Generalized-ensemble Monte Carlo method for systems with rough energy landscape

Ulrich H. E. Hansmann* and Yuko Okamoto[†]

Department of Theoretical Studies, Institute for Molecular Science, Okazaki, Aichi 444, Japan

(Received 9 December 1996; revised manuscript received 29 April 1997)

We present a Monte Carlo algorithm which enhances equilibrization of low-temperature simulations and allows sampling of configurations over a large range of energies. The method is based on a non-Boltzmann probability weight factor and is another version of the so-called generalized-ensemble techniques. The effectiveness of the approach is demonstrated for the system of a small peptide, an example of the frustrated system with a rugged energy landscape. [S1063-651X(97)03508-3]

PACS number(s): 02.70.Lq, 05.50.+q, 87.10.+e, 64.60.Fr

The energy landscape of many important physical systems is characterized by a huge number of local minima separated by high energy barriers. In the canonical ensemble with temperature T, the probability to cross an energy barrier of heights ΔE is proportional to $e^{-\Delta E/k_B T}$, where k_B is the Boltzmann constant. Hence, at low temperatures, canonical molecular dynamics and Monte Carlo simulations will get trapped in configurations corresponding to one of these local minima. Only small parts of the entire phase space can be explored, rendering the calculation of physical quantities unreliable.

In principle, one can think of two ways to overcome this difficulty. One way is to look for improved updates of configurations in the numerical simulation. The cluster algorithm [1] is an example of global updates that enhance thermalization and has been very successful in spin systems. However, for most other systems with frustration, no such updates are known. Another way to overcome the supercritical slowing down is to perform a simulation in a so-called *generalized ensemble*, which is based on a non-Boltzmann probability distribution. Multicanonical algorithm [2,3], $1/k$ sampling [4], and simulated tempering [5,6] are prominent examples of such an approach. Common to the three techniques is that a molecular dynamics or Monte Carlo simulation is performed in an artificial ensemble defined in such a way that a uniform (noncanonical) distribution of the chosen physical quantity is obtained. For instance, in the multicanonical algorithm the weight $w_{\mathrm{mu}}(E)$ is chosen so that the distribution of energy is uniform:

$$P(E) \propto n(E) w_{\mathrm{mu}}(E) = \mathrm{const}, \qquad (1)$$

where $n(E)$ is the density of states. A simulation based on this weight factor results in a free random walk in the energy space. Hence, the simulation can escape from any energy barrier, and even regions with small $n(E)$ can be explored in detail. Similarly, $1/k$ sampling yields a uniform distribution in (microcanonical) entropy, and simulated tempering a uniform distribution in temperature. The great advantage of these generalized-ensemble methods lies in the fact that from a single simulation run one can not only locate the energy global minimum but also obtain the canonical distribution for a wide temperature range by the reweighting techniques [7].

Despite their successful applications to systems with first-order phase transitions [2], spin glasses [8], and the protein folding problem [9,10], generalized-ensemble methods are not without problems. Unlike in the canonical ensemble, the probability weights are not *a priori* known. For instance, for the case of multicanonical algorithm, Eq. (1) implies

$$w_{\mathrm{mu}}(E) \propto n^{-1}(E), \qquad (2)$$

and the knowledge of the exact weight would be equivalent to obtaining the density of states $n(E)$, i.e., solving the system. Hence, one needs its estimator for a numerical simulation. The determination of the weight $w_{\mathrm{mu}}(E)$ is usually based on an iterative procedure first described in Ref. [3], and can be nontrivial and tedious. In this paper, we present a generalized-ensemble algorithm in which the determination of the weight is simple and straightforward.

Our aim is to develop a generalized-ensemble algorithm in which the determination of the probability weight factor is simpler. For this, we try to slightly modify the Boltzmann weight, whereas other generalized-ensemble approaches use drastically different weights. The weight should enhance the thermalization of low-temperature simulations and ensure sufficient sampling in the low-energy region. Hence, we are interested in an ensemble where not only the low-energy region can be sampled efficiently but also the high-energy states can be visited with finite probability. The latter feature ensures that energy barriers can be overcome and that the simulation can escape from local minima. The probability distribution of energy should resemble that of an ideal low-temperature Boltzmann distribution, but with a tail to higher energies. One choice is that the sampling of low-energy states is described by an exponential function (Boltzmann weight), while that of high-energy states follows a power law. Guided by these considerations, we propose the following as the new weight:

$$w(E) = \left(1 + \beta\, \frac{E - E_{\mathrm{GS}}}{m} \right)^{-m}, \qquad (3)$$

where $\beta \equiv 1/k_B T$, E_{GS} is the global-minimum energy, and $m(>0)$ is a free parameter. Here, we are shifting the zero of

---

*Electronic address: hansmann@ims.ac.jp

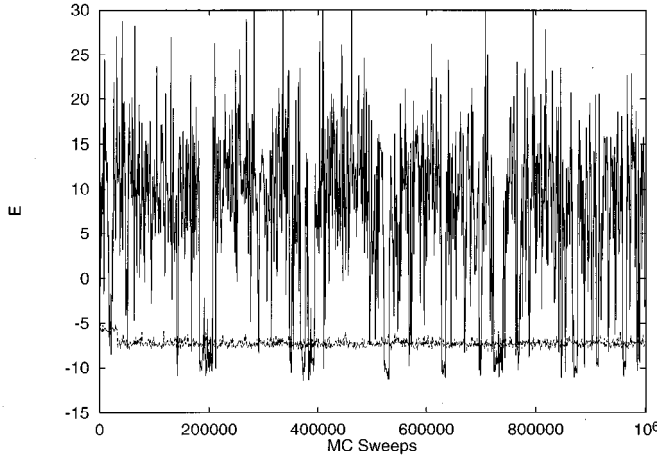[†]Electronic address: okamotoy@ims.ac.jp

FIG. 1. Time series of the total energy E_{tot} (in kcal/mol) from a regular canonical simulation at temperature $T = 50$ K (dotted curve) and from a simulation of the present method with the parameters: $E_0 = -12.2$ kcal/mol, $m = n_f = 19$, and $T = 50$ K (solid curve).

energy by E_{GS} in order to assure that energy is always non-negative. We remark that weights with the same mathematical structure also appear in the framework of Tsallis generalized statistical mechanics [11], which was developed for simulations of nonextensive systems (e.g., fractal random walks). An application to optimization problems can be found in Ref. [12].

Obviously, the new weight in Eq. (3) reduces to the canonical Boltzmann weight in the low-energy (and hence low-temperature) region for $[\beta(E - E_{GS})]/m \ll 1$. On the other hand, this weight at high energies is no longer exponentially suppressed, but only according to a power law with the exponent m. Note that our choice of the sign in Eq. (3) is important. From a mathematical point of view, $\{1 - \beta[(E - E_{GS})/m]\}^m$ is equally a good approximation to the canonical weight, but is not useful as a weight in numerical simulations, since the expression inside the curly brackets can become negative.

In this work we consider a system with continuous degrees of freedom. At low temperatures the harmonic approximation holds, and the density of states is given by

$$n(E) \propto (E - E_{GS})^{n_F/2}, \tag{4}$$

where n_F is the number of degrees of freedom of the system under consideration. Hence, by Eqs. (3) and (4) the probability distribution of energy at low temperatures for the present ensemble is given by

$$P(E) \propto n(E) w(E) \propto (E - E_{GS})^{(n_F/2)-m}, \tag{5}$$

for $\beta[(E - E_{GS})/m] \gg 1$. This implies that we need

$$m > \frac{n_F}{2}. \tag{6}$$

For, otherwise, the sampling of high-energy configurations will be enhanced too much. On the other hand, in the limit $m \to \infty$ our weight tends for all energies to the Boltzmann weight and high-energy configurations will not be sampled.

In order for low-temperature simulations to be able to escape from energy local minima, the weight should start
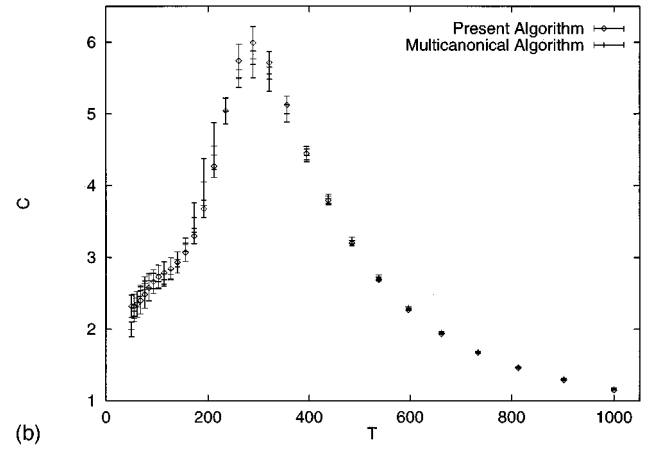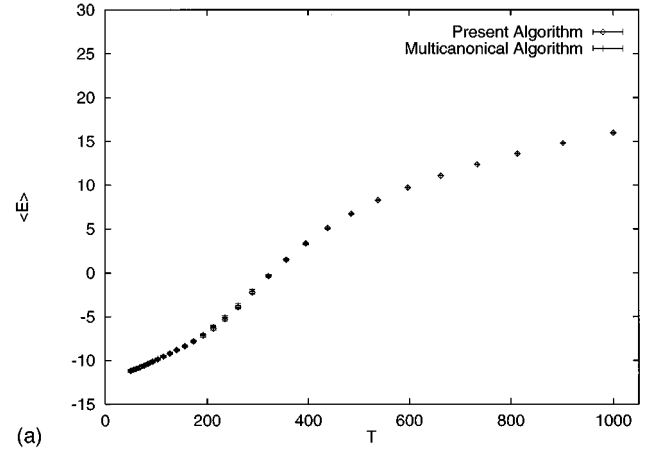

FIG. 2. Average energy (in kcal/mol) (a) and specific heat (b) as a function of temperature (in K). They were calculated by the re-weighting techniques from a single simulation run of the present method with the parameters: $E_0 = -12.2$ kcal/mol, $m = n_F = 19$, and $T = 50$ K. The results from a multicanonical simulation are also shown for comparison. In both simulations (by the present method and by the multicanonical algorithm) the total number of Monte Carlo sweeps was 10^6.

deviating from the (exponeβntially damped) Boltzmann weight at the energy near its mean value (because at low temperatures there are only small fluctuations of energy around its mean). This requires that $\beta(\langle E \rangle_T - E_{GS})/m = O(1)$ in Eq. (3). We may thus set

$$\beta \frac{\langle E \rangle_T - E_{GS}}{m} = \frac{1}{2}. \tag{7}$$

The mean value at low temperatures is given by the harmonic approximation

$$\langle E \rangle_T = E_{GS} + \frac{n_F}{2} k_B T = E_{GS} + \frac{n_F}{2\beta}. \tag{8}$$

Substituting this value into Eq. (7), we obtain the following optimal value for the exponent m:

$$m_{opt} = n_F. \tag{9}$$

Hence, the optimal weight factor is given by

$$w(E) = \left(1 + \beta \frac{E - E_0}{n_F}\right)^{-n_F}, \tag{10}$$

TABLE I. Lowest energy (in kcal/mol) obtained by the present method with several different choices of the free parameter m. The other free parameter E_0 was fixed at the value of the global-minimum energy $E_{GS}=-12.2$ kcal/mol. The temperature was set to $T=50$ K. The case for $m=\infty$ stands for a regular canonical run at $T=50$ K. For all cases, the total number of Monte Carlo sweeps per run was 50 000. $\langle E \rangle$ is the average of the lowest energy obtained by the ten runs (with the standard deviations in parentheses), and n_{GS} is the number of runs in which a conformation with $E \leqslant -11.0$ kcal/mol (the average energy at $T=50$ K) was obtained.

E_0	$E_{GS}=-12.2$	-12.2	-12.2	-12.2	-12.2	
m	$\frac{n_F}{2}=9.5$	14	$n_F=19$	50	100	∞
Run						
1	0.8	-5.2	-11.8	-6.9	-6.8	-4.2
2	-1.4	-2.6	-11.5	-7.1	-7.7	-5.2
3	0.1	-6.8	-11.5	-6.9	-4.9	-11.8
4	0.5	-5.5	-11.7	-8.2	-9.9	-7.1
5	-1.0	-3.4	-11.6	-7.4	-12.0	-3.3
6	1.1	-6.4	-11.6	-10.1	-8.8	0.9
7	-1.3	-5.1	-8.5	-8.7	-8.7	-5.3
8	0.4	-3.3	-9.7	-10.8	-9.5	-6.3
9	1.2	-8.1	-11.6	-12.0	-6.8	-6.4
10	1.2	-3.3	-11.9	-10.8	-9.5	-4.7
$\langle E \rangle$	0.2 (1.0)	-5.0 (1.8)	-11.1 (1.1)	-8.9 (1.9)	-8.5 (2.0)	-5.3 (3.2)
n_{GS}	0/10	0/10	8/10	1/10	1/10	1/10

where E_0 is the best estimate of the global-minimum energy E_{GS}.

We have tested our method in the system for the protein-folding problem, a long-standing problem in biophysics with rough energy landscape. Here, Met-enkephalin has become an often-used model to examine the performance of new algorithms, and we study the same system. Met-enkephalin has the amino-acid sequence Tyr-Gly-Gly-Phe-Met. The energy function E_{tot} (in kcal/mol) that we used is given by the sum of the electrostatic term E_C, 12-6 Lennard-Jones term E_{LJ} and hydrogen-bond term E_{HB} for all pairs of atoms in the peptide together with the torsion term E_{tor} for all torsion angles:

$$E_{tot}=E_C+E_{LJ}+E_{HB}+E_{tor}, \qquad (11)$$

$$E_C=\sum_{(i,j)} \frac{332 q_i q_j}{\epsilon r_{ij}}, \qquad (12)$$

$$E_{LJ}=\sum_{(i,j)} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} \right), \qquad (13)$$

$$E_{HB}=\sum_{(i,j)} \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right), \qquad (14)$$

$$E_{tor}=\sum_{l} U_l(1 \pm \cos(n_l \chi_l)). \qquad (15)$$

Here, r_{ij} is the distance (in Å) between the atoms i and j, and χ_l is the torsion angle for the chemical bond l. The parameters for the energy function and the molecular geometry (with fixed bond lengths and bond angles) were adopted from ECEPP/2 (Empirical Conformational Energy Program for Peptides) [13]. The dielectric constant ϵ was set equal to

2. Fixing the peptide bond angles ω to 180° leaves us with 19 torsion angles as independent degrees of freedom (i.e., $n_F=19$). The computer code KONF90 [14] was used. One Monte Carlo sweep updates every torsion angle of the peptide once.

It is known from our previous work that the global-minimum value of KONF90 energy for Met-enkephalin is $E_{GS}=-12.2$ kcal/mol [15]. The peptide has essentially a unique three-dimensional structure at temperatures $T \leqslant 50$ K, and the average energy is about -11 kcal/mol at $T=50$ K [9]. Hence, in the present work we always set $T=50$ K (or, $\beta=10.1[1/\text{kcal/mol}]$) in our probability weight factor. All simulations were started from completely random initial configurations (hot start).

To demonstrate that thermalization is greatly enhanced in our ensemble, we first compare the ''time series'' of energy as a function of Monte Carlo sweep. In Fig. 1 we show the results from a regular canonical Monte Carlo simulation at temperature $T=50$ K (dotted curve) and those from a generalized-ensemble simulation of the new algorithm (solid curve). Here, the weight we used for the latter simulation is given by Eq. (10) with $n_F=19$ and $E_0=E_{GS}=-12.2$ kcal/mol. For the canonical run the curve stays around the value $E=-7$ kcal/mol with small thermal fluctuations, reflecting the low-temperature nature. The run has apparently been trapped in a local minimum, since the mean energy at this temperature is $\langle E \rangle =-11.1$ kcal/mol as found by a multicanonical simulation in Ref. [15]. On the other hand, the simulation based on the new weight covers a much wider energy range than the canonical run. It is a random walk in energy space, which keeps the simulation from getting trapped in a local minimum. It indeed visits the ground-state region many times in 10^6 Monte Carlo sweeps. These properties are common features of generalized-ensemble methods.

Since the simulation by the present algorithm samples a large range of energies, we can use the reweighting techniques [7] to construct canonical distributions and calculate thermodynamic quantities over a wide temperature range. Following 10 000 sweeps for thermalization, we performed a single simulation of 10^6 Monte Carlo sweeps, storing the configuration information at every second sweep. We have set again $E_0 = -12.2$ kcal/mol and $n_F = 19$ in the weight of Eq. (10). From this production run one can calculate various thermodynamic quantities as a function of temperature. As examples we show the average energy and the specific heat in Figs. 2(a) and 2(b), respectively. The specific heat here is defined by the following equation:

$$ C \equiv \frac{1}{k_B} \frac{d\left(\frac{\langle E_{\text{tot}} \rangle_T}{N}\right)}{dT} = \beta^2 \frac{\langle E_{\text{tot}}^2 \rangle_T - \langle E_{\text{tot}} \rangle_T^2}{N}, \qquad (16) $$

where $N(=5)$ is the number of amino-acid residues in the peptide. The harmonic approximation holds at low temperatures, and by substituting $\langle E \rangle_T - E_{\text{GS}} = (n_F/2)k_B T$ in Eq. (8) into $\langle E_{\text{tot}} \rangle$ in the first equation of Eq. (16), we have

$$ C = \frac{n_F}{2N} = 1.9. \qquad (17) $$

Note that the curve in Fig. 2(b) approaches this value in the $T \to 0$ limit. The results from a multicanonical production run with the same statistics are also shown in the figures for comparison. The results from both methods are in complete agreement.

We now examine the dependence of the simulations on the values of the exponent m in our weight [see Eqs. (3) and (10)] and demonstrate that $m = n_F$ is indeed an optimal choice. Setting $E_0 = E_{\text{GS}} = -12.2$ kcal/mol, we performed ten independent simulation runs of 50 000 Monte Carlo sweeps with various choices of m. In Table I we list the lowest energies obtained during each of the ten runs for five choices of m values: $9.5(=n_F/2)$, 14, $19(=n_F)$, 50, and 100. The results from regular canonical simulations at $T = 50$ K with 50 000 Monte Carlo sweeps are also listed in the table for comparison. If m is chosen to be too small (e.g., $m = 9.5$), then the weight follows a power law in which the suppression for the higher-energy region is insufficient [see Eq. (5)]. As a result, the simulations tend to stay at high energies and fail to sample low-energy configurations. On the other hand, for too large a value of m (e.g., $m = 100$), the weight is too close to the canonical weight, and therefore the simulations will get trapped in local minima. It is clear from the table that $m = n_F$ is an optimal choice. In this case the simulations found the ground-state configurations 80% of the time (eight runs out of ten runs). This should be compared with 90%, 75%, 80%, and 40% for multicanonical annealing, $1/k$ annealing, simulated tempering annealing, and simulated annealing algorithms, respectively, in simulations with the same number of Monte Carlo sweeps [16].

To analyze the above results further, we calculated the actual probability distributions of energy for various values of m. This can be done by the reweighting techniques from the single production run of 10^6 Monte Carlo sweeps mentioned above [which is based on the weight of Eq. (10) with
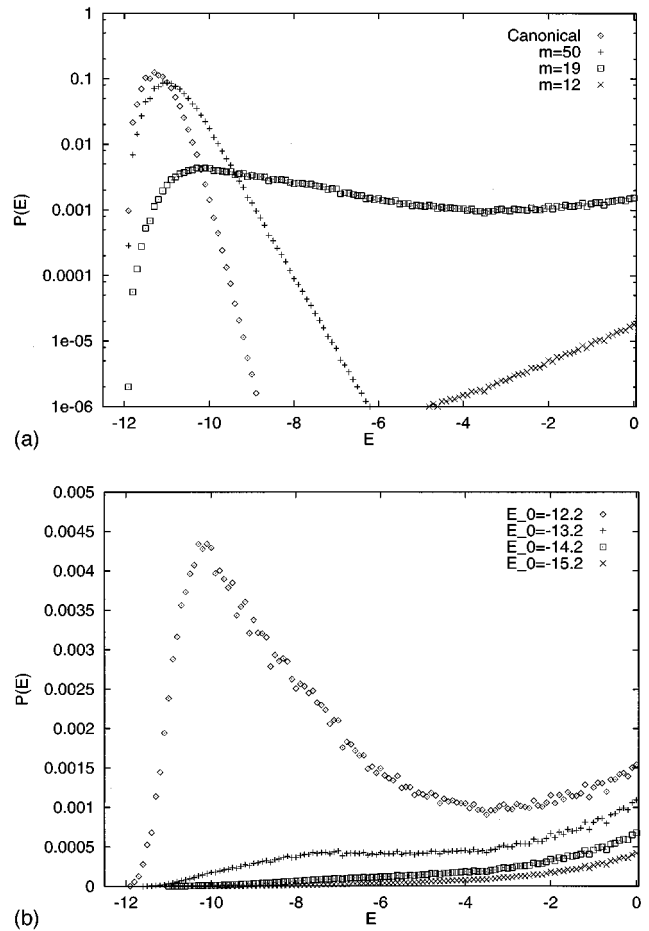

FIG. 3. Probability distributions of energy for various values of the exponent m (a) and the global-minimum energy estimate E_0 (b) in the present method. The ordinate for (a) is logarithmic. The results were obtained by the reweighting techniques from a single simulation run with the parameters: $E_0 = -12.2$ kcal/mol, $m = n_F = 19$, and $T = 50$ K. The total number of Monte Carlo sweeps was 10^6. For (a) the regular canonical distribution at $T = 50$ K as calculated by the reweighting techniques is also shown for comparison.

$E_0 = -12.2$ kcal/mol and $m = n_F = 19$]. The results are shown in Fig. 3(a). By examining the figure, we again find that $m = n_F$ is an optimal choice. It yields to an energy distribution which has a pronounced peak around the mean energy value ($\langle E \rangle = -11.1$ kcal/mol) at $T = 50$ K. At the same time, it has a tail to higher energies. This behavior is exactly what we were looking for and justifies our definition of weights in Eq. (10).

The greatest advantage of our method over other generalized-ensemble approaches is the simplicity of the weight factor. In multicanonical algorithms, $1/k$ sampling, or simulated tempering, the explicit functional forms of the weights are not known a priori and they have to be determined numerically by iterations of trial simulations. This can be a formidable task in many cases. On the other hand, the weight factor of the present algorithm just depends on the knowledge of the global-minimum energy E_{GS} [see Eq. (10)]. If its value is known, which is the case for some systems with frustration, the weight is completely determined. However, if E_{GS} is not known, we have to obtain its best estimate E_0. We can calculate the actual probability distributions of energy for various values of E_0 by the reweighting

TABLE II. Lowest energy (in kcal/mol) obtained by the present method with several different choices of the free parameter E_0. The other free parameter m was fixed at the optimal value of $n_F = 19$, the number of degrees of freedom. The temperature was set to $T = 50$ K. For all cases, the total number of Monte Carlo sweeps per run was 200 000. $\langle E \rangle$ is the average of the lowest energy obtained by the ten runs (with the standard deviations in parentheses), and n_{GS} is the number of runs in which a conformation with $E \leqslant -11.0$ kcal/mol (the average energy at $T = 50$ K) was obtained.

E_0	$E_{GS} = -12.2$	-13.2	-14.2	-15.2
m	$n_F = 19$	19	19	19
Run				
1	-11.8	-11.1	-10.5	-9.0
2	-11.9	-10.8	-8.3	-10.3
3	-11.9	-11.3	-11.6	-9.7
4	-11.9	-10.2	-10.9	-10.8
5	-11.8	-11.2	-6.9	-9.2
6	-11.3	-11.5	-10.8	-9.6
7	-11.9	-11.3	-8.3	-10.3
8	-11.8	-11.4	-5.9	-6.8
9	-12.0	-11.5	-10.6	-8.6
10	-11.7	-10.0	-10.3	-8.9
$\langle E \rangle$	-11.8 (0.2)	-11.0 (0.5)	-9.4 (1.9)	-9.3 (1.1)
n_{GS}	10/10	7/10	1/10	0/10

techniques again. The results are shown in Fig. 3(b). We see that for the system of Met-enkephalin, one needs the accuracy of about $1 \sim 2$ kcal/mol in the estimate of the global-minimum energy E_{GS} in order for our algorithm to be effective. This implication is supported by Table II where we list the lowest energies obtained during each of ten independent simulation runs of 200 000 Monte Carlo sweeps with $m = n_F = 19$. Four choices were considered for the E_0 value: -12.2, -13.2, -14.2, and -15.2 kcal/mol. We remark that E_0 has to underestimate E_{GS} to ensure that $E - E_0$ cannot become negative. Our data show again that an accuracy of $1 \sim 2$ kcal/mol in the estimate of the global-minimum energy is required for Met-enkephalin.

The use of our method therefore depends on the ability to find a good estimate for the ground-state energy E_{GS}, which is still much easier than the determination of the weights for other generalized-ensemble algorithms. In principle, such estimates can be found in an iterative way. Here, we give one of the effective iteration procedures. One first sets an initial guess of the optimal E_0 which should be lower than E_{GS}. One performs a simulation with the weight of the present method with small number of Monte Carlo sweeps. From this simulation one calculates the average energy $\langle E \rangle_T$ at the

chosen temperature T by the reweighting techniques. If $\langle E \rangle - E_0 \gg (n_F/2) k_B T$, one raises the value of E_0 by a certain amount and repeats the short simulation. One iterates this process until $\langle E \rangle - E_0 \approx (n_F/2) k_B T$. The search of the optimal E_0 can be further facilitated by information such as the average energy and the specific heat obtained from high temperature simulations. For Met-enkephalin the incorporation of such information gave a start value of $E_0 = -13.8$ kcal/mol, which is already within the 2 kcal/mol accuracy required by our method (see Ref. [17] for details).

In summary, we have introduced a generalized-ensemble algorithm for simulations of systems with frustration. We have demonstrated the effectiveness of the method by taking the example of the system of a small peptide Met-enkephalin, which has a rough energy landscape with a huge number of local minima. The advantage of the method lies in the fact that the determination of the probability weight factor is much simpler than in other generalized-ensemble approaches.

[1] R. H. Swendsen and J.-S. Wang, Phys. Rev. Lett. **58**, 86 (1987).

[2] B. A. Berg and T. Neuhaus, Phys. Lett. B **267**, 249 (1991); Phys. Rev. Lett. **68**, 9 (1992).

[3] B. A. Berg, Int. J. Mod. Phys. C **3**, 1083 (1992).

[4] B. Hesselbo and R. B. Stinchcombe, Phys. Rev. Lett. **74**, 2151 (1995).

[5] A. P. Lyubartsev, A. A. Martinovski, S. V. Shevkunov, and P. N. Vorontsov-Velyaminov, J. Chem. Phys. **96**, 1776 (1992).

[6] E. Marinari and G. Parisi, Europhys. Lett. **19**, 451 (1992).

[7] A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **61**, 2635 (1988); **63**, 1658(E) (1989), and references given in the erratum.

[8] B. A. Berg and T. Celik, Phys. Rev. Lett. **69**, 2292 (1992); B. A. Berg, T. Celik, and U. H. E. Hansmann, Phys. Rev. B **50**, 16 444 (1994).

[9] U. H. E. Hansmann and Y. Okamoto, J. Comput. Chem. **14**, 1333 (1993).

[10] M.-H. Hao and H. A. Scheraga, J. Chem. Phys. **98**, 4940 (1994).

[11] C. Tsallis, J. Stat. Phys. **52**, 479 (1988).

[12] D. A. Stariolo and C. Tsallis, in *Annual Reviews of Computational Physics II*, edited by D. Stauffer (World Scientific, Singapore, 1995), p. 343; I. Andricioaei and J. E. Straub, Phys. Rev. E **53**, R3055 (1996).

[13] M. J. Sippl, G. Némethy, and H. A. Scheraga, J. Phys. Chem. **88**, 6231 (1984), and references therein.

[14] H. Kawai, Y. Okamoto, M. Fukugita, T. Nakazawa, and T. Kikuchi, Chem. Lett. **1991**, 213 (1991); Y. Okamoto, M. Fukugita, T. Nakazawa, and H. Kawai, Protein Eng. **4**, 639 (1991).

[15] U. H. E. Hansmann and Y. Okamoto, J. Phys. Soc. Jpn. **63**, 3945 (1994); Physica A **212**, 415 (1994).

[16] U. H. E. Hansmann and Y. Okamoto, J. Comput. Chem. **18**, 920 (1997).

[17] U. H. E. Hansmann (unpublished).