JCTC Journal of Chemical Theory and Computation

Superposition State Molecular Dynamics

Arun Venkatnathan and Gregory A. Voth*

Department of Chemistry and Center for Biophysical Modeling and Simulation, University of Utah, 315 South 1400 East Room 2020, Salt Lake City, Utah 84112-0850

Received October 14, 2004

Abstract: The ergodic sampling of rough energy landscapes is crucial for understanding phenomena like protein folding, peptide aggregation, polymer dynamics, and the glass transition. These rough energy landscapes are characterized by the presence of many local minima separated by high energy barriers, where Molecular Dynamics (MD) fails to satisfy ergodicity. To enhance ergodic behavior, we have developed the Superposition State Molecular Dynamics (SSMD) method, which uses a superposition of energy states to obtain an effective potential for the MD simulation. In turn, the dynamics on this effective potential can be used to sample the configurational free energy of the real potential. The effectiveness of the SSMD method for a one-dimensional rough potential energy landscape is presented as a test case.

1. Introduction

The Molecular Dynamics (MD) simulation method is an extremely important tool in chemistry, materials science, and biology. A number of MD methods developed so far have effectively simulated processes such as protein folding, peptide aggregation, and surface deposition. Needless to say, for a large set of molecules (e.g, large proteins), there are many competing interactions (intermolecular and intramolecular), leading to an extremely rough energy landscape and thus making the process of MD simulation a challenging task. The difficulty in sampling these rough energy landscapes is severe due to the presence of many local minima separated by high energy barriers. At normal temperatures, conventional MD or Monte Carlo (MC) simulations will be largely trapped in one of the various local minima. In turn, only some parts of the entire phase space will be sampled, and consequently structural, dynamical, and thermodynamic properties of the system cannot be reliabily calculated.

A number of MC algorithms have been developed to overcome the ergodicity problems posed by rough energy landscapes, such as multicanonical MC sampling,[1] parallel tempering,[2] multicanonical parallel tempering,[3] and simulated tempering.[4] Multicanonical sampling replaces the Boltzmann weight by a non-Boltzmann weight to effectively obtain a flat energy distribution between neighboring local minima. Parallel tempering improves sampling by using noninteracting configuration (or replicas) with different temperatures. This method can be further enhanced through a parallel implementation by distributing the noninteracting configurations across computational nodes. Furthermore, a multicanonical parallel tempering method, combining the advantages of parallel tempering and multicanonical sampling, has been developed by Faller et al.[3]

In the simulated tempering method, a random walk in temperature space is used to sample the free energy space by escaping the local energy minima. This method has been applied to study protein folding.[5] Stolovitzky and Berne[6] have also developed the catalytic tempering method to reduce the free energy barriers and to accelerate the sampling in phase space of complex systems without disturbing the actual potential minima. The basin-hopping MC algorithm of Wales et al.[7] also reduces the free energy barriers as seen in their implementation to perform global optimization on various atomic clusters. Here, the actual potential energy surface has been modified by a staircase-like effective potential without disturbing the positions of any minima.

The methods discussed in the previous paragraph are primarily based on MC. On the other hand, the Replica Exchange Molecular Dynamics (REMD) method of Sugita and Okamoto[8] is similar to the original parallel tempering approach but instead uses MD. Voter and co-workers[9] have also developed a class of methods called accelerated MD methods, which includes hyperdynamics, temperature accelerated dynamics, and parallel replica dynamics (PRD). In the hyperdynamics method of Voter,[10] the acceleration of the MD simulation requires a computation of the gradients

10.1021/ct0499175 CCC: $30.25
Published on Web 12/08/2004

and Hessian of the potential. The PRD method can offer a large boost in simulation time, and its success has been seen, e.g., in a study of Cu(100) surface vacancy diffusion.[11] The accelerated MD method of Miron and Fichthorn[12] is similiar to the work of Voter where the potential energy close to the local minima is modified by a boost potential to obtain an effective potential. The success of this method has been shown for a surface diffusion problem as well. Unfortunately, the choice of the boost potential can lead to certain instablities in the numerical derivatives, giving rise to artificial energy peaks near the local minima. The conformational flooding method of Grubmüller[13] has been employed to predict structural transitions in irregular or disordered systems where conventional MD fails. This method uses a "fictitious" potential derived by defining a conformational substate which samples all regions pertaining to low free energy thereby reducing the free energy barriers in the original energy landscape. Also, the metadynamics approach[14,15] has been developed to sample rough energy landscapes by using a time-dependent Gaussian distribution to fill in the minima of the original potential as they are visited, thus biasing the dynamics to explore the more inaccessible regions of the energy landscape. The puddle-skimming method[16] and accelerated MD method of Hamelberg et al.[17] have also explored mechanisms to obtain an effective potential energy surface by using an appropriate choice of bias potential and boost energy. The puddle-skimming method[16] does not perform well for high values of boost energy and also has certain discontinuities in the computation of derivatives of the real potential along the potential energy surface. This flaw has been removed by Hamelberg et al.[17] through a better choice of the bias potential. Although, both of these methods seem to be promising, they involve a single choice of an effective potential surface.

The purpose of this paper is to introduce a simple and flexible MD method to help sample rough energy landscapes. Here, we will study the canonical ensemble generated by the Nose-Hoover Chain (NHC)[18] thermostat to sample free energy profiles, although any MD ensemble should be possible. The NHC has proven to be a reliable thermostat for many systems. However, as might be expected, the NHC does not sample ergodically in cases where there are rough energy landscapes.

The new method presented here is called Superposition State Molecular Dynamics (SSMD), which uses an effective superposition potential to help sample the entire phase space. The motivating concept behind this method is the way in which nonstationary states in quantum mechanics can explore the Hilbert space by undergoing transitions between basis states. The SSMD method also has some similiarities to the hyperdynamics approach used by Voter[10] to overcome the high energy barriers. The recent work using the puddle-skimming method[16] and the accelerated MD method of Hamelberg et al.[17] can also be largely recovered as one possible limit of the SSMD approach. However, in SSMD one performs a MD simulation on the effective potential obtained by using a superposition of energy states, thus allowing for a very large number of possible choices to aid in the energy landscape sampling.

To demonstrate the effectiveness of the SSMD method, we have chosen a one-dimensional potential formed from a distribution of Gaussian barriers in which standard canonical MD with a NHC fails to sample ergodically.

2. Method

Let us denote $V(x)$ as the real potential on which MD sampling fails to be ergodic due to the presence of a rough energy landscape. In SSMD, the MD simulation is carried out on an effective potential using a superposition of "states". The SSMD effective potential is constructed by using a combination of the real energy state $V(x)$ and a set of N "fictitious" states. To illustrate this N-state superposition concept, the simplest 2-state implementation is shown here, where $V(x)$ and $V_0(x)$ are the potential energy functions in the first and second state, respectively. It should be noted that $V_0(x)$ is *any fictitious potential* which can be freely chosen to enhance the MD sampling. The effective potential for the MD forces is then obtained by solving the 2×2 matrix eigenvalue problem, where the highest and lowest eigen-energies are the upper 2-state $[V_+(x)]$ and lower 2-state $[V_-(x)]$ surface on which the MD sampling can then be performed. It should be noted that one has the flexibility to run the MD on either $V_+(x)$ or $V_-(x)$, depending on the choice of $V_0(x)$ and/or the physical problem at hand.

Mathematically, the effective 2-state SSMD potential can be written as

$$V_{\pm}(x) = \frac{V(x) + V_0(x)}{2} \pm \frac{1}{2}\sqrt{4V_{12}^2(x) + (V(x) - V_0(x))^2} \quad (1)$$

where one chooses either $V_+(x)$ or $V_-(x)$ for the dynamics as discussed earlier.

The governing sampling equations for SSMD simulation are as follows: Consider, the exact canonical distribution function, given by

$$P(x) = \frac{e^{-\beta V(x)}}{Q} \quad (2)$$

where β is equal to $1/k_BT$. Here, Q is the partition function related to $V(x)$ which can be written as

$$Q = Q_{\pm}\langle e^{-\beta \Delta V(x)}\rangle_{\pm} \quad (3)$$

where $Q_{\pm}$ is the partition function corresponding to the superposition potential $V_{\pm}(x)$ and $\Delta V(x) = V(x) - V_{\pm}(x)$. Using eqs 2 and 3 and the definition of $\Delta V(x)$, we can write the exact probability distribution function as

$$P(x) = \frac{e^{-\beta V_{\pm}(x)} e^{-\beta \Delta V(x)}}{Q_{\pm}\langle e^{-\beta \Delta V(x)}\rangle_{\pm}} \quad (4)$$

or

$$P(x) = P_{\pm}(x)\frac{e^{-\beta \Delta V(x)}}{\langle e^{-\beta \Delta V(x)}\rangle_{\pm}} \quad (5)$$

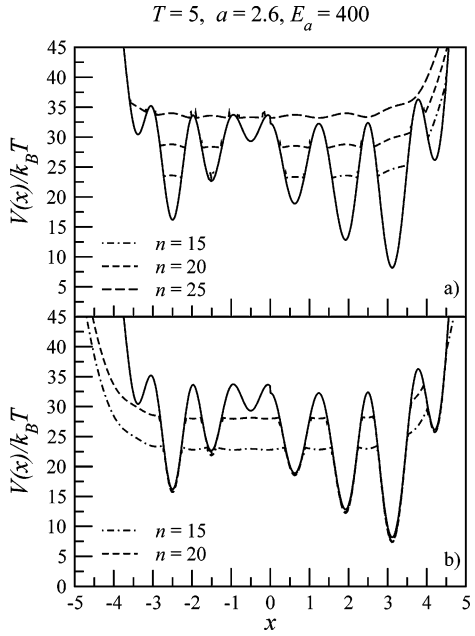

Figure 1. Real potential energies (solid lines) and effective SSMD potentials (dashed and dot−dashed lines) from an (a) upper 2-state SSMD surface and (b) lower 2-state SSMD surface.

where

$$P_{\pm}(x) = \frac{e^{-\beta V_{\pm}(x)}}{Q_{\pm}} \tag{6}$$

It should be noted that for the purposes of testing the SSMD sampling, the above equations can be rearranged to give

$$\frac{V(x)}{k_B T} = -\ln P_{\pm}(x) + \ln\langle e^{-\Delta V(x)/k_B T}\rangle_{\pm} - \ln Q + \frac{\Delta V(x)}{k_B T} \tag{7}$$

where the first two terms in the right-hand side of eq 7 are obtained from the SSMD simulation and the last two terms are easily calculated quantities for the simple potential used in this paper to test the SSMD method.

The choice of $V_0(x)$ is intended to enhance the SSMD sampling and also depends on the selection of effective potential surface, i.e., whether the SSMD is run on the upper surface, $V_+(x)$, or the lower surface, $V_-(x)$. If the MD sampling is performed on the upper 2-state surface, one simple choice is $V_0(x) = V_{\min} + nk_BT$, where $V_{\min}$ is an estimate of global minima of the real potential, while nk_BT corresponds to some user-defined height above V_{min}. For various choices of n, the different upper 2-state surfaces, $V_+(x)$, are shown in Figure 1a. As a result, when the real potential is lower than $V_0(x)$, the SSMD sampling is naturally performed on a "filled" potential energy surface with shallow wells. This approach is very similiar to that of the puddle-skimming method[16] and accelerated MD method of Hamelberg et al.,[17] although it is in principle much more flexible in the choice of $V_0(x)$.

Alternatively, the lower 2-state surface can be chosen as the SSMD potential, in which case a simple choice for $V_0(x)$ is for it to be equated to $V_{min} + nk_BT + V_{conf}$ to avoid being trapped in one of the local minima. Here V_{conf} is a confining potential to stop the trajectory from wandering away from the region of interest. For various values of n, the different lower 2-state surfaces obtained using the above criteria are shown in Figure 1b.

In both cases, the off-diagonal coupling elements (V_{12}) can be chosen as follows: When the difference between $V(x)$ and $V_0(x)$ is less than 1/2 k_BT, we chose V_{12} to be equal to k_BT. If the difference was more than 1/2 k_BT, then V_{12} was set as $d(V(x) - V_0(x))$, where $d = 1/k_BT$, though other choices of V_{12} are also possible.

The NHC equations of motion for a canonical ensemble are described in detail elsewhere.[18] However, for the sake of completeness, the NHC equations are written as follows:

$$\dot{r}_i = \frac{p_i}{m_i};\ \dot{p}_i = -\frac{\partial V(r)}{\partial r_i} - p_i\frac{p_{\eta_1}}{Q_1};\ \dot{\eta}_i = \frac{p_{\eta_i}}{Q_i}$$

$$\dot{p}_{\eta_1} = \left[\sum_{i=1}^{N_p}\frac{p_i^2}{m_i} - N_p k_B T\right] - p_{\eta_1}\frac{p_{\eta_2}}{Q_2}$$

$$\dot{p}_{\eta_j} = \left[\frac{p_{\eta_{j-1}}^2}{Q_{j-1}} - k_B T\right] - p_{\eta_j}\frac{p_{\eta_{j+1}}}{Q_{j+1}}$$

$$\dot{p}_{\eta_M} = \left[\frac{p_{\eta_{M-1}}^2}{Q_{M-1}} - k_B T\right] \tag{8}$$

where N_p is the number of particles, M is the number of thermostats, r's and p's are the particle positions and momenta, respectively, T is the kinetic temperature, and k_B is the Boltzmann constant. In eq 8, Q_i and m_i are masses of the ith thermostat and particle, respectively, while η and p_η are the thermostat positions and momenta, respectively. The conserved energy quantity corresponding to NHC dynamics is expressed as

$$H'(r, p, \eta, p_\eta) = V(r) + \sum_{i=1}^{N_p}\frac{p_i^2}{2m_i} + \sum_{i=1}^{M}\frac{p_{\eta_i}^2}{2Q_i} + N_p k_B T\eta_1 + \sum_{i=1}^{M} k_B T\eta_i \tag{9}$$

where $V(r)$ is the total potential energy of all particles in the system.

The Liouville operator for the equations of motion in eq 8 is given by

$$iL = \sum_{i=1}^{N_p} v_i\cdot\nabla_{r_i} + \sum_{i=1}^{N_p}\left[\frac{F_i(r)}{m_i}\right]\cdot\nabla_{v_i} - \sum_{i=1}^{N_p} v_{\eta_1} v_i\cdot\nabla_{v_i} + \sum_{i=1}^{M} v_{\eta_i}\frac{\partial}{\partial\eta_i} + \sum_{i=1}^{M-1}(\dot{p}_{\eta_i} - v_{\eta_i} v_{\eta_{i+1}})\frac{\partial}{\partial v_{\eta_i}} + \dot{p}_{\eta_M}\frac{\partial}{\partial v_{\eta_M}} \tag{10}$$

where v_i is the velocity of the ith particle and v_{η_M} is the velocity associated with the Mth thermostat. The first two terms on the right-hand side of eq 10 denotes a shift in the particle positions and velocities and is performed by using

Table 1. Parameters Used To Compute the Analytic Potential in Eq 11

i	C_i	σ_i	μ_i ($x \le 0$)	μ_i($x > 0$)
1	0.8	0.4	−9.0	11.25
2	0.8	0.4	−8.0	10.00
3	0.8	0.3	−7.0	8.75
4	0.6	0.3	−6.0	7.50
5	0.6	0.4	−5.0	6.25
6	0.6	0.4	−4.0	5.00
7	0.4	0.3	−3.0	3.75
8	0.4	0.3	−2.0	2.50
9	0.4	0.4	−1.0	1.25
10	0.4	0.4	0.0	0.00

the velocity verlet algorithm.[19] The multistep time propagation algorithm of Martyna et al.[20] was used to solve eq 10.

In the present work, the real potential is computed as follows

$$V(x) = E_a \sum_{i=0}^{10} C_i \, exp\{-(x-\mu_i)^2/2\sigma_i^2\} + (x/a)^l \quad (11)$$

where a and E_a are chosen constants and l is the order of the polynomial. For example, values of $a = 1.3$, $E_a = 300$, and $l = 4$ were used as well as $a = 2.6$, $E_a = 400$, when $l = 8$, as described later. Also, C_i and σ_i were chosen to be a set of random numbers between 0 and 1 as given in Table 1, whereas μ_i was set on a equi-spaced grid with a spacing length of 1.0 when $x \le 0$, and 1.25 when $x > 0$, as also shown in Table 1.

3. Results and Discussion

Results from the SSMD simulation using the NHC are for the analytic (real) potentials computed from eq 11. The following parameters were used for all calculations: The mass (m) of the particle was equated to 1.0, the number of thermostats (M) was set to 10, and the mass of the thermostats [$Q(1{:}M)$] was equated to 1.0. The starting thermostat positions (η) and velocities (ν_η) were initialized to 1.0. All simulations were performed for 1×10^8 time steps.

In the first test, the use of the upper 2-state surface as an effective SSMD potential is demonstrated. The upper 2-state surface is computed as the highest eigenvalue from eq 1. The analytic potential is computed from eq 11. Figure 2a−c shows the potential energies predicted using SSMD from eq 7 in comparison to the analytic potential. Here, the different effective SSMD potentials are computed using the choice of fictitious "filling potential", $V_0(x) = V_{min} + nk_BT$, with various values for the parameter n. It is seen that the use of various effective SSMD potentials predicts potential energies which are in very close agreement to the analytic potential energy. Also, this example demonstrates that the upper 2-state surface acts as a very effective SSMD potential surface provided a reasonable choice of effective potential $V_0(x)$ is made.

The choice of the lower 2-state surface as the effective SSMD potential is given by the lowest eigenvalue from eq 1. For $V_0(x)$ in this case, we use $V_0(x) = V_{min} + nk_BT + (x/a)^l$. Figure 3 shows the predicted potential energies from

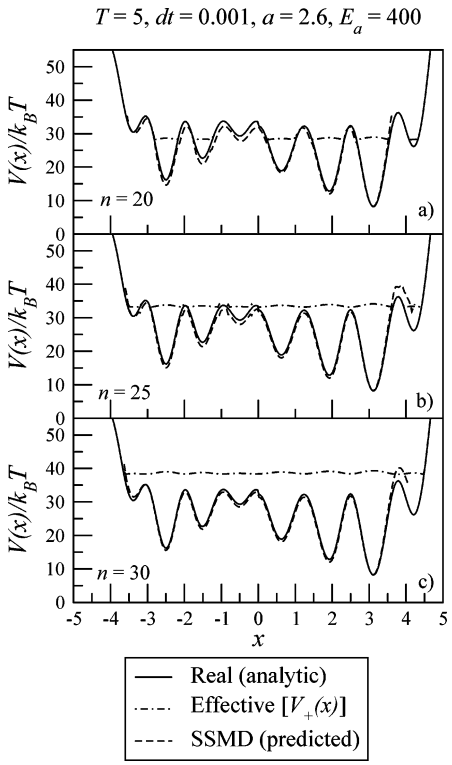

Figure 2. Predicted potential energies using the upper SSMD effective potential [$V_+(x)$], in comparison to the analytic potential, with the SSMD fictitious potential given by $V_0(x) = V_{min} + nk_BT$, at heights (a) $n = 20$, (b) $n = 25$, (c) $n = 30$, above the global minima [V_{min}].

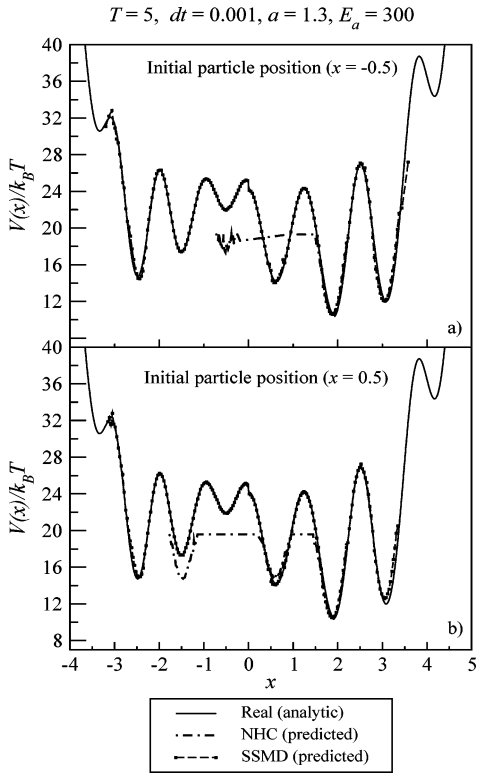

Figure 3. Predicted potential energies using NHC and the lower SSMD effective potential, [$V_-(x)$], in comparison to the analytic potential at $T = 5.0$ for different initial starting particle position at (a) $x = -0.5$ and (b) $x = 0.5$, as described in the text.

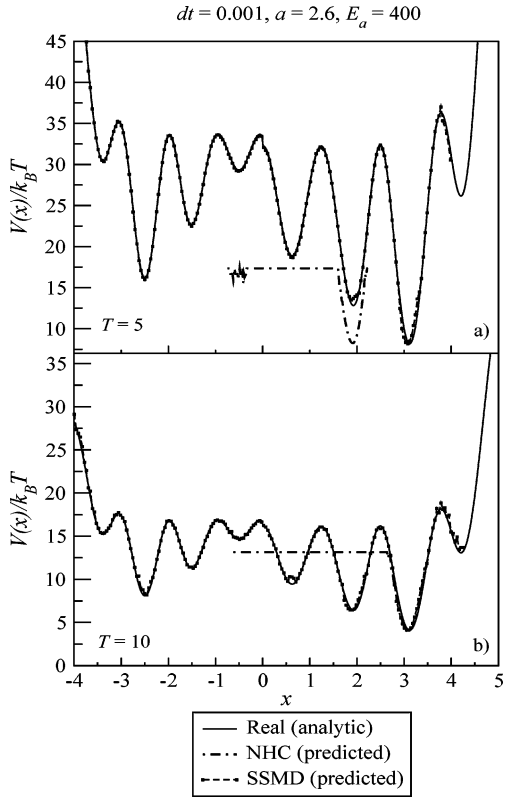

Figure 4. Predicted potential energies using NHC and the lower SSMD effective potential, $[V_{-}(x)]$, in comparison to a steeper walled analytic potential at (a) $T = 5$ and (b) $T = 10$, as described in the text.

NHC and SSMD in comparison to the analytic potential at simulation temperature $T = 5.0$, for $l = 4$ and $a = 1.3$. The MD sampling was performed with the choice of $n = 5$ in the fictitious potential. Here, different initial particle positions are used to test the SSMD method. As seen in Figure 3a,b, NHC on the actual potential $V(x)$ samples poorly as it gets trapped in one of the local minima. However, SSMD significantly increases ergodicity and predicts a result from eq 7 that resembles very closely the analytic potential.

The effectiveness of the SSMD method was further tested by choosing an analytic potential such that the walls rise more steeply ($l = 8$, $a = 2.6$, in eq 11; see also Table 1). As seen from Figure 4, this analytic potential shows more high barriers in the energy profile as compared to the example shown in Figure 3. Also, to ensure diversity in our tests and to further explore the performance of SSMD, simulations at two different temperatures are shown in Figure 4a,b. Again, the effectiveness of SSMD in sampling the rough energy landscape is observed, while standard NHC MD fails dramatically.

4. Concluding Remarks

The SSMD method has been presented in this paper and used to sample some model rough energy landscapes. It has been shown that the SSMD method is a simple and low cost computational method to ergodically sample potential energy surfaces. One possible advantage of SSMD is that it requires information primarily about global minima of the analytic potential. Also, the fictitious potential ($V_0(x)$ in SSMD) is completely general, and the choice between the upper 2-state or lower 2-state SSMD surface for the MD sampling can be made depending on the problem at hand. A mechanism to alternate between the upper 2-state and lower 2-state surface could also be implemented. In fact, in realistic situations the user may have a large number of possible choices for the SSMD fictitious potential $V_0(x)$ based on the need to overcome physical barriers in the actual potential energy surface. The choice of the off-diagonal element in SSMD should be carefully considered as well for multidimensional potentials, but here again there are many possible choices. SSMD therefore seems to be a promising and general tool to study systems with rough energy landscapes. Applications and further developments for realistic systems will be explored in the future.

Acknowledgment. This research is funded by the University of Utah Center for the Simulation of Accidental Fires and Explosions (C-SAFE), funded by the Department of Energy, Lawrence Livermore National Laboratory, under subcontract B341493. We acknowledge the Center for High Performance Computing, University of Utah for computer time support.

References

(1) Berg, B. A.; Neuhaus, T. *Phys. Rev. Lett*. **1992**, *68*, 9−12.
(2) Hansmann, U. H. E. *Chem. Phys. Lett*. **1997**, *281*, 140−150.
(3) Faller, R.; Yan, Q.; de Pablo, J. J. *J. Chem. Phys*. **2002**, *116*, 5419−5423.
(4) Lyubartsev, A. P.; Martsinovski, A. A.; Shevkunov, S. V.; Vorontsov- Velyaminov, P. N. *J. Chem. Phys*. **1992**, *96*, 1776−1783.
(5) Irbäcki, A.; Potthast, F. *J. Chem. Phys*. **1995**, *103*, 10298−10305.
(6) Stolovitzky, G.; Berne, B. J. *Proc. Nat. Sci*. **2000**, *97*, 11164−11169.
(7) Wales, D. J.; Doye, J. P. K. *J. Phys. Chem. A*. **1997**, *101*, 5111−5116.
(8) Sugita, Y.; Okamato, Y. *Chem. Phys. Lett*. **1999**, *314*, 141−151.
(9) Voter, A. F.; Montalenti, F.; Germann, T. C. *Annu. Rev. Mater. Res*. **2002**, *32*, 321−346.
(10) Voter, A. F. *J. Chem. Phys*. **1997**, *106*, 4665−4677.
(11) Voter, A. F. *Phys. Rev. B* **1998**, *57*, 13985−13988.
(12) Miron, R. A.; Fichthorn, K. A. *J. Chem. Phys*. **2003**, *119*, 6210−6216.
(13) Grubmüller, H. *Phys. Rev. E* **1995**, *52*, 2893−2906.
(14) Laio, A.; Parrinello, M. *Proc. Natl. Acad. Sci. U.S.A*. **2002**, *99*, 12562−12566.
(15) Micheletti, C.; Laio, A.; Parrinello, M. *Phys. Rev. Lett*. **2004**, *92*, 170601/1−170601/4.
(16) Rahman, J. A.; Tully, J. C. *J. Chem. Phys*. **2002**, *116*, 8750−8760.
(17) Hamelberg, D.; Mongan, J.; McCammon, J. A. *J. Chem. Phys*. **2004**, *120*, 11919−11929.
(18) Martyna, G. J.; Klein, M. L.; Tuckerman, M. *J. Chem. Phys*. **1992**, *97*, 2635−2643.
(19) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press: New York, 1987; pp 81−82.
(20) Martyna, G. J.; Tuckerman, M. E.; Tobias, D. J.; Klein, M. L. *Mol. Phys*. **1996**, *87*, 1117−1157.

CT0499175