An Analysis on Protein Folding Problem by Replica-Exchange Method

Yuji SUGITA*) and Yuko OKAMOTO**)

Department of Theoretical Studies, Institute for Molecular Science
Okazaki 444-8585, Japan

(Received October 11, 1999)

Because of the ruggedness of the energy landscape of protein systems, computer simulations of a protein folding process from the random-coil state to the native state is very difficult to achieve by the conventional molecular dynamics (MD) or Monte Carlo (MC) methods.[1] Recently, *Replica-Exchange MC method*[2] draws much attention as a promising algorithm that can overcome the multiple-minima problem. (This method is also referred to as *replica Monte Carlo method*,[3] *multiple Markov chain method*,[4] and *parallel tempering*.[5]) In this study, we have developed the MD version of *Replica-Exchange method* (*Replica-Exchange MD method*, REMD),[6] since in the system consisting of a protein and explicit water molecules, MD in the Cartesian coordinates is much easier to perform than MC.

In REMD, we consider a generalized-ensemble with M *non-interacting* copies (or, replicas) of the original system in the canonical ensemble at M different inverse temperatures β_m $(m = 1, \ldots, M)$. The weight factor for the state X in this generalized ensemble, $W_{\mathrm{REMD}}(X)$ is given by the product of Boltzmann factors for each replica, since the replicas are non-interacting. We now consider exchanging replicas i and j which are at β_m and β_n, respectively. 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 $w(X \to X')$:

$$W_{\mathrm{REMD}}(X)w(X \to X') = W_{\mathrm{REMD}}(X')w(X' \to X). \tag{1}$$

We have shown that the transition probability, $w(X \to X')$, which satisfies the detailed balance condition is as follows:[6]

$$w(X \to X') = \begin{cases} 1 & (\text{for } \Delta \le 0) \\ \exp(-\Delta) & (\text{for } \Delta > 0), \end{cases} \tag{2}$$

where

$$\Delta = (\beta_n - \beta_m)(E(q^i) - E(q^j)) \tag{3}$$

and $E(q^i)$ and $E(q^j)$ are the potential energies of replicas i and j, respectively. Note that this is exactly the same criterion that was originally derived for MC.[2)-5)]

The effectiveness of the algorithm has already been shown for a peptide (Met-enkephalin) in gas phase.[6] We also applied the algorithm to a peptide (a blocked

*) E-mail address: sugita@ims.ac.jp
**) E-mail address: okamotoy@ims.ac.jp

alanine dimer: Ace-Ala-Ala-Nme) in water. In this system, 228 water molecules are placed around the solute to fill a sphere of 12 Å radius. The force field parameters were the all atom version of AMBER by Cornell et al. [7] for the peptide and TIP3P [8] for water molecules, respectively. The unit time step was set to 0.5 fsec. For each replica, 0.5 nsec MD simulations were performed, starting from an extended conformation. We used the sixteen temperatures between 550 K and 250 K, which are distributed exponentially. By using this choice, the acceptance ratios of replica-exchange are almost uniform (all about 15 %). In Fig. 1, the canonical probability distributions obtained at the chosen sixteen temperatures are shown. We see that there are enough overlaps between all adjacent pairs of distributions. This indicates that our choice of temperatures is good enough to perform REMD for this system. In Fig. 2, the time series of potential energy for one of the replicas (Replica 1) is shown. We see that a random walk in the potential energy space between low and high energies is realized. Since each replica does not get trapped at any states of local-energy-minimum, the method samples much broader configuration space than the conventional canonical MD method. By combining the multiple histogram method [9], we can obtain various thermodynamic quantities as a function of temperature for a wide temperature range only from one REMD simulation. Hence, the new method is particularly useful for studying the protein folding problem.

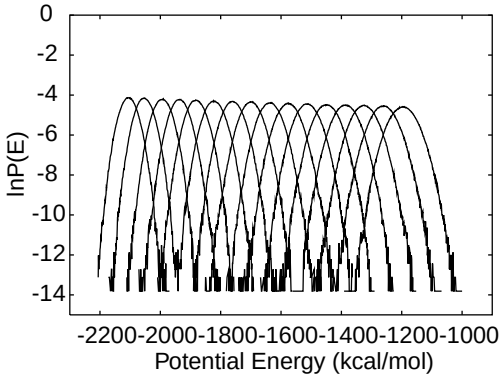

Fig. 1. The energy histograms for all the systems in the Replica-Exchange MD.

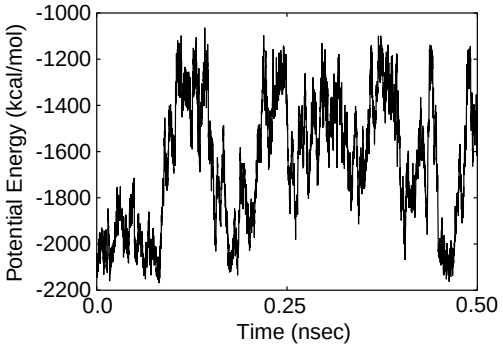

Fig. 2. Time series of potential energy in replica 1.

References

1) For a review, see U. H. E. Hansmann and Y. Okamoto, Curr. Opin. Struct. Biol. **9** (1999), 177.
2) K. Hukushima and K. Nemoto, J. Phys. Soc. Jpn. **65** (1996), 1604.
3) R. H. Swendsen and J.-S. Wang, Phys. Rev. Lett. **57** (1986), 2607.
4) C. J. Geyer, *Computing Science and Statistics: Proc. 23rd Symp. on the Interface*, ed. E. M. Keramidas (Interface Foundation, Fairfax Station, 1991), p. 156.
5) E. Marinari, G. Parisi and J. J. Ruiz-Lorenzo, in *Spin Glasses and Random fields*, ed. A.P. Young (World Scientific, Singapore, 1998), p. 140.
6) Y. Sugita and Y. Okamoto, Chem. Phys. Lett., in press.
7) W. D. Cornell et al., J. Am. Chem. Soc. **117** (1995), 5179.
8) W. L. Jorgensen et al., J. Chem. Phys. **79** (1983), 926.
9) A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **63** (1989), 1658.