the solute molecule is bulky, and we have more steric hindrance among the solute atoms by the Lennard-Jones interactions, and when it is small, the solute molecule can move more freely. We thus introduce a parameter λ that scales the van der Waals radius of each atom in the solute by

$$\sigma_{k\ell} \longrightarrow \lambda\sigma_{k\ell} \tag{4.37}$$

and write the Lennard-Jones energy term within E_p in Eq. 4.36 as follows:

$$V_\lambda\left(q_\mathrm{p}\right) = \sum_{k=1}^{N_\mathrm{p}-1} \sum_{\ell=k+1}^{N_\mathrm{p}} 4\epsilon_{k\ell} \left\{ \left(\frac{\lambda\sigma_{k\ell}}{r_{k\ell}}\right)^{12} - \left(\frac{\lambda\sigma_{k\ell}}{r_{k\ell}}\right)^{6} \right\}, \tag{4.38}$$

where $r_{k\ell}$ is the distance between atoms k and ℓ in the solute and $\epsilon_{k\ell}$ and $\sigma_{k\ell}$ are the corresponding Lennard-Jones parameters. The original potential energy is recovered when $\lambda = 1$, and the steric hindrance of solute conformations is reduced when $\lambda < 1$. We remark that this is the only λ-dependent term in E_λ in Eq. 4.36.

We prepare M values of λ, λ_m ($m = 1, \ldots, M$). Without loss of generality, we can assume that the parameter values are ordered as $\lambda_1 < \cdots < \lambda_M$. Here, we consider the case in which temperature is fixed to be $T_0 = 1/k_\mathrm{B}\beta_0$. The vWREM is realized by alternately performing the following two steps:

1. For each replica, a canonical MC or MD simulation at the corresponding parameter value λ_m is carried out simultaneously and independently for a certain steps with the corresponding Boltzmann factor of Eq. 4.3 for each replica.

2. We exchange a pair of replicas i and j which are at the neighboring parameter values λ_m and λ_{m+1}, respectively. The transition probability for this replica-exchange process is given by Eq. 4.33, where Δ in Eq. 4.35 now reads

$$\Delta = \beta_0 \left[\left(V_{\lambda_m} \left(q_{\mathrm{p}}^{[j]} \right) - V_{\lambda_m} \left(q_{\mathrm{p}}^{[i]} \right) \right) - \left(V_{\lambda_{m+1}} \left(q_{\mathrm{p}}^{[j]} \right) - V_{\lambda_{m+1}} \left(q_{\mathrm{p}}^{[i]} \right) \right) \right]. \tag{4.39}$$

Here, V_λ is the Lennard-Jones potential energy in Eq. 4.38 among the solute atoms only.

Note that because the λ dependence of E_λ exists only in V_λ, the rest of the terms have been canceled out in Eq. 4.35.

We see that Eq. 4.39 includes only the coordinates q_{p} of the atoms in the solute only and is independent of the coordinates q_{s} of solvent molecules. Because $N_{\mathrm{p}} \ll N$ usually holds, the difficulty in the usual REM that the number of required replicas increases with the number of degrees of freedom is much alleviated in this formalism.

We remark that in order to further enhance the conformational sampling, we can perform a two-dimensional REM in both temperature and λ, using Eqs. 4.34 and 4.35.

4.2.2.3 Reweighting Techniques

The results from MREM simulations with different parameter values can be analyzed by the reweighting techniques [54, 55]. Suppose that we have carried out a MREM simulation at a constant temperature T_0 with M replicas corresponding to M parameter values λ_m ($m = 1, \ldots, M$).

For appropriate reaction coordinates ξ_1 and ξ_2, the canonical probability distribution $P_{T,\lambda}(\xi_1, \xi_2)$ with any parameter value λ at any temperature T can be calculated from

$$P_{T,\lambda}(\xi_1, \xi_2) = \sum_{E_{\lambda_1}, \ldots, E_{\lambda_M}} \frac{\sum_{m=1}^{M} N_m(E_{\lambda_1}, \ldots, E_{\lambda_M}; \xi_1, \xi_2) e^{-\beta E_\lambda}}{\sum_{m=1}^{M} n_m e^{f_{T_0,\lambda_m} - \beta_0 E_{\lambda_m}}}, \tag{4.40}$$

and

$$e^{-f_{T_0,\lambda_m}} = \sum_{\xi_1, \xi_2} P_{T_0,\lambda_m}(\xi_1, \xi_2). \tag{4.41}$$

Here, $N_m(E_{\lambda_1}, \ldots, E_{\lambda_M}; \xi_1, \xi_2)$ is the histogram of the M-dimensional energy distributions at the parameter value λ_m and the reaction coordinate values (ξ_1, ξ_2), which was obtained by the MREM simulation, and n_m is the total number of samples obtained at the parameter value λ_m. Note that this probability distribution is not normalized. Equations 4.40 and 4.41 are solved self-consistently by iteration. Note also that these equations can be easily generalized to any reaction coordinates $(\xi_1, \xi_2, \ldots)$.

From the probability distribution $P_{T,\lambda}(\xi_1, \xi_2)$ in Eq. 4.40, the expectation value of a physical quantity A with any parameter value λ at any temperature T is given by

$$\langle A \rangle_{T,\lambda} = \frac{\sum_{\xi_1,\xi_2} A(\xi_1, \xi_2) P_{T,\lambda}(\xi_1, \xi_2)}{\sum_{\xi_1,\xi_2} P_{T,\lambda}(\xi_1, \xi_2)}. \tag{4.42}$$

We can also calculate the free energy (or the potential of mean force) as a function of the reaction coordinates ξ_1 and ξ_2 with any parameter value λ at any temperature T from

$$F_{T,\lambda}(\xi_1, \xi_2) = -k_B T \ln P_{T,\lambda}(\xi_1, \xi_2). \tag{4.43}$$

By utilizing these equations, therefore, we can obtain various physical quantities from the MREM simulations with the original and non-original parameter values. We remark that although we wrote *any* T in Eqs. 4.40, 4.42, and 4.43 above, the valid T value is limited in the vicinity of T_0. We also need the T-exchange process in Eq. 4.34 in order to have accurate average quantities for a wide range of T values.

4.2.3 Multicanonical Algorithm

The next generalized-ensemble algorithm that we present is the *multicanonical algorithm* (MUCA) [16, 17]. In the multicanonical ensemble, each state is weighted by a non-Boltzmann weight factor $W_{\mathrm{MUCA}}(E)$ (which we refer to as the *multicanonical weight factor*) so that a uniform potential energy distribution $P_{\mathrm{MUCA}}(E)$ may be obtained:

$$P_{\mathrm{MUCA}}(E) \propto n(E) W_{\mathrm{MUCA}}(E) \equiv \text{constant}. \tag{4.44}$$

The flat distribution implies that a free one-dimensional random walk in the potential energy space is realized in this ensemble. This allows the simulation to escape from any local-minimum-energy states and to sample the configurational space much more widely than the conventional canonical MC or MD methods.

The definition in Eq. 4.44 implies that the multicanonical weight factor is inversely proportional to the density of states, and we can write it as follows:

$$W_{\mathrm{MUCA}}(E) \equiv \exp\left[-\beta_0 E_{\mathrm{MUCA}}(E;T_0)\right] = \frac{1}{n(E)}, \tag{4.45}$$

where we have chosen an arbitrary reference temperature, $T_0 = 1/k_{\mathrm{B}}\beta_0$, and the "*multicanonical potential energy*" is defined by

$$E_{\mathrm{MUCA}}(E;T_0) \equiv k_{\mathrm{B}}T_0 \ln n(E) = T_0 S(E). \tag{4.46}$$

Here, $S(E)$ is the entropy in the microcanonical ensemble. Because the density of states of the system is usually unknown, the multicanonical weight factor has to be determined numerically by iterations of short preliminary runs [16, 17].

A multicanonical MC simulation is performed, for instance, with the usual Metropolis criterion [50]: The transition probability of state x with potential energy E to state x' with potential energy E' is given by

$$\begin{aligned} w(x \to x') &= \min\left(1, \frac{W_{\mathrm{MUCA}}(E')}{W_{\mathrm{MUCA}}(E)}\right) = \min\left(1, \frac{n(E)}{n(E')}\right) \\ &= \min\left(1, \exp\left(-\beta_0 \Delta E_{\mathrm{MUCA}}\right)\right), \end{aligned} \tag{4.47}$$

where

$$\Delta E_{\mathrm{MUCA}} = E_{\mathrm{MUCA}}(E';T_0) - E_{\mathrm{MUCA}}(E;T_0). \tag{4.48}$$

The MD algorithm in the multicanonical ensemble also naturally follows from Eq. 4.45, in which the regular constant temperature MD simulation (with $T = T_0$) is performed by replacing E by E_{MUCA} in Eq. 4.12 [19, 20]:

$$\dot{\boldsymbol{p}}_k = -\frac{\partial E_{\mathrm{MUCA}}(E;T_0)}{\partial \boldsymbol{q}_k} - \frac{\dot{s}}{s}\boldsymbol{p}_k = \frac{\partial E_{\mathrm{MUCA}}(E;T_0)}{\partial E}\boldsymbol{F}_k - \frac{\dot{s}}{s}\boldsymbol{p}_k. \tag{4.49}$$

Let $N_{\mathrm{MUCA}}(E)$ be the histogram of potential energy distribution $P_{\mathrm{MUCA}}(E)$ obtained by the production run. The best estimate of the density of states can then be given by the single-histogram reweighting techniques [57] as follows (see the proportionality relation in Eq. 4.44):

$$n(E) = \frac{N_{\mathrm{MUCA}}(E)}{W_{\mathrm{MUCA}}(E)}. \tag{4.50}$$

By substituting this quantity into Eq. 4.28, one can calculate ensemble averages of physical quantity $A(E)$ as a function of temperature. Moreover, the ensemble averages of any physical quantity A (including those that cannot be expressed as functions of potential energy) at any temperature T $(= 1/k_{\mathrm{B}}\beta)$ can also be obtained

as long as one stores the "trajectory" of configurations from the production run. Namely, we have

$$\langle A \rangle_T = \frac{\sum_{k=1}^{n_s} A(x_k) W_{\mathrm{MUCA}}^{-1}(E(x_k)) \exp\left[-\beta E(x_k)\right]}{\sum_{k=1}^{n_s} W_{\mathrm{MUCA}}^{-1}(E(x_k)) \exp\left[-\beta E(x_k)\right]}, \tag{4.51}$$

where x_k is the configuration at the kth MC (or MD) step and n_s is the total number of configurations stored.

4.2.4 *Extensions of Multicanonical Algorithm*

4.2.4.1 Multioverlap Algorithm and Multicanonical-Multioverlap Algorithm

While MUCA yields a flat distribution in potential energy and performs a random walk in potential energy space, we can, in principle, choose any other variable and induce a random walk in that variable. One such example is the *multioverlap algorithm* (MUOV) [33–35]. Here, we choose a protein system and define the overlap in the space of dihedral angles by [58]

$$O = 1 - d, \tag{4.52}$$

where d is the dihedral-angle distance given by

$$d = \frac{1}{n\pi} \sum_i d_a(\theta_i, \theta_i^0). \tag{4.53}$$

θ_i is the dihedral angle i, and θ_i^0 is the dihedral angle i of the reference conformation. The distance $d_a(\theta_i, \theta_i^0)$ between two dihedral angles is defined by

$$d_a(\theta_i, \theta_i^0) = \min(|\theta_i - \theta_i^0|, 2\pi - |\theta_i - \theta_i^0|). \tag{4.54}$$

The dihedral-angle distance d in Eq. 4.53 takes a value in the range $0 \leq d \leq 1$. If $d = 0$, all dihedral angles are coincident with those of the reference conformation. The dihedral-angle distance is thus an indicator of how similar the conformation is to the reference conformation. As one can see in Eq. 4.52, the dihedral-angle distance d is equivalent to the overlap O. We will deal with the dihedral-angle distance instead of the overlap hereafter.

In the multioverlap ensemble at a constant temperature T_0, the probability distribution is given by the following non-Boltzmann weight factor, which we refer to as the multioverlap weight factor:

$$W_{\mathrm{muov}}(d, E; T_0) = e^{-\beta_0 E_{\mathrm{muov}}}, \tag{4.55}$$

where E_{muov} is the "multioverlap potential energy" defined by

$$E_{\mathrm{muov}}(d, E; T_0) = E - k_{\mathrm{B}} T_0 f(d; T_0). \tag{4.56}$$

The function $f(d; T_0)$ 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

$$W_{\mathrm{muov}}(d_1, \ldots, d_L, E; T_0) = e^{-\beta_0 E_{\mathrm{muov}}} \equiv e^{-\beta_0 E + f(d_1, \ldots, d_L; T_0)}, \tag{4.57}$$

where L is the number of the reference conformations and d_i is the dihedral-angle distance, with respect to reference conformation i ($i = 1, \ldots, L$). The function $f(d_1, \ldots, d_L; T_0)$ is the dimensionless free energy with the fixed value of dihedral-angle distances $d_1, \cdots, d_L$. The dimensionless free energy $f(d_1, \ldots, d_L; T_0)$ is defined so that the probability distribution of dihedral-angle distances $P_{\mathrm{muov}}(d_1, \ldots, d_L; T_0)$ is flat:

$$\begin{aligned} P_{\mathrm{muov}}(d_1, \ldots, d_L; T_0) &= \int dE \; P_{\mathrm{muov}}(d_1, \ldots, d_L, E; T_0) \\ &\propto \int dE \; n(d_1, \ldots, d_L, E) W_{\mathrm{muov}}(d_1, \ldots, d_L, E; T_0) \\ &= \int dE \; n(d_1, \ldots, d_L, E) e^{-\beta_0 E + f(d_1, \cdots, d_L; T_0)} \\ &\equiv \text{constant}, \end{aligned} \tag{4.58}$$

where $P_{\mathrm{muov}}(d_1, \ldots, d_L, E; T_0)$ is the probability distribution of potential energy and dihedral-angle distances, and $n(d_1, \ldots, d_L, E)$ is its density of states.

The MD algorithm in the multioverlap ensemble also naturally follows from Eq. 4.57, in which the regular constant temperature MD simulation (with $T = T_0$) is performed by replacing E by E_{muov} in Eq. 4.12 [35, 36]:

$$\begin{aligned} \dot{\boldsymbol{p}}_k &= -\frac{\partial E_{\mathrm{muov}}}{\partial \boldsymbol{q}_k}(d_1, \ldots, d_L, E; T_0) - \frac{\dot{s}}{s} \boldsymbol{p}_k \\ &= \boldsymbol{F}_k + k_{\mathrm{B}} T_0 \frac{\partial f}{\partial \boldsymbol{q}_k}(d_1, \ldots, d_L; T_0) - \frac{\dot{s}}{s} \boldsymbol{p}_k. \end{aligned} \tag{4.59}$$

The multioverlap weight factor, or the dimensionless free energy, is not a priori known and has to be determined by the usual iterations of short simulations [2, 18]. Suppose that we have determined an appropriate dimensionless free energy $f(d_1, \ldots, d_L; T_0)$ at temperature T_0 and that we have made a production run at this temperature. The results of the multioverlap production run can then be analyzed by the reweighting techniques [57]. Namely, the expectation value of a physical quantity A at any temperature T is given by

$$\begin{aligned}\langle A\rangle_T &= \frac{\sum_{d_1,\cdots,d_L,E} A(d_1,\cdots,d_L,E) N_{\mathrm{muov}}(d_1,\cdots,d_L,E) W_{\mathrm{muov}}(d_1,\cdots,d_L,E;T_0)^{-1} e^{-\beta E}}{\sum_{d_1,\cdots,d_L,E} N_{\mathrm{muov}}(d_1,\cdots,d_L,E) W_{\mathrm{muov}}(d_1,\cdots,d_L,E;T_0)^{-1} e^{-\beta E}} \\ &= \frac{\sum_{d_1,\cdots,d_L,E} A(d_1,\cdots,d_L,E) N_{\mathrm{muov}}(d_1,\cdots,d_L,E) e^{-(\beta-\beta_0)E - f(d_1,\cdots,d_L;T_0)}}{\sum_{d_1,\cdots,d_L,E} N_{\mathrm{muov}}(d_1,\cdots,d_L,E) e^{-(\beta-\beta_0)E - f(d_1,\cdots,d_L;T_0)}},\end{aligned} \tag{4.60}$$

where $N_{\mathrm{muov}}(d_1, \ldots, d_L, E)$ is the histogram of the probability distribution $P_{\mathrm{muov}}(d_1, \ldots, d_L, E; T_0)$ of potential energy and dihedral-angle distances that was obtained by the multioverlap production run.

The multioverlap algorithm can further be combined with the multicanonical algorithm as follows (this method is referred to as the *multicanonical-multioverlap algorithm* (MUCA-MUOV)) [36]. In analogy with the multicanonical ensemble in Eq. 4.44 or the multioverlap ensemble in Eq. 4.58, by employing the non-Boltzmann weight factor $W_{\mathrm{mcmo}}(d_1, \ldots, d_L, E)$, which we refer to as the multicanonical-multioverlap weight factor, a uniform probability distribution with respect to the potential energy and dihedral-angle distances is obtained:

$$P_{\mathrm{mcmo}}(d_1, \ldots, d_L, E) \propto n(d_1, \ldots, d_L, E) W_{\mathrm{mcmo}}(d_1, \ldots, d_L, E) \equiv \text{constant}. \tag{4.61}$$

In this method, we obtain a random walk not only in the dihedral-angle distance space but also in the potential energy space.

4.2.4.2 Multibaric-Multithermal Algorithm

Besides the canonical ensemble, molecular simulations in the isobaric-isothermal ensemble are also commonly used. This is because most experiments are carried out under the constant pressure and constant temperature conditions. The canonical probability distribution $P_{\mathrm{B}}(E; T_0)$ in Eq. 4.6 is here replaced by the isobaric-isothermal distribution $P_{\mathrm{NPT}}(E, V; T_0, P_0)$ for potential energy E and volume V:

$$P_{\mathrm{NPT}}(E, V; T_0, P_0) \equiv n(E, V) \mathrm{e}^{-\beta_0 \mathcal{H}}. \tag{4.62}$$

Here, the density of states $n(E, V)$ is given as a function of both E and V, and $\mathcal{H}$ is the "enthalpy" (without the kinetic energy contributions):

$$\mathcal{H} = E + P_0 V, \tag{4.63}$$

where P_0 is the pressure at which simulations are performed. This weight factor produces an isobaric-isothermal ensemble at constant temperature (T_0) and constant pressure (P_0). This ensemble has bell-shaped distributions in both E and V.

As for the MD methods in this ensemble, we just present the Nosé-Andersen algorithm [51, 52, 59]. The equations of motion in Eqs. 4.11–4.14 are now generalized as follows:

$$\dot{\boldsymbol{q}}_k = \frac{\boldsymbol{p}_k}{m_k} + \frac{\dot{V}}{3V}\boldsymbol{q}_k, \tag{4.64}$$

$$\dot{\boldsymbol{p}}_k = -\frac{\partial \mathcal{H}}{\partial \boldsymbol{q}_k} - \left(\frac{\dot{s}}{s} + \frac{\dot{V}}{3V}\right)\boldsymbol{p}_k = \boldsymbol{F}_k - \left(\frac{\dot{s}}{s} + \frac{\dot{V}}{3V}\right)\boldsymbol{p}_k, \tag{4.65}$$

$$\dot{s} = s\,\frac{P_s}{Q}, \tag{4.66}$$

$$\dot{P}_s = \sum_{i=1}^{N} \frac{\boldsymbol{p}_i^2}{m_i} - 3Nk_{\mathrm{B}}T_0 = 3Nk_{\mathrm{B}}\left(T(t) - T_0\right), \tag{4.67}$$

$$\dot{V} = s\frac{P_V}{M}, \tag{4.68}$$

$$\dot{P}_V = s\left\{\frac{1}{3V}\left(\sum_{i=1}^{N}\frac{\boldsymbol{p}_i^2}{m_i} - \sum_{i=1}^{N}\boldsymbol{q}_i \cdot \frac{\partial \mathcal{H}}{\partial \boldsymbol{q}_i}\right) - \frac{\partial \mathcal{H}}{\partial V}\right\} = s\left[P(t) - P_0\right], \tag{4.69}$$

where M is the artificial mass associated with the volume, P_V is the conjugate momentum for the volume, and the "instantaneous pressure" $P(t)$ is defined by

$$\begin{aligned} P(t) &= \frac{1}{3V}\left(\sum_{i=1}^{N}\frac{\boldsymbol{p}_i(t)^2}{m_i} - \sum_{i=1}^{N}\boldsymbol{q}_i(t)\cdot\frac{\partial \mathcal{H}}{\partial \boldsymbol{q}_i}(t)\right) \\ &= \frac{1}{3V}\left(\sum_{i=1}^{N}\frac{\boldsymbol{p}_i(t)^2}{m_i} + \sum_{i=1}^{N}\boldsymbol{q}_i(t)\cdot\boldsymbol{F}_i(t)\right). \end{aligned} \tag{4.70}$$

We now introduce the idea of the multicanonical technique into the isobaric-isothermal ensemble method and refer to this generalized-ensemble algorithm as the *multibaric-multithermal algorithm* (MUBATH) [39, 40, 42, 43]. The molecular simulations in this generalized ensemble perform random walks both in the potential energy space and in the volume space.

In the multibaric-multithermal ensemble, each state is sampled by the multibaric-multithermal weight factor $W_{\rm mbt}(E,V) \equiv \exp\{-\beta_0 \mathcal{H}_{\rm mbt}(E,V)\}$ ($\mathcal{H}_{\rm mbt}$ is referred to as the multibaric-multithermal enthalpy) so that a uniform distribution in both potential energy and volume may be obtained:

$$P_{\rm mbt}(E,V) \propto n(E,V) W_{\rm mbt}(E,V) = n(E,V) \exp\{-\beta_0 \mathcal{H}_{\rm mbt}(E,V)\} \equiv \text{constant}. \tag{4.71}$$

In order to perform the multibaric-multithermal MD simulation, we just solve the above equations of motion (Eqs. 4.64–4.69) for the regular isobaric-isothermal ensemble (with $T = T_0$ and $P = P_0$), where the enthalpy $\mathcal{H}$ is replaced by the multibaric-multithermal enthalpy $\mathcal{H}_{\rm mbt}$ in Eqs. 4.65 and 4.69 [42].

The multibaric-multithermal weight factor is, however, not a priori known and has to be determined by the usual iterations of short simulations [2, 18]. After an optimal weight factor $W_{\rm mbt}(E,V)$ is obtained, a long production simulation is performed for data collection. We employ the reweighting techniques [57] for the results of the production run to calculate the isobaric-isothermal-ensemble averages. The probability distribution $P_{\rm NPT}(E,V;T,P)$ of potential energy and volume in the isobaric-isothermal ensemble at the desired temperature T and pressure P is given by

$$P_{\rm NPT}(E,V;T,P) = \frac{N_{\rm mbt}(E,V)\ W_{\rm mbt}(E,V)^{-1}\ \mathrm{e}^{-\beta(E+PV)}}{\displaystyle\sum_{E,V} N_{\rm mbt}(E,V)\ W_{\rm mbt}(E,V)^{-1}\ \mathrm{e}^{-\beta(E+PV)}}, \tag{4.72}$$

where $N_{\rm mbt}(E,V)$ is the histogram of the probability distribution $P_{\rm mbt}(E,V)$ of potential energy and volume that was obtained by the multibaric-multithermal production run. The expectation value of a physical quantity A at T and P is then obtained from

$$\langle A \rangle_{T,P} = \sum_{E,V} A(E,V)\ P_{\rm NPT}(E,V;T,P). \tag{4.73}$$

4.3 Examples of Simulation Results

We now present several examples of the simulation results by the generalized-ensemble algorithms described in the previous section.

The first example is a vWREM simulation of a small peptide [32]. In order to demonstrate the effectiveness of vWREM, in which we exchange pairs of the van der Waals radius parameter values, we applied the vWREM MD algorithm, which we refer to as the *vWREMD*, to the system of an alanine dipeptide in explicit water solvent and compared the results with those obtained by the replica-exchange MD (REMD) simulation [8] and conventional canonical MD simulations. The N-terminus and the C-terminus were blocked by the acetyl group and the N-methyl

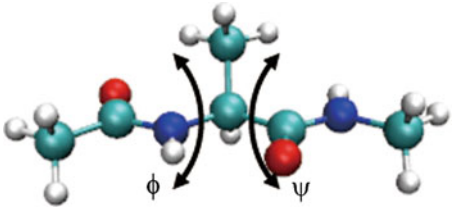

Fig. 4.1 The common initial conformation of the alanine dipeptide for the vWREMD, REMD, and canonical MD simulations. Reprinted from Ref. [32] with kind permission of © The American Institute of Physics (2010)

group, respectively. The number of water molecules was 67. The force field that we adopted was the AMBER parm96 parameter set [60], and the model for the water molecules was the TIP3P rigid-body model [61]. The vWREMD, REMD, and canonical MD simulations were carried out with the symplectic integrator with rigid-body water molecules, in which the temperature was controlled by the Nosé-Poincaré thermostat [44,45,62–65]. The system was put in a cubic unit cell with the side length of 13.4 Å, and we imposed the periodic boundary conditions.

In the vWREMD simulation, we needed only four replicas ($M = 4$). That is, we employed four different parameter values λ_m ($m = 1, \ldots, 4$), and their values were $\lambda_1 = 0.85$, $\lambda_2 = 0.9$, $\lambda_3 = 0.95$, and $\lambda_4 = 1.0$. The original potential energy corresponds to the scale factor $\lambda_4 = 1.0$. The temperature of the system T_0 was set to be 300 K for all the replicas in the vWREMD simulation. We also employed four replicas for the REMD simulation to compare the sampling efficiency with those of the vWREMD simulation, and the four different temperatures were 300 K, 315 K, 335 K, and 360 K, and these temperatures were determined so that exchanges between pairs of replicas were accepted sufficiently. Moreover, we carried out four canonical MD simulations at 300 K, and the difference among these four simulations was initial velocities. We employed the original parameter value $\lambda = 1.0$ for the REMD and canonical MD simulations. The initial conformations were the same for all the simulations, and the initial backbone dihedral angles ϕ and ψ of the alanine dipeptide were set $(\phi, \psi) = (180°, 180°)$, as shown in Fig. 4.1. The total time of the MD simulations was 2.5 ns per replica for the vWREMD and REMD simulations and 2.5 ns for each canonical simulation, including equilibration for 0.1 ns.

Figure 4.2 shows the time series of the backbone dihedral angles ϕ for the vWREMD, REMD, and the conventional canonical MD simulations. From the figure, we see that the samplings in the ϕ space in the vWREMD simulation were the most effective, then those in the REMD simulation, and the least effective in the conventional MD simulation.

The second example is a multioverlap MD simulation of the system of a pentapeptide, Met-enkephalin, in vacuum [34]. The amino-acid sequence is Tyr-Gly-Gly-Phe-Met. The N-terminus and the C-terminus were blocked with the acetyl group and the N-methyl group, respectively. The force field that we adopted is the CHARMM param 22 parameter set [66]. Our multioverlap MD simulations

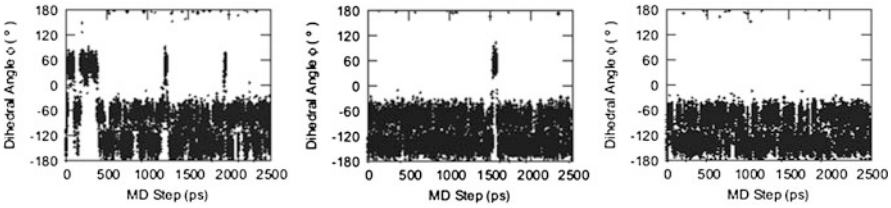

Fig. 4.2 Time series of the dihedral angle ϕ during the vWREMD simulation (*left*), REMD simulation (*center*), and canonical MD simulation (*right*)

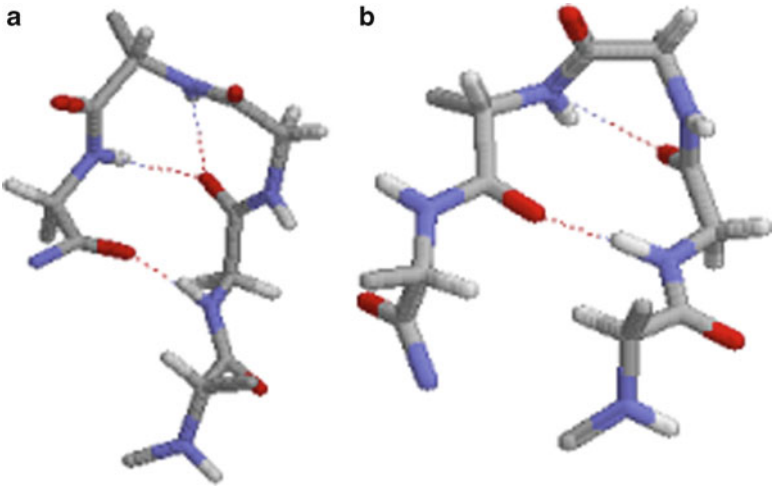

Fig. 4.3 (**a**) Reference conformation 1 and (**b**) reference conformation 2. The side chains are suppressed, and only backbone structures are shown. The *dotted lines* denote the hydrogen bonds. The N-terminus and the C-terminus are on the *right-hand side* and on the *left-hand side*, respectively

were performed by implementing the method in the CHARMM macromolecular mechanics program [67].

We considered two energy-local-minimum states of Met-enkephalin as reference conformations. In Fig. 4.3, we show these two reference conformations. We then set $L = 2$ in Eq. 4.57, and the dimensionless free energy is expressed as $f(d_1, d_2; T_0)$. The multioverlap MD simulation was carried out at $T_0 = 300\,\mathrm{K}$ with a time step of 0.5 fs.

Figure 4.4 shows the time series of the dihedral-angle distances with respect to each of the two reference conformations. While Fig. 4.4a, b shows the results of the conventional canonical MD simulation at $T_0 = 300\,\mathrm{K}$, Fig. 4.4c, d shows the results of the multioverlap MD simulation at the same temperature. When $d_1 = 0$, the values of dihedral angles of backbone completely coincide with those of reference conformation 1 and $d_2 = 0.122$. Conversely, when $d_2 = 0$, $d_1 = 0.122$. When d_1 (d_2) is near zero, the conformation is similar to reference conformation 1 (2). Therefore, Fig. 4.4 implies that the multioverlap MD simulation performed a

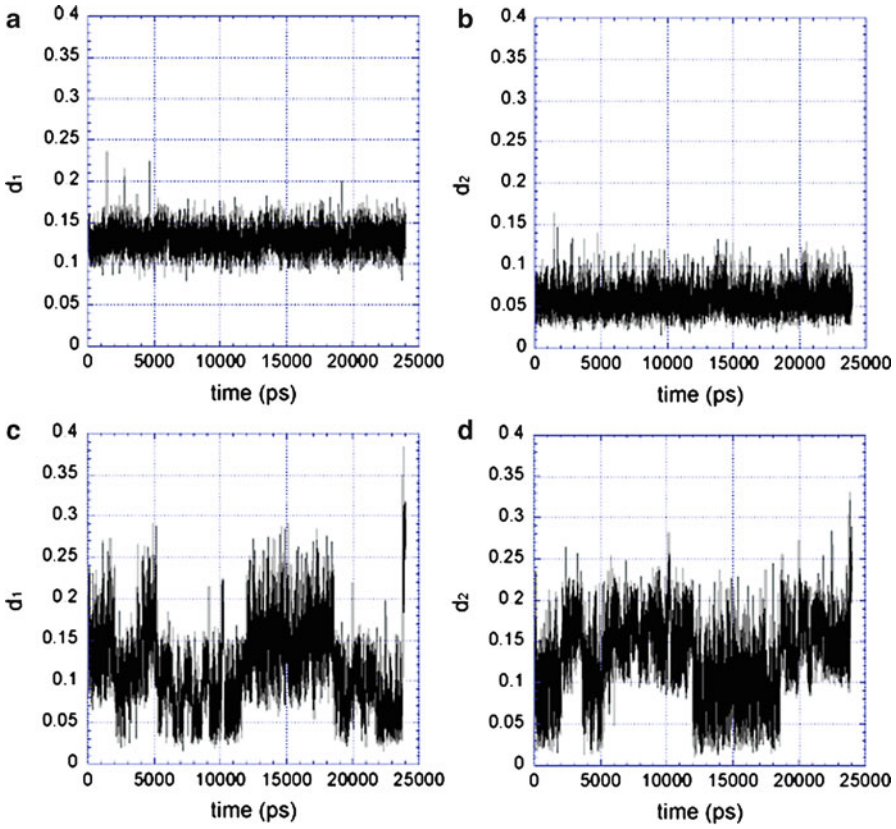

Fig. 4.4 The time series of the dihedral-angle distances d_1 and d_2. (**a**) and (**b**) are from the conventional canonical MD simulation, and (**c**) and (**d**) are the results from the multioverlap MD simulation at $T_0 = 300\,\mathrm{K}$

random walk in the dihedral-angle distance space between reference conformation 1 and reference conformation 2, whereas the usual canonical MD simulation got trapped in a local-minimum state near conformation 2.

The free energy $F(d_1, d_2; T)$ (or the potential of mean force) at temperature T is defined by

$$F(d_1, d_2; T) = -k_{\mathrm{B}} T \ln P_{\mathrm{B}}(d_1, d_2; T), \tag{4.74}$$

where $P_{\mathrm{B}}(d_1, d_2; T)$ is the reweighted canonical probability distribution of d_1 and d_2 at T and given by (see Eq. 4.60)

$$P_{\mathrm{B}}(d_1, d_2; T) = \frac{\sum_E N_{\mathrm{muov}}(d_1, d_2, E) e^{-(\beta-\beta_0)E - f(d_1,d_2;T_0)}}{\sum_{d_1,d_2,E} N_{\mathrm{muov}}(d_1, d_2, E) e^{-(\beta-\beta_0)E - f(d_1,d_2;T_0)}}. \tag{4.75}$$

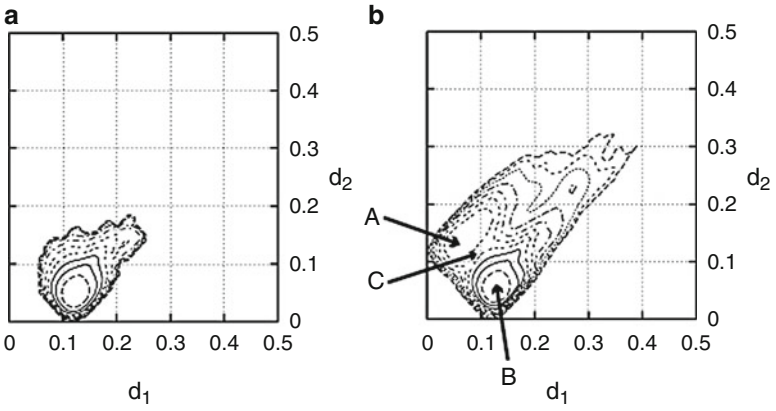

Fig. 4.5 The free-energy landscape obtained from (**a**) the conventional canonical MD simulation and (**b**) the multioverlap MD simulation at $T_0 = 300$ K. *Contour lines* are drawn every 1 kcal/mol. The labels A and B locate the local-minimum states. The label C stands for the *saddle point*, or the transition state, between these two local-minimum states

In Fig. 4.5, we illustrate the free-energy landscapes with respect to the dihedral-angle distances that were calculated from the results of the conventional canonical MD simulation and those of the multioverlap MD simulation. While in Fig. 4.5a only one local-minimum state exists near reference conformation 2, in Fig. 4.5b, we find a local-minimum state A and a local-minimum state B near reference conformation 1 and reference conformation 2, respectively. This result again implies that the canonical MD simulation got trapped in the latter local-minimum state. The local-minimum state B near reference conformation 2 corresponds to the global-minimum state at 300 K. The local-minimum state A near reference conformation 1 is another local-minimum state at 300 K. The free-energy difference between the global-minimum state (B) and the local-minimum state (A) is about 3 kcal/mol.

The saddle point C in Fig. 4.5b corresponds to the transition state between the global-minimum state (B) and the local-minimum state (A). The free-energy difference between B and C is about 5 kcal/mol and that between A and C is 2 kcal/mol. Because $k_B T \approx 0.6$ kcal/mol at $T = 300$ K, these barrier heights are rather high. This is why the conventional canonical MD simulation got trapped in the vicinity of the global-minimum state B.

Our next simulation is the multicanonical-multioverlap MD simulation of Alzheimer's amyloid-β (Aβ) peptide fragment [37]. The amino-acid sequence was Ace-GAIIGLMVGGVVIA-Nme. In multicanonical-multioverlap simulations, we must have a reference conformation. We adopted the conformation that was obtained from the corresponding part in the conformation whose PDB ID code is 2BEG. Here, we took into account only the backbone dihedral angles ϕ (the rotation angles around the N–C_α bonds) and ψ (the rotation angles around the C_α–C bonds) of the residues 30–41 of Aβ(29–42) as the reference dihedral angles in our simulations. The force field that we adopted is the CHARMM 22 parameter

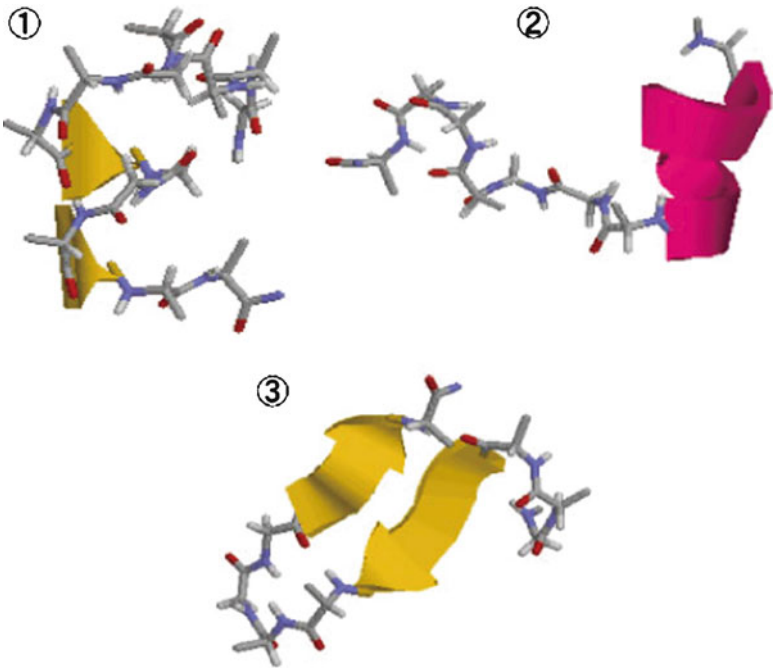

Fig. 4.6 Typical conformations of the Aβ(29–42) molecule at 300 K when two molecules are spatially separated. Reprinted from Ref. [37] with kind permission of © The American Chemical Association (2008)

set [66]. We employed the GB/SA model [68–70] as an implicit solvent model. We also introduced the harmonic constraint $k(r - r_0)^2/2$ when the distance between the center of mass of two Aβ(29–42) molecules exceeded 20 Å in order to avoid the states in which two molecules are too much spatially separated. Here, r is the distance between the center of mass of two molecules, and k is a force constant whose value is 200 kcal/(mol Å^2), and the value of r_0 is set 20 Å.

In Fig. 4.6, we show conformations of Aβ(29–42) monomer in the case when the distance between the center of mass of two peptides is more than 15 Å at 300 K. We identified three major metastable states. These states correspond to low concentrations of Aβ(29–42) peptides or to their monomeric states. Conformation 1 in Fig. 4.6 is a β-helix-like structure, conformation 2 is an α-helix (or sometimes π-helix) structure, and conformation 3 is an intramolecular antiparallel β-sheet (β-hairpin) structure. When the Aβ(29–42) peptide is in a monomeric state, therefore, it seems that the conformations of Aβ(29–42) peptides have the same structure as those in Fig. 4.6.

We show the free-energy landscape of the dimer system at 300 K in Fig. 4.7a. The free-energy landscape was obtained from the results of the multicanonical-multioverlap MD simulation by the reweighting techniques. The abscissa is the number of backbone C_α intermolecular contacts, and we regard a pair of C_α atoms

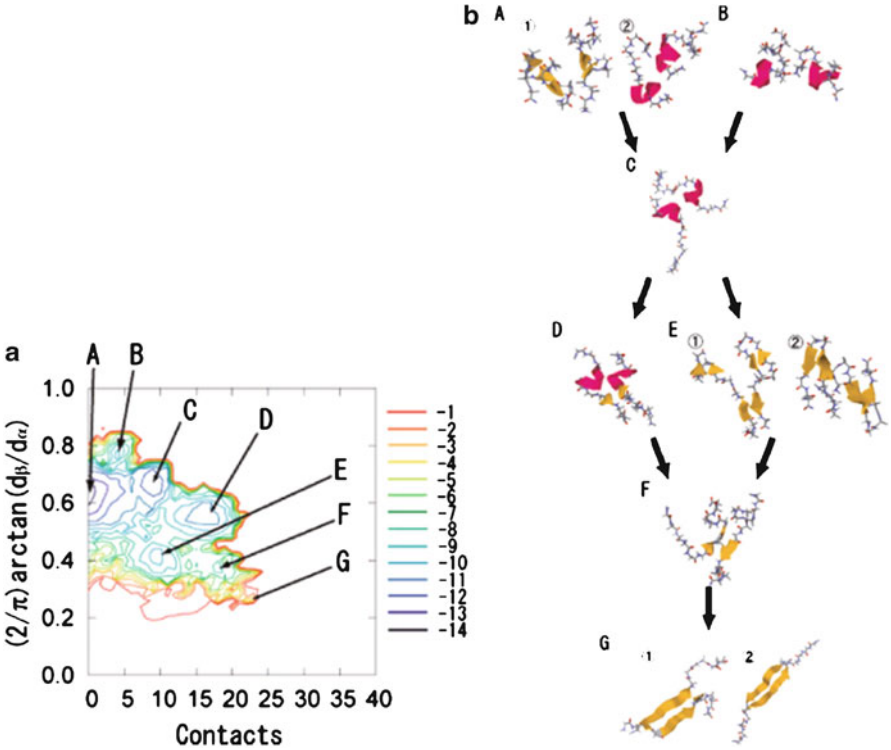

Fig. 4.7 (**a**) Free-energy landscape of an Aβ(29–42) dimer system at 300 K. The ordinate is an indicator of structure for helix and strand. When the value of the ordinate is close to 1, conformations of Aβ(29–42) become helical. Conversely, if the value is close to 0, the conformations have extended forms. The abscissa is the number of backbone C_α intermolecular contacts. *Contour lines* are drawn every 1 kcal/mol. (**b**) Typical structures in the corresponding local-minimum states in (**a**). The *arrows* indicate possible pathways of the early stages of amyloidogenesis. Reprinted from Ref. [37] with kind permission of © The American Chemical Association (2008)

as being in contact if the distance between the two atoms is within 6.5 Å. d_α and d_β in the label of the ordinate are dihedral-angle distances, which we introduced to set the reaction coordinates of the free-energy data analysis. When the value of d_α (d_β) is close to 0, the structures of Aβ(29–42) molecules are helical (extended strand). From the free-energy landscape in Fig. 4.7a, we identified seven local-minimum states. In Fig. 4.7b, we show typical conformations of the Aβ(29–42) in each local-minimum state.

From Figs. 4.6 and 4.7, we deduce the dimerization (oligomerization) process, which corresponds to a seeding process in amyloidogenesis, for Aβ(29–42) peptides as follows: Stage 1: When the Aβ(29–42) peptides are in the monomeric state, the peptides are mainly in one of the three conformational states in Fig. 4.6. Stage 2: Aβ(29–42) peptides come close to each other and create dimers (or oligomers) as a

result of hydrophobic effects. If the structures are intramolecular antiparallel β-sheet structures before dimerization, such as conformation 3 in Fig. 4.6, the conformation after dimerization will correspond to conformation 2 in the local-minimum state E in Fig. 4.7b. If the structures are like conformation 1 or 2 in Fig. 4.6, on the other hand, the Aβ(29–42) dimer will have structures like those of the conformations in A or B in Fig. 4.7b. Stage 3: If the conformations in stage 2 are in states A or B in Fig. 4.7b, then the peptides have helical conformations with extended parts like those in C. If the conformations in stage 2 are already in E in Fig. 4.7b, on the other hand, this corresponds to Stage 4 below. Stage 4: The extended parts will create intermolecular β-ladders such as those in D or E. Stage 5: The intramolecular secondary structures are broken, and the peptides will have a fully extended form such as those in F. Stage 6: The Aβ(29–42) dimer has intermolecular parallel or antiparallel β-sheet structure like those in G. These pathways are summarized in Fig. 4.7b (see the arrows). In the early process of amyloidogenesis, these intermolecular parallel or antiparallel β-sheet structure can be a seed of amyloid fibrils.

We now present the results of a multibaric-multithermal MD simulation [42]. We considered a Lennard-Jones 12–6 potential system. The length and the energy are scaled in units of the Lennard-Jones diameter σ and the depth of the potential ϵ, respectively. We use an asterisk $(*)$ for quantities reduced by σ and ϵ.

We used 500 particles ($N = 500$) in a cubic unit cell with periodic boundary conditions. We started the multibaric-multithermal weight factor determination from a regular isobaric-isothermal simulation at $T_0^* = 2.0$ and $P_0^* = 3.0$ (the multibaric-multithermal production run was also performed at this set of temperature and pressure values). These temperature and pressure values are respectively higher than the critical temperature T_c^* and the critical pressure P_c^* [71,72]. Recent reliable data are $T_c^* = 1.3207(4)$ and $P_c^* = 0.1288(5)$ [72]. The cutoff radius r_c^* was taken to be $r_c^* = 4.0$. A cutoff correction was added for the pressure and the potential energy.

In order to carry out the multibaric-multithermal MD simulation in Eqs. 4.64–4.69 with the replacement of $\mathcal{H}$ by $\mathcal{H}_{\rm mbt}$, we employed the Nosé-Poincaré formalism [44,45,62–65]. This gives the same equations of motion as the Nosé thermostat and provides a symplectic integrator. Therefore, it has an advantage that the secular deviation of the Hamiltonian is suppressed. We have recently shown that this integrator is also very effective for rigid-body molecules [64]. We performed a long production run of 10^6 MD steps.

In Fig. 4.8a, we show the probability distribution $P_{\rm NPT}(E^*/N, V^*/N)$ from the isobaric-isothermal simulation that was carried out first. It is a bell-shaped distribution. As the iteration of the multibaric-multithermal weight factor determination proceeds, $P_{\rm mbt}(E^*/N, V^*/N)$ will become flat and broad gradually. Figure 4.8b depicts the probability distribution $P_{\rm mbt}(E^*/N, V^*/N)$ from the multibaric-multithermal simulation that was finally performed. It shows a flat distribution, and the multibaric-multithermal MD simulation indeed sampled the conformational space in wider ranges of E^*/N and V^*/N than the conventional isobaric-isothermal MD simulation.

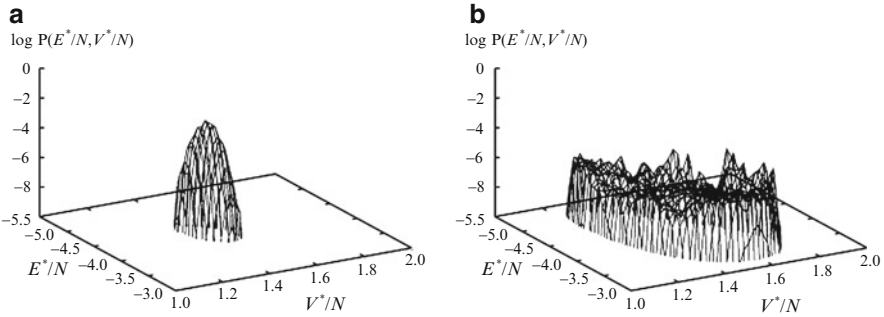

Fig. 4.8 (**a**) The probability distribution $P_{\mathrm{NPT}}(E^*/N, V^*/N)$ in the isobaric-isothermal MD simulation at $(T_0^*, P_0^*) = (2.0, 3.0)$ and (**b**) the probability distribution $P_{\mathrm{mbt}}(E^*/N, V^*/N)$ in the multibaric-multithermal MD simulation. Reprinted from Ref. [42] with kind permission of © Elsevier (2004)

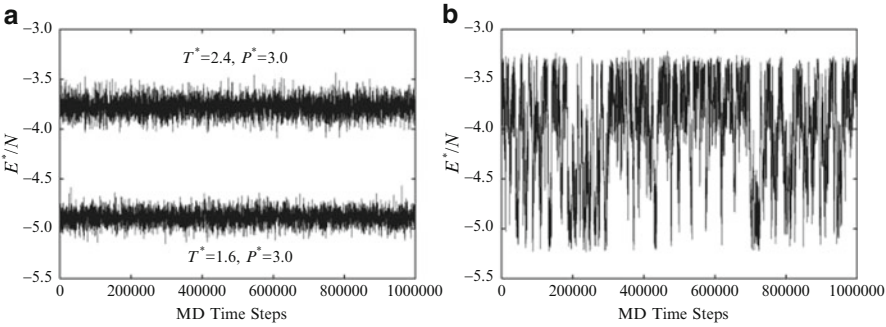

Fig. 4.9 The time series of E^*/N from (**a**) the conventional isobaric-isothermal MD simulations at $(T_0^*, P_0^*) = (2.4, 3.0)$ and at $(T_0^*, P_0^*) = (1.6, 3.0)$ and (**b**) the multibaric-multithermal MD simulation. Reprinted from Ref. [42] with kind permission of © Elsevier (2004)

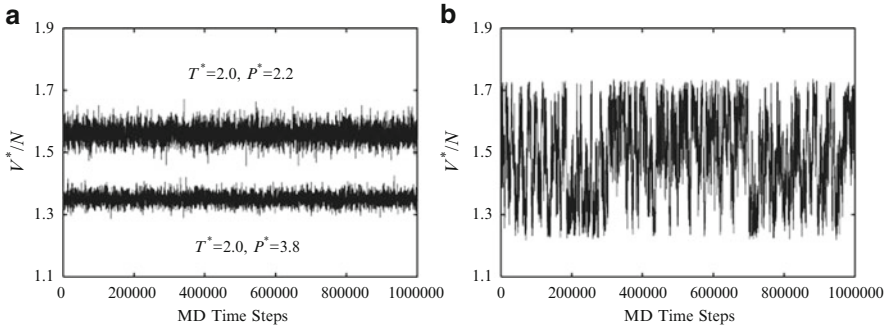

Fig. 4.10 The time series of V^*/N from (**a**) the conventional isobaric-isothermal MD simulations at $(T_0^*, P_0^*) = (2.0, 2.2)$ and at $(T_0^*, P_0^*) = (2.0, 3.8)$ and (**b**) the multibaric-multithermal MD simulation. Reprinted from Ref. [42] with kind permission of © Elsevier (2004)

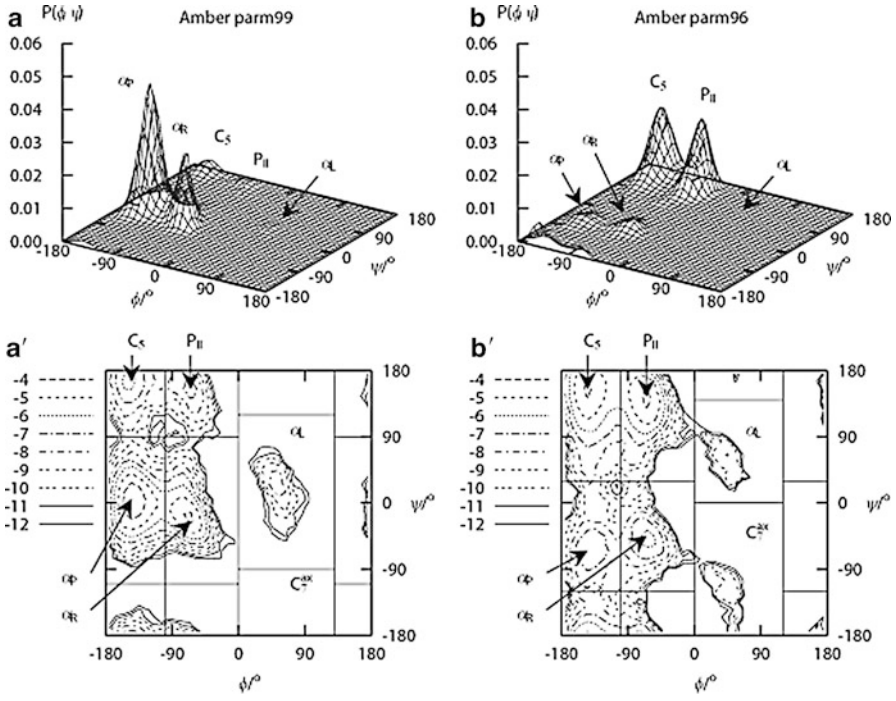

Fig. 4.11 The probability distributions $\mathcal{P}(\phi, \psi)$ of the backbone dihedral angles ϕ and ψ at $T = 298$ K and $P = 0.1$ MPa, which were obtained by the reweighting techniques from the results of the multibaric-multithermal MD simulation (**a**) in the AMBER parm99 force field and (**b**) in the AMBER parm96 force field. (**a′**) and (**b′**) are the contour map of (**a**) and that of (**b**), respectively. Reprinted from Ref. [45] with kind permission of © The American Chemical Society (2008)

The time series of E^*/N from two conventional isobaric-isothermal MD simulations at $(T_0^*, P_0^*) = (1.6, 3.0)$ and $(2.4, 3.0)$ are given in Fig. 4.9a. The potential energy fluctuates in narrow ranges of $E^*/N = -4.0 \sim -3.5$ at the higher temperature of $T_0^* = 2.4$ and in the ranges of $E^*/N = -5.1 \sim -4.7$ and at the lower temperature of $T_0^* = 1.6$. On the other hand, Fig. 4.9b shows that the multibaric-multithermal MD simulation realizes a random walk in the potential energy space and covers a wide energy range.

A similar situation is observed in V^*/N. In Fig. 4.10a the time series of two conventional isobaric-isothermal MD simulations at $(T_0^*, P_0^*) = (2.0, 2.2)$ and $(2.0, 3.8)$, is shown. The volume fluctuations are only in the range of $V^*/N = 1.3 \sim 1.4$ and $V^*/N = 1.5 \sim 1.6$ at $P_0^* = 3.8$ and at $P_0^* = 2.2$, respectively. On the other hand, the multibaric-multithermal MD simulation performs a random walk that covers even a wider volume range, as shown in Fig. 4.10b.

We applied the MUBATH MD algorithm to a system consisting of one alanine dipeptide molecule and 63 water molecules. We used enough water molecules so that the alanine dipeptide molecule was always held perfectly within the

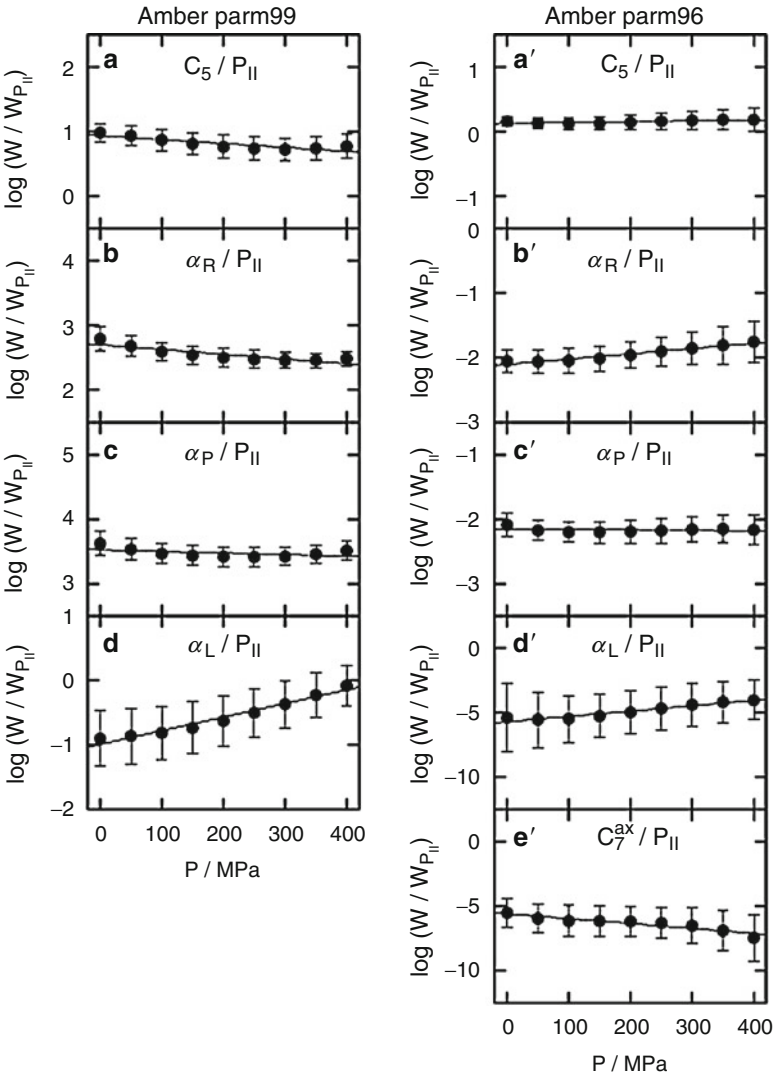

Fig. 4.12 The population ratios $W/W_{\mathrm{P_{II}}}$ against the $\mathrm{P_{II}}$ state as functions of pressure P at constant temperature of $T = 298$ K, which was obtained by the reweighting techniques from the results of the multibaric-multithermal MD simulation. Reprinted from Ref. [45] with kind permission of © The American Chemical Society (2008)

simulation box. We used both AMBER parm99 [73] and AMBER parm96 [60] force fields for the alanine dipeptide molecule and the TIP3P [61] rigid-body model for the water molecules. We employed a cubic unit cell with periodic boundary conditions. The electrostatic potential was calculated by the Ewald method. We calculated the van der Waals interaction, which is given by the Lennard-Jones 12–6

Table 4.1 Differences ΔV / ($\mathrm{cm^3 mol^{-1}}$) in partial molar volume of the $C_5, \alpha_R, \alpha_P, \alpha_L$, and C_7^{ax} states from that of the P_{II} state calculated by the MUBATH MD simulations. Raman experimental data are taken from Ref. [75]

State	AMBER parm99	AMBER parm96	Raman
C_5	1.5 ± 0.9	-0.3 ± 0.9	0.1 ± 0.3
α_R	1.8 ± 0.8	-2.0 ± 2.0	1.1 ± 0.2
α_P	0.6 ± 0.8	0.1 ± 1.2	–
α_L	-5.2 ± 1.0	-10.2 ± 5.6	–
C_7^{ax}	–	9.4 ± 4.1	–

term, for all pairs of the atoms within the minimum image convention instead of introducing the spherical potential cutoff. The time step was taken to be $\Delta t = 0.5$ fs.

Figure 4.11 shows $\mathcal{P}(\phi, \psi)$ obtained from the MUBATH MD simulations by the reweighting techniques at $T = 298$ K and $P = 0.1$ MPa. In the case of longer peptides or proteins, the α_R state corresponds to an α-helix structure, and the P_{II} and C_5 states correspond to a β-strand structure. It is known that, in general, the AMBER parm99 force field tends to form an α-helix structure, and the AMBER parm96 force field tends to form a β-sheet structure [74]. The distributions $\mathcal{P}(\phi, \psi)$ in Fig. 4.11 are consistent with this feature.

Figure 4.12 shows the population ratio of each state and the P_{II} state as a function of P at the constant temperature of $T = 298$ K. A pressure increase at constant temperature generally causes a decrease in the volume. The decreases in the population ratio of some state and the P_{II} state mean that the volume of that state is larger than that of the P_{II} state. The difference in partial molar volume ΔV of the C_5 state from that of the P_{II} state, for example, is calculated from the derivative of $\log(W_{C_5}/W_{P_{II}})$ with respect to P by

$$\Delta V = -RT \left[\frac{\partial \log(W_{C_5}/W_{P_{II}})}{\partial P} \right]_T . \tag{4.76}$$

The difference between the partial molar volume of the other states and that of the P_{II} state was also obtained in the same way. The values of ΔV are shown in Table 4.1. Note that all the experimental data lie in between the corresponding simulation results with the two force fields.

4.4 Conclusions

In this chapter, we described two powerful generalized-ensemble algorithms, namely, replica-exchange method (REM) and multicanonical algorithm (MUCA), which are effective for molecular simulations. We also introduced multidimensional/multivariable extensions of the two methods, namely, MREM, vWREM, MUOV, MUCA-MUOV, and MUBATH. These generalized-ensemble algorithms are particularly useful for biomolecular simulations.

Acknowledgements Some of the results were obtained by the computations on the supercomputers at the Institute for Molecular Science, Okazaki. This work was supported, in part, by Grants-in-Aid for Scientific Research on Innovative Areas ("Fluctuations and Biological Functions"), for the Next-Generation Super Computing Project, Nanoscience Program and Computational Materials Science Initiative from the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan.

References

1. Hansmann UHE, Okamoto Y (1999) New Monte Carlo algorithms for protein folding. Curr Opin Struct Biol 9:177–183
2. Mitsutake A, Sugita Y, Okamoto Y (2001) Generalized-ensemble algorithms for molecular simulations of biopolymers. Biopolymers 60:96–123
3. Sugita Y, Okamoto Y (2002) Free-energy calculations in protein folding by generalized-ensemble algorithms. In: Schlick T, Gan HH (eds) Lecture notes in computational science and engineering, computational methods for macromolecules: challenges and applications. Springer, Berlin, pp 304–332. e-print: arXiv:cond-mat/0102296
4. Itoh SG, Okumura H, Okamoto Y (2007) Generalized-ensemble algorithms for molecular dynamics simulations. Mol Simul 33:47–56
5. Okamoto Y (2009) Generalized-ensemble algorithms for studying protein folding. In: Kuwajima K, Goto Y, Hirata F, Kataoka M, Terazima M (eds) Water and biomolecules. Springer, Berlin, pp 61–95
6. Hansmann UHE, Okamoto Y (1993) Prediction of peptide conformation by multicanonical algorithm – new approach to the multiple-minima problem. J Comput Chem 14:1333–1338
7. Hukushima K, Nemoto K (1996) Exchange Monte Carlo method and application to spin glass simulations. J Phys Soc Jpn 65:1604–1608
8. Sugita Y, Okamoto Y (1999) Replica-exchange molecular dynamics method for protein folding. Chem Phys Lett 314:141–151
9. Sugita Y, Kitao A, Okamoto Y (2000) Multidimensional replica-exchange method for free-energy calculations. J Chem Phys 113:6042–6051
10. Fukunishi F, Watanabe O, Takada S (2002) On the Hamiltonian replica exchange method for efficient sampling of biomolecular systems: application to protein structure prediction. J Chem Phys 116:9058–9067
11. Liu P, Kim B, Friesner RA, Bern BJ (2005) Replica exchange with solute tempering: a method for sampling biological systems in explicit water. Proc Natl Acad Sci USA 102:13749–13754
12. Affentranger R, Tavernelli I, Di Iorio EE (2006) A novel Hamiltonian replica exchange MD protocol to enhance protein conformational space sampling. J Chem Theory Comput 2:217–228
13. Lou H, Cukier RI (2006) Molecular dynamics of apo-adenylate kinase: a distance replica exchange method for the free energy of conformational fluctuations. J Phys Chem B 110:24121–24137
14. Kannan S, Zacharias M (2007) Enhanced sampling of peptide and protein conformations using replica exchange simulations with a peptide backbone biasing-potential. Proteins 66:697–706
15. Mu Y (2009) Dissociation aided and side chain sampling enhanced Hamiltonian replica exchange. J Chem Phys 130:164107
16. Berg BA, Neuhaus T (19991) Multicanonical algorithms for 1st order phase transitions. Phys Lett B 267:249–253
17. Berg BA, Neuhaus T (1992) Multicanonical ensemble: a new approach to simulate first-order phase transitions. Phys Rev Lett 68:9–12
18. Berg BA (2004) Introduction to Monte Carlo simulations and their statistical analysis. World Scientific, Singapore

19. Hansmann UHE, Okamoto Y, Eisenmenger F (1996) Molecular dynamics, Langevin and hybrid Monte Carlo simulations in a multicanonical ensemble. Chem Phys Lett 259:321–330
20. Nakajima N, Nakamura H, Kidera A (1997) Multicanonical ensemble generated by molecular dynamics simulation for enhanced conformational sampling of peptides. J Phys Chem B 101:817–824
21. Berg BA, Hansmann UHE, Neuhaus T (1993) Simulation of an ensemble with varying magnetic field: a numerical determination of the order-order interface tension in the D = 2 Ising model. Phys. Rev. B 47:497–500
22. Janke W, Kappler S (1995) Multibondic cluster algorithm for Monte Carlo simulations of first-order phase transitions. Phys Rev Lett 74:212–215
23. Berg BA, Janke W (1998) Multioverlap simulations of the 3D Edwards-Anderson Ising spin glass. Phys Rev Lett 80:4771–4774
24. Kumar S, Payne P, Vásquez M (1996) Method for free-energy calculations using iterative techniques. J Comput Chem 17:1269–1275
25. Bartels C, Karplus M (1997) Multidimensional adaptive umbrella sampling: applications to main chain and side chain peptide conformations. J Comput Chem 18:1450–1462
26. Higo J, Nakajima N, Shirai H, Kidera A, Nakamura H (1997) Two-component multicanonical Monte Carlo method for effective conformation sampling. J Comput Chem 18:2086–2092
27. Iba Y, Chikenji G, Kikuchi M (1998) Simulation of lattice polymers with multi-self-overlap ensemble. J Phys Soc Jpn 67:3327–3330
28. Bachmann M, Janke W (2003) Multicanonical chain-growth algorithm. Phys Rev Lett 91:208105
29. Mitsutake A, Okamoto Y (2009) From multidimensional replica-exchange method to multidimensional multicanonical algorithm and simulated tempering. Phys Rev E 79:047701
30. Mitsutake A, Okamoto Y (2009) Multidimensional generalized-ensemble algorithms for complex systems. J Chem Phys 130:214105
31. Mitsutake A (2009) Simulated-tempering replica-exchange method for the multidimensional version. J Chem Phys 131:094105
32. Itoh SG, Okumura H, Okamoto Y (2010) Replica-exchange method in van der Waals radius space: overcoming steric restrictions for biomolecules. J Chem Phys 132:134105
33. Berg BA, Noguchi H, Okamoto Y (2003) Multioverlap simulations for transitions between reference configurations. Phys Rev E 68:036126
34. Itoh SG, Okamoto Y (2004) Multi-overlap molecular dynamics methods for biomolecular systems. Chem Phys Lett 400:308–313
35. Itoh SG, Okamoto Y (2006) Theoretical studies of transition states by the multioverlap molecular dynamics methods. J Chem Phys 124:104103
36. Itoh SG, Okamoto Y (2007) Effective sampling in the configurational space of a small peptide by the multicanonical-multioverlap algorithm. Phys Rev E 76:026705
37. Itoh SG, Okamoto Y (2008) Amyloid-β(29–42) dimer formations studied by a multicanonical-multioverlap molecular dynamics simulation. J Phys Chem B 112:2767–2770
38. Itoh SG, Tamura A, Okamoto Y (2010) Helix-hairpin transitions of a designed peptide studied by a generalized-ensemble simulation. J Chem Theor Comput 6:979–983
39. Okumura H, Okamoto Y (2004) Monte Carlo simulations in multibaric-multithermal ensemble. Chem Phys Lett 383:391–396
40. Okumura H, Okamoto Y (2004) Monte Carlo simulations in generalized isobaric-isothermal ensembles. Phys Rev E 70:026702
41. Okumura H, Okamoto Y (2004) Liquid-gas phase transitions studied by multibaric-multithermal Monte Carlo simulations. J Phys Soc Jpn 73:3304-3311
42. Okumura H, Okamoto Y (2004) Molecular dynamics simulations in the multibaric-multithermal ensemble. Chem Phys Lett 391:248–253
43. Okumura H, Okamoto Y (2006) Multibaric-multithermal ensemble molecular dynamics simulations. J Comput Chem 27:379–395
44. Okumura H, Okamoto Y (2007) Multibaric-multithermal molecular dynamics simulation of alanine dipeptide. Bull Chem Soc Jpn 80:1114–1123

45. Okumura H, Okamoto Y (2008) Temperature and pressure dependence of alanine dipeptide studied by multibaric-multithermal molecular dynamics simulations. J Phys Chem B 112:12038–12049
46. Nishikawa T, Ohtsuka H, Sugita Y, Mikami M, Okamoto Y (2000) Replica-exchange Monte Carlo method for Ar fluid. Prog Theor Phys Suppl 138:270–271
47. Okabe T, Kawata M, Okamoto Y, Mikami M (2001) Replica-exchange Monte Carlo method for the isobaric-isothermal ensemble. Chem Phys Lett 335:435–439
48. Paschek D, Garcia AE (2004) Reversible temperature and pressure denaturation of a protein fragment: a replica exchange molecular dynamics simulation study. Phys Rev Lett 93:238105
49. Mori Y, Okamoto Y (2010) Generalized-ensemble algorithms for the isobaric-isothermal ensemble. J Phys Soc Jpn 79:074003
50. Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E (1953) Equation of state calculations by fast computing machines. J Chem Phys 21:1087–1092
51. Nosé S (1984) A molecular dynamics method for simulations in the canonical ensemble. Mol Phys 52:255–268
52. Nosé S (1984) A unified formulation of the constant temperature molecular dynamics methods. J Chem Phys 81:511–519
53. Mori Y, Okamoto Y (2010) Replica-exchange molecular dynamics simulations for various constant temperature algorithms. J Phys Soc Jpn 79:074001
54. Ferrenberg AM, Swendsen RH (1989) Optimized Monte Carlo data analysis. Phys Rev Lett 63:1195–1198
55. Kumar S, Bouzida D, Swendsen RH, Kollman PA, Rosenberg JM (1992) The weighted histogram analysis method for free-energy calculations on biomolecules. 1. The method. J Comput Chem 13:1011–1021
56. Mitsutake A, Sugita Y, Okamoto Y (2003) Replica-exchange multicanonical and multicanonical replica-exchange Monte Carlo simulations of peptides. I. Formulation and benchmark test. J Chem Phys 118:6664–6675
57. Ferrenberg AM, Swendsen RH (1988) New Monte Carlo technique for studying phase transitions. Phys Rev Lett 61:2635–2638
58. Hansmann UHE, Masuya M, Okamoto Y (1997) Characteristic temperatures of folding of a small peptide. Proc Natl Acad Sci USA 94:10652–10656
59. Andersen HG (1980) Molecular dynamics simulations at constant pressure and/or temperature. J Chem Phys 72:2384–2393
60. Kollman PA, Dixon R, Cornell W, Fox T, Chipot C, Pohorille A (1997) The development/application of a 'minimalist' organic/biochemical molecular mechanic force field using a combination of ab initio calculations and experimental data. In: Wilkinson A, Weiner P, van Gunsteren WF (eds) Computer simulation of biomolecular systems, vol 3. Elsevier, Dordrecht, pp 83–96
61. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. J Chem Phys 79:926–935
62. Bond SD, Leimkuhler BJ, Laird BB (1999) The Nosé-Poincaré method for constant temperature molecular dynamics. J Comput Phys 151:114–134
63. Nosé S (2001) An improved symplectic integrator for Nosé-Poincaré thermostat. J Phys Soc Jpn 70:75–77
64. Okumura H, Itoh SG, Okamoto Y (2007) Explicit symplectic integrators of molecular dynamics algorithms for rigid-body molecules in the canonical, isothermal-isobaric, and related ensembles. J Chem Phys 126:084103
65. Okumura H (2008) Partial multicanonical algorithm for molecular dynamics and Monte Carlo simulations. J Chem Phys 129:124116
66. MacKerell AD Jr, Bashford D, Bellott M, Dunbrack RL Jr, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE III, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiórkiewicz-Kuczera J, Yin D, Karplus M (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. J Phys Chem B 102:3586–3616

67. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M (1983) CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. J Comput Chem 4:187–217
68. Still WC, Tempczyk A, Hawley RC, Hendrickson T (1990) Semianalytical treatment of solvation for molecular mechanics and dynamics. J Am Chem Soc 112:6127–6129
69. Dominy BN, Brooks CL III (1999) Development of a generalized-Born model parametrization for proteins and nucleic acids. J Phys Chem B 103:3765–3773
70. Feig M, Brooks CL III (2002) Evaluating CASP4 predictions with physical energy functions. Proteins 49:232–245
71. Okumura H, Yonezawa F (2000) Liquid-vapor coexistence curves of several interatomic model potentials. J Chem Phys 113:9162–9168
72. Okumura H, Yonezawa F (2001) Reliable determination of the liquid-vapor critical point by the NVT plus test particle method. J Phys Soc Jpn 70:1990–1994
73. Wang J, Cieplak P, Kollman PA (2000) How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? J Comput Chem Phys 21:1049–1074
74. Yoda T, Sugita Y, Okamoto Y (2004) Comparisons of force fields for proteins by generalized-ensemble simulations. Chem Phys Lett 386:460–467
75. Takekiyo T, Imai T, Kato M, Taniguchi Y (2004) Temperature and pressure effects on conformational equilibria of alanine dipeptide in aqueous solution. Biopolymers 73:283–290