Rigid-Body Dynamics in the Isothermal-Isobaric Ensemble: A Test on the Accuracy and Computational Efficiency

WATARU SHINODA, MASUHIRO MIKAMI

Research Institute for Computational Sciences (RICS) and Research Consortium for Synthetic Nano-Function Materials Project (SYNAF), National Institute for Advanced Industrial Science and Technology (AIST), Tsukuba Central 2, Umezono 1-1-1, Tsukuba, Ibaraki 305-8568, Japan

Received 18 November 2002; Accepted 4 December 2002

Abstract: We have developed a time-reversible rigid-body (rRB) molecular dynamics algorithm in the isothermal-isobaric (NPT) ensemble. The algorithm is an extension of rigid-body dynamics [Matubayasi and Nakahara, J Chem Phys 1999, 110, 3291] to the NPT ensemble on the basis of non-Hamiltonian statistical mechanics [Martyna, G. J. et al., J Chem Phys 1994, 101, 4177]. A series of MD simulations of water as well as fully hydrated lipid bilayer systems have been undertaken to investigate the accuracy and efficiency of the algorithm. The rRB algorithm was shown to be superior to the state-of-the-art constraint-dynamics algorithm SHAKE/RATTLE/ROLL, with respect to computational efficiency. However, it was revealed that both algorithms produced accurate trajectories of molecules in the NPT as well as NVT ensembles, as long as a reasonably short time step was used. A couple of multiple time-step (MTS) integration schemes were also examined. The advantage of the rRB algorithm for computational efficiency increased when the MD simulation was carried out using MTS on parallel processing computer systems; total computer time for MTS-MD of a lipid bilayer using 64 processors was reduced by about 40% using rRB instead of SHAKE/RATTLE/ROLL.

 J Comput Chem 24: 920–930, 2003

Key words: molecular dynamics simulations; isothermal-isobaric ensemble; holonomic constraint; rigid body; time-reversible integrator

Introduction

In recent years, molecular dynamics (MD) simulations have been extensively used to investigate the structure and dynamics of macromolecular systems such as proteins, polymers, liquid crystals, and so on.[1] In these complex systems, the scientific interest sometimes lies in the slow dynamics, which are often of vital importance in characterizing the properties or functions of the macromolecules. In an MD simulation, however, we have to choose a time step substantially shorter than the shortest relevant time scale in such a simulation.[2,3] This gives rise to a practical difficulty in studying the longer-term dynamics of macromolecules by MD simulation. To overcome the difficulty, holonomic constraints have been used. Especially for the macromolecular system, it is computationally efficient to build a model system by imposing holonomic constraints upon functionally unimportant, high-frequency degrees of freedom, such as bond stretching and angle bending relevant to hydrogen atoms.[2,3] This approximation allows us to use a long time step in our MD simulations. A multiple time-step (MTS) algorithm can be used to tackle the time-scale problem; the RESPA algorithm by Tuckerman et al.[4] provides an elegant MTS integration scheme without losing reversibility in time. In addition, an MD simulation of a system subjected to holonomic constraints can be carried out with an MTS algorithm fairly straightforwardly.

Complex fluids such as colloids, surfactants, and polymer systems generally show complex phase behavior, and a variety of phases sometimes appear within a narrow range of temperature and pressure.[5] In MD simulations for such systems, the isothermal-isobaric (NPT) condition is necessary in order to define clearly the thermodynamic state of the simulated system. In other words, NPT-MD simulation allows for proper comparison with experimental measurements. In addition, when precise structural data such as density or lattice constants are not available through experimentation, NPT-MD becomes advantageous because we can expect that a flexible cell will automatically adjust its overall structure (or density) in response to external pressure. For example, it is usually difficult to estimate the molecular areas for lipids

Correspondence to: W. Shinoda; e-mail: w.shinoda@aist.go.jp

Contract/grant sponsor: NEDO

and membrane proteins experimentally. Thus, several MD simulations of lipid bilayer systems have been carried out through the NPT ensemble,[6,7] especially in tandem with the Andersen method (three cell lengths are treated as independent variables respectively)[8] and the Parrinello-Rahman method (the fully flexible cell).[9] Furthermore, the use of NPT dynamics is essential when the first order phase transition is of interest in the study, because the phase transition involves a dramatic change in the density of the system. In fact, several novel molecular simulation techniques aimed at understanding of the phase behavior and/or phase transition have been developed in the NPT ensemble.[10,11]

In this work, we will consider an efficient and accurate algorithm for MD of a system subjected to holonomic constraints in the NPT ensemble. Actually, we have proposed time-reversible rigid-body (rRB) dynamics to realize a stable and computationally effective MD calculation in the NPT ensemble. Although rigid-body approximation is just a special case of holonomic constraints, we can utilize the approximation instead of more general constraint dynamics algorithm (e.g., SHAKE[12]) for many cases. In the rigid-body algorithm, macromolecules are treated as semiflexible molecules by introducing rigid-body approximation to the segments in the molecules. We have extended an rRB algorithm[13] to the NPT ensemble constructed on the basis of non-Hamiltonian statistical mechanics.[14–16] The resulting algorithm is explicitly reversible in time. To assess the performance of the rRB algorithm, a series of MD calculations has been undertaken for water as well as lipid bilayer systems. MTS integration of the rRB algorithm has also been examined. For the sake of comparison, we have carried out the same MD calculations using the SHAKE/RATTLE/ROLL algorithm.[12,15,17]

The article is organized as follows: In the Methods Section, after a description of the rRB algorithm in the NPT ensemble, the details of test MD calculations are presented. Results of a series of MD are discussed in the Results section. In the Discussion section, we discuss features of the rRB algorithm from the viewpoint of computational efficiency and applicability. The article ends with conclusions.

Methods

In the rRB algorithm, instead of degrees of freedom of atoms, the translational $\{\vec{r}_{gi}, \vec{p}_{gi}\}$ and rotational $\{\vec{q}_i, \vec{L}_i\}$ degrees of freedom of rigid bodies are used to describe the molecular dynamics. The quaternion ($\vec{q}_i$) representation was adopted to avoid the singularity of Euler angle dynamics.[2] Taking the molecular frame to diagonalize the inertia moment of the rigid body i, $\overleftrightarrow{I}_i$, the angular momentum $\vec{L}_i$ is combined with the angular velocity $\vec{\omega}_i$ as

$$L_{i,x} = I_{i,x}\omega_{i,x}$$

$$L_{i,y} = I_{i,y}\omega_{i,y}$$

$$L_{i,z} = I_{i,z}\omega_{i,z}$$

Non-Hamiltonian molecular dynamics schemes were employed to generate the NPT ensemble.[16,18] The Nosé-Hoover chain[19] was employed as a thermostat of the system, and a flexible cell was adopted to keep the pressure constant. The latter technique was originally developed by Andersen,[8] generalized by Parrinello and Rahman,[9] and recently extended to a modularly invariant form by Martyna et al.[14] It has already been shown that a stable integrator can be constructed for extended system dynamics on the basis of the Trotter factorization of the Liouville operator using the RESPA algorithm.[15]

The set of equations of motions to be solved for the rRB dynamics in the NPT ensemble is as follows:

$$\dot{\vec{r}}_{gi} = \frac{\vec{p}_{gi}}{M_i} + \frac{\overleftrightarrow{p}_h}{W_h}\vec{r}_{gi},$$

$$\dot{\vec{p}}_{gi} = \vec{F}_{gi} - \frac{\overleftrightarrow{p}_h}{W_h}\vec{p}_{gi} - \frac{1}{N_g}\frac{\mathrm{Tr}[\overleftrightarrow{p}_h]}{W_h}\vec{p}_{gi} - \frac{p_\xi}{Q}\vec{p}_{gi},$$

$$\dot{\vec{q}}_i = \frac{1}{2}\overleftrightarrow{A}(\vec{\omega})\vec{q}_i,$$

$$\dot{L}_{i,x} = \tau_{i,x} - \left(\frac{1}{I_{i,y}} - \frac{1}{I_{i,z}}\right)L_{i,y}L_{i,z},$$

$$\dot{L}_{i,y} = \tau_{i,y} - \left(\frac{1}{I_{i,z}} - \frac{1}{I_{i,x}}\right)L_{i,z}L_{i,x},$$

$$\dot{L}_{i,z} = \tau_{i,z} - \left(\frac{1}{I_{i,x}} - \frac{1}{I_{i,y}}\right)L_{i,x}L_{i,y},$$

$$\dot{\overleftrightarrow{h}} = \frac{\overleftrightarrow{p}_h}{W_h}\overleftrightarrow{h},$$

$$\dot{\overleftrightarrow{p}}_h = V(\overleftrightarrow{P}_{\mathrm{int}} - \overleftrightarrow{E}P_{\mathrm{ext}}) + \left(\frac{1}{N_g}\sum_{i=1}^{N}\frac{\vec{p}_{gi}^{\,2}}{M_i}\right)\overleftrightarrow{E} - \frac{p_{\xi_1}}{Q_1}\overleftrightarrow{p}_h,$$

$$\dot{\xi}_k = \frac{p_{\xi_k}}{Q_k} \qquad \text{for } k = 1, \ldots, M,$$

$$\dot{p}_{\xi_1} = \sum_{i=1}^{N}\frac{\vec{p}_{gi}^{\,2}}{M_i} + \sum_{i=1}^{N}\overleftrightarrow{I}^{-1}\vec{L}_i^{\,2} + \frac{1}{W_h}\mathrm{Tr}[\overleftrightarrow{p}_h'\overleftrightarrow{p}_h] - (N_f + d^2)kT_{\mathrm{ext}} - p_{\xi_1}\frac{p_{\xi_2}}{Q_2},$$

$$\dot{p}_{\xi_k} = \left(\frac{p_{\xi_{k-1}}^2}{Q_{k-1}} - kT_{\mathrm{ext}}\right) - p_{\xi_k}\frac{p_{\xi_{k+1}}}{Q_{k+1}} \qquad \text{for } k = 2, \ldots, M-1,$$

$$\dot{p}_{\xi_M} = \left(\frac{p_{\xi_{M-1}}^2}{Q_{M-1}} - kT_{\mathrm{ext}}\right) \tag{1}$$

where $\vec{r}_{gi}$ and $\vec{p}_{gi}$ are position and momentum of center of mass of rigid body i, $\overleftrightarrow{h}$ is the cell matrix, $\overleftrightarrow{p}_h$ is the modularly invariant form of the cell momenta, and ξ_k and p_{ξ_k} are the thermostat variable and the conjugated momentum of the kth thermostat of the Nosé-Hoover chain (length = M), respectively. M_i, W_h, and Q_k are the mass of rigid body i, barostat, and kth thermostat, respectively; the

latter two were used to tune the frequency at which those variables fluctuate.[15] The tensor $\overleftrightarrow{A}$ is defined as follows:

$$\overleftrightarrow{A}(\vec{\omega}) = \begin{pmatrix} 0 & -\omega_x & -\omega_y & -\omega_z \\ \omega_x & 0 & \omega_z & -\omega_y \\ \omega_y & -\omega_z & 0 & \omega_x \\ \omega_z & \omega_y & -\omega_x & 0 \end{pmatrix} \tag{2}$$

N_f is the number of degrees of freedom, and N_g is the number of translational degrees of freedom of rigid body. $\vec{F}_{gi}$ and $\vec{\tau}_i$ are the force and torque exerted on the rigid body i, respectively. The tensor $\overleftrightarrow{E}$ denotes the identity matrix. T_{ext} and P_{ext} are the desired pressure and temperature, respectively, and the internal pressure $\overleftrightarrow{P}_{\mathrm{int}}$ is defined as

$$(\overleftrightarrow{P}_{\mathrm{int}})_{\alpha\beta} = \frac{1}{V}\left\{\sum_{i=1}^{N}\frac{(\vec{p}_{gi})_\alpha(\vec{p}_{gi})_\beta}{M_i} + (\vec{F}_{gi})_\alpha(\vec{r}_{gi})_\beta - (\overleftrightarrow{\phi}'\overleftrightarrow{h}^t)_{\alpha\beta}\right\},$$

$$(\overleftrightarrow{\phi}')_{\alpha\beta} = \frac{\partial\phi(\vec{r},\overleftrightarrow{h})}{\partial(h)_{\alpha\beta}} \tag{3}$$

Note that, while the barostat is coupled only to the translational degrees of freedom of a rigid body, the first thermostat is coupled to both the translational and rotational degrees of freedom. The above equations of motions have the conserved quantity

$$H' = \sum_{i=1}^{N}\frac{\vec{p}_{gi}^{\,2}}{2M_i} + \sum_{i=1}^{N}\frac{\overleftrightarrow{I}_i^{-1}\vec{L}^2}{2} + \frac{1}{2W_h}\mathrm{Tr}[\overleftrightarrow{p}_h^t\overleftrightarrow{p}_h] + \sum_{k=1}^{M}\frac{p_{\xi k}^2}{2Q_k} + \phi(\vec{r},\vec{p}) + P_{\mathrm{ext}}|\overleftrightarrow{h}| + (N_f+d^2)kT_{\mathrm{ext}}\xi_1 + \sum_{k=2}^{M}kT_{\mathrm{ext}}\xi_k \tag{4}$$

In this work, the time-evolution operator of the extended system is factorized in the following manner:

$$\exp(iL\Delta t) = \exp\left(iL_{\mathrm{bath}}\frac{\Delta t}{2}\right)\exp\left(iL_2\frac{\Delta t}{2}\right)\exp(iL_1\Delta t)\exp\left(iL_2\frac{\Delta t}{2}\right) \times \exp\left(iL_{\mathrm{bath}}\frac{\Delta t}{2}\right) + O(\Delta t^3), \tag{5}$$

where iL is the Liouville operator, which is decomposed into three components, iL_1, iL_2, and iL_{bath}. Each component of the operator is defined as follows:

$$iL = iL_1 + iL_2 + iL_{\mathrm{bath}}, \tag{6}$$

$$iL_1 = \sum_{i=1}^{N}[\vec{v}_{gi} + \overleftrightarrow{v}_h\vec{r}_{gi}]\cdot\nabla_{\vec{r}_i} + \sum_{\alpha,\beta}(\overleftrightarrow{v}_h)_{\alpha,\beta}\frac{\partial}{\partial(\overleftrightarrow{h})_{\alpha,\beta}} + \sum_{i=1}^{N}\left[\frac{1}{2}\overleftrightarrow{A}(\vec{\omega}_i)\vec{q}_i\right]\cdot\nabla_{\vec{q}_i},$$

$$iL_2 = \sum_{i=1}^{N}\left[\frac{\vec{F}_{gi}}{M_i}\right]\cdot\nabla_{\vec{v}_{gi}} + \sum_{i=1}^{N}\left\{[\vec{\tau}_i]\cdot\nabla_{I_i} - \left[\left(\frac{1}{I_{i,y}}-\frac{1}{I_{i,z}}\right)L_{i,y}L_{i,z}\right]\frac{\partial}{\partial L_{i,x}} - \left[\left(\frac{1}{I_{i,z}}-\frac{1}{I_{i,x}}\right)L_{i,z}L_{i,x}\right]\frac{\partial}{\partial L_{i,y}} - \left[\left(\frac{1}{I_{i,x}}-\frac{1}{I_{i,y}}\right)L_{i,x}L_{i,y}\right]\frac{\partial}{\partial L_{i,z}}\right\},$$

$$iL_{\mathrm{bath}} = \sum_{i=1}^{N}\left[-\left\{\overleftrightarrow{v}_h + \left(\frac{1}{N_g}\right)\mathrm{Tr}[\overleftrightarrow{v}_h] + v_{\xi_1}\right\}\vec{v}_{gi}\right]\cdot\nabla_{\vec{v}_{gi}} + \sum_{\alpha,\beta}\left[\frac{1}{W_h}\left(\sum_{i=1}^{N}M_i(\vec{v}_{gi})_\alpha(\vec{v}_{gi})_\beta + \sum_{i=1}^{N}(\vec{F}_{gi})_\alpha(\vec{r}_{gi})_\beta - (\overleftrightarrow{\phi}'\overleftrightarrow{h}')_{\alpha\beta} + \left(\frac{1}{N_g}\sum_{i=1}^{N}M_i\vec{v}_{gi}^{\,2} - P_{\mathrm{ext}}V\right)\delta_{\alpha\beta}\right) - v_{\xi_1}(\overleftrightarrow{v}_h)_{\alpha\beta}\right]\frac{\partial}{\partial(\overleftrightarrow{v}_h)_{\alpha\beta}} + \sum_{k=1}^{M}v_{\xi k}\frac{\partial}{\partial\xi_k} + \left[\frac{1}{Q_1}\left(\sum_{i=1}^{N}M_i\vec{v}_{gi}^{\,2} + \sum_{i=1}^{N}\overleftrightarrow{I}_i^{-1}\vec{L}_i^2 + W_h\mathrm{Tr}[\overleftrightarrow{v}_h^t\overleftrightarrow{v}_h] - (N_f+d^2)kT_{\mathrm{ext}}\right) - v_{\xi_1}v_{\xi_2}\right] \times\frac{\partial}{\partial v_{\xi_1}} + \sum_{k=2}^{M-1}\left[\frac{1}{Q_k}(Q_{k-1}v_{\xi_{k-1}}^2 - kT_{\mathrm{ext}}) - v_{\xi k}v_{\xi_{k+1}}\right]\frac{\partial}{\partial v_{\xi_k}} + \left[\frac{1}{Q_M}(Q_{M-1}v_{\xi_{M-1}}^2 - kT_{\mathrm{ext}})\right]\frac{\partial}{\partial v_{\xi_M}} \tag{7}$$

where $\vec{v}_{gi} = \vec{p}_{gi}/M_i \neq \dot{\vec{r}}_{gi}$, $\overleftrightarrow{v}_h = \overleftrightarrow{p}_h/W_h$, $\vec{v}_{\xi_i} = \vec{p}_{\xi_i}/Q_i$. Operator iL_1 involves positions and quaternions of rigid bodies and cell matrix, iL_2 involves translational and angular momenta of rigid bodies due to interatomic force, and iL_{bath} involves the rest, including dynamics of barostat, thermostat, and the correction of momenta of rigid bodies due to the dynamics of the bath variables. To solve the angular equations of motions (the Euler equations, included in the operator iL_2) of rigid bodies, we employed a time-reversible integration scheme, where the time evolution of angular momentum is solved as a superposition of a rotation due to torque and one or two free rotations.[13]

Modifying the above operators to MTS form is fairly straightforward. In the extended system, there are typically two factorization schemes with regard to extended variables: XO-RESPA (extended system outside—reference system propagator algorithm) and XI-RESPA (extended system inside—reference system propagator algorithm).[15] Here let us consider the case in which the force exerted on atom i, $\vec{F}_i$, can be decomposed into two different components, say short-range and long-range forces

$$\vec{F}_i = \vec{F}_i^{\text{short}} + \vec{F}_i^{\text{long}} \tag{8}$$

and the operator iL_2 will be divided such that

$$iL_2 = iL_2' + iL_3'$$

$$iL_2' = \sum_{i=1}^{N} \left[\frac{\vec{F}_i^{\text{short}}}{m_i} \right] \cdot \nabla_{\vec{v}_i}$$

$$iL_3' = \sum_{i=1}^{N} \left[\frac{\vec{F}_i^{\text{long}}}{m_i} \right] \cdot \nabla_{\vec{v}_i} \tag{9}$$

For simplicity, the operators relevant to angular momenta are omitted in the notation. Then, a small time step, $\delta t = \Delta t/n (n > 1)$, is used to update velocity of atoms due to the short-range force. In XO-RESPA, the time-evolution operator is factorized as

$$\begin{aligned} \exp(iL\Delta t) = {} & \exp\left(iL_{\text{bath}} \frac{\Delta t}{2}\right) \exp\left(iL_3' \frac{\Delta t}{2}\right) \\ & \times \left[\exp\left(iL_2' \frac{\delta t}{2}\right) \exp(iL_1 \delta t) \exp\left(iL_2' \frac{\delta t}{2}\right) \right]^n \\ & \times \exp\left(iL_3' \frac{\Delta t}{2}\right) \exp\left(iL_{\text{bath}} \frac{\Delta t}{2}\right) + O(\Delta t^3) \end{aligned} \tag{10}$$

On the other hand, in the XI-RESPA, the following factorization is used:

$$\begin{aligned} \exp(iL\Delta t) = {} & \exp\left(iL_{\text{bath}} \frac{\delta t}{2}\right) \exp\left(iL_3' \frac{\Delta t}{2}\right) \exp\left(-iL_{\text{bath}} \frac{\delta t}{2}\right) \\ & \times \left[\exp\left(iL_{\text{bath}} \frac{\delta t}{2}\right) \exp\left(iL_2' \frac{\delta t}{2}\right) \exp(iL_1 \delta t) \right. \\ & \left. \times \exp\left(iL_2' \frac{\delta t}{2}\right) \exp\left(iL_{\text{bath}} \frac{\delta t}{2}\right) \right]^n \exp\left(-iL_{\text{bath}} \frac{\delta t}{2}\right) \\ & \times \exp\left(iL_3' \frac{\Delta t}{2}\right) \exp\left(iL_{\text{bath}} \frac{\delta t}{2}\right) + O(\Delta t^3) \end{aligned} \tag{11}$$

Martyna et al.[15] reported that, while the XO-RESPA algorithm gave good energy conservation for all choices of the number of inner time steps, n, the XI-RESPA algorithm showed variant behaviors: energy conservation improved with increasing n, and then degraded.

As mentioned previously, one of our objectives was to compare the rRB algorithm and the SHAKE/RATTLE/ROLL algorithm in terms of numerical stability, accuracy, and computational efficiency. The latter algorithm extends SHAKE,[12] which is the most widely used method for imposing holonomic constraints upon a system, to the NPT ensemble, where not only constraint force but also constraint virial is taken into account in the iterative calculations during the RESPA integration. It should be noted here that when SHAKE/RATTLE/ROLL is used, only XI-RESPA algorithm may be used for the MTS integration because virial correction due to the constraint must be taken into account for every time step in the SHAKE/RATTLE/ROLL algorithm. The rRB algorithm is free of this restriction, so that both XO- and XI-RESPA may be used to perform the MTS integration.

Test Calculations

To evaluate the numerical stability and accuracy of the algorithms, a series of MD simulations has been undertaken for the following two systems: (1) liquid water, and (2) fully hydrated dipalmitoyl phosphatidylcholine (DPPC) lipid bilayer. The potential model used for lipid molecules was the all-atom CHARMM force field, version 27,[20] in which the modified TIP3P model[21] was used for water molecules. The Lennard-Jones interactions were truncated between 10 and 12 Å with a smooth switching of the potential. Electrostatic interactions were calculated with the Ewald method.[2] Note that, when the rigid-body model is used, the internal virial of rigid bodies is eliminated explicitly in the Ewald summation.[22,23] System 1 consisted of 512 water molecules, and an Andersen-type barostat was used to control the pressure in the cubic cell. The length of the Nosé-Hoover chain was fixed at 5 for all examinations. The pressure was held at 0.1 MPa, and the temperature was maintained at room temperature (298 K). A series of MD in the canonical (NVT) ensemble has also been carried out to clarify the effect of cell dynamics on energy conservation and the properties of the system. In the case of system 2, a series of MD simulations of the system, including 72 DPPC and 2,088 water molecules in a fully flexible cell, was undertaken using a Parrinello-Rahman-type barostat. In this case, the pressure was fixed at 0.1 MPa, and the temperature was set at 323 K to maintain the DPPC lipid bilayer in the liquid crystal phase. Not only water molecules, but also all the local groups including hydrogen atoms in the DPPC molecules, such as —CH_3, —CH_2, and —CH, were treated as rigid bodies. As a result, the DPPC molecule was modeled as a semiflexible molecule. During the test calculations, we checked the effect of the time steps (multiple time steps) on energy conservation and static and dynamic properties of the system. MD simulations and all analyses were carried out on a 10-nodes AlphaStation XP1000 (Alpha 21264 667 MHz) and a 32-nodes Compaq Evo workstation W6000 (Intel Xeon 1.8 GHz-dual CPU) by an originally developed MD package "MPDyn",[24] in which parallel directives were implemented using public domain Message Passing Interface (MPI).

Results

Water System

We performed a 100 ps NPT-MD run from a randomly generated configuration for 512 water molecules to generate the initial configuration for our series of tests. Then, a series of MD calculations

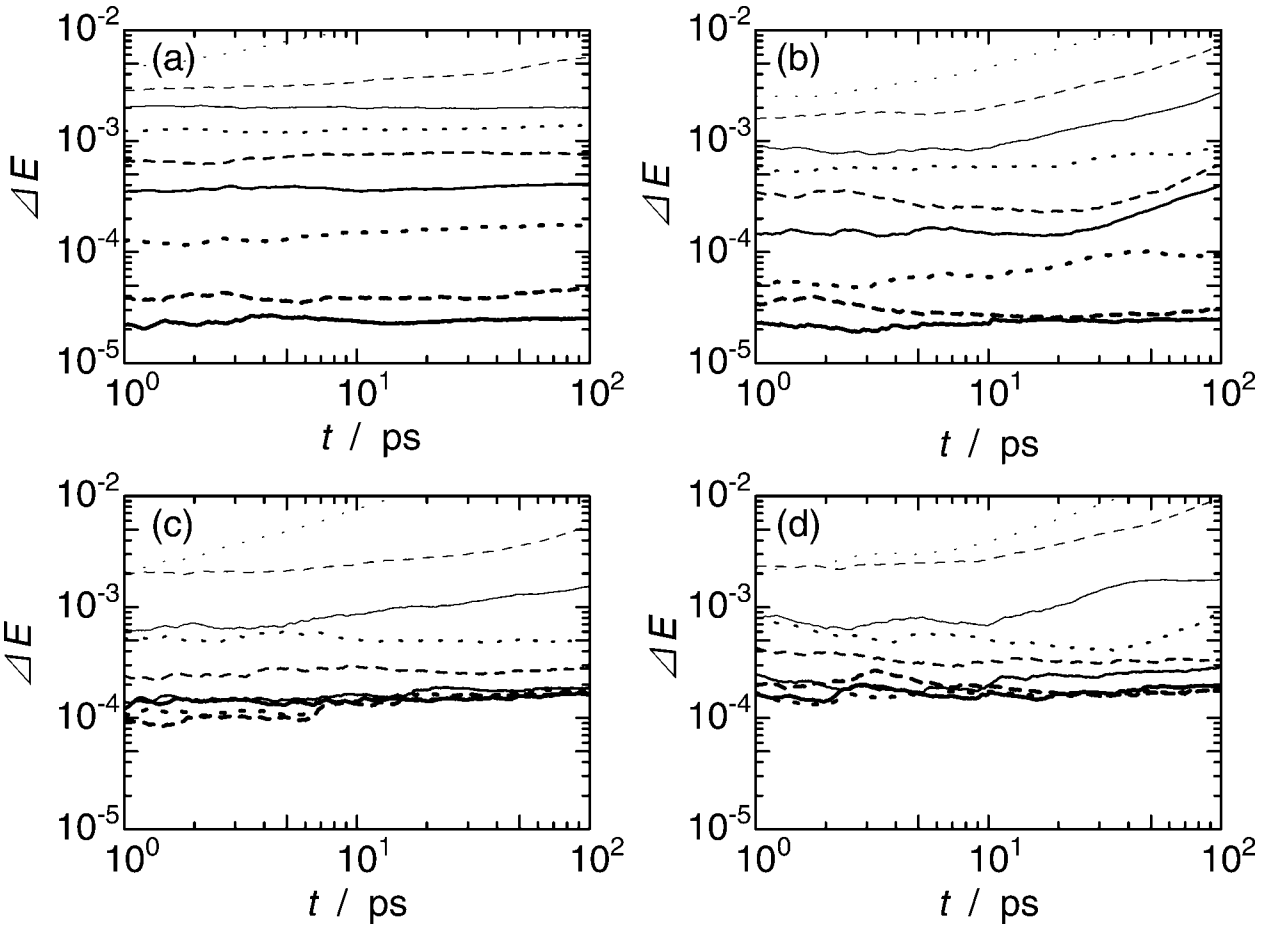

Figure 1. Average drift of the conserved quantity, ΔE. (a) SHAKE/RATTLE constraint method in the NVT ensemble, (b) reversible rigid-body method in the NVT ensemble, (c) SHAKE/RATTLE/ROLL method in the NPT ensemble, and (d) reversible rigid-body method in the NPT ensemble. Thick solid line: $\Delta t = 0.5$ fs; thick dashed line: $\Delta t = 1$ fs; thick dotted line: $\Delta t = 2$ fs; solid line: $\Delta t = 3$ fs; dashed line: $\Delta t = 4$ fs; dotted line: $\Delta t = 5$ fs; thin solid line: $\Delta t = 6$ fs; thin dashed line: $\Delta t = 7$ fs; and thin dotted line: $\Delta t = 8$ fs.

for 100 ps was carried out with a single time step of 0.5, 1, 2, 3, 4, 5, 6, 7, and 8 fs, respectively. To evaluate energy conservation during the MD simulations, we used the average energy deviation, ΔE, defined as

$$\Delta E(t) = \frac{1}{N_{\text{step}}} \sum_{i=1}^{N_{\text{step}}} \left| \frac{H'(i\Delta t) - H'(0)}{H'(0)} \right| \tag{12}$$

where H' was the conserved energy of the system. In Figure 1, the average energy deviations were plotted as a function of time for a series of MDs by the rRB and SHAKE/RATTLE/ROLL algorithms in both the NPT and the NVT ensembles. The ROLL procedure was skipped in the NVT ensemble, because a correction for the constraint virial did not affect the trajectories any more. In the NVT ensemble, energy conservation of rRB and SHAKE/RATTLE were comparable to each other, and the energy conservation systematically degraded with increasing time steps. Both algorithms showed energy conservation with an error less than 0.05% for 100 ps when we used a time step shorter than about 4 fs. The results clearly show the numerical stability of the present integrators constructed by the RESPA algorithm. The same quality of energy conservation could not be achieved by the conventional Gear predictor-corrector algorithm.[13] When a flexible cell was used for the pressure control, motions of the basic cell by which the interatomic distances were scaled may be introduced numerical error (local fluctuation) of the conserved quantity, which could limit the quality of energy conservation of the NPT dynamics. In fact, the energy deviation of only about 0.01% was detected for the NPT-MD, even when the time step as short as 0.5 fs was used. However, except for the inevitable limitation, both rRB and SHAKE/RATTLE(/ROLL) algorithms showed comparable energy conservation for the two ensembles. It should also be noted here that the quality of the energy conservation of the SHAKE/RATTLE/ROLL algorithm is sensitive to the tolerance (relative

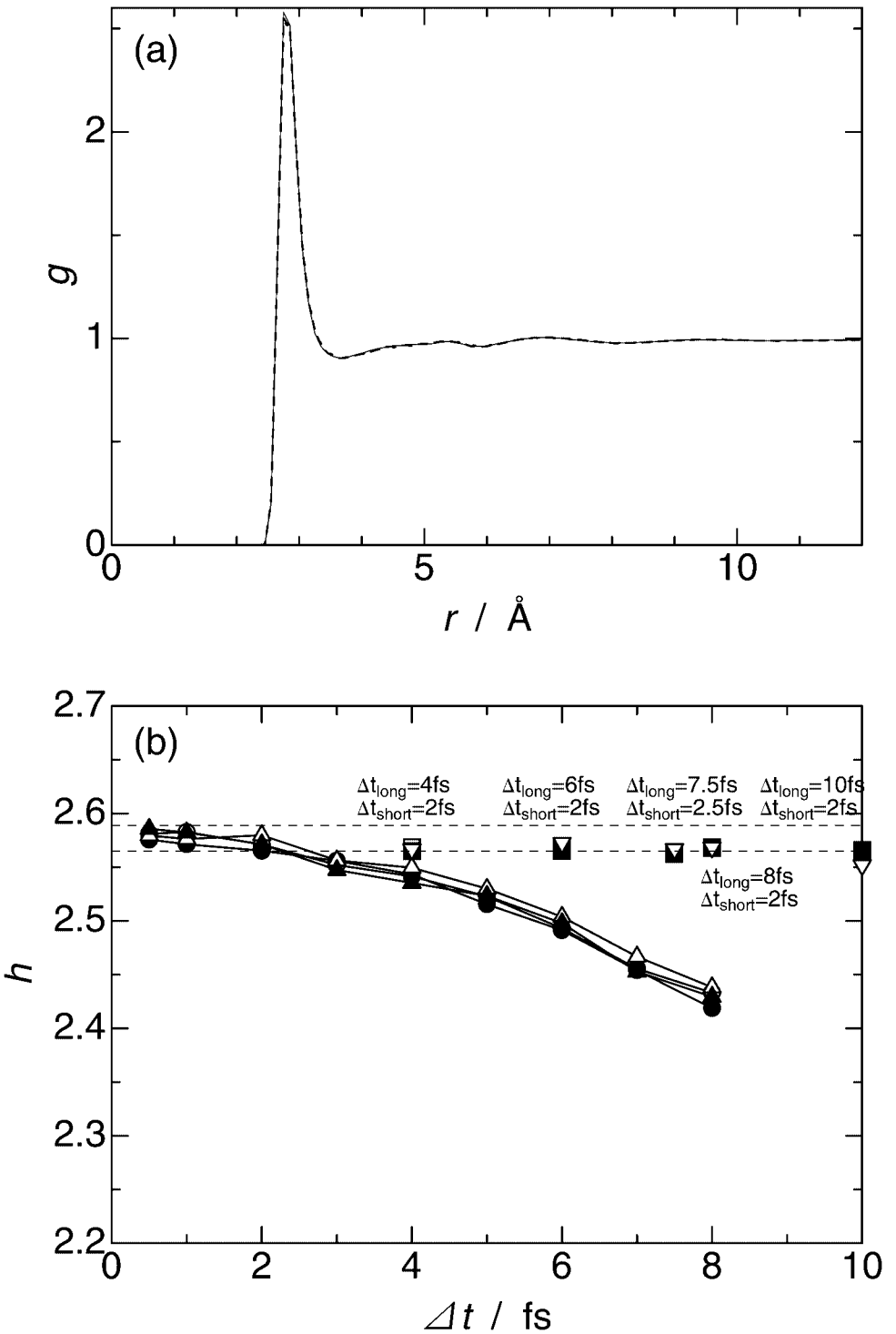

Figure 2. (a) Radial distribution function between oxygen and oxygen atoms in a water molecule. Thin solid line: Δt = 0.5 fs with the reversible rigid-body (rRB) algorithm; dashed line: Δt = 4 fs with the rRB method; and dotted line: Δt = 4 fs with the SHAKE/RATTLE/ROLL method. (b) The height of the first peak of the O—O radial distribution function as a function of the time step length, Δt. Open circle: the rRB method in the NVT ensemble; closed circle: the SHAKE/RATTLE method in the NVT ensemble; open triangle: the rRB method in the NPT ensemble; closed triangle: the SHAKE/RATTLE/ROLL method in the NPT ensemble; open inverted triangle: the rRB method in the NPT ensemble with the XO-RESPA algorithm; open square: the rRB method in the NPT ensemble with the XI-RESPA algorithm; and closed square: the SHAKE/RATTLE/ROLL method in the NPT ensemble with the XI-RESPA algorithm.

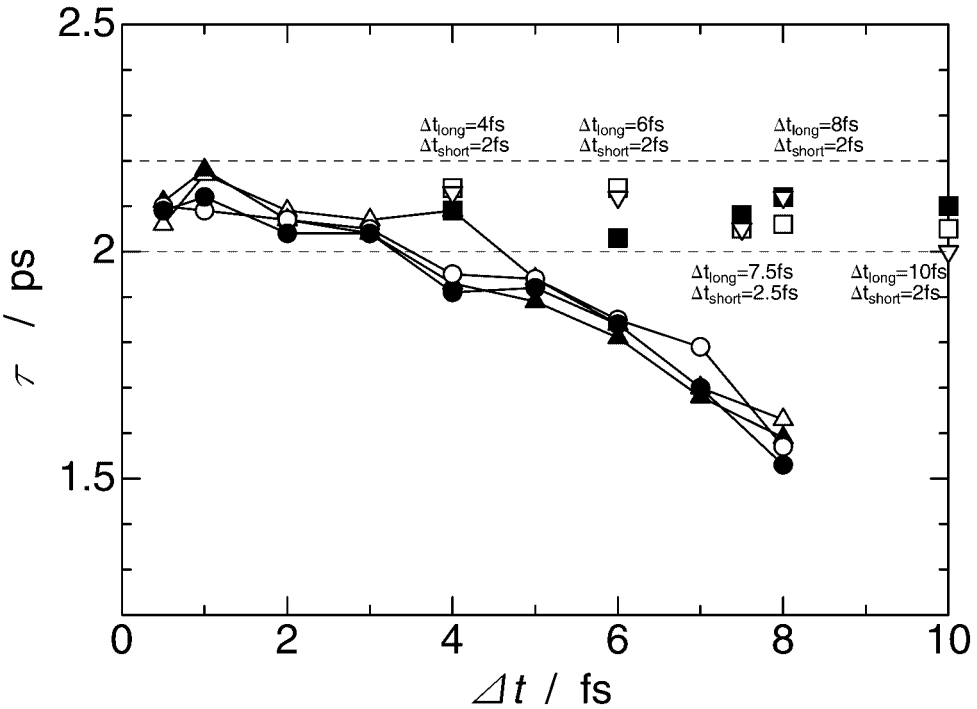

Figure 3. Plots of the rotational relaxation time of the water dipole axis as a function of time steps. Symbols are the same as in Figure 2b.

error) of the fixed bond length. We used a tolerance of 10^{-8} for the NPT ensemble and 10^{-6} for the NVT ensemble. With a tolerance of 10^{-6} in the NPT ensemble, a global energy drift was found. In this case, even when we selected a short time step such as Δt = 0.5 or 1 fs, a global increase of the average energy with time was observed. The computer time of the SHAKE/RATTLE/ROLL algorithm inevitably depends on the required accuracy. In fact, the number of self-consistent iterative loops doubles when the relative tolerance is increased from 10^{-6} to 10^{-8}.

Next, the effect of the time step on the static and dynamic properties of the water system was investigated. The radial distribution functions between water molecules were calculated. In Figure 2a, the radial distribution functions between oxygen and oxygen atoms, g_{oo}, in the NPT ensemble were plotted for three representative conditions: rRB with Δt = 0.5 fs, Δt = 4 fs, and SHAKE/RATTLE/ROLL with Δt = 4 fs. All three lines of g_{oo} plotted here were almost identical, and the errors were almost within the line width. A closer look at the height of the first peak of g_{oo}, however, alerted us to the existence of a weak dependence of g_{oo} on the time step. A glance at Figure 2b reveals that the height of the first peak of g_{oo} became lower with increasing time steps. The trend was also found in the NVT ensemble, and the differences among four plots were within the statistical errors. The agreement between the NPT and NVT ensembles proves that the effect of cell motions on the radial distribution function of the molecular system is negligible.

Next, the rotational relaxation time of the water dipole axis, τ, was investigated to check the quality of MD calculations. The relaxation time was estimated by using the rotational correlation function, defined as

$$C(t) = P1(t) = \langle \cos \theta(t) \rangle = \langle \vec{d}_i(t) \cdot \vec{d}_i(0) \rangle \tag{13}$$

where $\vec{d}_i(t)$ is the dipole axis of ith water molecule at time t, and the brackets denote averaging over all water molecules and time origins. In the plot of $C(t)$ as a function of time, after a fast decay due to the librational motion of the molecule in the cage formed by the surrounding molecules, an exponential decay due to rotational diffusion was observed. Relaxation time was estimated by the numerical integration of the correlation function; a correction was made by fitting the tail region of the correlation function to a single exponential one. Rotational relaxation time was plotted in Figure 3 as a function of time step, Δt, for both the rRB and SHAKE/RATTLE/ROLL algorithms. Here, the results by NVT-MD were also plotted. No significant difference in relaxation time was detected between two algorithms. For all MD calculations at the same time step, the same relaxation time was reproduced within a statistical error. It was also found that the relaxation time was underestimated with a longer time step.

In order to evaluate translational dynamics, the self-diffusion coefficient D was estimated. Self-diffusion coefficients were calculated from the mean square displacement (MSD) of all oxygen atoms by using the Einstein relation:

$$6Dt = \lim_{t \to \infty} \langle |\vec{r}(t) - \vec{r}(0)|^2 \rangle \tag{14}$$

where $\vec{r}(t)$ is the position of the oxygen atom of the water molecule at time t. The diffusion coefficient was estimated from the slope of

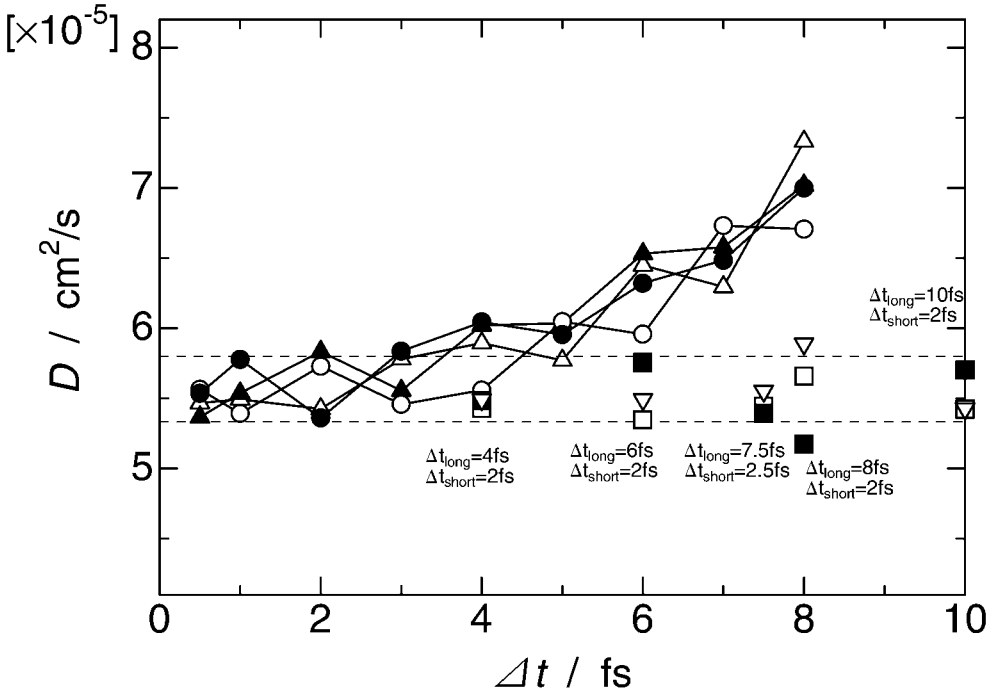

Figure 4. Plots of the diffusion coefficient of a water molecule as a function of time steps. Symbols are the same as in Figure 2b.

the linear part of elongated stretches of the MSD plot. Here we have used the MSD data of the time range of 10 to 20 ps for the estimation. Calculated diffusion coefficients were plotted in Figure 4 as a function of time step. Diffusion coefficients increased with increasing time step. Translational motions were artificially activated by using too large a time step, as in the case of rotational motions. Anyway, it was clearly shown that the two different algorithms (rRB and SHAKE/RATTLE/ROLL) produced the same molecular dynamics at the same time step, and the choice of the NPT ensemble, where the cell matrix was treated as a variable, did not affect the average dynamical properties of molecules.

Because the choice of the time step influenced the system properties significantly, now we shift our focus to the use of the MTS algorithm to achieve accurate and efficient calculations. For the water system, double time steps, Δt_{long} and Δt_{short}, were employed. The former was the time step for a numerical integration of the motion associated with nonbonded interactions in the reciprocal space, while the latter was the time step for a numerical integration of motion associated with nonbonded interactions in the real space. We examined five sets of time steps: (Δt_{long}, Δt_{short}) = (4 fs, 2 fs), (6 fs, 2 fs), (8 fs, 2 fs), (7.5 fs, 2.5 fs), and (10 fs, 2 fs). Although both XO- and XI-RESPA were used for numerical tests of the rRB algorithm, only XI-RESPA was used for the SHAKE/RATTLE/ROLL algorithm. Figure 5 shows the average energy drift, ΔE, for rRB with the XO-RESPA algorithm as a function of time t. The rRB-XO-RESPA showed good energy conservation with a set of long time steps; even when a set of (Δt_{long} = 8 fs, Δt_{short} = 2 fs) was used, no global drift of the conserved quantity was observed. The XI-RESPA algorithm in combination with rRB as well as with SHAKE/RATTLE/ROLL showed almost the same quality of energy conservation as rRB-XO-RESPA did. The degradation of energy conservation of XI-RESPA with a large number of inner time steps, n, reported by Martyna et al.,[15] was not clearly observed in the range of time steps used here.

We have also analyzed the structure and dynamics of the water molecules for all cases of MTS integration (SHAKE/RATTLE/ROLL-XI-RESPA, rRB-XI-RESPA, and rRB-XO-RESPA), and the results are shown in Figures 2–4. At first, as shown in Figure 2b, though the height of the first peak of radial distribution function, g_{oo}, decreased slightly with increasing time step, the decrease was almost negligible, and the deviation was comparable to the statistical error even with the largest time steps (Δt_{long} = 10 fs, Δt_{short} = 2 fs) used here. We could not find any significant difference in the values calculated by the three MTS algorithms used here. The same statements are true for the rotational (Fig. 3) and translational dynamics (Fig. 4) of water molecules; when we used MTS integration, the results were not significantly influenced by any set of time steps used here and were comparable to those obtained by a single time step shorter than 2 fs. Thus, all combinations of MTS and the holonomic constraint algorithms used here improved the computational efficiency without violating molecular properties.

DPPC Lipid Bilayer System

The SHAKE/RATTLE(/ROLL) algorithm has been extensively used for a variety of macromolecular systems to apply holonomic constraints to local bonds. Similarly, we can make use of the rRB algorithm to introduce local rigid bodies in a macromolecule.[25,26] We have applied the rigid body approximation for all —CH_3,

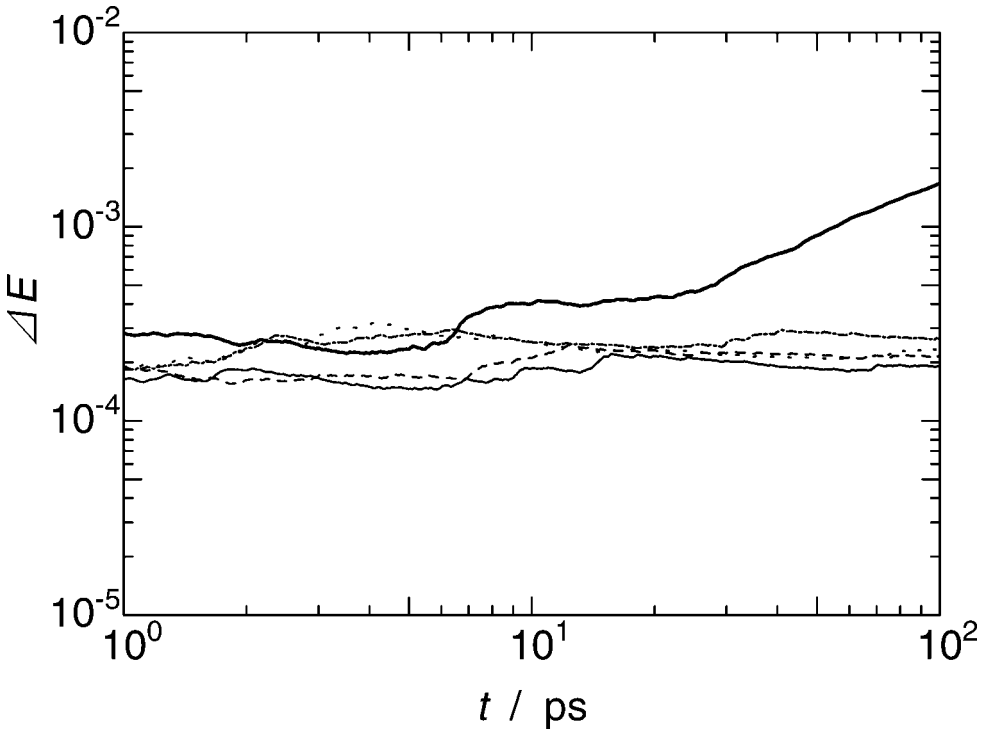

Figure 5. Average drift of the conserved quantity, ΔE, in the case of reversible rigid-body XO-RESPA algorithm in the NPT ensemble. Solid line: $\Delta t_{\mathrm{long}} = 4$ fs and $\Delta t_{\mathrm{short}} = 2$ fs; dashed line: $\Delta t_{\mathrm{long}} = 6$ fs and $\Delta t_{\mathrm{short}} = 2$ fs; dotted line: $\Delta t_{\mathrm{long}} = 8$ fs and $\Delta t_{\mathrm{short}} = 2$ fs; dash-dotted line: $\Delta t_{\mathrm{long}} = 7.5$ fs and $\Delta t_{\mathrm{short}} = 2.5$ fs; and thick solid line: $\Delta t_{\mathrm{long}} = 10$ fs and $\Delta t_{\mathrm{short}} = 2$ fs.

—CH_2, and —CH groups in the DPPC molecule. Special treatment is needed for the linear rigid body, —CH, where the rotational degree of freedom is 2, and one of the three inertial moments is 0. When we take molecular coordinates of the linear rigid body on the x-axis, the x component of the inertial moment becomes 0 or $I_x = 0$, and the Euler equations of motion are simplified, and then the quaternion representation is no longer needed. Although a special treatment for linear rigid bodies, such as an algorithm based on leap-frog integration,[2] may be desirable to minimize the computational load, we treated —CH in a quaternion framework with $L_x = 0$. On the other hand, for a rigid body composed of two interaction sites like —CH, we can make use of SHAKE algorithm without losing time reversibility because the constraint force is evaluated analytically.[27] However, the same argument does not hold for a linear rigid body composed of more than three bodies.

The average deviations of the conserved quantity were calculated for the DPPC lipid bilayer system in the NPT ensemble and plotted as a function of time in Figure 6. Both rRB (Fig. 6a) and SHAKE/RATTLE/ROLL (Fig. 6b) showed good energy conservation with a short time step, and energy conservation systematically improved with decreasing time steps. When the time step was less than 2.5 fs, no accumulative energy drift was found in the time range from 1 to 100 ps. Because strong intramolecular forces due to bond stretching, angle bending, and torsion interactions were exerted on the rigid bodies in the macromolecular system, the magnitude of the energy drift was larger than that of a molecular rigid-body (like water) system at the same time step (e.g., $\Delta t = 4$ fs). Figure 6c shows the time course of the average energy drift, when the rRB-XO-RESPA algorithm was used. In this case, we used three time steps: Δt_{long} was used to integrate the dynamics advanced by nonbonded interactions in the reciprocal space, $\Delta t_{\mathrm{middle}}$ was used to integrate the dynamics advanced by nonbonded interactions in the real space, and $\Delta t_{\mathrm{short}}$ was used to integrate the dynamics advanced by intramolecular interactions (bond stretching, angle bending, and torsion). As long as $\Delta t_{\mathrm{short}}$ was selected to be less than $2 \sim 2.5$ fs, the relation between the energy conservation and the time steps was similar to that of the water system. The intramolecular interaction can be calculated at order-N scaling in computer time, which is typically much shorter than that needed for the calculation of nonbonded interactions. Therefore, using the MTS integration, we can perform an MD simulation for macromolecular systems as accurately and efficiently as the simulation for a water system.

Discussion

During an MD simulation, the average drift of the conserved quantity is usually used to ensure that the equations of motions are solved correctly. In this study, when the rRB algorithm was employed in the NPT ensemble, the quality of energy conservation was improved with time steps decreasing to the limit, which may

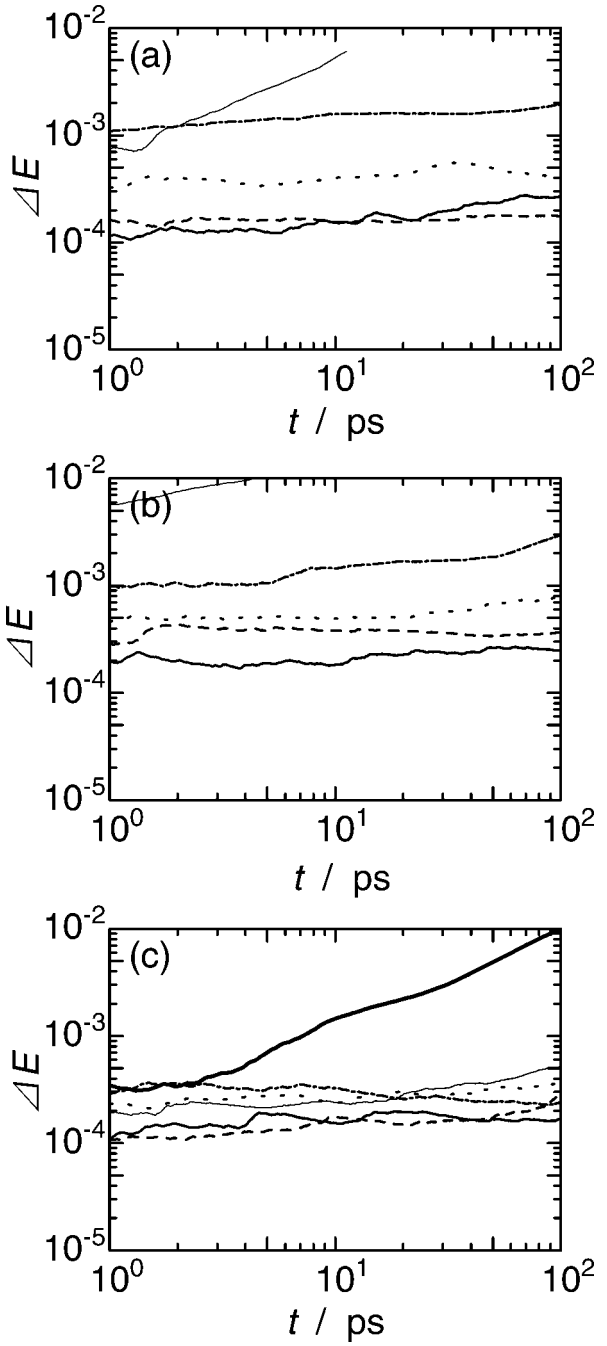

Figure 6. Average drift of the conserved quantity, ΔE, of the DPPC lipid bilayer system. (a) Single time step results with the SHAKE/RATTLE/ROLL algorithm, (b) single time step results with the rRB method, and (c) multiple time step results with the rRB-XO-RESPA algorithm. For single time steps, Δt = 1 fs (solid lines), 2 fs (dashed lines), 2.5 fs [dotted line (only for the rRB method)], 3 fs (dash-dotted lines), and 4 fs (thin solid lines). For multiple time steps, $\{\Delta t_{\text{long}}, \Delta t_{\text{middle}}, \Delta t_{\text{short}}\}$ = {4 fs, 2 fs, 1 fs} (solid line), {8 fs, 2 fs, 1 fs} (dashed line), {6 fs, 2 fs, 2 fs} (dotted line), {8 fs, 2 fs, 2 fs} (dot-dashed line), {7.5 fs, 2.5 fs, 2.5 fs} (thin solid line), and {8 fs, 4 fs, 2 fs} (thick solid line).

Table 1. Total Execution Times of 1ps-NPT-MD of the Water System.

	rRB		SHAKE	
Δt [fs]	16 CPUs	64 CPUs	16 CPUs	64 CPUs
0.5	262	139	325	196
1	131	70	154	101
2	66	35	85	52
3	45	24	58	36
4	34	18	40	27
5	27	15	35	22
6	23	12	30	19
7	20	10	26	16
8	18	9	23	14
(4, 2)	47	28	64	46
(6, 2)	41	25	56	44
(8, 2)	37	24	54	42
(10, 2)	35	23	52	42
(7.5, 2.5)	33	20	45	35

Total execution times are in seconds.

be bounded by the inevitable numerical error (local fluctuation), $\Delta E \sim 10^{-4}$, due to the coupling between the cell and particle motions. Apart from the limitation, the energy conservation in the NPT ensemble was comparable to that in the NVT ensemble. A further important point was that no accumulated error of the average energy drift was found during 1 to 100 ps. In addition, the simple relation between energy conservation and time step means that the accuracy of the MD trajectories can be evaluated by the degree of the energy drift, from which we can select the time step for efficient and precise simulation. The same thing will be true for the SHAKE/RATTLE/ROLL, if and only if the accurate tolerance for the fixed bond distance is chosen. The greater accuracy of a lower tolerance, say 10^{-8}, in turn demanding a higher computational load, was necessary to produce trajectories with good energy conservation in the NPT ensemble. The rRB algorithm is free from this problem. Another merit of rRB is that the XO-RESPA algorithm can be employed to construct the MTS integrator. The XO-RESPA algorithm is superior to the XI-RESPA with respect to computational efficiency. However, the MTS integration of SHAKE/RATTLE/ROLL, in which the virial correction due to the constraint force must be taken into account during the self-consistent iterative loop, must be carried out by XI-RESPA. In the case of MD simulations on parallel processing systems, the estimation of the virial correction demands a global communication among processors, which must be situated in the inner loop of the MTS integration. Thus, the rRB-XO-RESPA algorithm, which is free from the iterative constraint procedure, becomes advantageous by increasing the number of iterations, the number of inner time steps n, and the number of processors to be used. The elapsed time for the 1 ps-MD simulations listed in Table 1 (water) and 2 (DPPC bilayer) clearly showed this tendency. In addition, we can expect that the rRB algorithm is more advantageous when the rigid-body approximation is applied to a larger group of atoms. When, for example, rigid-body approximation is applied to a benzene group, which is one of the components frequently found in polymers and biological macromolecules, the rRB algorithm will be much more

Table 2. Total Execution Times of 1ps-NPT-MD of the DPPC Bilayer System.

	rRB		SHAKE	
Δt [fs]	16 CPUs	64 CPUs	16 CPUs	64 CPUs
1	2053	910	2221	1180
2	1065	465	1159	609
3	742	315	807	424
4	578	242	631	327
(4, 2, 1)	875	483	1159	801
(8, 2, 1)	734	444	1031	763
(8, 2, 2)	657	322	803	492
(7.5, 2.5, 2.5)	581	274	696	412
(8, 4, 2)	472	251	626	424

Total execution times are in seconds.

efficient than SHAKE/RATTLE/ROLL, because the latter needs a large number of iterative calculations to achieve an accurate integration for large rigid bodies.

The rRB algorithm, because of its reversibility in time, can be employed in the hybrid Monte Carlo,[13] whereas SHAKE/RATTLE/ROLL cannot. In addition, the former has a convenient form for application to some of the novel molecular simulation techniques that have been developed to enhance sampling efficiency. For example, the rRB algorithm can be easily implemented to self-guided molecular dynamics (SGMD), which was shown to effectively improve the conformational sampling efficiency when it was applied for the protein folding[28] as well as the crystallization of the bulk LJ fluid.[11] It is expected that the SGMD in conjunction with the rRB methods will provide an effective conformational sampling technique, where not only the guiding force but also the guiding torque is taken into account. Fine tuning of the parameters, such as bias parameters and local sampling time, for translational and rotational motions, respectively, will make the method even more efficient, as there is usually a gap of correlation times between these motions.

Conclusion

In order to produce long term trajectories of macromolecules with good accuracy, an efficient MD calculation without accumulative global error must be undertaken. By applying reasonable holonomic constraints to the system, we achieved an efficient dynamical description of molecular and macromolecular systems. The rRB algorithm was shown to be accurate with respect to energy conservation in the NPT ensemble, and more advantageous than SHAKE/RATTLE/ROLL in terms of computational efficiency. For a multiple-time-step integration, the rRB model can be solved by the XO-RESPA algorithm, which demands less computational time than XI-RESPA, while SHAKE/RATTLE/ROLL cannot. The validity of the rRB algorithm in the NPT ensemble was confirmed by a series of MD calculations of water and DPPC lipid bilayer systems. Finally, the rRB method has a form suitable for application to a group of novel molecular simulation techniques such as hybrid MC and self-guided MD.

Acknowledgments

This work is partly supported by NEDO under the Nanotechnology Materials Program.

References

1. Schlick, T. Molecular Modeling and Simulation; Springer-Verlag: New York, 2002.
2. Allen, M. P.; Tildesley, D. J. Computer Simulation of Liquids; Clarendon Press: Oxford, UK, 1987.
3. Frenkel, D.; Smit, B. Understanding Molecular Simulation, 2nd ed.; Academic Press: San Diego, 2002.
4. Tuckerman, M.; Martyne, G. J.; Berne, B. J. J Chem Phys 1992, 97, 1990.
5. (a) Daoud, M.; Williams, C. E. Eds. Soft Matter Physics; Springer-Verlag: Berlin, 1999; (b) Jönsson, B.; Lindman, B.; Holmberg, K.; Kronberg, B. Surfactants and Polymers in Aqueous Solution; John Wiley & Sons: Chichester, 1998; p 61.
6. Shinoda, W.; Namiki, N.; Okazaki, S. J Chem Phys 1997, 106, 5731.
7. Tu, K.; Tobias, D. J.; Klein, M. L. Biophys J 1995, 69, 2558.
8. Andersen, H. C. J Chem Phys 1980, 72, 2384.
9. Parrinello, M.; Rahman, A. Phys Rev Lett 1980, 45, 1196.
10. Okabe, T.; Kawata, M.; Okamoto, Y.; Mikami, M. Chem Phys Lett 2001, 335, 435.
11. Shinoda, W.; Mikami, M. Chem Phys Lett 2001, 335, 265.
12. Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. J Comput Phys 1977, 23, 327.
13. Matubayasi, N.; Nakahara, M. J Chem Phys 1999, 110, 3291.
14. Martyna, G. J.; Tobias, D. J.; Klein, M. L. J Chem Phys 1994, 101, 4177.
15. Martyna, G. J.; Tuckerman, M. E.; Tobias, D. J.; Klein, M. L. Mol Phys 1996, 87, 1117.
16. Tuckerman, M. E.; Mundy, C. J.; Martyna, G. J. Europhys Lett 1999, 45, 149.
17. Andersen, H. C. J Comput Phys 1983, 52, 24.
18. Tuckerman, M. E.; Martyna, G. J. J Phys Chem B 2000, 104, 159.
19. Martyna, G. J.; Tuckerman, M. E.; Klein, M. L. J Chem Phys 1992, 97, 2638.
20. (a) MacKerell Jr., A. D.; Bashford, D.; Bellott, M.; Dunbrack Jr, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher III, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. J Phys Chem B 1998, 102, 3586; (b) MacKerell Jr., A. D.; http://www.pharmacy.umaryland.edu/~alex/research.html
21. Jorgensen, W. L.; Chandrashekar, J.; Madura, J. D.; Impey, R.; Klein, M. L. J Chem Phys 1983, 79, 926.
22. Nosé, S.; Klein, M. L. Mol Phys 1983, 50, 1055.
23. Fang, Z.; Haymet, A. D. J.; Shinoda, W.; Okazaki, S. Comput Phys Commun 1999, 116, 295.
24. Shinoda, W. MPDyn: A software package for molecular simulations, manuscript in preparation.
25. Wu, X.-W.; Sung, S.-S. J Comput Chem 1998, 19, 1555.
26. Forester, T. R.; Smith, W. J Comput Chem 1998, 19, 102.
27. Miller III, T. F.; Eleftheriou, M.; Pattnaik, P.; Ndirango, A.; Newns, D.; Martyna, G. J. J Chem Phys 2002, 116, 8649.
28. Wu, X.-W.; Wang, S. J Phys Chem B 1998, 102, 7238.