Modified Replica Exchange Simulation Methods for Local Structure Refinement

Xiaolin Cheng,† Guanglei Cui,† Viktor Hornak,‡ and Carlos Simmerling†‡*

Department of Chemistry and Center for Structural Biology, Stony Brook University, Stony Brook, New York 11794-3400

Received: October 6, 2004; In Final Form: February 17, 2005

Parallel tempering, also known as replica exchange molecular dynamics (REMD), has recently been successfully used to study the structure and thermodynamic properties of biomolecules such as peptides and small proteins. For large systems, however, applying REMD can be costly since the number of replicas needed increases as the square root of the number of degrees of freedom in the system. Often, enhanced sampling is only needed for a subset of atoms, such as a loop region of a large protein or a small ligand binding to a receptor. In such applications, it is often reasonable to assume a weak dependence of the structure of the larger region on the instantaneous conformation of the smaller region of interest. For these cases, we derived two variant replica exchange methods, partial replica exchange molecular dynamics (PREMD) and local replica exchange molecular dynamics (LREMD). The Hamiltonian for the system is separated, with replica exchange carried out only for terms involving the subsystem of interest while the remainder of the system is maintained at a single temperature. The number of replicas required for efficient exchange thus depends on the number of degrees of freedom in the fragment needing refinement rather than on the size of the full system. The method can be applied to much larger systems than was previously practical. This also provides a means to preserve the integrity of the structure outside the refinement region without introduction of restraints. LREMD takes this weak coupling approximation a step further, employing only a single representation of the large fragment that simultaneously interacts with all of the replicas of the subsystem of interest. This is obtained by combining replica exchange with the locally enhanced sampling approximation (LES), reducing the computational expense of replica exchange simulations to near that of a single standard molecular dynamics (MD) simulation. Use of LREMD also permits the use of LES without requiring the specification of a single temperature, a known difficulty for standard LES simulations. We tested these two methods on the loop region of an RNA hairpin model system and find significant advantages over standard MD and REMD simulations.

Introduction

The potential energy surfaces of biological systems have long been recognized to be rugged,[1−3] which hampers the efficiency of conformational transitions between the various local minima. Due to this property of the free energy landscape, efficient computational approaches for searching for low-energy minima in these complex systems present a great challenge. This sampling problem can preclude success even when a sufficiently accurate Hamiltonian of the system is used in the simulations. Thus, numerous algorithms have been developed to improve the sampling of phase space for molecular simulations (conformational sampling was recently reviewed[4] and is also the subject of a recent special issue[5] of the *Journal of Molecular Graphics and Modelling*).

A general approach to surface flattening is obtained by the application of mean-field theory. Among these, the locally enhanced sampling (LES) method[6−12] is of particular interest for structure optimization due to the equivalence of the LES global energy minimum to that of the original system,[13,14] thus avoiding cumbersome mapping procedures. The LES method has been successfully applied to many biomolecular problems such as structure prediction,[13,15−19] free energy calculations,[11,20] and ligand design.[21] A recent quantum mechanics (QM)/molecular mechanics (MM) study used LES-like replicas for the QM region interacting with a single MM bath.[22] One problem that has been noted with LES is that the choice of a temperature for simulations is nontrivial,[7,8] often necessitating the use of trial and error to determine an optimal temperature for use with structure prediction simulations.[19,23] However, we will employ certain aspects of LES in one of the REMD variants that we describe below.

Another category of methods that has seen a recent increase in use is often referred to as generalized algorithms, including multicanonical methods,[24,25] simulated tempering,[26,27] and the replica exchange method (REM).[28−30] Multicanonical methods and simulated tempering improve the sampling by replacing the Boltzmann factor $\exp(-\beta E)$ with the multicanonical probability $n(E)^{-1}$. This allows the system to sample freely as a one-dimensional random walk in energy or temperature space. A difficulty in applying these two methods is in a priori determination of the multicanonical probability functions.[27]

In parallel tempering[29] and replica exchange molecular dynamics (REMD),[31] several noninteracting copies (replicas) are independently and simultaneously simulated at different temperatures. At intervals during the otherwise standard simulations, conformations of the system being sampled at different temperatures are exchanged based on a Metropolis-type criterion that considers the probability of sampling each conformation at the alternate temperature. In this way, REMD is hampered

* Author to whom correspondence should be addressed. Phone: (631) 632-1336. Fax (631) 632-1555. E-mail: carlos.simmerling@stonybrook.edu.
† Department of Chemistry.
‡ Center for Structural Biology.

10.1021/jp045437y CCC: $30.25
Published on Web 04/06/2005

to a lesser degree by the local minima problem, since the low-temperature simulations (replicas) have the potential to escape kinetic traps by jumping to minima that are being sampled by the higher-temperature replicas where kinetic trapping is less prevalent. However, the high-energy regions of conformational basins often sampled by the high-temperature replicas can be relaxed in a way similar to the temperature annealing method. Moreover, the transition probability is constructed such that the canonical ensemble properties are maintained during each simulation, thus providing potentially useful information about conformational probabilities as a function of temperature. Due to these advantages, REMD has been widely applied to studies of peptide and small protein folding.[29,31−41]

For large systems, however, application of REMD can require significant computational resources, thus limiting its advantages. In particular, the number of replicas needed to cover a given temperature range increases as $O(f^{1/2})$ for a system with f degrees of freedom.[29,42] Several promising techniques have been proposed to deal with this apparent disadvantage to the REM. In a method proposed by Sugita, the REM was coupled to the multicanonical method to take advantage of both techniques.[43,44] Pak proposed combining the REM with the generalized effective potential method.[45] Takada et al. developed an alternative REM called Hamiltonian REM. In their method, the Hamiltonian was separated into two parts, one assumed to be tightly coupled to temperature space and the other not.[42] REMD was then performed on the former part of Hamiltonian, including the torsion angle and repulsive van der Waals terms, which are mainly responsible for the rugged energy surface. By doing so, the number of replicas needed was reduced by excluding some degrees of freedom such as solvent, bond length, and bond angle terms. Two of their examples, called scaled hydrophobicity REM and phantom chain REM, demonstrated the strength of the approach.

Motivated by our desire to focus the enhanced sampling on a subset of the degrees of freedom in the system, we propose a different partitioning of the Hamiltonian for a large system. We define a subset of the system as the "focus" region where increased sampling is required. The remainder is often the part of system that we are not specifically interested in but which must be present during the simulation. Typical examples would be division into a protein core built on a homology model and a surface loop for which experimental structural data is unavailable or a ligand binding to a receptor. We desire improved sampling for the loop or ligand, but we typically do *not* want to significantly enhance sampling for the remaining degrees of freedom (and potentially lose the "good" structure for the protein core).

Through this reasoning, the system is divided into two parts in one of our proposed methods, partial replica exchange molecular dynamics (PREMD). Under the assumption of no strong coupling between them, the two regions are coupled to separate temperature baths. Replicas of this dual thermostat system are then created. The bath temperature of the bath for the region of interest is varied between replicas, while the other bath has the same temperature for all replicas. An acceptable exchange ratio can thus be obtained for the system with fewer replicas than needed with traditional REMD.

PREMD assumes that changes in conformation of the focus region do not require or introduce correspondingly large changes in the remainder of the system (which is maintained at a lower temperature). However, this region need not be restrained and can respond (albeit slowly) to changes in conformation of the focus region. For cases where this coupling is expected to be particularly weak, we introduce "local" replica exchange molecular dynamics (LREMD) in which all of the replicas of the focus region share identical coordinates for the remainder of the system. We note that this approach shares a common philosophy with LES, and we thus use LES to build the LREMD system. Only the atoms in the focus region are replicated, and a *single* set of atoms for the remainder of the system interacts in a mean-field manner with *all* replicas of the focus region. The LREMD system is therefore much smaller in size than the full set of replicated systems obtained using REMD or PREMD. However, LREMD represents a further approximation beyond PREMD in that the system cannot react individually to differences in conformation among the focus region replicas.

From the perspective of replica exchange methods, pairing with LES reduces the computational cost of applying REMD to large systems. The reduction in barrier heights intrinsically provided by the LES component of the method also suggests that a smaller temperature range may be sufficient to achieve adequate sampling of the relevant conformations. From the LES viewpoint, however, applying REMD to the LES copies can relieve the user from the difficulty of choosing a single temperature for the LES region, since the copies are coupled to separate temperature baths that span a range of values. Standard replica exchange is then carried out, with the exception that we need exchange only a small portion of the system, not the entire system.

The remainder of the text will be organized as follows. In the Methods section, we will discuss the algorithm and detailed implementation of the two variant REMD methods, which will be available for use in AMBER[46] version 9. In the Results and Discussion section, we describe application of both methods to a model system and compare the results to those obtained from standard REMD and LES simulations. The consistency among various methods will be further verified by projecting multiple molecular dynamics (MD) trajectories to a free energy landscape derived from simulation.

Methods

Replica Exchange Molecular Dynamics (REMD). In standard parallel tempering or replica exchange molecular dynamics,[29,31] the simulated system consists of M noninteracting copies (replicas) at M different temperatures. The positions, momenta, and temperature for each replica are denoted by $\{q^{[i]},p^{[i]},T_m\}$, $i = 1, ..., M$; $m = 1, ..., M$. The equilibrium probability for this generalized ensemble is

$$W(p^{[i]},q^{[i]},T_m) = \exp\left\{-\sum_{i=1}^{M}\frac{1}{k_B T_m}H(p^{[i]},q^{[i]})\right\} \qquad (1)$$

where the Hamiltonian $H(p^{[i]},q^{[i]})$ is the sum of kinetic energy $K(p^{[i]})$ and potential energy $E(q^{[i]})$. For convenience, we denote $\{p^{[i]},q^{[i]}\}$ at temperature T_m by $x_m^{[i]}$ and further define $X = \{x_1^{[i(1)]}, ..., x_M^{[i(M)]}\}$ as one state of the generalized ensemble. We now consider exchanging a pair of replicas. Suppose we exchange replicas i and j, which are at temperatures T_m and T_n, respectively

$$X = \{...; x_m^{[i]}; ...; x_n^{[j]}; ...\} \rightarrow X' = \{...; x_m^{[j]}; ...; x_n^{[i]}; ...\} \quad (2)$$

To maintain detailed balance of the generalized system, microscopic reversibility has to be satisfied, thus giving

$$W(X)\rho(X \rightarrow X') = W(X')\rho(X' \rightarrow X) \qquad (3)$$

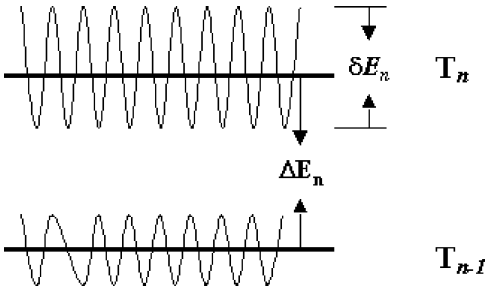

Figure 1. Schematic diagram illustrating the energy fluctuations for simulations at two temperatures for neighboring replicas. To obtain high exchange probabilities, the energy fluctuations δE in each simulation should be of comparable magnitude to the mean energy difference ΔE.

where $\rho(X \rightarrow X')$ is the exchange probability between two states X and X'. With the canonical ensemble, the potential energy E rather than total Hamiltonian H will be used simply because the momentum can be integrated out. Inserting eq 1 into eq 3, we obtain the following equation for the Metropolis criterion for the exchange probability

$$\rho = \min\left(1, \exp\left\{\left(\frac{1}{k_B T_m} - \frac{1}{k_B T_n}\right)(E(q^{[i]}) - E(q^{[j]}))\right\}\right) \quad (4)$$

In practice, several replicas at different temperatures are simulated simultaneously and independently for a chosen number of MD steps. Exchange between a pair of replicas is then attempted with a probability of success calculated from eq 4. If the exchange is accepted, then the bath temperatures of these replicas will be swapped, and the velocities will be scaled accordingly. Otherwise, if the exchange is rejected, then each replica will continue on its current trajectory with the same bath.

As we described above, one of the major limitations of REM is that the number of replicas needed to span a temperature range grows proportionally to the square root of the number of degrees of freedom in the simulated system. While a more rigorous analysis of the acceptance probability in REM trials has been given recently using a Gaussian energy distribution model,[47,48] one can also approximate from eq 4 that the overall exchange probability P_{acc} is proportional to $\exp(-\Delta T^2/T^2)$, which implies that a greater acceptance ratio requires a smaller temperature gap ΔT or a more dense temperature distribution to reach. However, ΔT should be as large as possible so as to span a maximum temperature range by using a fixed number of replicas. The relationship can be estimated through consideration of potential energy fluctuations of two replicas sampling at the target temperature T_n and T_{n-1} (Figure 1). The instantaneous energy fluctuation δE in a given simulation at temperature T scales as $\sqrt{f}T$, and the average energy gap ΔE between two neighboring replicas is proportional to $f\Delta T$, where f is the number of degrees of freedom and $\Delta T = T_n - T_{n-1}$. Obtaining a reasonable acceptance ratio relies on keeping the replica energy gap comparable to the energy fluctuations, thus $\Delta E/\delta E$ should be near unity. Since $\Delta E/\delta E$ is proportional to $\Delta T\sqrt{f}/T$, the acceptable temperature gap between neighboring replicas therefore decreases with larger systems as $\Delta T \approx 1/\sqrt{f}$, and more simultaneous simulations are needed to cover the desired temperature range. For large systems with thousands of degrees of freedom, this presents a practical problem even with large and relatively inexpensive PC clusters.

Partial Replica Exchange Molecular Dynamics (PREMD). In PREMD, the system is divided into a "focus" region of interest and the (typically larger) "remainder" of the system. Under the weak coupling assumption between these two regions, they are coupled to separate thermostats. For the new system composed of focus atoms F and remainder atoms R, the kinetic energy is taken as

$$K = \frac{1}{2}\sum_F \frac{|p_F|^2}{m_F} + \frac{1}{2}\sum_R \frac{|p_R|^2}{m_R} \quad (5)$$

and the potential energy is

$$E = E_{RR}(q_R) + [E_{FR}(q_F, q_R) + E_{FF}(q_F)] \quad (6)$$

where E_{RR} stands for the potential energy between remainder atoms and depends only on coordinates R, E_{FF} is the energy for focus atoms, and E_{FR} is the energy of interaction between the focus and remainder regions. These two equations are essentially the same as that of the original system, with the expected average kinetic energy related to temperatures as follows

$$\langle K \rangle = \frac{k_B}{2}(3N_F T_F + 3N_R T_R) \quad (7)$$

where T_F and T_R are thermostat temperatures and N_F and N_R are degrees of freedom for the focus and remainder regions, respectively. Once separate thermostats are introduced for the two regions, the PREMD system consists of n noninteracting replicas of the original system with focus regions at n different temperatures while all of the remainder regions maintain the same temperature.

If we assume that the behavior of the remainder regions are not highly dependent on the conformations of focus regions at various temperatures, then the time-dependent potential energy term E_{RR} in eq 6 will approximately cancel between replicas. The values would cancel exactly if the remainder regions adopted identical coordinates in all replicas, as is the case for the LREMD approach described later in this paper. The remaining E_{PREMD} includes all terms that depend on coordinates of the focus region (which can adopt quite different conformations), and thus these will not cancel among replicas. From eq 6, the energy difference between a pair of replicas i and j becomes

$$\Delta E \approx E_{PREMD,i} - E_{PREMD,j} = [E_{FR}(q_F, q_R)_i + E_{RR}(q_R)_i] - [E_{FR}(q_F, q_R)_j + E_{RR}(q_R)_j] \quad (8)$$

and analogous to eq 4, the exchange probability $\rho(X \rightarrow X')$ for exchanging replicas i and j is given by

$$\rho = \min\left[1, \exp\left(\left(\frac{1}{k_B T_i} - \frac{1}{k_B T_j}\right)(E_{PREMD}(q_F^{[i]}, q_R^{[i]}) - E_{PREMD}(q_F^{[j]}, q_R^{[j]}))\right)\right] \quad (9)$$

During the simulation, as shown in Figure 2, only the focus regions are simulated over a range of temperatures, with periodic exchanges performed according to eq 9 while the remainder regions are maintained at the same temperature for all replicas.

There are several potential advantages to separation of the system into two regions with different thermostats. If the number of degrees of freedom f in the focus region is relatively small compared to that for the remainder (r), then the number of replicas needed is reduced in PREMD as compared to REMD. Furthermore, heating the remainder region may not be desirable during local optimization. By introducing and varying only local

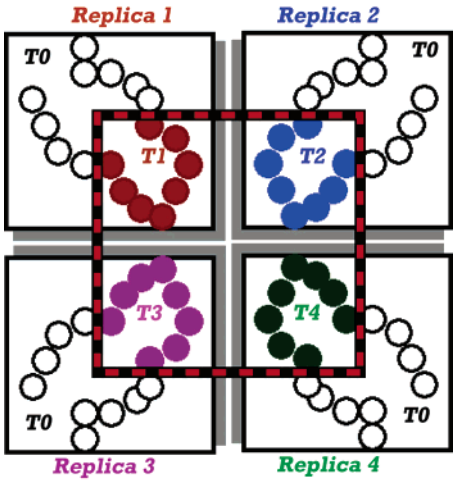

Figure 2. Schematic illustration of the PREMD method, with four replicas. Focus regions are at temperatures T_1, T_2, T_3, and T_4 (shown inside red square), and the remainder of the systems are at temperature T_0. During the simulation, target temperatures for only the focus regions are exchanged with a predefined frequency based on the Metropolis-type criterion given in eq 9.

temperature, we can target the enhanced sampling to the desired part of the system while maintaining the remainder at a lower temperature that will preserve its structural integrity. Similar results could be obtained with REMD by restraining the coordinates of atoms in the remainder region, but the number of replicas in that case would remain higher than needed for PREMD, and the restraints could prevent even weak coupling to changes in the focus region.

Local Replica Exchange Molecular Dynamics (LREMD). Motivated by the similarities in the "exchange part of the Hamiltonian" PREMD variant to the LES approach, we further extend PREMD to the LREMD method. In LREMD, instead of replicating the entire system, we use LES to construct a new system in which only a subset of atoms is replicated. The LES-replicated region corresponds to the focus region in PREMD. Similar to LES, each of these replicas interacts with the nonreplicated remainder of the system. Similar to standard REMD, the subsystem LES replicas do not interact with each other and are coupled to thermostats at different temperatures.

For a standard LES system composed of nonreplicated atoms of type A and n LES copies for atoms of type B, the kinetic energy is taken as

$$K = \frac{1}{2}\sum_{A}\frac{|p_A|^2}{m_A} + \frac{1}{2}\sum_{i=1}^{n}\sum_{B}\frac{|p_{iB}|^2}{m_B} \tag{10}$$

with the temperature related to the average kinetic energy K as

$$\langle K\rangle = \frac{k_B T}{2}(3N_A + 3nN_B) \tag{11}$$

where N_A and N_B are total degrees of freedom in the non-LES and LES region (before making LES copies), respectively. The resulting energy for the new LES system is

$$E = E_{AA}(q_A) + \frac{1}{n}\sum_{i=1}^{n}[E_{AB}(q_A,q_{iB}) + E_{BB}(q_{iB})] \tag{12}$$

which can be written as an average over n full copies of the single-copy "reference" systems obtained by combining all of the atoms belonging to one LES copy with the noncopied atoms. It is important to note that in the LES method interactions involving LES atoms are scaled by $1/n$ to maintain balance between (a) interactions involving a non-LES atom and the multiple copies of a LES atom and (b) interactions completely outside the LES region. This also maintains the energy identity between a non-LES system and the LES system in which all copies occupy the same coordinates as the corresponding atoms in the non-LES system.

Up to this point, we have described a standard LES system. For LREMD, we extend this treatment by allowing coupling of non-LES and each LES region to different thermostats (as we showed for PREMD). The following equation therefore holds

$$\langle K\rangle = \frac{k_B}{2}(3N_A T_A + 3N_B\sum_{i=1}^{N} T_{Bi}) \tag{13}$$

where T_A and T_{Bi} are thermostat temperatures for the non-LES and individual LES copies, respectively. If we assume there is no strong coupling between the LES and non-LES atoms, then this single system represents the behavior of the full set of reference systems over a range of temperatures.

We rewrite eq 12 with notation of the thermostat temperatures

$$E = E_{AA}(q_A)_{\{T_0\}} + \sum_{i=1}^{n}[E_{BB}(q_{iB}) + E_{A,B}(q_A,q_{iB})]_{T_i} \tag{14}$$

where q_{iB} stands for the coordinates of the ith LES copy of region B coupled to a bath at T_i and q_A are the coordinates of the non-LES region A at temperature T_0. Since all of the copies have the same non-LES atoms, the potential energy of the non-LES region exactly cancels among the reference systems (an approximate cancellation was assumed for PREMD), and thus the probability of exchanging LES copies i and j is defined as

$$\rho = \min[1, \exp(-\Delta_{ij})] \tag{15}$$

where

$$\Delta_{ij} = \left(\frac{1}{k_B T_i} - \frac{1}{k_B T_j}\right)[(E_{BB}(q_{iB}) + E_{A,B}(q_A,q_{iB})) - (E_{BB}(q_{jB}) + E_{A,B}(q_A,q_{jB}))] \tag{16}$$

This expression is analogous to eq 9 for PREMD and includes the internal energy of the relevant LES copies as well as the interaction of these copies with the noncopied region shared among all replicas.

During molecular dynamics, LES copies are simulated over a range of temperatures, with periodic exchanges (Figure 3) performed with a probability given by eq 15.

Simulation Details. We demonstrate the utility of the modified replica exchange methods by testing on the RNA UUCG tetraloop system ($G_1G_2A_3C_4[U_5U_6C_7G_8]G_9U_{10}C_{11}C_{12}$), for which structures have been determined by NMR.[49,50] This makes an excellent model due to its small size and since several previously reported theoretical studies explored the conversion of an incorrect conformation for the loop region into the correct one.[11,51,52] Most relevant is that we recently reported[23] improved loop conformational sampling and multiple transition pathways for this system using LES and standard MD in a continuum solvent. Application of the modified REMD approaches to the same system will allow direct comparison with standard REMD as well as our previous LES simulations.

In Figure 4, the correct and incorrect structures of the RNA tetraloop derived from NMR studies are shown. The 12 bases

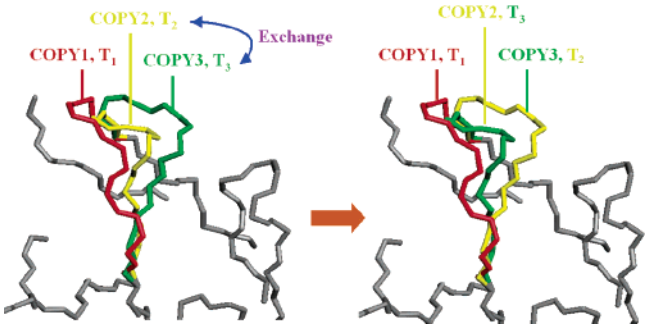

Figure 3. Simple model system to clarify the exchange scheme of the LREMD method, with a single simulation including one non-LES region (grey) and three LES copies (red, yellow, and green). The copies are coupled to thermostats at T_1, T_2, and T_3. After a predefined period of time, the target temperatures of two copies (in this case copy 2 and copy 3) are swapped based on a Metropolis-type criterion using the transition probability defined in eqs 17 and 18.

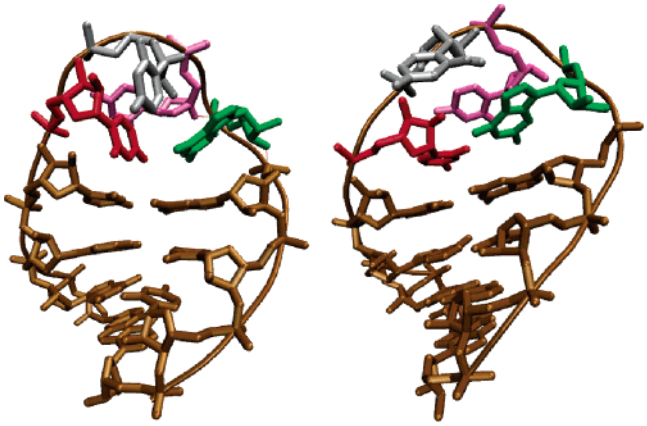

Figure 4. Incorrect (left) and correct (right) structures of the RNA tetraloop derived from NMR studies, with U5 in red, C7 in purple, G8 in green, and the flexible U6 ring in gray.

of this single-stranded RNA fold back to form a double-helical stem capped by a UUCG loop. The major difference between the two structures of the RNA stem-loop system is the hydrogen bond pattern between the bases U5 and G8 in the tetraloop region. A bifurcated hydrogen bond is present between one of the U5 carboxyl oxygen atoms and the imino and amino groups of the G8 in the correct structure. In contrast, this base pair forms a reverse wobble pattern in the incorrect structure. One of our measures of the difference between these structures and those sampled in simulations will be the root-mean-square deviation (rmsd) for non-hydrogen atoms in the UUCG tetraloop (residues 5−8) except the base of U6, which does not form specific contacts and shows higher mobility. This rmsd selection was chosen to be consistent with a previous theoretical study of this system.[11]

Number of Replicas Required for Simulations. Since only part of the system is coupled to the effective temperature space in both modified REMD approaches, the total number of replicas needed should be smaller than that for conventional REMD. We performed several test simulations on the RNA tetraloop with standard REMD and both modified REMD approaches. Results listed in Table 1 show that our PREMD does require fewer replicas to cover the same temperature range as compared to conventional REMD. For example, standard REMD requires 8 replicas for the tetraloop with the generalized Born (GB)[53] continuum solvation model, with ~40 required for the RNA in explicit TIP3P[54] water. However, the number of replicas is reduced to 5 if PREMD is applied to the four nucleotides of the UUCG loop. In addition, this model system is relatively small, and the advantages to using PREMD would be expected to be even more dramatic for larger systems such as proteins with hundreds of amino acids.

REMD Setup. The standard REMD simulations were run using our implementation in AMBER (version 8). The temperatures of the 8 replicas were 266, 282, 300, 318, 338, 359, 381, and 405 K. These temperatures were optimized to give an uniform exchange acceptance ratio of about 10%. The original incorrect and correct RNA tetraloop NMR structures were used as starting structures for the simulations. The time step was 2 fs, and SHAKE[55] was applied to all bonds involving hydrogen. Berendesen temperature coupling[56] was used with a relaxation time of 0.5 ps^{-1}. The AMBER ff94 force field[57] was used for all calculations. Solvent effects were included through use of the GB continuum model[53] as implemented in AMBER. The intrinsic Born radii were adopted from Bondi[58] with modification of hydrogen,[59] and the scaling factors for the Born radii were taken from the TINKER modeling package.[60] No cutoff on nonbonded interactions was used. Before the production runs, each replica was equilibrated at its target temperature for 100 ps. Exchange was attempted every 1 ps, and the data was also saved every 1 ps for later analysis. A total of 60 000 exchange attempts were performed (corresponding to a 60 ns simulation for each replica). The first 2000 of the 60 000 snapshots were discarded to minimize the influence of the starting structures. The free energy was then calculated as $F = -k_B T \ln P(\text{rmsd},\chi_{U5})$, where $P(\text{rmsd},\chi_{U5})$ is the probability distribution function for each pair of rmsd and glycosidic angle values.

Distance restraints on Watson−Crick hydrogen bonds between the first three base pairs in the stem region (residues 1:12, 2:11, and 3:10) were required during standard REMD to preserve the experimental conformation for this region at the high T used on all atoms. No such restraints were needed for PREMD or LREMD, since these methods allowed coupling of this region to a thermostat at a lower temperature.

PREMD Setup. The REMD module in SANDER was modified to permit separation of the system into two parts, with coupling of each part to a separate thermostat so that replica exchange could be performed on a subset of system. The UUCG loop residues were defined as the focus region. The remainder (stem) region temperature was maintained at 300 K while five different target temperatures were used for the focus regions (5 replicas): 266, 295, 327, 363, and 403 K. The temperature distribution was also optimized to give a desired exchange ratio of approximately 10%. Compared to the 8 replicas needed in standard REMD, these 5 replicas can effectively cover the same temperature range with PREMD. In this system, $^1/_3$ of the residues are replicated; the difference in the number of replicas required for REMD and PREMD would be greater for large systems with smaller fractions replicated (as discussed above). All of the other control parameters used in PREMD simulations were the same as described above for standard REMD.

LREMD Setup. For LREMD simulation, the AMBER module ADDLES was employed to construct an LES system. SANDER was modified to permit coupling of each LES copy to a separate temperature bath. LREMD employed our GB + LES module in AMBER. The entire UUCG loop was replaced by five LES copies. All LES copies of individual atoms were initially assigned identical coordinates but unique velocities according to the target temperatures. We used the following target temperatures for the LES copies, 80, 88, 99, 108, and 120 K, whereas the non-LES region was maintained at 100 K. Similar to PREMD, the temperature distribution was also

TABLE 1: REM Related Variables for Various Applications to the RNA Tetraloop[a]

focus region	REMD degrees of freedom	E_{pot}/T slope	$\delta^2 T$ at 300 K	$\delta^2 E_{pot}$ at 300 K	$N_{replica}$	resulting acceptance ratio
RNA 2b (GB)	192	0.18	29.44	3.94	3	0.12
RNA 4b (GB)	375	0.35	22.06	6.73	5	0.11
RNA (GB)	1146	1.13	9.28	12.98	8	0.13
RNA (explicit water)	27 432	26.82	2.57	69.61	38	

[a] All example systems are built on the RNA tetraloop. $N_{replica}$ indicates the number of replicas that would be required to span a temperature range of 260−420 K. RNA 2b stands for two nucleotides as the focus region for PREMD simulation; RNA 4b stands for four focus nucleotides for PREMD simulation; RNA stands for standard REMD with the entire RNA strand replicated. The undetermined value in the acceptance ratio column for RNA (explicit water) means that the data was not obtained due to the computational requirements.

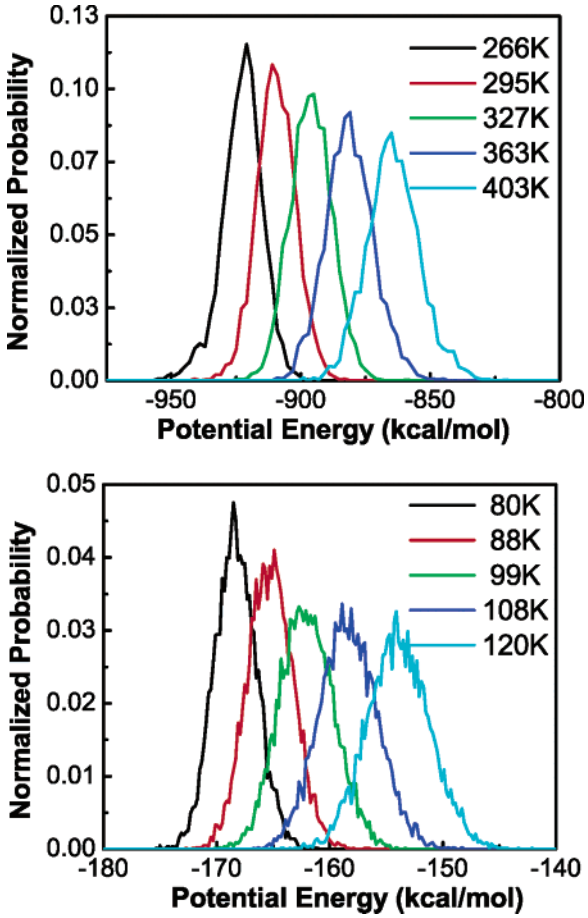

Figure 5. Probability distribution of potential energy for the atoms in the focus region, at each target temperature. The upper plot shows the distribution sampled from PREMD simulation, the lower from LREMD simulation. Overlap between the distributions facilitates efficient exchange.

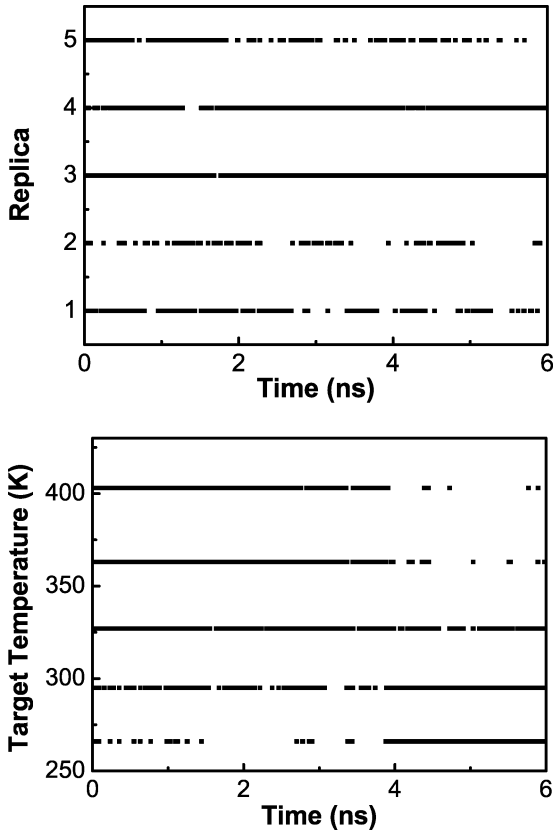

Figure 6. Time series of replicas at 295 K (upper) and target temperatures of replica 2 (lower) for PREMD simulations of the RNA tetraloop.

optimized to give a desired exchange ratio of approximately 10%. All other control parameters for the simulations were the same as described above for REMD and PREMD.

Multiple Baths for Temperature Coupling. In both REMD variants, we couple various regions of the system to baths at different temperatures, such as the focus and remainder regions in PREMD and the non-LES region and each LES region in LREMD. To verify that the target temperature could be properly maintained for each temperature subspace, we first performed an equilibration simulation on the RNA LES system. This system was constructed by replacing the entire UUCG loop with five LES copies. During the simulation, the non-LES and each LES region were coupled to different temperature baths while no exchange was attempted (standard LES MD). The non-LES region target temperature was 300 K, and the LES region target temperatures were evenly distributed from 200 to 360 K. Both the non-LES and LES regions maintained the target temperatures (data not shown), an essential first step for both modified REMD approaches.

In an attempt to further validate the performance of coupling the system to multiple thermostats, we monitored potential energies and temperatures of replicas during the simulations. To have sufficient exchanges between neighboring replicas, the potential energy probability distributions for each temperature should demonstrate acceptable overlap. As shown in Figure 5 for both modified REMD methods, the distributions have the desired behavior with enough overlap between each adjacent temperature pairs to allow exchange to occur. We also find that the potential energy at each temperature follows a Gaussian distribution, with linear temperature dependence for the mean and standard deviation (data not shown). We thus conclude that the multiple baths in each simulation are capable of supporting the requirements of PREMD and LREMD.

Results and Discussion

Figures 6 and 7 show the time series of replicas sampling a particular target temperature in PREMD and LREMD, respec-

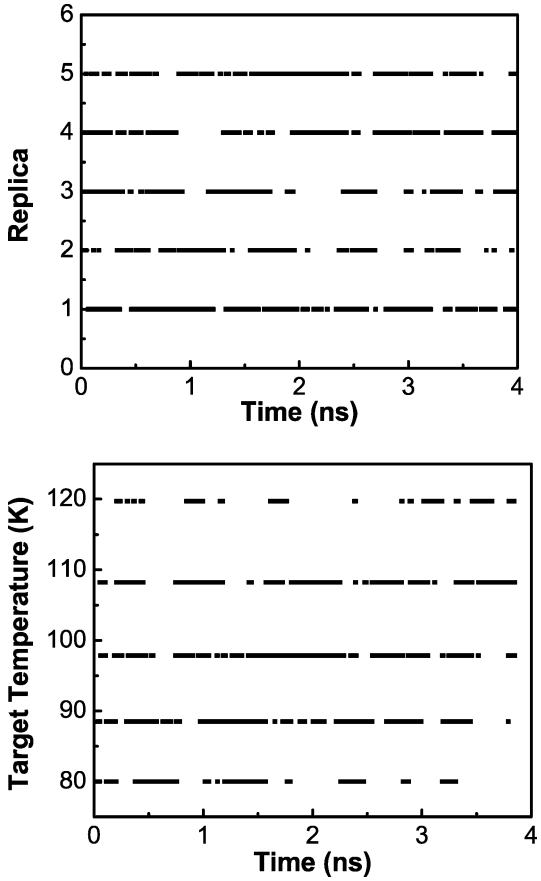

Figure 7. Time series of replicas at 99 K (upper) and target temperatures of replica 2 (lower) for LREMD simulations of the RNA tetraloop.

TABLE 2: Acceptance Ratios for the PREMD Simulation

temperature pairs	exchange acceptance ratio
266 K ↔ 295 K	0.11
295 K ↔ 327 K	0.11
327 K ↔ 363 K	0.11
363 K ↔ 403 K	0.12

TABLE 3: Acceptance Ratios for the LREMD Simulation

temperature pairs	exchange acceptance ratio
80 K ↔ 88 K	0.10
88 K ↔ 99 K	0.13
99 K ↔ 108 K	0.13
108 K ↔ 120 K	0.10

tively, as well as the time series of target temperatures for one of the replicas. We observe a free random walk in temperature space in both REMD variants, as expected for correctly evolving REMD simulations. This provides further evidence that the use of the partitioned Hamiltonian in each simulation type does not affect the ability to carry out efficient exchanges.

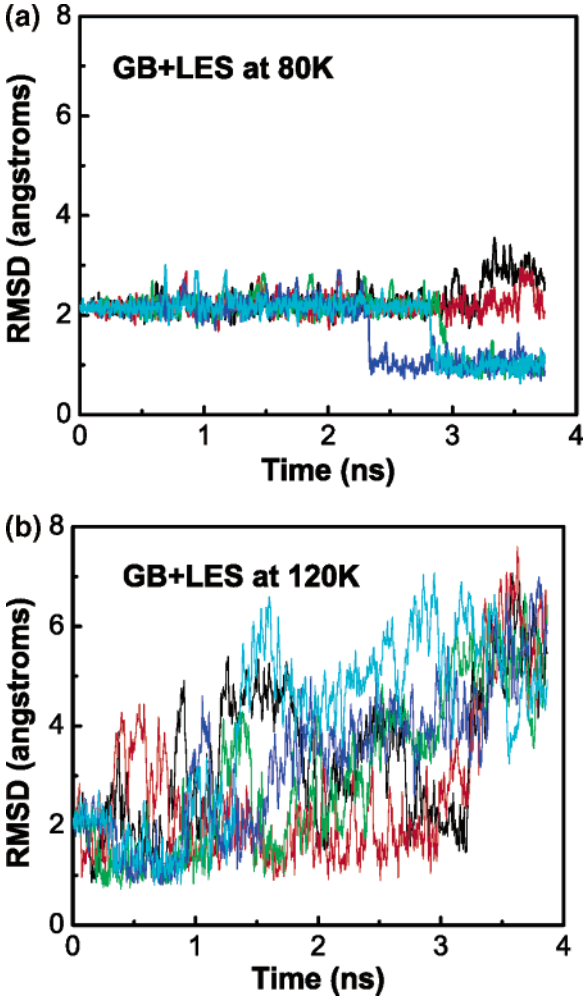

Figure 8. Loop rmsd as a function of time for two standard LES simulations: (a) 80 K; (b) 120 K. These are the boundaries of the temperature range used in LREMD.

Tables 2 and 3 provide the overall exchange acceptance ratio for PREMD and LREMD simulations, respectively. The choice of target temperatures was designed to result in an exchange acceptance ratio of 0.1. The resulting exchange frequencies are sufficiently close to the predicted acceptance ratio, and exchanges are uniform between pairs.

A strength of REMD is that the improved sampling possible at higher T can be employed to predict conformations that are populated at a lower T where standard MD simulations tend to be kinetically trapped. Experimentally, the single-stranded RNA is found to adopt a well-defined hairpin structure with a double-helical stem capped by a UUCG loop. A bifurcated hydrogen bond between U5 and G8 is present in the loop region. The similarity between structures is measured by several metrics, with detailed results provided in Table 4.

TABLE 4: Comparison of Representative Structures of the Most Populated Cluster Sampled by Various Simulation Methods[a]

system	NMR-based	explicit solvent MD[11]	GB MD[23]	REMD	PREMD	LREMD
rmsd (Å)		0.9	1.0	1.0	1.0	1.1
U5:O2−G8:N1 (Å)	2.7	2.8	2.9	2.9	2.9	3.0
U5:O2−G8:N2 (Å)	3.7	3.7	3.4	3.4	3.4	4.0
U5 χ angle (deg)	−157	−161	−153	−152	−150	−144
G8−C7 stacking (Å)	5.9	6.2	6.7	6.9	7.2	7.8

[a] Rows 2 and 3 correspond to distances for hydrogen bonds between the U5 and G8 bases present in the correct and incorrect structures, respectively; the U5 χ angle is the glycosidic angle defined as the torsion angle formed by O4′−C1′−N−C2; G8−C7 stacking is calculated as the distance between mass centers of the two bases.

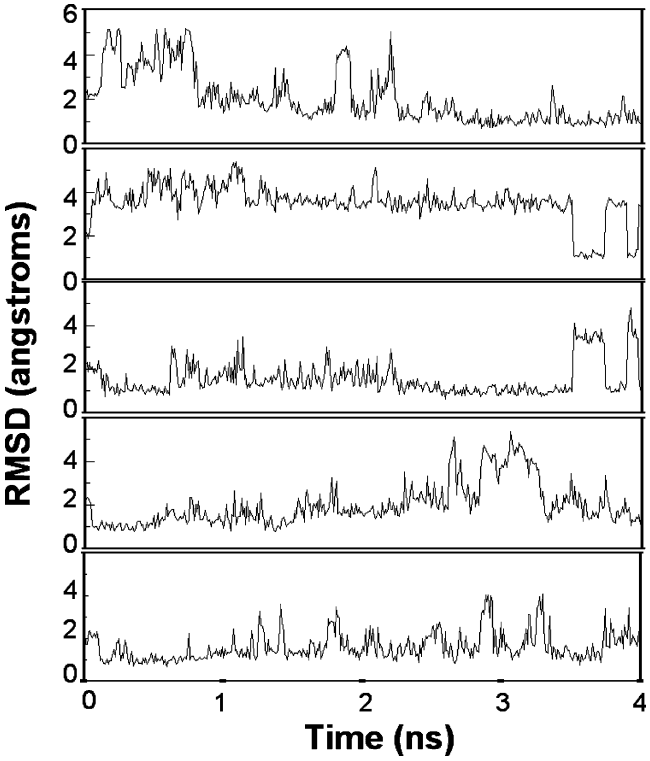

Figure 9. Loop rmsd as a function of time during LREMD simulation. Each graph represents a single replica (which samples a range of temperatures). All replicas successfully undergo the transition from incorrect to correct conformation.

All of the REMD, PREMD, and LREMD simulations started from the *incorrect* structure, but in every case the most populated structures are in agreement with the *correct* conformation. The results confirm that PREMD and LREMD both predict the same global free energy minimum as standard REMD.

Since LREMD is built on top of LES and designed to ameliorate its temperature limitation, it is useful to directly compare results obtained from these two methods. We initiated two LES simulations from the incorrect structure, with all five LES copies having identical initial coordinates. As shown in Figure 8a, at 80K the loop appears to be stable as monitored by rmsd. In approximately 2.4 ns, one of the five copies converted to the correct structure. At 4 ns, two copies remain close to the original incorrect basin, and those that located the correct structure have remained there. We therefore elevated the bath temperature to 120 K for all LES copies (Figure 8b). In this case, the simulation resulted in a fully extended conformation (unfolded) of the RNA. The rmsd rose to 4.0 Å, and all base pairs were lost during the simulation. Similar unfolding was observed when starting from the correct conformation. Therefore, as we reported previously,[19,23] the choice of appropriate temperature for LES simulations is a major difficulty of applying this efficient mean-field method to many other systems.

We hypothesized that the sensitivity to the temperature in LES simulations would be remedied through the use of the LREMD approach. We therefore investigated whether our approach would be successful in this regard. In Figure 9, we show the rmsd value for each of the five copies as a function of time during 4 ns LREMD simulation at temperatures ranging from 80 to 120 K (the temperatures tested above for GB + LES). It is evident that most replicas undergo a faster I → C transition than seen with LES at 80 K, yet the hairpin is stable. In contrast to the simulation with GB + LES at 80 K, copies that convert to the correct conformation do not always remain there; instead there are multiple transitions between the correct and other structures during the entire simulation, thus possibly providing a temperature-dependent sampling of nonnative conformations.

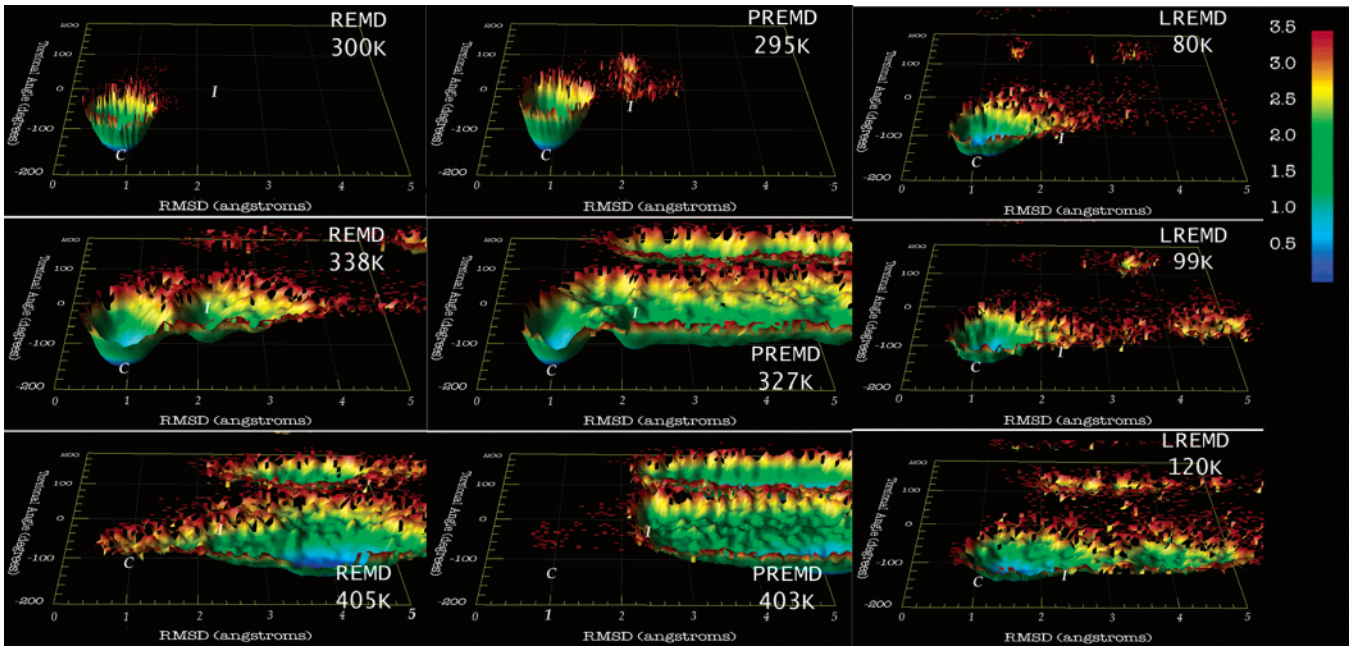

Figure 10. Free energy landscapes of the RNA tetraloop constructed from various REMD simulations using the loop region rmsd (compared to the correct structure) and U5 glycosidic angles as two reaction coordinates: (left) standard REMD; (middle) PREMD; (right) LREMD. I and C labels denote the locations of the basins for incorrect and correct structure families, respectively. Despite being initiated in the incorrect conformation, all three surfaces show the correct NMR-based structure as the global minimum at low *T*, demonstrating the utility of our new methods for structure refinement. REMD and PREMD show high similarity at all temperatures, with differences possibly due to the need for stem region restraints in standard REMD. While LREMD appears somewhat different, the temperature range chosen for LES does not necessarily match that used for REMD/PREMD.

At this point, it is necessary to ask to what extent the results from these modified approaches resemble those that would be obtained from standard REMD. Since the canonical ensemble properties are maintained in REMD, the free energy landscape can be reconstructed through histogram analysis along chosen reaction coordinates. This analysis employs structure ensembles corresponding to a given temperature rather than the trajectories for each replica that were shown in Figure 9. We can do the same analysis on the data obtained from our modified approaches; nevertheless, one should bear in mind that even if the correct thermodynamic properties can be reproduced from PREMD under the weak coupling assumption, the resulting free energy landscape constructed from LREMD will surely not quantitatively match the original system due to the modification of the potential function inherent in the LES approximation. We speculate that it might be possible for LREMD to rapidly (and inexpensively) give a free energy landscape that reproduces the general features of the real one, including the location of the global free energy minimum. This would prove to be highly useful for structure prediction or refinement purposes.

To compare our simulation results with standard REMD, we constructed the free energy landscape using the loop region rmsd (as compared to the correct structure) and U5 glycosidic angle (χ_{U5}) as two reaction coordinates for each of the three simulations types. The free energy surfaces calculated from PREMD and LREMD are compared with those from REMD in Figure 10. Different rows are from different temperatures. Three temperatures are shown, the closest to 300 K (to match previous MD studies[23,51,52]), the middle temperature of the range, and the highest. For LREMD, the lowest, middle, and highest temperatures are shown. All of the surfaces show a very similar pattern. Only one deep free energy minimum is observed at low temperature in all REMD methods, which corresponds to the NMR-based correct structure (recall that all simulations were initiated with the incorrect structure). This shows that both of our variant methods can locate the same free energy minimum as REMD (which also matches the experimental conformation, but this accuracy relates to the force field, not the sampling method). As temperature increases, more disordered structures are sampled. At the middle temperatures, both REMD and PREMD surfaces show two major minima. The correct structure is highly dominant while the incorrect structure is only a metastable structure corresponding to a much shallower minimum on the free energy surface.

While REMD and PREMD show similar behavior at similar temperatures, the temperature range chosen for LREMD appears to not reach as high an *effective* temperature as the ~400 K shown for PREMD and REMD, since significant population of the native basin occurs in LREMD even at 120 K. This is not unexpected, since a direct correspondence between the LES temperature and the effective temperature corresponding to a non-LES system cannot be exactly derived,[7,8] highlighting the difficulty in choosing a single temperature for a standard LES simulation.

Notably, however, the LREMD surfaces are much smoother as compared to REMD and PREMD ones, with a shallower global minimum and nearly no barrier to transition to the native basin even at the highest temperature where it is still apparent (120 K). These results are not surprising given the mean-field approximation employed by LREMD and highlight the smoothing properties of LES that have led to its successful use in many previous studies.

We further demonstrate that the free energy landscape resulting from PREMD is consistent with our previous MD simulations of this RNA hairpin,[23] in which three distinct transition pathways from the incorrect to the correct conformation were observed. In Figure 11, we show the free energy surface calculated from the PREMD data, with white spheres showing the sampling of this landscape during representative standard MD simulations for the three transition pathways. The

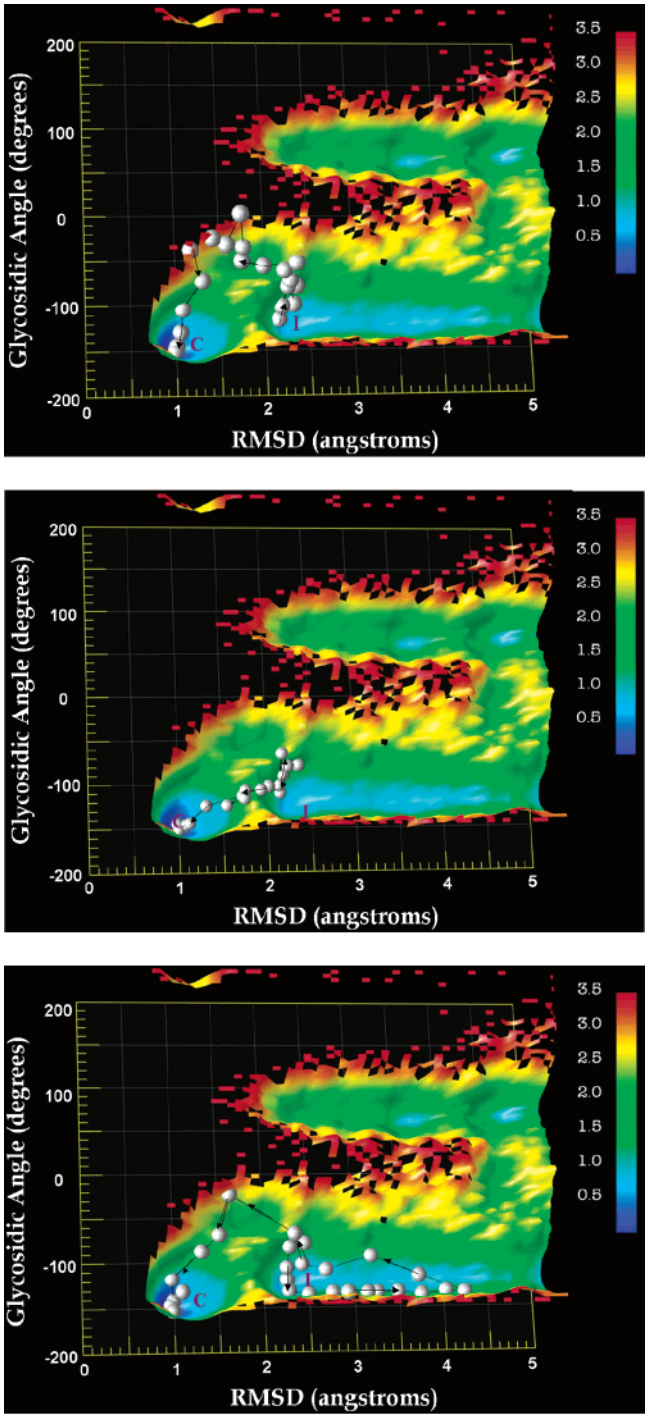

Figure 11. Projection of structures (white spheres) from three different transition pathways sampled during MD[23] onto the free energy surface constructed from PREMD simulation at 327 K. Reaction coordinates are the same as those in Figure 10, but the view is shifted to above the surface: (top) base-flipping pathway; (middle) minor reorganization pathway; (bottom) partial unfolding then refolding pathway. All three transition pathways from nonequilibrium MD are consistent with the PREMD-derived free energy landscape. The upper basin is not sampled in any of the three reorganization pathways (see text for further discussion).

first two pathways appear to be similar and sample only structures close to the incorrect and correct basins. The transitions involve direct crossing of the barrier between two minima, but careful examination shows that the two pathways cross at different locations. The first path involves slight flipping out of the U5 base, rotation of its glycosidic bond, and conversion to the correct structure. The second path only involves minor reorganization in the loop region with no flipping. On the basis of the PREMD energy landscape, we estimate the free energy barrier for the first path involving significant rotation of glycosidic bond to be ~1.7 kcal/mol at 327 K, while that for the second path to be ~2.2 kcal/mol. At lower temperatures, the barrier separating the basins is not populated enough to permit reliable estimation of the free energy, and techniques such as umbrella sampling may be more efficient (though determination of an appropriate reaction coordinate for the umbrella potential may also present challenges).

The third pathway is markedly different from the first two and samples a much broader region on the energy landscape. At an early point, it starts to show great deviation from both loop structures yet eventually adopts the correct structure. This indirect path is consistent with the broad basin of unfolded loop structures sampled in this region during PREMD.

We note that the basin in the upper region of the landscape is not sampled in any of the three reorganization pathways, yet it is also present in the standard REMD landscape. It appears to indicate an unproductive avenue through which the loop region adopts a conformation close to the correct one but with an incorrect U5 glycosidic angle. Rotation about this bond is impossible in the compact state, and reverting to the unfolded state would be required before the correct conformation could be adopted.

Conclusions

Replica exchange of many nontrivial systems has been demonstrated as a powerful tool to optimize structure, sample equilibrium thermodynamic quantities, and estimate pathways that could be sampled in nonequilibrium simulations (such as peptide folding). With the recent implementation of REMD in several MD simulation programs, the use of this method has been gaining popularity. However, one of the problems that prevent the method from straightforward extension to larger systems is due to the fact that when the system size increases, the number of replicas needed to effectively span a given temperature range also grows. This increase, along with the inherent growth in computational requirements for performing MD for each replica of the larger system, can rapidly make application of REMD prohibitive. To address this difficulty, we introduced two variant replica exchange methods called PREMD and LREMD, which employ a similar strategy to reduce the number of replicas needed for the simulation.

In both PREMD and LREMD, replica exchange was performed on part of the system after partitioning of the Hamiltonian. Since only the degrees of freedom of the focus or LES regions are coupled to the temperature space used in the calculation of exchange probability, the total number of replicas needed for either REMD variant is much smaller than that for standard REMD. For example, in a protein system of 4000 atoms where only a loop region of less than 100 atoms are of interest, we estimate that the number of replicas needed in PREMD or LREMD would be 6 times fewer than that for standard REMD.

LREMD is based on the idea of the LES method, representing a further approximation beyond PREMD. It is worth noting that the Hamiltonian of a LES system is scaled as compared to the original one, through which the energy surface is flattened through a mean-field effect to improve overall sampling. This smoothing is apparent in the free energy landscape we derived with LREMD. Therefore, it is impossible for LREMD to provide accurate quantities of thermodynamic properties. Otherwise, in comparison to PREMD, LREMD has several additional advantages. First, since the entire set of replicas shares the same non-LES region, a large fraction of the calculations that would be duplicated using standard REMD replicas need only be calculated once for all LREMD replicas. In other words, a full set of replica exchange results can be obtained at nearly no additional cost compared to a single, standard MD simulation. Another advantage of LREMD is inherited from the LES approximation. Since the energy surface of the LES system is smoother, it becomes easier for each copy to overcome energy barriers. This has a significant implication in the practical use of LREMD, which is that a smaller temperature range can be used with LREMD than those in standard REMD and PREMD, resulting in even fewer replicas needed to achieve the same sampling ability. LREMD now provides a more straightforward approach to application of LES, since a single optimal temperature for the LES region need not be chosen.

An important assumption used in both modified REMD methods is that the conformation of the majority of the system is relatively insensitive to the smaller region of interest to which replica exchange is applied. This is expected to be true in many cases of interest, such as the refinement of a portion of a structure for which data is missing or unreliable (as shown in this study) or for calculations of structure and/or thermodynamic properties of ligand−receptor binding (work in progress). In the latter case, inclusion of selected receptor atoms in the focus region would allow for coupling of the receptor to changes to ligand conformation.

Acknowledgment. C.S. thanks Adrian Roitberg for useful discussions. Support for this project was provided by the National Institutes of Health (GM6167803) and by the National Computational Science Alliance (MCA02N028), which provided computational resources at the NCSA. C.S. is a Cottrell Scholar of the Research Corporation.

References and Notes

(1) Straub, J. E.; Rashkin, A. B.; Thirumalai, D. *J. Am. Chem. Soc.* **1994**, *116*, 2049−2063.
(2) Berne, B. J.; Straub, J. E. *Curr. Opin. Struct. Biol.* **1997**, *7*, 181−189.
(3) Wolynes, P. G.; Onuchic, J. N.; Thirumalai, D. *Science* **1995**, *267*, 1619−1620.
(4) Tai, K. *Biophys. Chem.* **2004**, *107*, 213−220.
(5) Roitberg, A.; Simmerling, C. *J. Mol. Graphics Modell.* **2004**, *22*, 317−317.
(6) Elber, R.; Karplus, M. *J. Am. Chem. Soc.* **1990**, *112*, 9161−9175.
(7) Straub, J. E.; Karplus, M. *J. Chem. Phys.* **1991**, *94*, 6737−6739.
(8) Ulitsky, A.; Elber, R. *J. Chem. Phys.* **1993**, *98*, 3380−3388.
(9) Simmerling, C.; Elber, R. *J. Am. Chem. Soc.* **1994**, *116*, 2534−2547.
(10) Zheng, W. M.; Zheng, Q. *J. Chem. Phys.* **1997**, *106*, 1191−1194.
(11) Simmerling, C.; Miller, J. L.; Kollman, P. A. *J. Am. Chem. Soc.* **1998**, *120*, 7149−7155.
(12) Hixson, C. A.; Wheeler, R. A. *Phys. Rev. E* **2001**, *64*, 026701.
(13) Roitberg, A.; Elber, R. *J. Chem. Phys.* **1991**, *95*, 9277−9287.
(14) Stultz, C. M.; Karplus, M. *J. Chem. Phys.* **1998**, *109*, 8809−8815.
(15) Keasar, C.; Elber, R. *J. Phys. Chem.* **1995**, *99*, 11550−11556.
(16) Zheng, Q.; Rosenfeld, R.; Delisi, C.; Kyle, D. J. *Protein Sci.* **1994**, *3*, 493−506.
(17) Simmerling, C.; Fox, T.; Kollman, P. A. *J. Am. Chem. Soc.* **1998**, *120*, 5771−5782.
(18) Cui, G.; Simmerling, C. *J. Am. Chem. Soc.* **2002**, *124*, 12154−12164.

(19) Hornak, V.; Simmerling, C. *Proteins: Struct., Funct., Genet.* **2003**, *51*, 577.
(20) Verkhivker, G.; Elber, R.; Nowak, W. *J. Chem. Phys.* **1992**, *97*, 7838−7841.
(21) Joseph-McCarthy, D.; Tsang, S. K.; Filman, D. J.; Hogle, J. M.; Karplus, M. *J. Am. Chem. Soc.* **2001**, *123*, 12758−12769.
(22) Woodcock, H. L.; Hodoscek, M.; Sherwood, P.; Lee, Y. S.; Schaefer, H. F.; Brooks, B. R. *Theor. Chem. Acc.* **2003**, *109*, 140−148.
(23) Cheng, X. L.; Hornak, V.; Simmerling, C. *J. Phys. Chem. B* **2004**, *108*, 426−437.
(24) Neuhaus, B. A. B. a. T. *Phys Lett. B* **1991**, *267*, 249−253.
(25) Hansmann, U. H. E.; Okamoto, Y.; Eisenmenger, F. *Chem. Phys. Lett.* **1996**, *259*, 321−330.
(26) Lyubartsev, A. P.; Martsinovski, A. A.; Shevkunov, S. V.; Vorontsov-Velyaminov, P. N. *J Chem Phys* **1992**, *96*, 1776−1783.
(27) Mitsutake, A.; Sugita, Y.; Okamoto, Y. *Biopolymers* **2001**, *60*, 96−123.
(28) Swendsen, R. H.; Wang, J. S. *Phys. Rev. Lett.* **1986**, *57*, 2607−2609.
(29) Hansmann, U. H. E. *Chem. Phys. Lett.* **1997**, *281*, 140−150.
(30) Tesi, M. C.; van Rensburg, E. J. J.; Orlandini, E.; Whittington, S. G. *J. Stat. Phys.* **1996**, *82*, 155−181.
(31) Sugita, Y.; Okamoto, Y. *Chem. Phys. Lett.* **1999**, *314*, 141−151.
(32) Sugita, Y.; Kitao, A.; Okamoto, Y. *J. Chem. Phys.* **2000**, *113*, 6042−6051.
(33) Garcia, A. E.; Sanbonmatsu, K. Y. *Proteins* **2001**, *42*, 345−354.
(34) Zhou, R. H.; Berne, B. J.; Germain, R. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 14931−14936.
(35) Garcia, A. E.; Sanbonmatsu, K. Y. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 2782−2787.
(36) Rhee YM, P. V. *Biophys. J.* **2003**, *84*, 775−786.
(37) Pitera, J. W.; Swope, W. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 7587−7592.
(38) Karanicolas, J.; Brooks, C. L. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 3954−3959.
(39) Feig, M.; Karanicolas, J.; Brooks, C. L. *J. Mol. Graphics Modell.* **2004**, *22*, 377−395.
(40) Nymeyer, H.; Garcia, A. E.; Woolf, T. B. *Biophys. J.* **2003**, *84*, 381A.
(41) Kinnear, B. S.; Jarrold, M. F.; Hansmann, U. H. E. *J. Mol. Graphics Modell.* **2004**, *22*, 397−403.
(42) Fukunishi H, W. O., Takada S *J. Chem. Phys.* **2002**, *116*, 9058−9067.
(43) Sugita, Y.; Okamoto, Y. *Chem. Phys. Lett.* **2000**, *329*, 261−270.
(44) Mitsutake A, S. Y., Okamoto Y *J. Chem. Phys.* **2003**, *118*, 6664−6688.
(45) Jang, S.; Shin, S.; Pak, Y. *Phys. Rev. Lett.* **2003**, *91*, 58305.
(46) Case, D. A.; Darden, T. A.; Cheatham, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Wang, B.; Pearlman, D. A.; Crowley, M.; Brozell, S.; Tsui, V.; Gohlke, H.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Schafmeister, C.; Caldwell, J. A.; Ross, W. S.; Kollman, P. A. *AMBER*, version 9; University of California: San Francisco, 2004.
(47) Kofke, D. A. *J. Chem. Phys.* **2002**, *117*, 6911−6914.
(48) Kofke, D. A. *J. Chem. Phys.* **2004**, *120*, 10852−10852.
(49) Varani, G.; Cheong, G.; Tinoco, I. J. *Biochemistry* **1991**, *30*, 3280−3289.
(50) Allain, F. H. T.; Varani, G. *J. Mol. Biol.* **1995**, *250*, 333−353.
(51) Miller, J. L.; Kollman, P. A. *J. Mol. Biol.* **1997**, *270*, 436−450.
(52) Williams, D. J.; Hall, K. B. *Biophys. J.* **1999**, *76*, 3192−3205.
(53) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127−6129.
(54) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(55) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.
(56) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Di Nola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(57) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(58) Bondi, A. *J. Phys. Chem.* **1964**, *68*, 441−451.
(59) Tsui, V.; Case, D. A. *J. Am. Chem. Soc.* **2000**, *122*, 2489−2498.
(60) Ponder, J. W.; Richards, F. M. *J. Comput. Chem.* **1987**, *8*, 1016−1024.