13 July 2001

ELSEVIER

Chemical Physics Letters 342 (2001) 382–386

CHEMICAL PHYSICS LETTERS

www.elsevier.com/locate/cplett

# Sampling efficiency of molecular dynamics and Monte Carlo method in protein simulation

Hiroshi Yamashita [a], Shigeru Endo [b], Hiroshi Wako [c], Akinori Kidera [a,*]

[a] *Department of Chemistry, Graduate School of Science, Kyoto University, Kitashirakawa, Sakyo-ku, Kyoto 606-8502, Japan*
[b] *Department of Physics, School of Science, Kitasato University, Kitasato, Sagamihara 228-8555, Japan*
[c] *School of Social Sciences, Waseda University, Nishi-Waseda, Shinjuku-ku, Tokyo 169-8050, Japan*

Received 15 February 2001; in final form 11 May 2001

**Abstract**

Molecular dynamics (MD) and Monte Carlo (MC) method were compared in terms of the sampling efficiency in protein simulations. In the comparison, both methods use torsion angles as the degrees of freedom and the same force field, ECEPP/2. The MC method used here is the force-bias scaled-collective-variable Monte Carlo (SCV MC) [A. Kidera, Int. J. Quant. Chem. 75 (1999) 207], which corresponds to a finite step size extension to Brownian dynamics. It is shown that MD has about 1.5 times larger sampling efficiency. This difference is attributed to the inertia force term in MD, which does not exist in MC. 

## 1. Introduction

Molecular dynamics (MD) and Metropolis Monte Carlo (MC) method are the standard techniques for molecular simulations, and play almost the same role in sampling configurational space of liquids [1]. It is recognized that the sampling efficiency of the MC method is comparable, or sometimes superior to MD in liquid simulations [2].

On the other hand, in protein simulations, a straight forward application of the MC method results in extremely low efficiency, because the random MC steps are incompatible with the collective nature of protein motions, which are originated from the covalent structures and the tightly packed globular shape [3–5]. To improve the sampling efficiency in the MC simulation of proteins, Noguti and Go proposed the scaled-collective-variable Monte Carlo (SCV MC) method [4]. The SCV MC method uses the instantaneous normal modes [6] in torsion angle space as the random variables for determining a trial MC step, and succeeded to avoid unfavorable motions caused by the random MC steps lowering the acceptance ratio. In the SCV MC method, the increments of $N$ torsion angles, $\Delta\boldsymbol{\theta}$, for a trial step is defined by,

$$\Delta\boldsymbol{\theta} = \beta^{-1/2}\boldsymbol{\Omega}|\boldsymbol{\Lambda}|^{-1/2}\boldsymbol{\sigma}, \tag{1}$$

where $\beta = 1/k_{\mathrm{B}}T$, $\boldsymbol{\sigma}$ is a vector of $N$ random numbers with unit variance corresponding to the instantaneous normal modes of the current structure, and $\boldsymbol{\Lambda}$ and $\boldsymbol{\Omega}$ are the $N \times N$ eigenvalue and eigenvector matrices for the normal modes,

* Corresponding author. Fax: +81-75-753-3669.
*E-mail address:* kidera@qchem.kuchem.kyoto-u.ac.jp (A. Kidera).

0009-2614/01/$ - see front matter 
PII: S0009-2614(01)00613-3

respectively. Later, a force-bias form of the SCV MC method was developed to improve the sampling efficiency by us [5]. Since then, no proper comparison of the sampling efficiency between MD and SCV MC has not been made yet.

As explained in [5], the force-bias SCV MC method is a natural extension of Brownian dynamics to an algorithm with a finite and heterogeneous step size [7,8]. The equations of Brownian dynamics in Cartesian coordinates, $\mathbf{r}$,

$$-\mathbf{\Gamma}\dot{\mathbf{r}} + \mathbf{f} + \mathbf{a} = \mathbf{0} \tag{2}$$

with

$$\langle \mathbf{a} \rangle = \mathbf{0} \quad \text{and} \quad \langle \mathbf{a}\mathbf{a}^{\mathrm{t}} \rangle = 2\beta^{-1}\mathbf{\Gamma}, \tag{3}$$

can be rewritten directly in the force-biased form of the SCV MC method in torsion angle space:

$$\Delta\boldsymbol{\theta} = \beta^{-1/2}\mathbf{\Omega}|\mathbf{\Lambda}|^{-1/2}\boldsymbol{\sigma} + (1/2)|\mathbf{F}|^{-1}\mathbf{f}, \tag{4}$$

where $\mathbf{\Gamma}$ is the friction matrix, $\mathbf{f}$ is the force vector, the vector $\mathbf{a}$ represents the random fluctuation with the variances specified in Eq. (3), and $\mathbf{F}$ is the Hessian matrix of the potential energy [9]. Therefore, the major difference between the MD and the force-bias SCV MC methods is in the inertia force term, $\mathbf{M}\ddot{\mathbf{r}}$ ($\mathbf{M}$ being the diagonal mass weight matrix) which exists in the MD, but not in Brownian dynamics or the force-bias SCV MC method. In the force-bias SCV MC simulation, a trial step immediately loses the memory of velocity at the previous step. The question in the comparison is thus how much the inertia force term affects the sampling efficiency.

For a proper comparison of MD and MC, both methods have to use the same degrees of freedom and the same force fields. Here, we used a torsion angle MD algorithm [10] with the force field ECEPP/2 [11] for the comparison.

## 2. Methods

The force-bias SCV MC method and the torsion angle MD algorithm are explained very briefly. Details of each algorithm are given elsewhere [5,10].

### 2.1. Force-bias SCV MC method

In the force-bias SCV MC method, the increments of $N$ torsion angles, $\Delta\boldsymbol{\theta}$, for a trial step is calculated by Eq. (4). Since the Hessian matrix of the instantaneous normal modes is not positive definite, the eigenvalues are replaced by their absolute values with a suitable lower bound ($= 1.0$ kcal mol$^{-1}$ rad$^{-2}$ in this Letter) to avoid the floating exception. The Metropolis criterion to accept a trial step from an old configuration $m$ to a new configuration $n$ is

$$p_{mn} = \min(1, Q_{mn}) \tag{5}$$

with

$$Q_{mn} = \exp\{-\beta[E_n - E_m]\}(|\mathbf{\Lambda}_n|/|\mathbf{\Lambda}_m|)^{1/2} \times \exp\left\{ -\frac{1}{2}[(\boldsymbol{\sigma}_n - \boldsymbol{\sigma}_m)\cdot(\boldsymbol{\gamma}_m + \boldsymbol{\gamma}_n) + \frac{1}{4}(\boldsymbol{\gamma}_n^2 - \boldsymbol{\gamma}_m^2)]\right\}, \tag{6}$$

where $E_m$ is the potential energy of a configuration in state $m$, and $\boldsymbol{\gamma}_m$ is the force vector along the normal mode axes defined by

$$\boldsymbol{\gamma}_m = \beta^{1/2}|\mathbf{\Lambda}_m|^{-1/2}\mathbf{\Omega}_m{}^t\mathbf{f}_m. \tag{7}$$

The Hessian matrix can be calculated by Noguti and Go's rapid calculation algorithm [9,12,13]. The eigenvalue and eigenvector matrices were updated every $10^3$ steps.

### 2.2. Torsion angle MD

In torsion angle space, MD describes time evolution of protein structure following the equations of motion. Angular accelerations for $N$ torsion angles were computed with a recursive algorithm using the Newton–Euler inverse mass operator [14] from angular velocities as well as torques around torsion angles. Torques, the first derivatives of the potential energy in torsion angle space, were computed by Noguti and Go's algorithm [9,12,13]. Numerical integration of the equations of motion was done by the implicit leap frog integrator [15] and the Nosé–Hoover thermostat [16,17] with the step size of 2 fs. Detail

description of the MD program will be published elsewhere [10].

## 3. Computations

Both the MD and force-bias SCV MC simulations were carried out in torsion angle space on phage 434 Cro protein [18], a small protein consisting of 71 amino acids (PDB entry code: 2CRO), in vacuo at 300 K. As references, a simple MC simulation ($\mathbf{\Omega} = \mathbf{\Lambda} = \mathbf{I}$ in Eq. (1), $\mathbf{I}$ being the identity matrix) and SCV MC simulation (Eq. (1)) were also carried out. Force field parameters used in both the MD and MC simulations were taken from the parameter set of ECEPP/2 [11]. Starting with a well-equilibrated configuration generated by torsion angle MD of 100 ps, both the MD and MC simulations generates five runs in configuration space for $10^5$ steps. Resultant statistics for each method were given by averaging over five runs.

## 4. Results and discussion

The sampling efficiency of each method is measured by the rate at which the configurational space is sampled. Here, we adopted the root-mean-square (RMS) fluctuation of the protein structure in Cartesian coordinates as the measure. Table 1 summarizes the results of the MD and MC simulations.

It is confirmed in Table 1 that the MD and MC simulations sampled the configurations with almost the same mean potential energy, which are resulted from the probability distribution function corresponding to the canonical ensemble at 300 K. The larger value of energy fluctuation in MD may reflect the difference in the size of the sampled space. When comparing the values of RMS fluctuation, we found that the sampling power increases with the order, simple MC, SCV MC, force-bias SCV MC, and torsion angle MD. The SCV algorithm adequately reflects the cooperative nature of protein dynamics to increase the sampling efficiency. The force-bias simulation succeeds to increase the step size, $\langle \Delta\boldsymbol{\theta}^2 \rangle^{1/2}$, without decreasing the acceptance ratio. Table 1 shows that the step size in force-bias SCV MC does not affect the efficiency, because the increase in the step size immediately lowers the acceptance ratio. In conclusion, the highest efficiency was achieved by the torsion angle MD.

The simulation step dependence of RMS fluctuation in Fig. 1 demonstrates the sampling power of the MD algorithm more clearly. The RMS fluctuation of the MD exceeds that of the force-bias SCV MC by about 50–60% through all the simulation steps. Fig. 2 shows the residue profiles of the fluctuations. The profile of the MD simulation similar to that of the force-bias SCV MC indicates that the MD algorithm enhances the sampling space not only in the solvent exposed region where the atoms can move freely, but also

Table 1
Results of MD and MC simulations[a]

| Method | $\langle \Delta\boldsymbol{\theta}^2 \rangle^{1/2}$[b] (radian) | Acceptance ratio | Mean potential energy (kcal/mol) | RMS fluctuation[c] (Å) | |
|---|---|---|---|---|---|
| | | | | All atoms | Main-chain atoms |
| Torsion angle MD | 0.017 | – | $-512.9 \pm 10.4$ | $0.96 \pm 0.09$ | $0.60 \pm 0.04$ |
| Force-bias SCV MC | 0.038 | $0.12 \pm 0.08$ | $-511.8 \pm 8.8$ | $0.66 \pm 0.02$ | $0.41 \pm 0.02$ |
| | 0.028 | $0.32 \pm 0.09$ | $-510.7 \pm 9.2$ | $0.66 \pm 0.02$ | $0.41 \pm 0.02$ |
| | 0.022 | $0.51 \pm 0.11$ | $-510.7 \pm 9.1$ | $0.69 \pm 0.01$ | $0.43 \pm 0.01$ |
| | 0.017 | $0.74 \pm 0.09$ | $-511.4 \pm 8.9$ | $0.67 \pm 0.01$ | $0.42 \pm 0.01$ |
| SCV MC | 0.009 | $0.48 \pm 0.05$ | $-511.4 \pm 8.7$ | $0.58 \pm 0.03$ | $0.36 \pm 0.01$ |
| Simple MC | 0.002 | $0.55 \pm 0.06$ | $-511.0 \pm 8.0$ | $0.38 \pm 0.01$ | $0.24 \pm 0.01$ |

[a] The results were calculated from five runs of $10^5$ step simulation, and given as the averages and the S.D.s.
[b] The RMS torsion angle change between successive simulation steps. $\langle \cdots \rangle$ denotes the average over $10^5$ steps (only accepted steps for MC) for all rotatable torsion angles.
[c] The RMS deviation from the average coordinates.

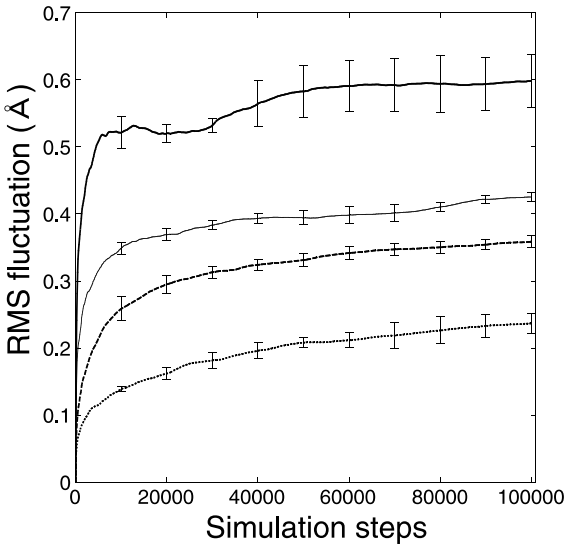


Fig. 1. Extension of sampling space with simulation steps. The curves and the error bars are the average and the S.D. for five runs of the simulations, respectively. The thick solid, thin solid, broken, and dotted curves are for the simulations by torsion angle MD, force-bias SCV MC, SCV MC, and simple MC, respectively.

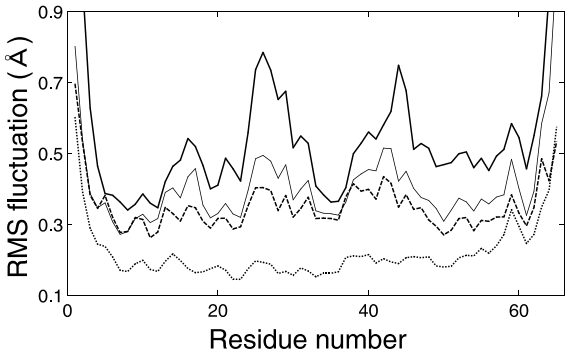


Fig. 2. RMS fluctuations of the main-chain atoms are plotted against the residue number. The thick solid, thin solid, broken, and dotted curves for the simulations by torsion angle MD, force-bias SCV MC, SCV MC, and simple MC, respectively.

in the buried region where the atoms are tightly packed.

As explained above, the major difference between the MD and the force-bias SCV MC methods is in the inertia force term, which exists in the MD, but not in the force-bias SCV MC method. Since the force-bias SCV MC method is a finite step size version of Brownian dynamics, a trial step in the simulation immediately loses the memory of velocity at the previous step. Actually, as shown in Fig. 3, the auto-correlation functions of the increment of the torsion angles ($\phi$ and $\psi$), $\Delta\boldsymbol{\theta}$, and of the displacement vectors of coordinates of $C^{\alpha}$ atoms, $\Delta\mathbf{r}$, show that the motions in the MC simulation have almost no correlations, but the MD trajectory has a long correlation time. Therefore, it is reasonable to conclude that the higher sampling efficiency in MD is due to the long correlation time caused by the inertia force term.

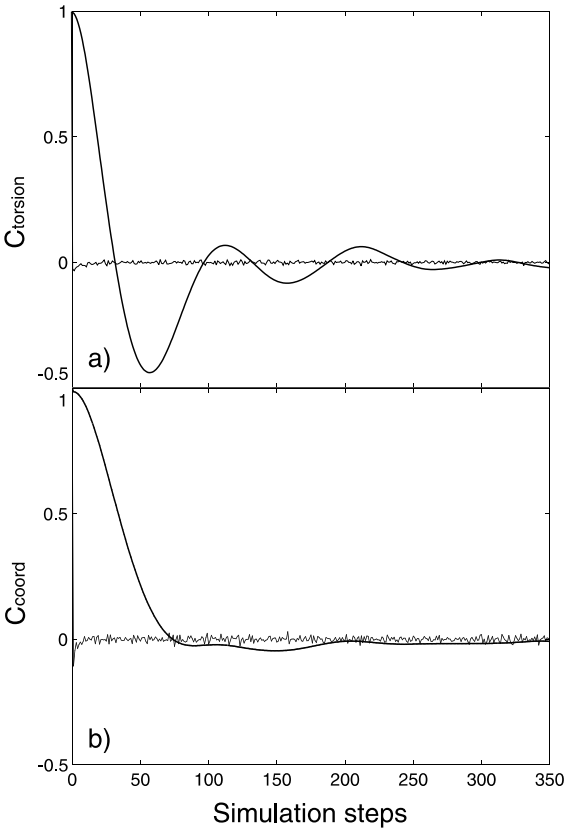


Fig. 3. Auto-correlation functions of (a) the increment of the torsion angles ($\phi$ and $\psi$), $\Delta\boldsymbol{\theta}$, between successive simulation steps and of (b) the displacement vectors of coordinates of $C^{\alpha}$ atoms, $\Delta\mathbf{r}$, between successive simulation steps defind by $C_{\text{torsion}}(s) \equiv \langle \Delta\boldsymbol{\theta}(0) \cdot \Delta\boldsymbol{\theta}(s) \rangle / \langle \Delta\boldsymbol{\theta}(0)^2 \rangle$ and $C_{\text{coord}}(s) \equiv \langle \Delta\mathbf{r}(0) \cdot \Delta\mathbf{r}(s) \rangle / \langle \Delta\mathbf{r}(0)^2 \rangle$, respectively. The thick and thin curves are for the simulations by torsion angle MD and force-bias SCV MC, respectively.

## Acknowledgements

The computations were done in the Computer Center of Institute for Molecular Science, Center for Promotion of Computational Science and Engineering of JAERI. This work was supported by the Grant-in-Aid for Scientific Research from MEXT to A.K.

## References

[1] M.P. Allen, D.J. Tildesley, Computer Simulation of Liquids, Oxford University Press, Oxford, 1989.
[2] W.L. Jorgensen, J. Tirado-Rives, J. Phys. Chem. 100 (1996) 14508.
[3] S.H. Northrup, J.A. McCammon, Biopolymers 19 (1980) 1001.
[4] T. Noguti, N. Go, Biopolymers 24 (1985) 527.
[5] A. Kidera, Int. J. Quant. Chem. 75 (1999) 207.
[6] B.-C. Xu, R.M. Stratt, J. Chem. Phys. 92 (1990) 1923.
[7] C. Pangali, M. Rao, B.J. Berne, Chem. Phys. Lett. 55 (1978) 413.
[8] P.J. Rossky, J.D. Doll, H.L. Friedman, J. Chem. Phys. 69 (1978) 4628.
[9] T. Noguti, N. Go, J. Phys. Soc. Jpn. 52 (1983) 3685.
[10] H. Wako, S. Endo, in preparation.
[11] G. Némethy, M.S. Pottle, H.A. Scheraga, J. Phys. Chem. 87 (1983) 1883.
[12] H. Abe, W. Braun, T. Noguti, N. Go, Comput. Chem. 8 (1984) 239.
[13] H. Wako, N. Go, J. Comput. Chem. 8 (1987) 625.
[14] N. Jain, N. Vaidehi, G. Rodriguez, J. Comput. Phys. 106 (1993) 258.
[15] A.K. Mazur, J. Comp. Chem. 18 (1997) 1354.
[16] S. Nosé, J. Chem. Phys. 81 (1984) 511.
[17] W.G. Hoover, Phys. Rev. A 31 (1984) 1695.
[18] A. Mondragón, C. Wolberer, S.C. Harrison, J. Mol. Biol. 205 (1989) 179.