Sugita, Y. & Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. Chem. Phys. Lett. 314...

Article *in* Chemical Physics Letters · November 1999

DOI: 10.1016/S0009-2614(99)01123-9

CITATIONS
2,148

READS
714

2 authors:

Yuji Sugita
RIKEN
208 PUBLICATIONS **6,081** CITATIONS
SEE PROFILE

Yuko Okamoto
Nagoya University
268 PUBLICATIONS **8,897** CITATIONS
SEE PROFILE

26 November 1999

ELSEVIER

Chemical Physics Letters 314 (1999) 141–151

CHEMICAL PHYSICS LETTERS

www.elsevier.nl/locate/cplett

Replica-exchange molecular dynamics method for protein folding

Yuji Sugita [a,1], Yuko Okamoto [a,b,*]

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

Received 7 July 1999; in final form 9 September 1999

Abstract

We have developed a formulation for molecular dynamics algorithm for the replica-exchange method. The effectiveness of the method for the protein-folding problem is tested with the penta-peptide Met-enkephalin. The method can overcome the multiple-minima problem by exchanging non-interacting replicas of the system at several temperatures. From only one simulation run, one can obtain probability distributions in canonical ensemble for a wide temperature range using multiple-histogram reweighting techniques, which allows the calculation of any thermodynamic quantity as a function of temperature in that range.

1. Introduction

In protein-folding simulations, it is usually difficult to obtain accurate canonical distributions at low temperatures by conventional simulation methods because simulations at low temperatures tend to get trapped in one of a huge number of local minimum-energy states. One way to overcome this multiple-minima problem is to perform a simulation based on non-Boltzmann probability weight factors so that a random walk in energy space may be realized. Random walks allow the simulation to pass any energy barrier and to sample a much wider phase space than by conventional methods. Monitoring the energy in a single simulation run, one can obtain not only the global minimum-energy state but also any thermodynamic quantities as a function of temperature for a wide temperature range. One such well-known method is the *multicanonical algorithm* [1]. This method and its generalizations have already been used in many applications in protein and related systems (see, e.g., Refs. [2–15]). While a simulation in multicanonical ensemble performs a free 1-dimensional (1D) random walk in energy space, that in *simulated tempering* [16,17] performs a free random walk in temperature space. This random walk, in turn, induces a random walk in energy space and allows the simulation to escape from local minima-energy states. Simulated tempering has also been applied to the protein-folding problem [18,19]. These methods which perform random walks in energy space due to non-Boltzmann weight factors are

* Corresponding author. Fax: +81-564-53-4660; e-mail: okamotoy@ims.ac.jp
[1] E-mail: sugita@ims.ac.jp

0009-2614/99/$ - see front matter
PII: S0009-2614(99)01123-9

now given a generic name: *generalized-ensemble algorithm* [19]. (For a review of the generalized-ensemble approach in the protein-folding problem, see, e.g., Ref. [20].)

The generalized-ensemble method is powerful but, in the above two methods, the probability weight factors are not a priori known and have to be determined by iterations of short trial simulations. This process can be non-trivial and very tedious. In the present work, we develop a molecular dynamics (MD) algorithm based on a new generalized-ensemble algorithm, the *replica-exchange method* [21–25]. (The method is also referred to as the *replica Monte Carlo* method [22], *multiple Markov chain method* [24], and *parallel tempering* [25].) In this method, the weight factor is essentially known and there is no complication in its determination. The Monte Carlo (MC) (and MD algorithms in dihedral space) in this generalized ensemble has been applied to an oligopeptide system [26]. Details for the MD algorithm (in Cartesian coordinates) have yet to be worked out, and it is the purpose of the present Letter to do so. The performance of the new algorithm is tested with the system of a penta-peptide, Met-enkephalin, in gas phase.

2. Methods

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 $E(q)$:

$$H(q,p) = K(p) + E(q)\,, \tag{1}$$

where

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

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:

$$W_{\mathrm{B}}(x;T) = \mathrm{e}^{-\beta H(q,p)}, \tag{3}$$

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

$$\langle K(p)\rangle_T = \left\langle \sum_{k=1}^{N} \frac{\boldsymbol{p}_k^2}{2m_k} \right\rangle_T = \tfrac{3}{2}Nk_{\mathrm{B}}T\,. \tag{4}$$

In the original version of the *replica-exchange method* (REM) [21–25], the MC algorithm was used. Here, we describe the method in the context of MD algorithm.

The generalized ensemble for REM consists of M *non-interacting* 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 is 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:

$$\begin{cases} i = i(m) & \equiv f(m)\,, \\ m = m(i) & \equiv f^{-1}(i)\,, \end{cases} \tag{5}$$

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. Here, the superscript and the subscript in $x_m^{[i]}$ label the replica and the temperature, respectively. The state X is specified by the M sets of coordinates $q^{[i]}$ and momenta $p^{[i]}$ of N atoms in replica i at temperature T_m:

$$x_m^{[i]} \equiv \left(q^{[i]}, p^{[i]} \right)_m . \tag{6}$$

Because the replicas are non-interacting, 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):

$$W_{\mathrm{REM}}(X) = \exp\left\{ - \sum_{i=1}^{M} \beta_{m(i)} H\left(q^{[i]}, p^{[i]} \right) \right\} = \exp\left\{ - \sum_{m=1}^{M} \beta_m H\left(q^{[i(m)]}, p^{[i(m)]} \right) \right\}, \tag{7}$$

where $i(m)$ and $m(i)$ are the permutation functions in Eq. (5).

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

$$X = \left(\ldots, x_m^{[i]}, \ldots, x_n^{[j]}, \ldots \right) \rightarrow X' = \left(\ldots, x_m^{[j]'}, \ldots, x_n^{[i]'}, \ldots \right). \tag{8}$$

Here, i, j, m, and n are related by the permutation functions in Eq. (5), and the exchange of replicas introduces a new permutation function f':

$$\begin{cases} i = f(m) \rightarrow j = f'(m), \\ j = f(n) \rightarrow i = f'(n). \end{cases} \tag{9}$$

The exchange of replicas can be written in more detail as

$$\begin{cases} x_m^{[i]} \equiv \left(q^{[i]}, p^{[i]} \right)_m \rightarrow x_m^{[j]'} \equiv \left(q^{[j]}, p^{[j]'} \right)_m, \\ x_n^{[j]} \equiv \left(q^{[j]}, p^{[j]} \right)_n \rightarrow x_n^{[i]'} \equiv \left(q^{[i]}, p^{[i]'} \right)_n, \end{cases} \tag{10}$$

where the definitions for $p^{[i]'}$ and $p^{[j]'}$ will be given below. We remark that this process is equivalent to exchanging a pair of temperatures T_m and T_n for the corresponding replicas i and j as follows:

$$\begin{cases} x_m^{[i]} \equiv \left(q^{[i]}, p^{[i]} \right)_m \rightarrow x_n^{[i]'} \equiv \left(q^{[i]}, p^{[i]'} \right)_n, \\ x_n^{[j]} \equiv \left(q^{[j]}, p^{[j]} \right)_n \rightarrow x_m^{[j]'} \equiv \left(q^{[j]}, p^{[j]'} \right)_m. \end{cases} \tag{11}$$

In the original implementation of the REM [21–25], MC algorithm was used, and only the coordinates q (and the potential energy function $E(q)$) had to be taken into account. Here, in our implementation by MD algorithm, we also have to deal with the momenta p. We propose the following momentum assignment in Eq. (10) (and in Eq. (11)):

$$\begin{cases} p^{[i]'} \equiv \sqrt{\dfrac{T_n}{T_m}}\, p^{[i]}, \\ p^{[j]'} \equiv \sqrt{\dfrac{T_m}{T_n}}\, p^{[j]}, \end{cases} \tag{12}$$

which we believe is the simplest and 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) may be satisfied.

In order for this exchange process to converge towards an equilibrium distribution, it is sufficient to impose the detailed balance condition on the transition probability $w(X \to X')$:

$$W_{\mathrm{REM}}(X)\, w(X \to X') = W_{\mathrm{REM}}(X')\, w(X' \to X)\,. \tag{13}$$

From Eqs. (1), (2), (7), (12) and (13), we have

$$\begin{aligned}\frac{w(X \to X')}{w(X' \to X)} &= \exp\Big\{-\beta_m\big[K(p^{[j]'}) + E(q^{[j]})\big] - \beta_n\big[K(p^{[i]'}) + E(q^{[i]})\big] \\ &\quad + \beta_m\big[K(p^{[i]}) + E(q^{[i]})\big] + \beta_n\big[K(p^{[j]}) + E(q^{[j]})\big]\Big\}, \\ &= \exp\Bigg\{-\beta_m\frac{T_m}{T_n}K(p^{[j]}) - \beta_n\frac{T_n}{T_m}K(p^{[i]}) + \beta_m K(p^{[i]}) + \beta_n K(p^{[j]}) \\ &\quad - \beta_m\big[E(q^{[j]}) - E(q^{[i]})\big] - \beta_n\big[E(q^{[i]}) - E(q^{[j]})\big]\Bigg\}, \\ &= \exp(-\Delta)\,,\end{aligned} \tag{14}$$

where

$$\Delta \equiv [\,\beta_n - \beta_m]\big(E(q^{[i]}) - E(q^{[j]})\big)\,, \tag{15}$$

and i, j, m, and n are related by the permutation functions (in Eq. (5)) before the exchange:

$$i = f(m)\,, \quad j = f(n)\,. \tag{16}$$

This can be satisfied, for example, by the usual Metropolis criterion:

$$w(X \to X') \equiv w\big(x_m^{[i]} \,\big|\, x_n^{[j]}\big) = \begin{cases} 1\,, & \text{for } \Delta \leqslant 0\,, \\ \exp(-\Delta)\,, & \text{for } \Delta > 0\,, \end{cases} \tag{17}$$

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. Note that this is exactly the same criterion that was originally derived for MC algorithm [21–25].

Without loss of generality, we can assume $\beta_1 < \beta_2 < \ldots < \beta_M$. A simulation of the REM [21–25] 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 $x_{m+1}^{[j]}$, are exchanged with the probability $w\big(x_m^{[i]} \,\big|\, x_{m+1}^{[j]}\big)$ in Eq. (17).

In the present approach, we employ the MD algorithm for Step (1). Note that in Step (2) we exchange only pairs of replicas corresponding to neighboring temperatures, because the acceptance ratio of the exchange decreases exponentially with the difference of the two βs (see Eqs. (17) and (19)). Note also that whenever a replica exchange is accepted in Step (2), the permutation functions in Eq. (5) are updated.

The major advantage of REM over other generalized-ensemble methods such as multicanonical algorithm [1] and simulated tempering [16,17] lies in the fact that the weight factor is a priori known (see Eq. (7)), while in the latter algorithms the determination of the weight factors can be very tedius and time-consuming. For the optimal performance of REM, however, one still has to choose an appropriate temperature distribution. There exists an iterative procedure for this [21], and we have modified it further. The details of our procedure for the determination of the optimal temperature distribution will be given elsewhere (Y. Sugita and Y. Okamoto, in preparation).

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 as follows:

$$\langle A \rangle_{T_m} = \frac{1}{N_{\mathrm{sim}}} \sum_{t=1}^{N_{\mathrm{sim}}} \sum_{i=1}^{M} A\big(x_{f^{-1}(i;t)}^{[i]}(t)\big)\, \delta_{f^{-1}(i;t),m}\,, \tag{18}$$

where $N_{\rm sim}$ is the total number of measurements made for each replica, $f^{-1}(i;t)$ is the permutation function in Eq. (5) at tth measurement, and $\delta_{k,l}$ is Kronecker's delta function. When the temperature-exchange view point is taken, this equation can also be written as

$$\langle A\rangle_{T_m} = \frac{1}{N_{\rm sim}} \sum_{t=1}^{N_{\rm sim}} A\left(x_m^{[f(m;t)]}(t)\right). \tag{19}$$

For the expectation value at any intermediate temperature, we use the multiple-histogram reweighting techniques [27,28] (an extension of which is also referred to as WHAM [28]) as follows. Suppose we have made R-independent simulation runs at R different temperatures. Let $N_m(E)$ and n_m be the energy histogram and the total number of samples obtained in the mth run, respectively. (In REM we have $n_m = N_{\rm sim}$.) The expectation value of a physical quantity A at any intermediate temperature $T = 1/k_{\rm B}\beta$ is given by

$$\langle A\rangle_T = \frac{\sum_E A(E)P(E;\beta)}{\sum_E P(E;\beta)}, \tag{20}$$

where

$$P(E;\beta) = \frac{\sum_{m=1}^{R} g_m^{-1} N_m(E)\,\mathrm{e}^{-\beta E}}{\sum_{m=1}^{R} n_m g_m^{-1}\,\mathrm{e}^{f_m-\beta_m E}}, \tag{21}$$

and

$$\mathrm{e}^{-f_m} = \sum_E P(E;\beta_m). \tag{22}$$

Here, $g_m = 1 + 2\tau_m$, and τ_m is the integrated autocorrelation time at temperature T_m. Note that $P(E;\beta)$ and f_m in Eqs. (21) and (22) are solved self-consistently by iteration [27,28].

3. Results and discussion

The effectiveness of the algorithm presented in the previous section was tested for the system of a penta-peptide, Met-enkephalin, in gas phase. This peptide has the amino-acid sequence Tyr–Gly–Gly–Phe–Met. The N and C termini of the peptide was blocked with acetyl and N-methyl groups, respectively. The force-field parameters were taken from the all-atom version of AMBER [29], and the dielectric constant was set equal to 1. The temperature during the MD simulations was controlled by the constraint method [30,31]. The computer code developed in Ref. [32,33], which is based on PRESTO [34], was used. The unit timestep was set to 0.5 fs, and we made an MD simulation of 2×10^6 timesteps (or 1.0 ns) for each replica, starting from an extended conformation. (Before taking the data, we made regular canonical MD simulations for 100 ps at each temperature and then a replica-exchange simulation of 100 ps for thermalization.)

We used the following eight temperatures ($M = 8$ in Eq. (7), etc.): 700, 585, 489, 409, 342, 286, 239, and 200 K, which are distributed exponentially, following the annealing schedule of simulated annealing simulations [35]. As is shown below, this choice already gave an optimal temperature distribution, and thus we did not need the elaborate process for the determination of the temperature distribution for this peptide. The replica exchange was tried every 10 fs, and the data were stored just before the replica exchange for later analyses. We thus have $N_{\rm sim} = 10^5$ in Eq. (18) for each replica.

As is apparent from the description of REM in the previous section, a replica-exchange simulation is particularly suitable for parallel computers. Because one can minimize the amount of information exchanged among nodes, it is best to assign each replica to each node (exchanging pairs of temperature values among nodes is much faster than exchanging coordinates and momenta). This means that we keep track of the permutation function $m(i;t) = f^{-1}(i;t)$ in Eq. (5) during the simulation. After every 10 fs of parallel MD simulations, four pairs of replicas corresponding to neighboring temperatures were exchanged, and the pairing was alternated between the two possible choices.

As for expectation values of physical quantities at various temperatures, we used Eqs. (20)–(22). In the present work we set $R = M = 8$ in these equations, and took into account the runs of all the replicas. For biomolecular systems the integrated autocorrelation times are approximately equal [28]. We thus set $g = \text{const}$ in Eq. (21) for simplicity. We also have $n_m = N_{\text{sim}} = 10^5$ in Eqs. (18) and (21).

We first examine whether the present replica-exchange simulation indeed performed properly. There are three points to check. (a) Were the temperatures optimally distributed? (b) Was the number of replicas (temperatures) sufficient? (c) Was the highest temperature sufficiently high so that no trapping in an energy-local-minimum state occurs? The first two points can be checked by examining the acceptance ratios of replica exchange corresponding to the adjacent pairs of temperatures. For the first point, the optimal temperature distributions imply that all the acceptance ratios are the same, resulting in a free random walk in the replica (temperature) space. For the second point, the number of replicas (temperatures) is sufficient if the acceptance ratios are not too small (say, greater than 0.1). In Table 1 we list the acceptance ratios of replica exchange. The values are indeed uniform (all about 15% of acceptance probability) and large enough ($> 10\%$). Hence, the two of the above criteria ((a) and (b)) for optimal performance are met.

It is not as straightforward to check the third point as in the previous two points. Observed random walks among replicas (and temperatures) will not be sufficient for the third point. This is because we cannot exclude the following possibility. If all the replicas happen to be in the same energy-local-minimum state and the energy barrier to escape from this state is very high (with respect to the highest temperature), then we will still observe random walks among all the replicas (and temperatures) but they will stay in the same local minimum. Supporting evidence for the third point being met can be obtained by comparing the results with those of regular canonical simulations, as discussed below further.

The results in Table 1 imply that one should observe a free random walk in the 'replica (and temperature) space'. In Fig. 1 we show the time series of replica exchange at the lowest temperature ($T = 200$ K). We indeed observe a random walk in the 'replica space'. The complimentary picture to this is the temperature exchange for each replica. The results for one of the replicas (Replica 6) are shown in Fig. 2a. We again observe a random walk in the 'temperature space' between the lowest and highest temperatures. In Fig. 2b the corresponding time series of the total potential energy is shown. We see that a random walk in the potential energy space between low and high energies is realized. Note that there is a strong correlation between the behaviors in Fig. 2a,b, as there should.

Table 1
Acceptance ratios of replica exchange corresponding to pairs of neighboring temperatures

Pair of temperatures	Acceptance ratio
200 ↔ 239 K	0.160
239 ↔ 286 K	0.149
286 ↔ 342 K	0.143
342 ↔ 409 K	0.139
409 ↔ 489 K	0.142
489 ↔ 585 K	0.146
585 ↔ 700 K	0.146

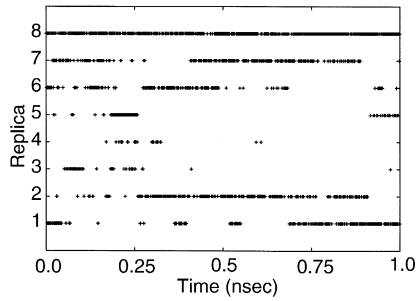

Fig. 1. Time series of replica exchange at $T = 200$ K.

In Fig. 3a the canonical probability distributions obtained at the chosen eight temperatures from the replica-exchange simulation are shown. We see that there are enough overlaps between all neighboring pairs of distributions, indicating that there will be sufficient numbers of replica exchanges between pairs of replicas (see Table 1). In Fig. 3b we compare the above canonical probability distributions at three temperatures ($T = 200$, 239, and 700 K) obtained by the replica-exchange MD simulation with those obtained by the regular canonical MD simulations made separately at the corresponding temperatures. The canonical simulations were performed with the same initial conditions and simulation time (1 ns) as the replica-exchange simulation. We observe the expected behavior that the distributions agree at higher temperatures and that they tend to deviate at lower temperatures. The fact that the distributions obtained by the regular canonical simulations at low temperatures tend to shift to the right with respect to those obtained by the replica-exchange ones is the signal that the canonical simulations got trapped in states of energy local minima at these temperatures. This point will be further elucidated below.

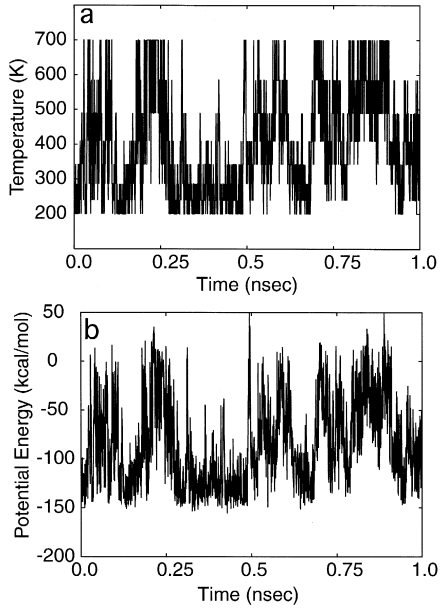

Fig. 2. Time series of temperature exchange (a) and the total potential energy (b) for one of the replicas.

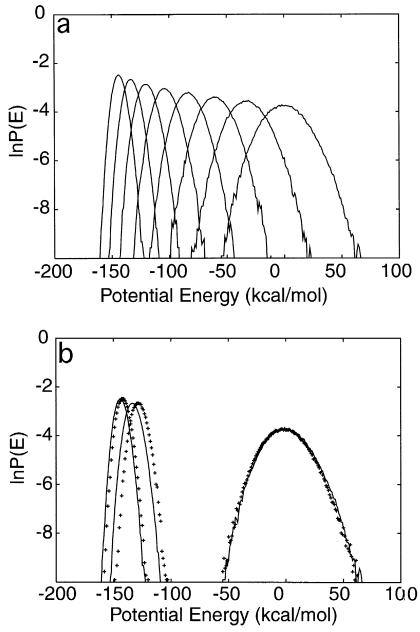

Fig. 3. The canonical probability distributions of the total potential energy of Met-enkephalin obtained from the replica-exchange MD simulation at the eight temperatures (a) and comparison of canonical probability distributions obtained from the replica-exchange MD simulation (solid curves) and the conventional canonical MD simulations (crosses) at three temperatures (b). The distributions in (a) correspond to the following temperatures (from left to right): 200, 239, 286, 342, 409, 489, 585, and 700 K. Pairs of distributions in (b) correspond to the following temperatures (from left to right): 200, 239, and 700 K.

All these results give enough evidence that the above (at least first two) criteria are met and that the present replica-exchange simulation indeed performed properly and effectively.

We further compare the results of the replica-exchange simulation with those of a single canonical MD simulation (of 1 ns) at the corresponding temperatures. In Fig. 4 we compare the distributions of a pair of dihedral angles (ϕ,ψ) of Gly-2 at two extreme temperatures ($T = 200$ and 700 K). While the results at $T = 200$ K from the regular canonical simulation are localized with only one dominant peak, those from the replica-exchange simulation have several peaks (compare Fig. 4a and Fig. 4b). Hence, the replica-exchange run samples much broader configuration space than the conventional canonical run at low temperatures. Note that the sets of peaks observed in the distribution from the replica-exchange simulation include those from the canonical simulation as a subset. However, the latter peak is not the highest one in the former, suggesting that the canonical run did not end up in the ground state but got trapped in one of other energy-local-minimum states. The average potential energy at 200 K of the conformation corresponding to the highest peak in distributions for the canonical run (Fig. 4a) is by about 2 kcal/mol higher than that for the replica-exchange simulation (Fig. 4b) (-141 versus ca. -143 kcal/mol). The results at $T = 700$ K (Fig. 4c,d), on the other hand, are similar, implying that a regular canonical simulation can give accurate thermodynamic quantities at high temperatures.

Incidentally, the fact that the distribution obtained from the replica-exchange simulation has several peaks even at low temperatures gives a partial support that the third criterion above for the optimal performance of replica-exchange simulations is met. Namely, the highest temperature is sufficiently high so that wide conformational space is sampled and the distribution is not forced to converge to a single conformation even at low temperatures, where regular canonical simulations fall in a single energy-local-minimum state. (We should

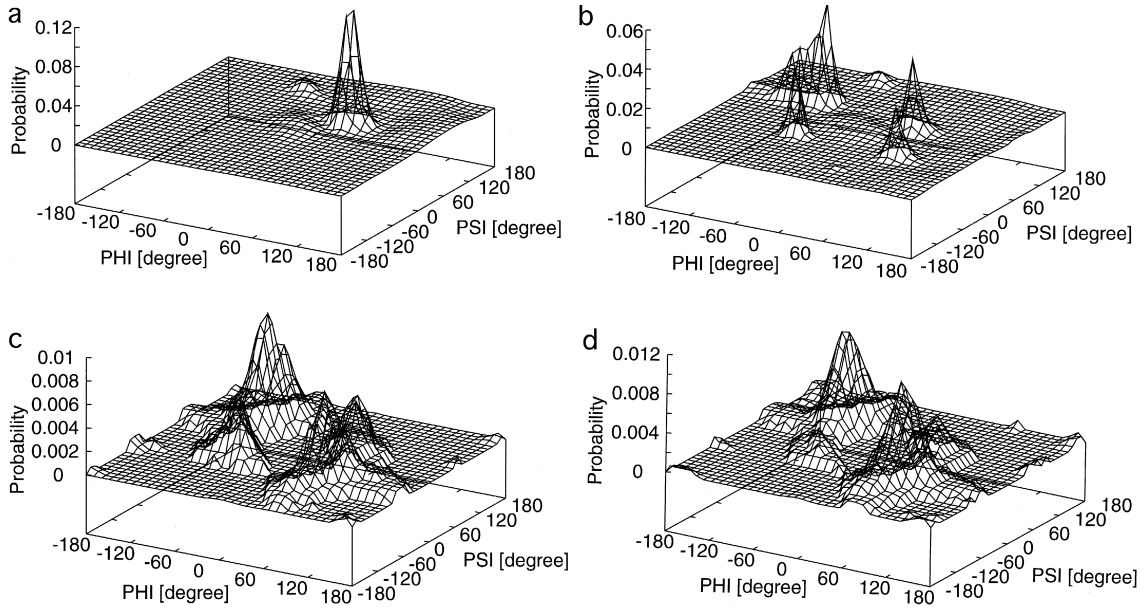

Fig. 4. Distributions of a pair of dihedral angles (ϕ,ψ) of Gly-2 for: (a) $T = 200$ K from a regular canonical MD simulation, (b) $T = 200$ K from the replica-exchange MD simulation, (c) $T = 700$ K from a regular canonical MD simulation, and (d) $T = 700$ K from the replica-exchange MD simulation.

of course see only the ground-state conformation at zero temperature, which implies that the lowest temperature (200 K) we considered is not low enough to see a single distribution peak of the ground state.)

As all simulations in generalized ensembles should, a single replica-exchange simulation can give any thermodynamic quantity as a function of temperature. For this we use the multiple-histogram reweighting techniques [27,28] (see Eqs. (20)–(22)). About 10–100 iterations were necessary for the convergence of Eqs. (21) and (22). In Fig. 5 we show the average total potential energy as a function of temperature. As expected from the results of Figs. 3 and 4, we observe that the canonical simulations at low temperatures got trapped in states of energy local minima, resulting in the discrepancies in average values between the results from the canonical simulations and those from the replica-exchange simulation. Note that the canonical simulations start getting trapped already near 300 K (and below), which is an experimentally relevant temperature. This implies

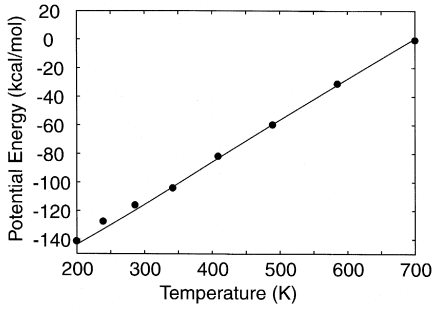

Fig. 5. Average total potential energy as a function of temperature. The solid curve is the result from the replica-exchange MD simulation and the dots are those of regular canonical MD simulations at eight temperatures.

that much longer simulation runs are required to obtain accurate thermodynamic averages at these temperatures by conventional MD methods based on canonical ensemble. As expected, we do have complete agreement at higher temperatures between the results from the canonical simulations and those from the replica-exchange simulation.

4. Conclusions

In this Letter we have presented a formulation for MD algorithm for the replica-exchange method [21–25]. In this method the weight factor is essentially known and there is no complication in its determination, while in other generalized-ensemble algorithms (such as multicanonical algorithm [1]) the determination of the (non-Boltzmann) weight factor can be very tedious and time-consuming. The effectiveness of the method was tested with the penta-peptide Met-enkephalin. It was shown that from a single simulation run one can obtain various thermodynamic quantities as a function of temperature for a wide temperature range. Hence, the new method is particularly useful for studying the protein-folding problem where information of a wide conformational space (from a random-coil state to the native folded state) is required.

Acknowledgements

Our simulations were performed on the Hitachi and other computers at the IMS Computer Center. This work is supported, in part, by a grant from the Research for the Future Program of the Japan Society for the Promotion of Science (JSPS-RFTF98P01101).

References

[1] B.A. Berg, T. Neuhaus, Phys. Lett. B267 (1991) 249.
[2] U.H.E. Hansmann, Y. Okamoto, J. Comp. Chem. 14 (1993) 1333.
[3] M.H. Hao, H.A. Scheraga, J. Phys. Chem. 98 (1994) 4940.
[4] Y. Okamoto, U.H.E. Hansmann, J. Phys. Chem. 99 (1995) 11276.
[5] A. Kolinski, W. Galazka, J. Skolnick, Proteins 26 (1996) 271.
[6] N. Urakami, M. Takasu, J. Phys. Soc. Jpn 65 (1996) 2694.
[7] S. Kumar, P. Payne, M. Vásquez, J. Comp. Chem. 17 (1996) 1269.
[8] N. Nakajima, H. Nakamura, A. Kidera, J. Phys. Chem. 101 (1997) 817.
[9] H. Noguchi, K. Yoshikawa, Chem. Phys. Lett. 278 (1997) 184.
[10] J. Higo, N. Nakajima, H. Shirai, A. Kidera, H. Nakamura, J. Comp. Chem. 18 (1997) 2086.
[11] C. Bartels, M. Karplus, J. Phys. Chem. B 102 (1998) 865.
[12] Y. Iba, G. Chikenji, M. Kikuchi, J. Phys. Soc. Jpn 67 (1998) 3327.
[13] M. Schaefer, C. Bartels, M. Karplus, J. Mol. Biol. 284 (1998) 835.
[14] U.H.E. Hansmann, Y. Okamoto, J. Phys. Chem. B 103 (1999) 1595.
[15] A. Mitsutake, Y. Okamoto, Chem. Phys. Lett. 309 (1999) 95.
[16] A.P. Lyubartsev, A.A. Martinovski, S.V. Shevkunov, P.N. Vorontsov-Velyaminov, J. Chem. Phys. 96 (1992) 1776.
[17] E. Marinari, G. Parisi, Europhys. Lett. 19 (1992) 451.
[18] A. Irbäck, F. Potthast, J. Chem. Phys. 103 (1995) 10298.
[19] U.H.E. Hansmann, Y. Okamoto, J. Comp. Chem. 18 (1997) 920.
[20] U.H.E. Hansmann, Y. Okamoto, Curr. Opin. Struct. Biol. 9 (1999) 177.
[21] K. Hukushima, K. Nemoto, J. Phys. Soc. Jpn 65 (1996) 1604.
[22] R.H. Swendsen, J.-S. Wang, Phys. Rev. Lett. 57 (1986) 2607.
[23] C.J. Geyer, in: E.M. Keramidas (Ed.), Computing Science and Statistics: Proc. 23rd Symp. on the Interface, Interface Foundation, Fairfax Station, 1991, p. 156.

[24] M.C. Tesi, E.J.J. van Rensburg, E. Orlandini, S.G. Whittington, J. Stat. Phys. 82 (1996) 155.
[25] E. Marinari, G. Parisi, J.J. Ruiz-Lorenzo, in: A.P. Young (Ed.), Spin Glasses and Random Fields, World Scientific, Singapore, 1998, p. 59.
[26] U.H.E. Hansmann, Chem. Phys. Lett. 281 (1997) 140.
[27] A.M. Ferrenberg, R.H. Swendsen, Phys. Rev. Lett. 63 (1989) 1195.
[28] S. Kumar, D. Bouzida, R.H. Swendsen, P.A. Kollman, J.M. Rosenberg, J. Comp. Chem. 13 (1992) 1011.
[29] S.J. Weiner, P.A. Kollman, D.T. Nguyen, D.A. Case, J. Comp. Chem. 7 (1986) 230.
[30] W.G. Hoover, A.J.C. Ladd, B. Moran, Phys. Rev. Lett. 48 (1982) 1818.
[31] D.J. Evans, G.P. Morris, Phys. Lett. A98 (1983) 433.
[32] Y. Sugita, A. Kitao, Proteins 30 (1998) 388.
[33] A. Kitao, S. Hayward, N. Go, Proteins 33 (1998) 496.
[34] K. Morikami, T. Nakai, A. Kidera, M. Saito, H. Nakamura, Comput. Chem. 16 (1992) 243.
[35] Y. Okamoto, M. Fukugita, T. Nakazawa, H. Kawai, Protein Eng. 4 (1991) 639.