THE JOURNAL OF PHYSICAL CHEMISTRY B

ARTICLE

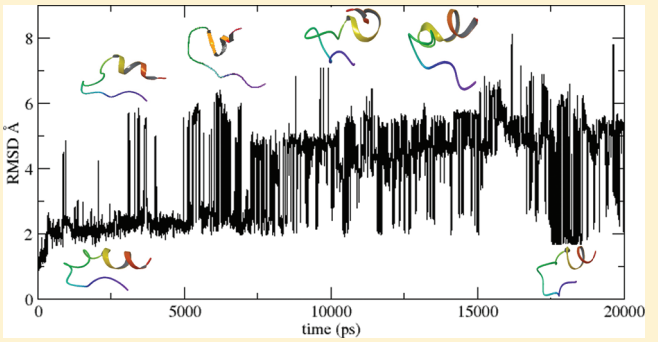
pubs.acs.org/JPCB

Replica Exchange with Solute Scaling: A More Efficient Version of Replica Exchange with Solute Tempering (REST2)

Lingle Wang, Richard A. Friesner, and B. J. Berne*

Department of Chemistry, Columbia University, 3000 Broadway, New York, New York 10027, United States

ABSTRACT: A small change in the Hamiltonian scaling in Replica Exchange with Solute Tempering (REST) is found to improve its sampling efficiency greatly, especially for the sampling of aqueous protein solutions in which there are large-scale solute conformation changes. Like the original REST (REST1), the new version (which we call REST2) also bypasses the poor scaling with system size of the standard Temperature Replica Exchange Method (TREM), reducing the number of replicas (parallel processes) from what must be used in TREM. This reduction is accomplished by deforming the Hamiltonian function for each replica in such a way that the acceptance probability for the exchange of replica configurations does not depend on the number of explicit water molecules in the system. For proof of concept, REST2 is compared with TREM and with REST1 for the folding of the trpcage and β-hairpin in water. The comparisons confirm that REST2 greatly reduces the number of CPUs required by regular replica exchange and greatly increases the sampling efficiency over REST1. This method reduces the CPU time required for calculating thermodynamic averages and for the ab initio folding of proteins in explicit water.



INTRODUCTION

Sampling the conformational space of complex biophysical systems, such as proteins, remains a significant challenge, because the barriers separating the local energy minima are usually much higher than k_BT, leading to kinetic "trapping" for long periods of time and quasi-ergodicity in the simulations. The Temperature Replica Exchange Method (TREM) has attracted attention recently as a means for overcoming the problem of quasi-ergodicity.[1−6] However, the number of replicas required to get efficient sampling in normal TREM scales as $f^{1/2}$, where f is the number of degrees of freedom of the whole system, which often limits the applicability of TREM for large systems. To overcome this problem, we recently devised the method "Replica Exchange with Solute Tempering" (REST1),[7] in which only the solute biomolecule is effectively heated up while the solvent remains cold in higher temperature replicas, so that the number of the replicas required is greatly reduced. It has been shown that the required number of replicas in REST1 scales as $(f_p)^{1/2}$, where f_p is the number of degrees of freedom of the solute, and the speedup versus the TREM, in terms of converging to the correct underlying distribution, is $O(f/f_p)^{1/2}$ for small solutes such as alanine dipeptide.[7] However, when applying REST1 to large systems involving large conformational changes, such as the trpcage and β hairpin, it was found that REST1 can be less efficient than TREM.[8] For example, we observed that the lower temperature replicas stayed in the folded structure, the higher temperature replicas stayed in the extended structure, and the exchange between those two conformations was very low.[8] Moors et al.[9] and Terakawa et al.[10] independently modified our REST1 scaling factor for E_{pw} so that the approach could be easily run in GROMACS. Moors et al. included only part of the protein in the "hot region", keeping the rest of it "cold", and called their method "Replica Exchange with Flexible Tempering" (REFT). Interestingly, they observed an improved sampling efficiency in sampling a particular reaction coordinate involving the opening and closing of the binding pocket in T4 lysozyme and suggested that the improved sampling efficiency for their method over REST1 occurred because in REST1 all of the protein degrees of freedom contribute to the acceptance probability for replica exchange, whereas in REFT only those degrees of freedom involved in the opening and closing of the pocket contribute. Thus the acceptance probability for replica exchange is larger in REFT than in REST1. As we shall see, this is not the only reason for the observed improvement.

In this paper we use the modified scaling of the Hamiltonians suggested by Moors et al.[9] and Terakawa et al.[10] instead of the original scaling of our REST1, to see if it samples the folded and unfolded conformations of proteins more efficiently than REST1, although all of the protein degrees of freedom are allowed to be hot in this study. For simplicity, we call REST with this new scaling REST2. Application of REST2 to the trpcage and the β-hairpin systems, the same systems that were problematic when sampled by REST1, indicates that REST2 is much more efficient than REST1 in sampling the conformational

Received: May 11, 2011
Revised: June 16, 2011
Published: June 29, 2011

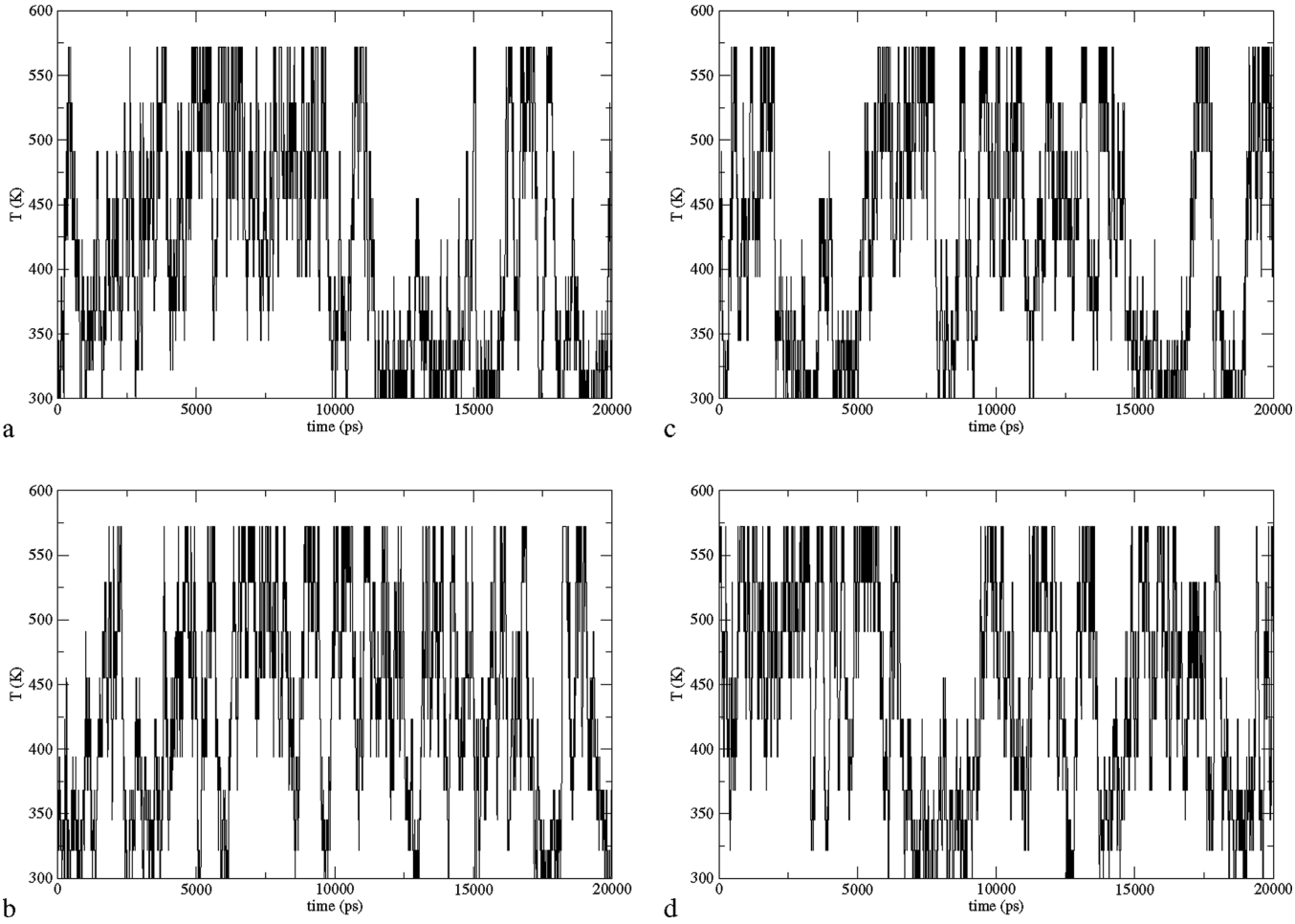


Figure 1. Temperature trajectories of four representative replicas with the effective temperature of the protein started at 300 K (a), 368 K (b), 455 K (c), and 572 K (d) for the trpcage system starting from the native structure. It should be noted that the temperatures referred to are the effective temperatures of the protein, which arise from the scaling of the force field parameters of the protein, while the actual simulation is done at temperature T_0.

space of large systems undergoing large conformation changes. In what follows, we will present the scaling, its connection with our original scaling of REST1, and the results for the trpcage and β-hairpin systems for REST1, REST2, and TREM. We will also discuss the reasons for the improvement found with REST2.

METHODOLOGY

In REST1, the total interaction energy of the system was decomposed into three components: the protein intramolecular energy, E_{pp}; the interaction energy between the protein and water, E_{pw}; and the self-interaction energy between water molecules, E_{ww}. Replicas running at different temperatures then evolve through different Hamiltonians involving relative scalings of these three components. To be specific, the replica running at temperature T_m has the following potential energy:

$$E_m^{\mathrm{REST1}}(X) = E_{pp}(X) + \frac{\beta_0 + \beta_m}{2\beta_m} E_{pw}(X) + \frac{\beta_0}{\beta_m} E_{ww}(X) \quad (1)$$

Here, X represents the configuration of the whole system, $\beta_m = 1/k_B T_m$, and T_0 is the temperature in which we are interested. The potential for replica running at T_0 reduces to the normal potential.

Imposing the detailed balance condition, the acceptance ratio for the exchange between two replicas m and n depends on the following energy difference:

$$\Delta_{mn}(\mathrm{REST1}) = (\beta_m - \beta_n)\left[\left(E_{pp}(X_n) + \frac{1}{2} E_{pw}(X_n)\right) - \left(E_{pp}(X_m) + \frac{1}{2} E_{pw}(X_m)\right)\right] \quad (2)$$

Note that the water self-interaction energy, E_{ww}, does not appear in the acceptance ratio formula, and this is the reason only a relatively small number of replicas are sufficient to achieve good exchange probabilities in REST1.

In REST1, both the potential energy and the temperature are different for different replicas. According to the law of corresponding states, the thermodynamic properties of a system with potential energy E_m at temperature T_m, are the same as those for a system with potential energy $(T_0/T_m)E_m$ at temperature T_0. So instead of using different potential energies and different temperatures for different replicas, we can run all the replicas at the same temperature albeit on different potential energy surfaces using the Hamiltonian Replica Exchange Method (H-REM).[11,12]

To be specific, in REST2, all of the replicas are run at the same temperature T_0, but the potential energy for replica m is scaled differently:

$$E_m^{\mathrm{REST2}}(X) = \frac{\beta_m}{\beta_0} E_{\mathrm{pp}}(X) + \sqrt{\frac{\beta_m}{\beta_0}} E_{\mathrm{pw}}(X) + E_{\mathrm{ww}}(X) \quad (3)$$

In REST1, enhanced sampling of the protein conformations is achieved by increasing the temperature of the protein, but between attempted exchanges with neighboring replicas, replica m moves on the full intramolecular protein potential energy surface with high energy barriers, although the other energy terms are scaled. In REST2, enhanced sampling is achieved through scaling the intramolecular potential energy of the protein by (β_m/β_0), a number smaller than 1, so that the barriers separating different conformations are lowered. Thus between attempted replica exchanges, replica m moves on a modified potential surface where the barriers in the intra protein force field are reduced by the scaling. We call REST with this new scaling "Replica Exchange with Solute Scaling" (REST2). Thus REST1 and REST2 arrive at the final distribution at temperature T_0 by different but rigorously correct routes. The acceptance criteria for replica exchanges are different in REST1 and REST2, but the Hamiltonians for the molecular dynamics (MD) trajectories are also different in such a way that the long-time sampling at T_0 should converge to the same ensemble for REST1 and REST2, albeit with different rates of convergence for the two methods. In REST2, the differences between different replicas are the different scaling factors used, but to make connections with REST1, we will keep using the term "temperature" for replica m, which means the effective temperature of the protein with the unscaled potential energy.

Note that the scaling factor used in REST2 for the interaction energy between the solute and water for replica m is $(\beta_m/\beta_0)^{1/2}$, which is different from $(\beta_0 + \beta_m)/2\beta_0$ used in REST1 (eq 1). The interaction energy in eq 3 can be easily achieved by scaling the bonded interaction energy terms, the Lennard-Jones (LJ) ε parameters, and the charges of the solute atoms by (β_m/β_0), (β_m/β_0), and $(\beta_m/\beta_0)^{1/2}$, respectively, and the scaling factor for the E_{pw} term, $(\beta_m/\beta_0)^{1/2}$, follows naturally from standard combination rules for LJ interactions. This minor change of the scaling factor for the E_{pw} term, suggested in the original REST paper but not appreciated at that time, proves to be important for the better performance of the REST2. In addition, we find that scaling the bond stretch and bond angle terms does not help the sampling, so in practice only the dihedral angle terms in the bonded interaction of the solute are scaled, and this makes the transition between different conformations of the solute faster.

Another consequence of the different scaling factors used for the E_{pw} term in REST1 and REST2 is the different acceptance ratio formulas in these two methods. It is easy to show by imposing detailed balance conditions that the acceptance ratio for exchange between replicas m and n in REST2 is determined by

$$\Delta_{mn}(\mathrm{REST2}) = (\beta_m - \beta_n)\left[(E_{\mathrm{pp}}(X_n) - E_{\mathrm{pp}}(X_m)) + \frac{\sqrt{\beta_0}}{\sqrt{\beta_m} + \sqrt{\beta_n}}(E_{\mathrm{pw}}(X_n) - E_{\mathrm{pw}}(X_m))\right] \quad (4)$$

For replica m, the exchanges to neighboring replicas $m - 1$ and $m + 1$, are determined by the fluctuation of $E_{\mathrm{pp}} + (\beta_0)^{1/2}/((\beta_m)^{1/2} + (\beta_{m-1})^{1/2})E_{\mathrm{pw}}$ and $E_{\mathrm{pp}} + (\beta_0)^{1/2}/((\beta_m)^{1/2} + (\beta_{m+1})^{1/2})E_{\mathrm{pw}}$, respectively. Thus for discussion purposes, but not in the simulations, the fluctuation of $E_{\mathrm{pp}} + (1/2)(\beta_0/\beta_m)^{1/2} E_{\mathrm{pw}}$ can be thought to determine the acceptance ratios for exchanges of the replica at temperature T_m to neighboring replicas because, to a good approximation, $\beta_{m-1} \approx \beta_m \approx \beta_{m+1}$. Note then that the difference in the acceptance ratio formulas between REST1 and REST2 lies in the replacement of the factor 1/2 by the factor $(1/2)(\beta_0/\beta_m)^{1/2}$ multiplying the term E_{pw}. This difference is also partly responsible for the improvement of REST2 over REST1 due to an approximate cancellation of E_{pp} and the scaled E_{pw} in the acceptance probability or equivalently in Δ_{nm} of REST2 but not in REST1, as we shall see.

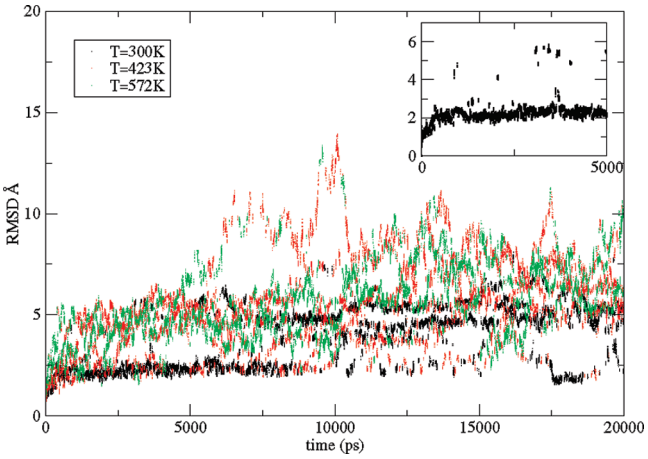


Figure 2. Protein heavy atom rms deviation from the native structure as a function of simulation time for replicas with different effective temperatures of the protein for the trpcage system. Inset of the figure highlights the rmsd for replica at effective temperature 300 K in the first 5 ns of the simulation.

RESULTS AND DISCUSSION

Using REST2, we simulated the trpcage system with DESMOND[13] using 10 replicas with effective temperatures of the solute at 300, 322, 345, 368, 394, 423, 455, 491, 529, and 572 K. The OPLS-AA force field[14] was used for the protein, and the Tip4p model[15] was used for water. All the replicas were started from the "native" NMR structure (PDB ID 1L2Y)[16] and the simulation lasted for 20 ns. Conformations of the protein were saved every 0.5 ps, and exchange of configurations between neighboring replicas are attempted every 2 ps with an average acceptance ratio of about 30%.

Four representative temperature trajectories for the trpcage replicas started at 300, 368, 455, and 572 K in the folded state are displayed in Figure 1. It can be seen that the temperature trajectory for each replica visits all of the temperatures many times, even during the first 5 ns of the simulation, and all of the replicas visit any given temperature many times during the simulation. This is a good indication of the efficiency of the sampling. By comparison, none of the temperature trajectories using REST1 were able to visit all of the temperatures during a 5 ns simulation for the same system (see Figure 6b in ref 8). In REST2 the time interval for attempted exchange was 2 ps, while in the REST1 simulation 0.4 ps was used. We expect that even more rapid diffusion in temperature space could be achieved if

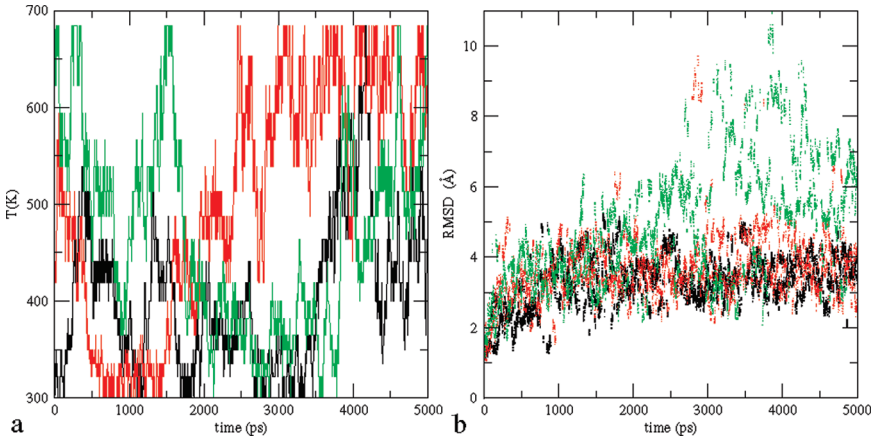


Figure 3. (a) Temperature trajectories for three representative replicas with the effective temperature of the protein initially at low (T = 310 K), intermediate (T = 419 K), and high (T = 684 K) temperatures for the β-hairpin system. (b) The protein heavy atom rmsd versus time at each of the above temperatures when replicas visit those temperatures (black, T = 310 K; red, T = 419 K; green, T = 684 K).

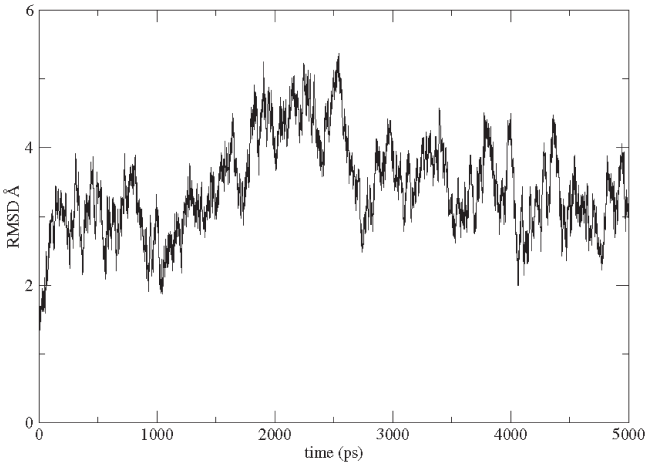


Figure 4. Heavy atom rmsd from the native structure of the β-hairpin as a function of simulation time with the effective temperature of the protein at 600 K using the scaled Hamiltonian of REST2 without attempted replica exchanges. Both rmsd > 4 Å and < 4 Å are sampled; by comparison, in REST1 only rmsd > 4 Å is sampled at high temperatures (Figure 4 of ref 7).

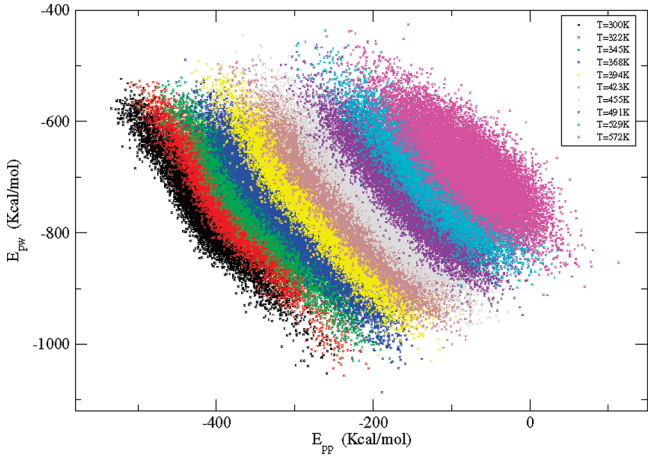


Figure 5. Anticorrelation between the intramolecular potential energy of the protein and the interaction energy between the protein and water for replicas with different effective temperatures of the protein for the trpcage system.

shorter time intervals between attempted exchanges were used in REST2.

In the REST1 simulations, it was observed that only the folded structures were sampled at the lower temperatures while the folded structures were rarely sampled at higher temperatures such as 572 K after an initial equilibration phase (see Figure 7a in ref 8). The REST2 simulation of the protein heavy atom deviation (root-mean-square deviation (rmsd)) from the native structure is displayed in Figure 2 for replicas with effective temperature of the protein at 300, 423, and 572 K. It is clear that both the folded structure and the unfolded structures are sampled at 300 K even in the first 5 ns of the simulation (inset of Figure 2). At the intermediate temperature (423 K), the folded and unfolded structures are sampled with almost equal probability, and the unfolded structures dominate at high temperature (572 K). However, unlike in REST1, the folded structures are also sampled at 572 K after the initial equilibration phase.

The β-hairpin system is likewise more efficiently sampled by REST2. For the same number of replicas and the same temperature levels used in REST1,[8] the temperature trajectories for three representative replicas, initially at low (T = 310 K), intermediate (T = 419 K), and high (T = 684 K) temperatures, are shown in Figure 3a. We also determined the protein heavy atom rmsd versus time at each of the above temperatures when replicas visited those temperatures that are shown in Figure 3b. With a time interval of 2 ps for attempted exchange, each replica is able to visit all the temperatures within 5 ns, and both the folded and unfolded structures are sampled at low and high temperatures. By comparison, using REST1, none of the replicas were able to visit all the temperatures (see Figure 3b in ref 8), whereas the low temperature replicas stayed folded and the high temperature replicas stayed unfolded after the initial stage (see Figure 4 in ref 8). Thus REST2 is clearly superior to REST1 for both the trpcage and the β-hairpin.

The different scaling factors used for the E_{pw} term in REST1 and REST2 are responsible for the improvement of REST2 over REST1 as expected from the discussion given in the

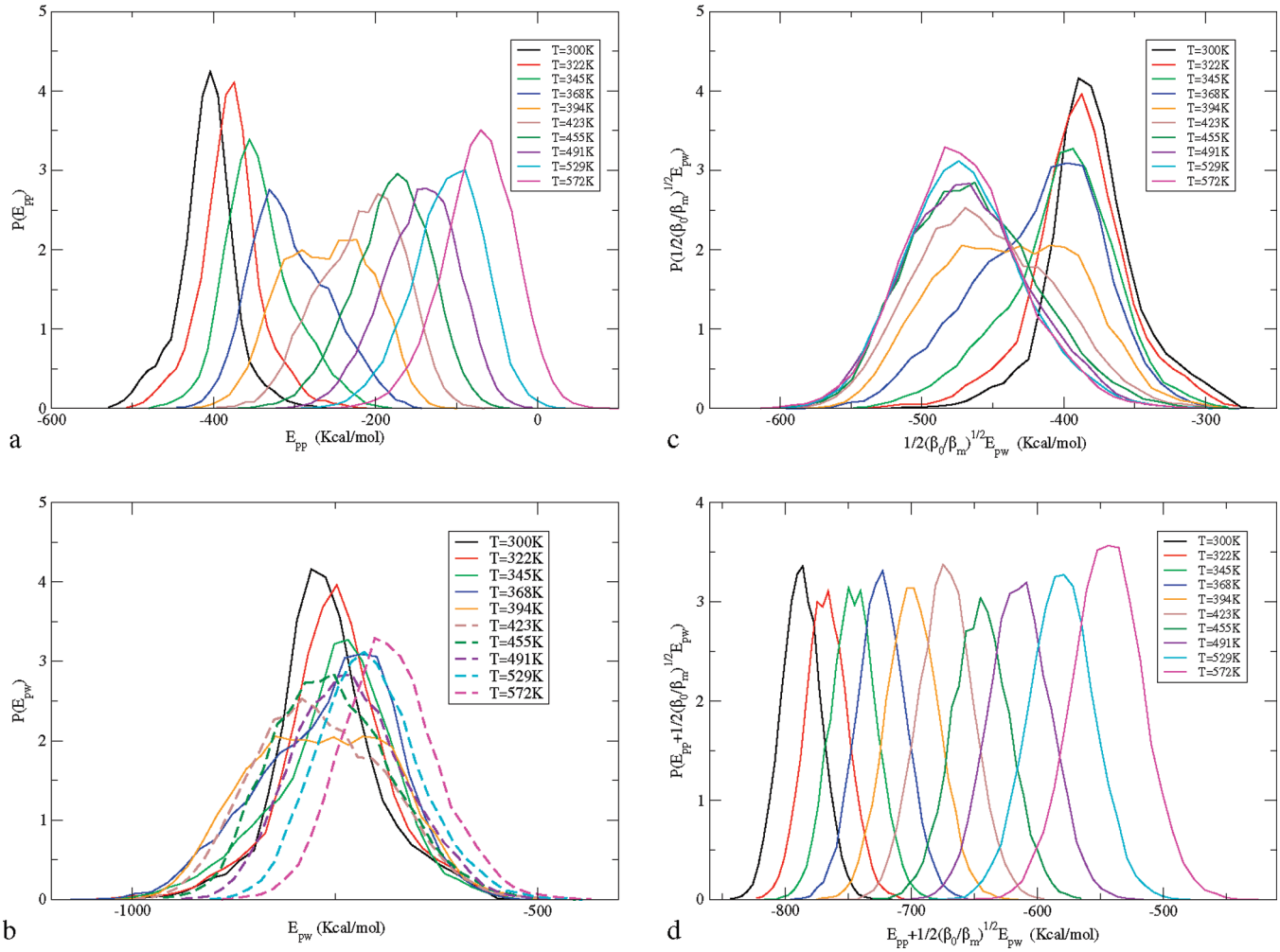


Figure 6. (a) Distribution of intramolecular potential energy of the protein for replicas with different effective temperatures of the protein. (b) Distribution of interaction energy between protein and water for replicas with different effective temperatures of the protein. (c) Distribution of $(1/2)(\beta_0/\beta_m)^{1/2}E_{pw}$ for replicas with different effective temperatures of the protein. (d) Distribution of $E_{pp} + (1/2)(\beta_0/\beta_m)^{1/2}E_{pw}$ for replicas with different effective temperatures of the protein.

previous section. Consider the constant temperature MD trajectory between attempted replica exchanges. In REST1, the scaling factor for the E_{pw} term was $(\beta_0 + \beta_m)/2\beta_m$. In the limit when $T_m \rightarrow \infty$, REST1 will effectively sample the distribution $\exp(-\beta_0 (E_{pw}/2 + E_{ww}))$. Since the unfolded structure has more favorable solute water interactions than the folded structure, replicas at higher temperature will sample the unfolded structure with dominating probability in REST1, and this was indeed observed in REST1 simulations for the trpcage as well as for the β-hairpin.[8] In REST1, the replicas at high temperatures can not sample the whole conformational space efficiently, and replicas at high and low temperatures sample completely different regions of conformation space. This is one of the reasons for the observed inefficient sampling in REST1. By comparison, in REST2, we use a scaling factor $(\beta_m/\beta_0)^{1/2}$ for the E_{pw} term. In the limit when $T_m \rightarrow \infty$, REST2 will effectively sample the distribution $\exp(-\beta_0 E_{ww})$. So both the folded and unfolded structures are sampled efficiently during the trajectories between attempted replica exchanges for the higher temperature replicas in REST2, and this is one of the reasons why REST2 is more efficient than REST1. This is shown in Figure 4, where it can be seen that in a constant temperature MD simulation using the scaled Hamiltonian of REST2 at high temperature, the heavy atom rmsd for the β-hairpin fluctuates from the native structure from values close to 2.5 Å to 5.0 Å and back again to 2.5 Å, whereas in Figure 4 in ref 8, it stayed above 4 Å after short times.

Figure 5 displays the relation between the intramolecular potential energy of the trpcage system, E_{pp}, and the interaction energy between the trpcage and water, E_{pw}, for replicas with different effective temperatures of the protein in REST2. At each temperature, there is a strong anticorrelation between those two terms. This is easy to understand: the more extended the structure is, the less favorable the intramolecular potential energy of the protein, and the more favorable the interaction energy between the protein and water (which scales with the surface area of the protein). With increasing temperature, the probability for the unfolded structure gets larger, and the intramolecular potential energy of the protein becomes less negative. For the interaction energy between the protein and water, there are two counterbalancing effects. On one hand, the higher the temperature, the more favorable the unfolded structure, and the more favorable the interaction between water and protein. On the

dx.doi.org/10.1021/jp204407d |J. Phys. Chem. B 2011, 115, 9431–9438

other hand, every single component of the potential energy would increase with increasing temperature because of the generalized equi-partition theorem. This is exactly what we observe in Figure 5: with increasing temperature, the E_{pp} term becomes less negative, while the E_{pw} term increases very slowly because of the compensation of the two effects mentioned above. This is clearly demonstrated in Figure 6a,b, where the distribution of E_{pp} and E_{pw} are displayed for replicas at different temperatures. At 394 K, both E_{pp} and E_{pw} are binomially distributed with the folded and unfolded structures with almost equal probability. Below 394 K, the folded structure dominates, and above 394 K the unfolded structure dominates. This is the reason why replicas running below 394 K and above 394 K were not able to exchange efficiently in REST1.[8] While E_{pp} increases monotonically with increasing temperature, the behavior of E_{pw} is more complicated. Below 394 K, the center of the distribution for E_{pw} gets less negative, but the distribution gets broader in the left tail of the distribution because of the increased probability of extended structures. Above 394 K, E_{pw} increases with temperature because of the equipartition theorem.

The absence of a compensating term E_{ww} in the replica exchange probability of REST1 was suggested in our previous paper[8] to explain the observed better performance of TREM for the exchange between folded and unfolded structures where there is a big difference in the energies of these two states,[8] but in the Appendix we show why we now think that this is not the reason for this difference. Actually the compensation or lack of compensation between E_{pp} and the scaled E_{pw} is more important than the loss of any compensation between E_{ww} and E_{pp}.

In REST1, it is the fluctuation of $E_{pp} + (1/2)E_{pw}$ that determines the acceptance ratio. While the two terms can compensate each other to some extent, they both tend to increase with increasing temperature of the solute. By comparison, in REST2, it is the fluctuation of $E_{pp} + (1/2)(\beta_0/\beta_m)^{1/2}E_{pw}$ that determines the acceptance ratio. Since $(\beta_0/\beta_m)^{1/2}$ increases with increasing temperature of the solute, it will compensate the decrease of the magnitude of E_{pw}. (The E_{pw} term is negative, and the magnitude of it decreases with increasing temperature.) Figure 6c displays the distribution of $(1/2)(\beta_0/\beta_m)^{1/2}E_{pw}$ for replicas at different temperatures. It is quite clear that the factor $(\beta_0/\beta_m)^{1/2}$ perfectly compensates the increase of E_{pw}. Below 394 K, the distribution is centered at about −380 kcal/mol corresponding to the folded structure; above 394 K, the distribution is centered at about −495 kcal/mol corresponding to the unfolded structure. At 394 K, the folded and unfolded structures are almost equally distributed. With increasing temperature, the probability of the unfolded structure increases and the probability of folded structure decreases. The difference in the E_{pp} term between the folded and unfolded structures is compensated by the difference in $(1/2)(\beta_0/\beta_m)^{1/2}E_{pw}$ term, which makes the distribution of $E_{pp} + (1/2)(\beta_0/\beta_m)^{1/2}E_{pw}$ have sufficient overlap for neighboring replicas (Figure 6d). The approximate cancellation of the contributions E_{pp} and E_{pw} in REST2 and their smaller cancellation in REST1 makes the acceptance ratio for replica exchange larger in REST2 than in REST1, and this is part of the reason for the more efficient sampling in REST2 than in REST1. For the trpcage system studied here, with the same number of replicas and the same temperature levels, we obtained an average acceptance ratio for REST2 of 30%, while in REST1 only 20% was obtained. In addition, the more frequent barrier crossings in the MD trajectories of REST2 than in REST1 contributes considerably to the better efficiency of REST2.

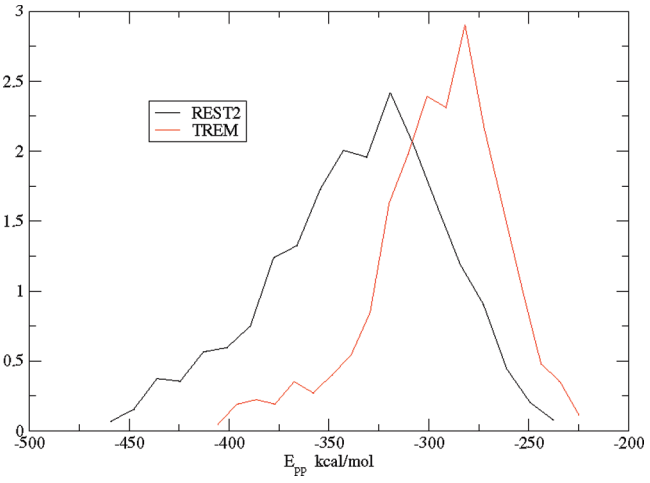


Figure 7. The distribution of intramolecular potential energy of the protein at the lowest temperature replica using TREM and REST2 starting from an almost fully extended structure.

As mentioned, the rate of convergence of REST1 (relative to TREM) to the correct underlying distribution was shown to scale as $O((f/f_p)^{1/2})$ for small solutes such as alanine dipeptide; however, for systems involving large conformation changes, REST1 fails to achieve this expected speed up.[7] The results presented in the above sections clearly demonstrate that REST2 is much more efficient than REST1 for sampling systems with large conformational change, but does REST2 do better compared to TREM for these problematic systems? To answer this question, we simulated the trpcage system starting from an almost fully extended configuration using both TREM and REST2. As before, 10 replicas were used for REST2, and 48 replicas were needed in TREM to maintain an appropriate acceptance ratio. It should be noted that the replica exchange ratio for TREM is 10%, whereas for REST2 it is 30% so that we could have used fewer replicas in REST2 to get the same exchange ratio as in TREM. Within 2 ns simulations, none of the replicas in TREM were able to visit all the temperatures, while in REST2 all 10 replicas were able to visit all of the temperatures, indicating that REST2 is much more efficient in diffusing through temperature space than TREM. The distribution of intramolecular energy of the trpcage for the lowest level replica calculated from TREM and REST2 is shown in Figure 7. For trajectories of the same length, REST2 samples a broader region in conformation space than TREM, and in addition the CPU cost of generating equal length trajectories is greater for TREM than for REST2 (see the Appendices).

CONCLUSION

We find that REST2 more efficiently samples the conformation space than REST1. We used a different scaling factor for the interaction energy between the protein and water, E_{pw}, than we used in REST1. Application of REST2 to the trpcage and β-hairpin systems results in an improvement over REST1 in sampling large systems involving large conformational energy changes. The better efficiency of REST2 over REST1 arises because there is a greater cancellation between the scaled terms E_{pp} and E_{pw} in REST2 than in REST1. This gives rise to REST2's larger replica exchange probability than REST1's, and also to its better sampling between replica exchanges at high

temperature, as we now discuss. For example, for $T_0 = 300$ K and $T_m = 600$ K, the deformed potential for REST2 is $E_m^{(\mathrm{REST2})} = 0.5E_{pp} + 0.71E_{pw} + E_{ww}$, but it is run at $T_0 = 300$ K, whereas for REST1 it is $E_m^{(\mathrm{REST2})} = E_{pp} + 1.5E_{pw} + 2E_{ww}$, but it is run at $T_m = 600$ K. The exponents in the Boltzmann factors for these two cases, are

$$\beta_0 E_m^{\mathrm{REST2}} = \beta_m[E_{pp} + 1.41E_{pw} + 2.0E_{ww}]$$

and

$$\beta_m E_m^{(\mathrm{REST1})} = \beta_m[E_{pp} + 1.5E_{pw} + 2.0E_{ww}]$$

The only difference between these is due to the different scaling factors of the E_{pw} term, which for $T_0 = 300$ K and $T_{max} = 600$ K, is $(1/2)(\beta_0 + \beta_m)/\beta_m = 1.5$ versus $(\beta_0/\beta_m)^{1/2} \approx 1.41$. We have seen in Figure 5 that the E_{pw} term is usually much larger in magnitude than the E_{pp} term, so a small change in the scaling factor of the E_{pw} term leads to better sampling efficiencies for the high temperature MD between replica exchanges for REST2 than REST1. For the trpcage and β-hairpin systems studied here, in REST2 both folded and unfolded conformations are sampled at higher temperature replicas, whereas in REST1 only the unfolded conformational space of the solute are sampled at higher temperature replicas. Because of the larger replica exchange probability and because of better constant temperature sampling, these folded and unfolded conformations filter down to the replica at the temperature of interest, T_0. In addition, since all the replicas are running at the same temperature in REST2, there is no need to rescale the velocity during the exchange process, which will save some computer time and makes it easier to implement in various MD programs. We also found REST2 to be more efficient in sampling the trpcage than TREM because of the much smaller number of replicas and faster CPU times required to generate the MD trajectories in REST2 compared to TREM. In addition, the lowest level replica was found to explore a larger region of energy space for REST2 than for TREM for the same MD trajectory lengths for each replica. Thus, REST2 speeds up the sampling of the trpcage, by at least a factor of 9.6 over TREM.

We believe that REST2 should be used for investigating large protein−water systems especially when there are large conformation energy changes in the protein. The improvement comes from (a) the larger replica exchange probabilities and concomitantly the smaller number of replicas that can be used, and (b) the more efficient MD sampling of the conformational states between replica exchanges on the upper replica potential energy surfaces.

APPENDIX 1: CAN TREM BE MORE EFFICIENT THAN REST1?

In a previous paper we noted that REST1 can sometimes be less efficient than TREM in systems in which there are large conformational energy changes between folded and unfolded structures. We attributed this to the absence of E_{ww} in the REST1 acceptance ratio formula (eq 2), a term that might be able to compensate for the large differences of $(E_{pp} + (1/2)E_{pw})$ between the folded and unfolded conformations.[8] On further analysis it appears that this may not be the reason for this behavior of REST1 in these systems. On one hand, the acceptance ratio for an exchange from a folded structure to an unfolded structure is much lower in REST1 than in normal TREM if $\Delta(E_{pp} + (1/2)E_{pw})$ is much larger than $\Delta(E_{pp} + E_{pw} + E_{ww})$ between the folded and unfolded structures. (The acceptance ratio in normal TREM depends on $\Delta(E_{pp} + E_{pw} + E_{ww})$.) On the other hand, the unfolded structure is sampled with much larger probability in REST1 than in normal TREM at higher temperatures. This is expected because the potential energy for water is scaled in REST1 making the unfolded structure more favorable. If high temperature replicas sample the whole conformation space efficiently, both unfolded and folded structures should be observed. Detailed balance then shows that the unfolding and folding rates at T_0 for TREM and REST1 will be identical and the correct distribution at T_0 should be maintained for TREM and REST1. So the absence of E_{ww} is not responsible for the inefficient sampling of REST versus TREM as was thought previously. The problem is that in REST1 the replicas at higher temperatures sample the unfolded structure with dominating probability and the overlap of conformation space for replicas at lower and higher temperatures is very small. This is not a problem in REST2 where there is a smaller scaling factor of the E_{pw} term.

APPENDIX 2: WHY REST2 IS MORE EFFICIENT THAN TREM

There are two important factors that make REST2 more efficient than TREM. First, TREM uses far more replicas than REST2. Second, the higher temperature trajectories in TREM are much slower with respect to CPU time than in REST2. This occurs because in the higher levels the atoms in TREM move much faster than in the higher levels of REST2, thus requiring that, for the same skin thickness, its nearest neighbor list must be updated much more frequently (or alternatively, for fixed update frequency, the skin thickness must be increased). In either case, the TREM trajectory requires longer CPU times. Because replica exchange is attempted at a constant time interval, the speed in TREM is limited by the longer CPU times for the high temperature replicas. For example, for benchmark MD trajectories of the same duration run for the trpcage at 600 K and at 300 K, using DESMOND, the 300 K trajectory requires half the CPU time that the 600 K trajectory requires. Thus we save a factor of 48/10 from the smaller number of replicas and a factor of 2 for each trajectory (from the above difference in CPU times for trajectories of the same length) for a total speed up of at least $4.8 \times 2 = 9.6$ of REST2 over TERM for the trpcage. Although TREM and REST2 are rigorous sampling methods, TREM will converge much more slowly in CPU time than REST2.

ACKNOWLEDGMENT

We thank Dr. Byungchan Kim of Schrodinger, Inc. for showing us how to run REMD using DESMOND. This work was supported by NIH grants to B.J.B. (NIH GM 43340) and to R.A.F. (NIH GM 40526). B.J.B. and R.A.F. acknowledge that this work was also supported in part by the National Science Foundation through TeraGrid resources provided by NCSA and ABE (MCA08X002).

REFERENCES

(1) Swendsen, R. H.; Wang, J. S. *Phys. Rev. Lett.* **1986**, *57* (21), 2607–2609.

(2) Hukushima, K.; Nemoto, K. *J. Phys. Soc. Jpn.* **1996**, *65* (6), 1604–1608.

(3) Garcia, A. E.; Sanbonmatsu, K. Y. *Proteins: Struct., Funct., Bioinf.* **2001**, *42* (3), 345–354.
(4) Zhou, R.; Berne, B. J.; Germain, R. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98* (26), 14931–14936.
(5) Rhee, Y. M.; Pande, V. S. *Biophys. J.* **2003**, *84* (2), 775–786.
(6) Zheng, W.; Andrec, M.; Gallicchio, E.; Levy, R. M. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104* (39), 15340–15345.
(7) Liu, P.; Kim, B.; Friesner, R. A.; Berne, B. J. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102* (39), 13749–13754.
(8) Huang, X.; Hagen, M.; Kim, B.; Friesner, R. A.; Zhou, R.; Berne, B. J. *J. Phys. Chem. B* **2007**, *111* (19), 5405–5410.
(9) Moors, S. L. C.; Michielssens, S.; Ceulemans, A. *J. Chem. Theory Comput.* **2011**, *7* (1), 231–237.
(10) Terakawa, T.; Kameda, T.; Takada, S. *J. Comput. Chem.* **2011**, *32* (7), 1228–1234.
(11) Fukunishi, H.; Watanabe, O.; Takada, S. *J. Chem. Phys.* **2002**, *116* (20), 9058–9067.
(12) Affentranger, R.; Tavernelli, I.; Di Iorio, E. E. *J. Chem. Theory Comput.* **2006**, *2* (2), 217–228.
(13) Bowers, K. J.; Chow, E.; Xu, H.; Dror, R. O.; Eastwood, M. P.; Gregersen, B. A.; Klepeis, J. L.; Kolossvary, I.; Moraes, M. A.; Sacerdoti, F. D.; Salmon, J. K.; Shan, Y.; Shaw, D. E. In *SC '06: Proceedings of the 2006 ACM/IEEE Conference on Supercomputing*; Association for Computing Machinery: New York, 2006; p 84 .
(14) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. B* **2001**, *105* (28), 6474–6487.
(15) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79* (10), 926–935.
(16) Neidigh, J. W.; Fesinmeyer, R. M.; Andersen, N. H. *Nat. Struct. Mol. Biol.* **2002**, *9* (6), 425.