Article

Evaluation of Enhanced Sampling Provided by Accelerated Molecular Dynamics with Hamiltonian Replica Exchange Methods

Daniel R. Roe, Christina Bergonzo, and Thomas E. Cheatham, III*

Department of Medicinal Chemistry, College of Pharmacy, University of Utah, 2000 South 30 East Room 105, Salt Lake City, Utah 84112, United States

Ⓢ Supporting Information

ABSTRACT: Many problems studied via molecular dynamics require accurate estimates of various thermodynamic properties, such as the free energies of different states of a system, which in turn requires well-converged sampling of the ensemble of possible structures. Enhanced sampling techniques are often applied to provide faster convergence than is possible with traditional molecular dynamics simulations. Hamiltonian replica exchange molecular dynamics (H-REMD) is a particularly attractive method, as it allows the incorporation of a variety of enhanced sampling techniques through modifications to the various Hamiltonians. In this work, we study the enhanced sampling of the RNA tetranucleotide r(GACC) provided by H-REMD combined with accelerated molecular dynamics (aMD), where a boosting potential is applied to torsions, and compare this to the enhanced sampling provided by H-REMD in which torsion potential barrier heights are scaled down to lower force constants. We show that H-REMD and multidimensional REMD (M-REMD) combined with aMD does indeed enhance sampling for r(GACC), and that the addition of the temperature dimension in the M-REMD simulations is necessary to efficiently sample rare conformations. Interestingly, we find that the rate of convergence can be improved in a single H-REMD dimension by simply increasing the number of replicas from 8 to 24 without increasing the maximum level of bias. The results also indicate that factors beyond replica spacing, such as round trip times and time spent at each replica, must be considered in order to achieve optimal sampling efficiency.
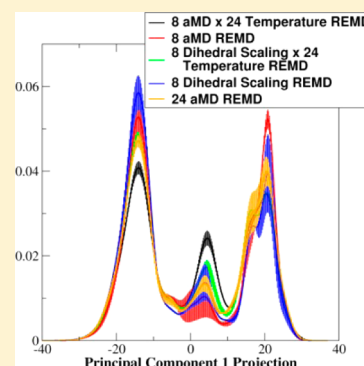
INTRODUCTION

Obtaining a well-converged ensemble of structures is a key challenge within molecular dynamics simulation approaches, where "well-converged" implies adequate sampling of the conformational space of the system of interest such that all important conformational states are observed with populations relatively close to their Boltzmann-weighted ones. From a well-converged ensemble, one can calculate various thermodynamic quantities with a reasonable degree of accuracy, which in turn allows in-depth validation of force field parameters. Advances in computational power have enabled molecular dynamics (MD) simulations to reach increasingly longer time scales, with simulations on the microsecond time scale becoming routine and the millisecond time scale accessible through the use of massively parallel clusters such as Blue Waters, specialized hardware such as Anton,[1] and/or GPUs.[2,3]

Even on these time scales, it is difficult to obtain well-converged ensembles of structures using MD.[4,5] Because of this, simulation techniques which enhance conformational sampling, such as replica exchange MD (REMD),[6] self-guided Langevin dynamics,[7] adaptively biased MD,[8] and accelerated MD,[9] to name a few examples, have become quite popular. Of these, arguably the most commonly used is REMD, most likely since the method is relatively simple to implement and by its nature almost embarrassingly parallel. In REMD, N noninteracting replicas of the system are made. Each replica is simulated for a certain amount of time, after which exchanges are attempted between certain replicas. The probability of exchanging is defined as

$$P_{exch} = \min(1, e^{\beta_1 \Delta H_1 - \beta_0 \Delta H_0})$$

where β_1 and β_0 are the inverse of temperature times the Boltzmann's constant for each of the exchanging replicas, and

$$\Delta H_N = H_N(X_1) - H_N(X_0)$$

Here H_N is the Hamiltonian of one of the exchanging replicas and X_1 and X_0 are the coordinates of the exchanging replicas. The exchange is accepted or rejected using the Metropolis criterion. In the original and most common form of REMD, the only difference between replicas is temperature, which simplifies the exchange probability to

$$P_{exch} = \min(1, e^{\Delta \beta \Delta H})$$

This form (referred to as T-REMD) has the advantage of being quite simple to implement from a computational standpoint, since only temperature and potential energy ever need to be communicated between neighboring replicas.

While T-REMD has proven to be extremely useful in enhancing sampling, its use does not guarantee convergence. One issue with T-REMD is that, if the major barriers to conformational sampling in a given system are not temperature-dependent (e.g., the folding rate of peptides[10−12]), the overall sampling enhancement may be limited. In fact, our recent study has shown that, even for a relatively small RNA tetranucleotide, complete convergence could not be obtained even after 2 μs of simulation per replica.[4] In that study, better convergence was obtained by using reservoir REMD[13−15] (R-REMD), in which the highest temperature replica is replaced by a pregenerated reservoir of structures. However, R-REMD results can be sensitive to how the reservoir is generated; if the reservoir itself does not cover adequate conformational space (i.e., is not representative of the true converged ensemble), the resulting distribution may be missing key conformations.[4] The more general form of REMD involves exchanges between different Hamiltonians (referred to as H-REMD). Several H-REMD schemes have been developed in which various parts of the Hamiltonian are modified in order to enhance sampling in dimensions other than temperature.[15−22]

Another enhanced sampling method is accelerated MD[9] (aMD). In aMD, increased conformational sampling is obtained by applying a so-called "boost" potential to part or all of the system when the potential energy drops below a user-specified energy cutoff, Ethreshold.

$$V*(r) = \begin{cases} V(r) & V(r) \geq \text{Ethreshold} \\ V(r) + \text{Eboost}(r) & V(r) < \text{Ethreshold} \end{cases}$$

Here r represents the coordinates of the system, $V*(r)$ is the "boosted" potential, $V(r)$ is the unmodified potential, and Eboost(r) is the "boost" potential, for which several forms have been proposed.[9,23−25] The original form of the "boost" potential[9] is

$$\text{Eboost}(r) = \frac{(\text{Ethreshold} - V(r))^2}{\alpha + (\text{Ethreshold} - V(r))}$$

The "boost" allows the system to more easily escape minima and cross the potential barriers which slow conformational sampling. The resulting conformational ensemble is biased; in theory, a given property of the biased ensemble A* can be reweighted using the "boost" energies to obtain the unbiased ensemble average:

$$\langle A \rangle = \frac{\int \partial r A*(r) e^{-\beta V*(r)} e^{\text{Eboost}(r)}}{\int \partial r e^{-\beta V*(r)} e^{\text{Eboost}(r)}}$$

However, in practice, reweighting biased ensembles such as those obtained from aMD simulations can be challenging[26,27] due to what is referred to by Markwick and McCammon as *statistical noise error* and *statistical mechanical sampling error*.[28] The former arises from distortions to the system's energy hypersurface caused by the biasing potential, and the latter arises from incomplete sampling of the biased energy hypersurface. The sources of these errors are in direct opposition to one another: a low biasing potential will decrease the statistical noise error but increase the sampling time, while a high biasing potential will decrease the sampling time but increase the statistical noise.[28] Variations of aMD have been introduced that apply the biasing potential only to certain degrees of freedom[29] or to rotatable torsions,[30] thus keeping the overall boost low. However, if a sizable boost is desired, the statistical issues encountered during reweighting will remain.

The issue of reweighting is intrinsically related to the issue of how to space replicas in a REMD simulation. Although there are various methods for choosing replica spacing (in T-REMD, this is choosing an "optimal" distribution of temperatures),[31,32] the ultimate requirement is that any two neighboring replicas be spaced close enough so that there is a non-negligible probability of exchange.[33] During REMD simulations, the exchange probability is implicitly driving the reweighting of each replica. Thus, statistical noise error can be avoided even if the difference between the energy hypersurfaces of the "lowest" and "highest" replicas is large as long as the differences between neighboring replicas are small. This makes the combination of aMD and H-REMD (AMD-HREMD) very attractive, since the replicas can have varying levels of "boost". The "highest" replica can have a large "boost" to provide enhanced sampling, while the "lowest" replica can have no "boost" to provide the unbiased ensemble. This concept was first introduced by Fajer et al. in the context of thermodynamic integration calculations,[34] and has been used as a means of improving conformational sampling in (and hence the convergence of) free energy calculations.[35,36] Furthermore, since by definition all replicas are noninteracting, the replica exchange framework is generalizable to any number of dimensions.[37] This makes it possible to further enhance sampling by, e.g., combining an aMD Hamiltonian dimension with a temperature dimension.

In this study, we evaluate the ability of AMD-HREMD (with the aMD boost applied to torsions only) and multidimensional REMD using aMD and temperature dimensions (AMD-MREMD) to enhance the conformational sampling of the tetranucleotide r(GACC). This molecule is a good test system because it provides surprising conformational complexity, yet is small enough to obtain a good amount of simulation data in a reasonable amount of time, and has been previously studied both via NMR and computational methods.[4,5,38] We compare the sampling enhancement provided by adding more replicas to the Hamiltonian dimension to that provided by the addition of a temperature dimension. We also compare the aMD results to previously run H-REMD and MREMD simulations that used scaling of dihedral force constants (DFC) to enhance sampling in the Hamiltonian dimension. We show that, no matter which Hamiltonian scheme is used (aMD or DFC), agreement of populations as determined from cluster analysis between independent simulations is much better for the conformational ensemble generated from M-REMD with 192 replicas than from H-REMD with 8 replicas, and that convergence is much faster for the M-REMD simulations. Certain conformations present in the M-REMD ensembles are seen rarely or not at all in the H-REMD ensembles. We also show that faster convergence can be obtained by increasing the number of replicas in H-REMD from 8 to 24 without increasing the maximum applied boost, and attempt to explain this phenomenon in terms of a given structure's ability to move efficiently through replica space.

■ METHODS

Model System. The system studied was the RNA tetranucleotide r(GACC) solvated by 2497 TIP3P[39] explicit water molecules and neutralized with three sodium cations using parameters from Joung and Cheatham.[40] Although the salt concentration is minimal, effectively net-neutralizing ions with no excess salts, our previous work shows very little

difference in the sampled conformations with ~100 mM excess NaCl or KCl salt.[5] The system was built with the ff12SB[41−44] force field parameter set using LEaP from AmberTools 12[45] and equilibrated using the protocol previously described by Henriksen et al.[4]

All simulations were run using the MPI version of PMEMD from a developmental version of Amber 12,[46] which had to be modified so that the aMD boost energy was included in the total system potential energy, which is necessary for proper calculation of the exchange probability in H-REMD. All simulations were run under constant temperature and volume (*NTV*) conditions, using a Langevin thermostat with a collision frequency of 5 ps^{-1}, with the random number generator seed based on the date and time of each run. Long range electrostatic interactions were handled using the particle mesh Ewald[47] scheme with a cutoff of 8.0 Å and default parameters. Bonds to hydrogen were constrained using the SHAKE[48] algorithm, and a MD time step of 2 fs was used.

Hamiltonian and Multidimensional REMD Simulations. For the AMD-HREMD simulation, eight replicas were used at 300 K, with the first replica having no aMD dihedral boost active and the remaining replicas using the aMD dihedral boost parameters shown in Table 1. Values of Ethreshold and α

Table 1. aMD Dihedral Boost Parameters Used for Eight-Replica AMD-HREMD Simulations

replica	Ethreshold (kcal/mol)	α (kcal/mol)
1	n/a	n/a
2	112.0	50.0
3	116.0	20.0
4	120.0	10.0
5	124.0	5.0
6	128.0	2.5
7	132.0	1.25
8	136.0	1.0

for replica 2 were chosen by trial and error so that there would be overlap between the dihedral energy distributions of replicas 1 and 2. An additional consideration for Ethreshold was that it be greater than the maximum dihedral energy (109.4 kcal/mol, determined from two independent MD simulations of 1.5 ns each at 300 K) so that the boost is always active. This was done to ensure the Hamiltonian of each replica would be different enough to prevent the acceptance rate from becoming 100%. The values of Ethreshold and α for the remaining replicas were chosen so that replicas would be roughly evenly spaced in potential energy overlap. We found that a constant interval of 4.0 kcal for Ethreshold (which is approximately one standard deviation of the dihedral energy over the course of the two aforementioned 1.5 ns simulations) and a roughly exponential interval for α provided good potential energy overlap between neighboring replicas (see the Supporting Information). The exchange acceptance ranged from 14 to 29% with an average of 20%. Trajectory frames for the H-REMD simulations were saved every 1 ps.

In order to ascertain the effect of the number of replicas on conformational sampling and convergence, AMD-HREMD simulations with 24 replicas were also run (referred to hereafter as AMD-H24). The aMD dihedral boost parameters were chosen so that the minimum and maximum parameters would match the minima and maxima of the AMD-HREMD simulations, with Ethreshold using a linear interval and α

using an exponential interval; actual values used can be found in the Supporting Information.

For the AMD-MREMD simulations, a temperature dimension of size 24 corresponding to the same temperature range used by Henriksen et al.[4] (see the Supporting Information for temperatures) was added to the aMD Hamiltonian dimension used in the AMD-HREMD simulations, for a total of 192 replicas. Exchanges were attempted separately in each dimension, alternating between the temperature and Hamiltonian dimensions. Trajectory frames for the M-REMD simulations were saved every 10 ps.

Analysis. Clustering of trajectories was performed with CPPTRAJ[49] using the DBSCAN[50] clustering algorithm with the minimum number of points set to 25 and epsilon set to 0.9. The distance metric was coordinate RMSD using atoms N2, O6, C1′, and P of residue G, atoms H2, N6, C1′, and P of residue A, and atoms O2, H5, C1′, and P of the C residues. In order to ensure that clusters found would be consistent across all simulations, clustering was performed on the combined trajectories from each run at 300.0 K and using nonmodified Hamiltonians, i.e., using the "lowest" trajectory from each REMD simulation (6 103 495 frames total). To conserve memory, the initial clustering used a "sieve" value of 305 (i.e., every 305th frame was used), resulting in 20 012 initial frames. Sieved frames were then added into a cluster only if they were within epsilon of a frame originally in that cluster. The combined clustering results were then parsed to obtain results for each individual simulation. The total time to complete the clustering was 751 s using an Intel Core i7-2600 CPU (with the pairwise distance and sieve calculations being OpenMP parallelized to use all cores).

Principal component analysis (PCA) was performed with CPPTRAJ. The coordinate covariance matrix was calculated for all heavy atoms (82 atoms total, 246 coordinates). In order to ensure the eigenvectors obtained would be consistent across all simulations, the coordinate covariance matrix was calculated on the same combination of trajectories that was used in the cluster analysis. All frames were RMS-fit to the overall average structure prior to calculating the coordinate covariance matrix. The coordinates from the trajectories of each simulation were then separately projected along each eigenvector to obtain a time series of principal component projection values for each principal component. A CPPTRAJ script for performing this calculation can be found in the Supporting Information.

For each simulation type, the Kullback−Leibler divergence (KLD) between the principal component histograms from the two independent simulations over time was calculated using a development version of CPPTRAJ that will be released with AmberTools 14. The KLD has been used to assess convergence between independent simulations in our previous study,[5] and will be described briefly here. The time-dependent KLD is calculated as

$$\mathrm{KLD}(t) = \sum_i P(t, i) \ln\left(\frac{P(t, i)}{Q(t, i)}\right)$$

Here $P(t, i)$ and $Q(t, i)$ represent different probability distributions normalized to 1.0, with i representing a histogram bin index and t representing the time at which the histogram is being constructed (i.e., it includes all data from time 0 to t). The KLD between two distributions P and Q is undefined if $P(i)$ is populated but $Q(i)$ is not (where i indicates the bin number) and vice versa; KLD values were not calculated for

Table 2. Clustering Results for the Top Six Clusters from Combined Trajectories at 300 K with Non-Modified Hamiltonians (Davies−Bouldin Index = 0.88)[a]

#cluster	frames	%	AvgDist[b]	Stdev[b]	AvgCDist[c]	name
0	1 230 717	20.2	1.83	0.66	4.64	A-form-minor_NMR
1	976 585	16.0	1.54	0.62	3.76	intercalcated-anti
2	776 618	12.7	1.51	0.67	4.62	A-form-major-NMR
3	289 784	4.7	0.90	0.30	4.33	inverted-syn
4	244 102	4.0	1.09	0.35	3.72	intercalated-syn
5	117 082	1.9	1.62	0.56	4.55	1_3-stack

[a]See the Methods for details. [b]AvgDist and Stdev are the average distance and standard deviation (in Å) between all frames in the cluster, respectively. [c]AvgCdist is the average distance (in Å) of the cluster to every other cluster.

such points. To reduce the number of bins with no population, histograms were constructed using a Gaussian kernel density estimator. For each histogram, a total of 400 bins were used, and the bandwidth was determined from the normal distribution approximation.

■ RESULTS

Two independent 8-replica AMD-HREMD simulations (1153 and 1128 ns per replica), two independent 24-replica AMD-HREMD simulations (465 and 545 ns per replica), and two independent 192-replica AMD-MREMD simulations (838 and 1005 ns per replica) were performed. These were compared with previously run 8-replica HREMD with dihedral force constant scaling (DFC) and 192-replica MREMD with DFC scaling and temperature,[5] referred to as DFC-HREMD (1020 ns per replica, each) and DFC-MREMD (289 and 300 ns per replica), respectively.

Cluster Analysis. Cluster analysis was performed on the combined trajectories (see the Methods for further details) at 300.0 K with no aMD boost active from all simulations in order to determine the major conformations observed. Table 2 shows clustering results for the top 6 clusters; full results (24 clusters total) are available in the Supporting Information. Roughly 30% of the combined trajectory was considered "noise" by the DBSCAN algorithm. Representative structures were assigned names using the nomenclature of Henriksen et al.[4] by choosing the closest reference structure to the cluster representative structure via RMSD; the representative was required to have a RMSD to reference <1.0 Å in order to be assigned.

The three most populated clusters (accounting for ∼49% of all structures) correspond to the "A-form-minor-NMR", "intercalated-anti", and "A-form-major-NMR" structures. These conformations were also found in MD simulations performed by Yildirim et al. with a modified force field,[38] although with different populations. The next three most populated clusters account for 4.7, 4.0, and 1.9% of the total structures, respectively. The remaining 18 clusters together make up only 9.9% of the total structures, with each accounting for less than 1.4% of the total population, and in individual trajectories, none are populated above 3.0%. The only exception to this is the 1_3-basepair structure, which is populated only in both sets of M-REMD runs; the 1_3-basepair structure has no population in any of the H-REMD runs.

Figure 1 shows percent populations of the top six clusters from each of the different simulation types (calculated as the average over two independent simulations, with error bars representing the difference between the simulations). The populations for all clusters can be found in the Supporting Information. The populations obtained for the two independent AMD-MREMD simulations and two independent DFC-
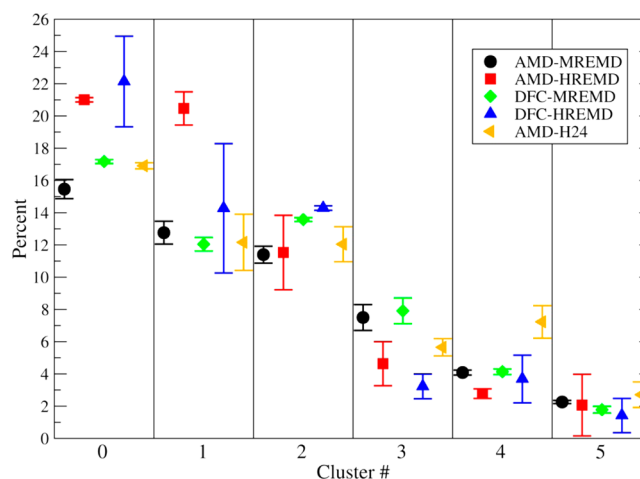

Figure 1. Average percent population present from each simulation type (shown in different symbols and colors) for the top six clusters calculated as $(X_1 + X_2)/2$, where X_1 and X_2 are the populations from simulations 1 and 2, respectively; error bars are calculated as $|X_1 - X_2|/2$.

MREMD simulations agree well with each other, as evidenced by the small error bars. In contrast, the populations obtained from the two AMD-HREMD and two DFC-HREMD simulations do not agree as well overall, and in some cases are quite different (e.g., the cluster 1 intercalated-anti population difference between the two DFC-HREMD simulations is 8%). The populations obtained for the two AMD-H24 simulations agree reasonably well, although not quite as well as the M-REMD simulations, but clearly better than the other eight-replica H-REMD simulations.

Table 3 shows the average absolute difference and standard deviation of cluster populations (in percentage) between the two independent simulations for each simulation type, both across all 24 clusters and just for the top 3 clusters (A-form-minor-NMR, intercalated-anti, and A-form-major-NMR). It should be noted that certain clusters (from cluster 10 on) were not found in some of the AMD-HREMD, DFC-HREMD, and AMD-H24 simulations; if either independent simulation for a given type did not find a certain cluster, the delta (which would be 0.0) was not included in the calculation of averages/standard deviations, as this would serve to artificially lower them.

The AMD-MREMD and DFC-MREMD simulations have similar average differences (0.38 ± 0.52 and 0.37 ± 0.53%, respectively) when all clusters are considered. However, if just the three most populated clusters are considered, the average differences in the AMD-MREMD simulations are slightly larger than the differences in the DFC-MREMD simulations (1.22 ± 0.19% versus 0.43 ± 0.34%), indicating that the AMD-

Table 3. Average Absolute Difference and Standard Deviation of Cluster Populations (in Percentage) Both across All Clusters and Just for the Top 3 Clusters

	AMD-MREMD	AMD-HREMD[a]	DFC-MREMD	DFC-HREMD[a]	AMD-H24[a]
avg delta	0.38	0.93	0.37	1.13	0.70
SD delta	0.52	1.23	0.53	1.98	0.87
top 3 avg	1.22	2.32	0.43	4.64	2.01
top 3 SD	0.19	2.19	0.34	3.97	1.56

[a]The average/standard deviation does not include clusters for which no population was found in one or both of the independent simulations.

Table 4. Average, Standard Deviation, and Maximum Starting Time Taken to Find Clusters (in ns) across All Clusters for Each Type of Simulation

	AMD-MREMD	AMD-HREMD[a]	DFC-MREMD	DFC-HREMD[a]	AMD-H24[a]
average	21.2	126.9	9.0	112.8	45.6
stdev	25.6	172.0	8.5	116.5	83.0
max	115.6	917.3	42.8	536.6	435.4

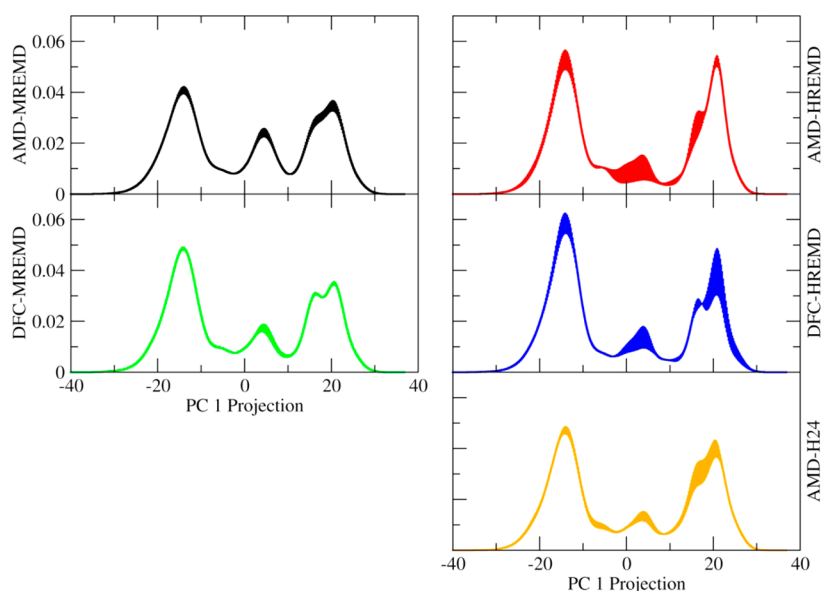[a]Not all clusters were found for this simulation type.

Figure 2. Histograms of the projection of principal component 1 (accounting for ~35% of the total motion; see the Supporting Information for eigenvalues) for each simulation type. Histograms were calculated using a Gaussian kernel density estimator (see the Methods for details). Each histogram as shown is actually the average of the histograms from the two simulations of each type, with the width of the error bars representing the differences between the simulations.

MREMD results are not quite as well-converged. This may be why the populations from the AMD-MREMD simulations do not quite agree with the populations from the DFC-MREMD simulations, in particular the A-form-minor-NMR and A-form-major-NMR structures.

The AMD-HREMD and DFC-HREMD simulations have much larger differences in population between their independent runs, both across all clusters and for the top three clusters, indicating these simulations are not converged. The AMD-H24 simulations have slightly smaller differences in population than the eight-replica HREMD runs but still larger than any of the MREMD simulations. It is interesting to note however that for the AMD-H24 simulations the average populations for the top three clusters are relatively close to those of the M-REMD simulations. This indicates that the AMD-H24 simulations are better converged than either of the eight-replica H-REMD simulations.

A rough measure of the sampling efficiency in a simulation is how long it takes to "find" individual clusters; more efficient sampling should lead to more rapid discovery of clusters. Table 4 shows the average, standard deviation, and maximum time (in frames) needed to find all clusters for each type of simulation (calculated from the combination of two independent simulations of each type). Individual values for each simulation can be found in the Supporting Information. Both M-REMD-type simulations are on average an order of magnitude faster in finding clusters than the eight-replica H-REMD-type simulations. The AMD-H24 simulations are on average slower than the M-REMD-type simulations in finding clusters but still 2−3 times faster than the eight-replica H-REMD-type simulations.

Principal Component Analysis. Principal component analysis (PCA) provides a measure of the dynamic nature of a structure. PCA is a good complement to cluster analysis; clustering provides information on what conformations are adopted over the course of a simulation, while PCA provides information on what motions the structure undergoes to adopt those conformations. Figure 2 shows histograms for the projection of coordinates along the first eigenvector (i.e., the

eigenvector with the highest eigenvalue) for each simulation type, calculated as the average from histograms of two independent simulations, with error bars representing the difference between the independent simulations. As with the clustering results, the largest differences between the independent runs is found in the AMD-HREMD, DFC-HREMD, and AMD-H24 simulation sets, while the AMD-MREMD and DFC-MREMD simulations show much smaller differences. These trends also apply for the next four principal components (see the Supporting Information).

The overlap of principal component (PC) projection histograms can be a useful measure of convergence between two independent runs. As two independent runs with different initial conditions proceed, they will start to sample the same conformational space and undergo similar motions, leading to an increase in their overlap in PC space. Until there is a certain degree of overlap between PC projection histograms for every "essential" PC (i.e., a PC eigenvector with a non-negligible eigenvalue), the two simulations are not well-converged. Kullback−Leibler divergence[51] (KLD) is a measure of how well two probability distributions overlap. Measuring the KLD as a function of time between histograms from two independent simulations can serve as a quantitative measure of convergence.

Figure 3 shows the KLD between PC projection histograms of PCs 1, 2, and 3 (which comprise ∼58% of the total motion)
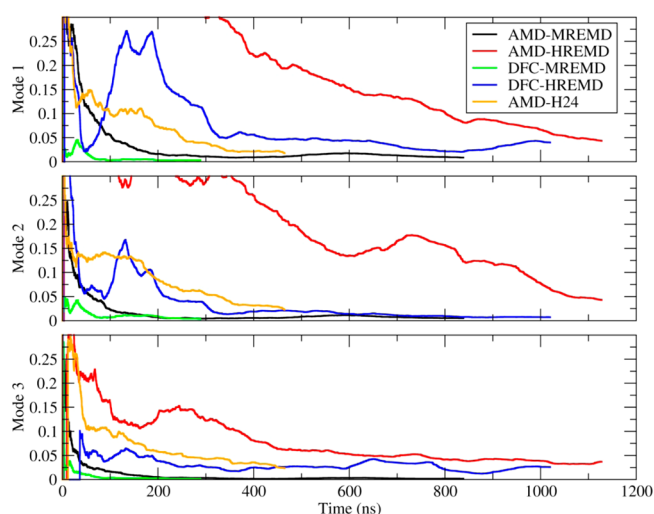

Figure 3. Kullback−Leibler divergence of principal component projection histograms from independent simulations vs time for the first three principal components.

from two independent runs for every simulation type. Consistent with previous results, the AMD-MREMD and DFC-MREMD runs converge relatively quickly. The slope of the KLD for the AMD-MREMD runs stops changing significantly by ∼200 ns. The DFC-MREMD runs converge even faster, with the slope of each line not changing significantly after 100 ns. The H-REMD runs all converge slower, and have more variation in the rate at which they converge than the M-REMD runs. The AMD-HREMD runs in particular have the slowest convergence.

Note that, in the case of the DFC-HREMD runs, the initial KLD values for PCs 1 and 2 drop rapidly to ∼0.021 at around 48 ns before rising back up to ∼0.269. This illustrates a potential pitfall of using KLD as a metric of convergence when sampling is limited, such as at the beginning of an MD simulation: if two otherwise independent runs begin in the same conformational space of a certain PC, they may both initially explore the same limited subspace along that PC and the overlap of their PC projection histograms may actually be quite good. Once sampling increases and one or both of the runs can explore outside of their initial regions, the KLD better reflects the convergence between the two simulations. Therefore, the KLD of PC projection histograms is a good measure of convergence between two simulations only if at least one simulation has explored along most or all of a given PC. In other words, there is a minimum sampling time required before this metric can be considered viable.

Table 5 shows the average, minimum, and maximum times at which the KLD for the PC projection histograms of the first 20 PCs (accounting for ∼96% of the total motion) reaches and remains below 0.02 for each simulation type; values for each individual PC can be found in the Supporting Information. The 0.02 cutoff was chosen on the basis of the observation that in each simulation the slope of the KLD plot no longer changes significantly once it is below this value. For the AMD-MREMD and DFC-MREMD simulations, all 20 PCs reach 0.02 by 178.5 and 49.1 ns per replica, respectively. The PCs that take the longest to reach 0.02 are PC 1 for AMD-MREMD and PC 2 for DFC-MREMD, consistent with the fact that these PCs correspond to the lowest frequency motions. For the H-REMD runs, however, not all PCs reach 0.02. In fact, only one of the top three PCs (accounting for ∼57% of the total motion) reaches 0.02 in the AMD-H24 (PC 1) and DFC-HREMD (PC 2) runs, and none of the top three PCs do in the AMD-HREMD runs. This is consistent with the clustering results that showed certain conformations populated in the M-REMD runs are not populated in the H-REMD runs, and indicates that convergence of the lowest frequency motions is essential for obtaining correct structural populations.

Table 5. Simulation Time (in ns) Needed for the KLD of PC Projection Histograms of the First 20 PCs to Reach 0.02 and Number of PCs for Which KLD Never Reached 0.02

	AMD-MREMD	AMD-HREMD	DFC-MREMD	DFC-HREMD	AMD-H24
average[a]	84.3	816.1	25.2	586.5	222.2
SD[a]	37.2	256.0	15.1	162.2	105.5
min[a]	31.6	176.4	6.2	311.3	83.5
max[a]	178.5	1126.5	49.1	759.6	460.4
KLD > 0.02	0	5	0	6	3
top 3 KLD > 0.02	0	3	0	2	2

[a]Values for which KLD never reached 0.02 are not included.

aMD Reweighting of Individual Replicas. As mentioned in the Introduction, direct reweighting of aMD simulations can be problematic due to statistical noise error and statistical mechanical sampling error,[28] particularly when the boost is high. To illustrate this point, we calculated the population-based free energy for rotation around the alpha backbone dihedral angle of A2 using the unbiased replica, the reweighted data from the lowest-boosted replica, and the reweighted data from the highest-boosted replica of an AMD-HREMD run (Figure 4). Reweighting of the lowest-boosted replica is able to
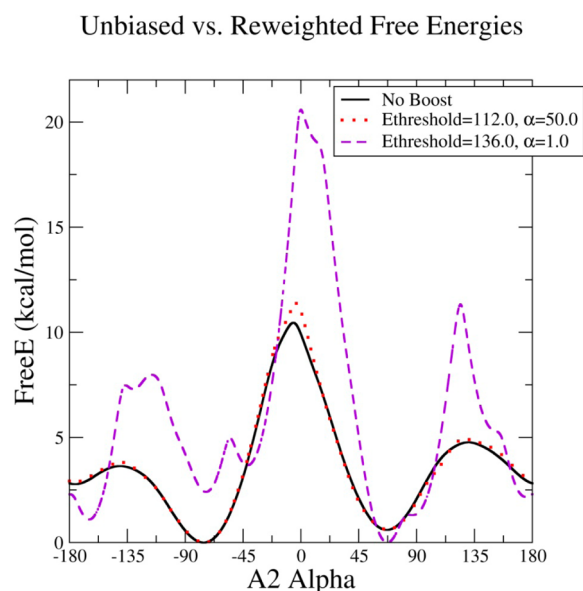
Unbiased vs. Reweighted Free Energies

Figure 4. Population-based free energies for rotation around the A2 alpha dihedral calculated from population histograms for the unbiased replica (solid line), the reweighted population of the lowest aMD-boosted replica (dotted line), and the reweighted population of the highest aMD-boosted replica (dashed line) from an AMD-HREMD run. All histograms were calculated using a Gaussian KDE with a bandwidth of 6.8.

successfully recapture the free energy profile of the unbiased replica. However, reweighting of the highest-boosted replica results in an extremely different free energy profile compared to the unbiased replica. This serves to highlight the benefits of aMD combined with REMD; one can have the sampling benefits of using a high boost without worrying about reweighting artifacts.

Impact of Number of Replicas on Convergence. Overall, the results clearly show that, in REMD simulations, the rate of convergence of cluster population and PC component projections between independent simulations depends on the number of replicas used, with more replicas increasing the rate. It has already been shown that convergence of an H-REMD run could be increased by adding a temperature dimension[5] (i.e., turning it into an M-REMD run). In this work, that remains true, but here it is also shown that convergence of an H-REMD run can also be improved somewhat by simply adding more replicas in the same biasing/Hamiltonian space (i.e., without increasing the maximum boost used). The AMD-H24 simulations with 24 replicas found more clusters in a shorter amount of time than the AMD-HREMD simulations with 8 replicas, and had a faster rate of convergence in PC space. Importantly, since the maximum boost used in both the

AMD-HREMD and AMD-H24 runs was equal, the improvement results solely from the additional replicas.

To explain what causes this increased convergence, we have compared average exchange acceptance, replica round trip times, and average time spent at each replica between one of the AMD-HREMD runs and one of the AMD-H24 runs. Figure 5 shows the fraction of exchange attempts to the next-highest
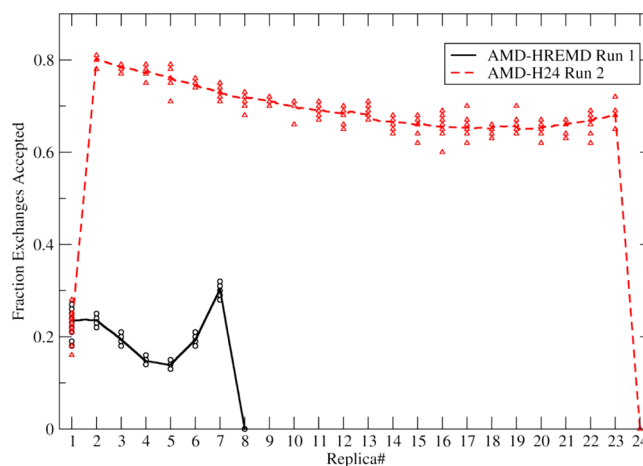

Figure 5. Fraction exchange attempts to the next-highest replica accepted for one of the 8-replica AMD-HREMD runs (black circles and solid line) and one of the 24-replica AMD-H24 runs (red triangles and dashed line). The boost of AMD-HREMD replicas 2 and 8 matches the boost of AMD-H24 replicas 2 and 24, respectively; replica 1 is not boosted in either case. Points are the fraction accepted calculated at several intervals along each simulation, and the lines represent the average of the points. The highest replica for each simulation always has an exchange acceptance of 0.00, since it is trying to exchange with the lowest replica (which is never accepted).

replica accepted for both simulations. As mentioned in the Methods, the boost levels for the AMD-H24 simulations were chosen so that the aMD parameters for the lowest boosted replica (replica 2) and highest boosted replica (replica 24) would match the lowest and highest boosted replicas of the AMD-HREMD run (replicas 2 and 8, respectively). Therefore, the spacing between replicas 1 and 2 in both simulations is the same, which is reflected by the similar average acceptance for exchanging from replica 1 to 2 (~0.23 for both). However, because the spacing for the remaining replicas is necessarily closer in the AMD-H24 simulation, the average exchange acceptance for these replicas is much higher, ranging from ~0.65 to ~0.80, than in the AMD-HREMD simulation, which ranges from ~0.14 to ~0.30. Although the disparate exchange acceptance between replicas 1 and 2 and the remaining replicas in the AMD-H24 simulation may seem like cause for concern, it will be shown through analysis of replica round trip times and time spent at each replica that there is little discernible effect on overall exchange efficiency.

In order for a replica exchange scheme to be efficient, coordinates should have the opportunity to visit all replicas multiple times. One metric for this introduced by Abraham and Gready is the "transit number", which is the number of replica exchange attempts required for a 95% probability that at least one replica has visited the maximum before visiting the minimum.[52] A metric directly related to the transit number is the "round-trip time" (RT), which is simply the time needed to transition from the lowest replica to the highest and back.

Figure 6 shows the number of round trips per ns for each set of starting coordinates, as well as the average, minimum, and
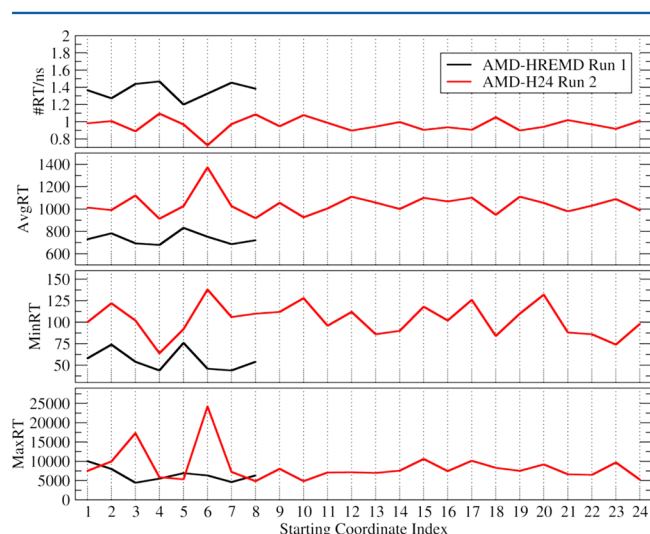

Figure 6. Number of round trips per ns (#RT/ns), as well as the average, minimum, and maximum round-trip times in number of exchanges (1 exchange/ps) for an AMD-HREMD and an AMD-H24 simulation. The starting coordinate index indicates the initial replica that a given set of coordinates started at.

maximum RT in exchanges (one exchange per ps) for each simulation. Since the AMD-HREMD simulation has fewer replicas to traverse for a round trip than the AMD-H24 simulation, it might be expected that the RT for the AMD-HREMD simulation would be shorter. While this is indeed the case (in the AMD-HREMD simulation, it takes on average 735 ns to make a round trip vs 1042 ns for the AMD-H24 simulation), the change in round time (71%) is not directly proportional to the change in the number of replicas (30%). This is most likely because, although there are more replicas in the AMD-H24 simulation, the higher exchange acceptance means that a given set of coordinates can transition between replicas at an effectively higher rate. It is particularly interesting to note the extreme difference between the minimum and maximum RT in both simulations, which can range from hundreds of picoseconds to tens of thousands of picoseconds (>10 ns). While the minimum RTs for the AMD-HREMD simulation are somewhat shorter than the AMD-H24 simulation, the maximum RTs (with two exceptions) for both simulations are around 10 000 exchanges. Again, although the number of replicas increases from 8 to 24, on average, the longest time it takes for a structure to make a round trip is the same for both simulations.

Figure 7 shows the average time that coordinates spend at each replica for both simulations. The closer the slope of this line is to zero, the better a given set of coordinates has transitioned through replica space. The average time a given set of coordinates spends at each replica for the AMD-HREMD simulation ranges from 8 to 18 ps, and the somewhat large slope of the lines indicates certain coordinates spend more time at certain replicas than others. In contrast, the average time spent at each replica for the AMD-H24 simulation ranges from only 2 to 7 ps, and the slope of the lines is correspondingly smaller. This provides some insight as to why convergence is faster in the AMD-H24 simulation; effectively any given set of coordinates in the AMD-H24 simulation has a better chance of
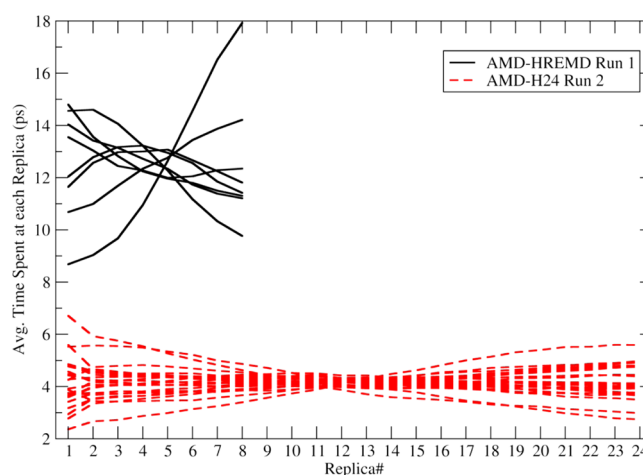

Figure 7. Average time spent at each replica for each set of starting coordinates (1 line per coordinate set) for the AMD-HREMD and AMD-H24 simulations.

reaching a higher boost in a shorter amount of time than in the AMD-HREMD simulation.

As previously mentioned, because of the disparity in exchange acceptance rates in the AMD-H24 simulation, ∼0.2 between replica 1 (the unbiased Hamiltonian) and replica 2 (the first aMD-boosted Hamiltonian) and ∼0.7 for all other replicas, one might expect structures to perhaps spend more time at the unbiased replica. However, this is clearly not the case, suggesting that either the disparity between the exchange acceptance rates does not matter or perhaps having only one disparity is not enough to significantly impact the ability of coordinates to transition through replica space. It is important to note that, although coordinates are spending less time per replica in the AMD-H24 simulation, there is still on average more than 2 ps between each exchange. This is important because exchanges must not be too close temporally;[52] otherwise, the potential energies of a replica pair may still be correlated at the time of the next exchange, violating the central premise of REMD: that all replicas are independent.

■ CONCLUSIONS

A comparison of the enhanced sampling provided by HREMD (using either aMD or scaling of DFC for the Hamiltonians) with 8 replicas to MREMD with 192 replicas (8 Hamiltonian, 24 temperature) showed that, in terms of convergence, MREMD was superior to HREMD for both Hamiltonian types. Certain conformations present in the M-REMD ensembles are seen rarely or not at all in the H-REMD ensembles. Although some of this may be due to poor convergence, one conformation in particular in which a base pair is formed between G1 and C3 (the "1_3-basepair" conformation[4]) is never seen in any of the H-REMD-type simulations, indicating that the addition of a temperature dimension in the M-REMD simulations may help to sample rare conformations. This also implies that boosting the dihedral energy does not always translate into easier rotations around torsions. This is clear from examination of free energies for rotation around torsions from one of the AMD-MREMD simulations (see the Supporting Information). While the barriers for rotation around certain torsions (such as alpha, gamma, zeta, and sugar pucker pseudorotation) are indeed reduced by ∼2 kcal/mol or more in some cases, the barriers for

3550

dx.doi.org/10.1021/jp4125099 | J. Phys. Chem. B 2014, 118, 3543–3552

other torsions (such as epsilon and chi) are not significantly reduced, indicating that there are other factors contributing to these barriers. Indeed, it should be noted that, although barriers to rotation around dihedrals play a role in structural transitions, other factors such as van der Waals and electrostatic interactions play important roles as well. The addition of the temperature dimension may serve to overcome barriers that cannot be surmounted by boosting torsions alone. It is likely that additional Hamiltonian dimensions specifically targeted at these other factors may increase convergence rates even more.

It was also shown that convergence of AMD-HREMD simulations could be improved by adding more replicas to the Hamiltonian dimension (increasing from 8 to 24 total replicas) without increasing the maximum bias. The improved convergence appears to arise from an increased ability of coordinates to travel through replica space via higher replica exchange acceptance, and therefore get to replicas with higher levels of aMD boost where potential barriers may be crossed in a shorter amount of time.

It is noted that the DFC-MREMD simulations were particularly fast to converge. This may not necessarily reflect that DFC scaling is inherently faster than aMD in terms of sampling. Instead, it may just be a consequence of energy; the difference in torsion energy between the lowest and highest replicas for the aMD simulations is on the order of ~40 kcal/mol, while the difference for the DFC simulations is on the order of ~70 kcal/mol. It may be that if the aMD boost level is increased the convergence speed will approach that of the DFC simulations. However, this would require an increased number of replicas in the Hamiltonian dimension in order to maintain a decent exchange rate (>0.2), so it appears as though DFC scaling is overall more efficient in terms of sampling when in an H-REMD framework, at least for this system. It may also be possible that other forms of the aMD boost than the one used in this study would be more efficient.[25]

Overall, the results indicate there is a delicate balance to strike in H-REMD simulations between the number of replicas used and how those replicas are spaced. For example, adding more replicas may increase the probability that a given set of coordinates can get to a "higher" replica where sampling may be enhanced, but adding too many replicas may increase the time to get to the highest replica so much that any benefits are lost. Choosing parameters to optimize this balance is of key importance and is being explored in future work.

ASSOCIATED CONTENT

Ⓢ Supporting Information

AMD-HREMD replica dihedral energy overlap. Temperatures used for temperature dimension in MREMD simulations. aMD parameters used in AMD-H24 simulations. Full clustering results from combined trajectories. Fraction cluster populations for each simulation. Cluster starting frames for each simulation. CPPTRAJ scripts for cluster/PC analysis. Principal component histograms. Eigenvalues and fraction of total for each. Time for KLD of principal components to reach 0.02 for each simulation type. Free energies for various torsions. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: tec3@utah.edu.

Author Contributions

The manuscript was written through contributions of all authors. D.R.R. performed all aMD-related simulations and all analyses, and C.B. performed all DFC-related simulations. All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

REFERENCES

(1) Shaw, D. E.; Deneroff, M. M.; Dror, R. O.; Kuskin, J. S.; Larson, R. H.; Salmon, J. K.; Young, C.; Batson, B.; Bowers, K. J.; Chao, J. C. Anton, a special-purpose machine for molecular dynamics simulation. *Commun. ACM* **2008**, *51* (7), 91−97.

(2) Götz, A. W.; Williamson, M. J.; Xu, D.; Poole, D.; Le Grand, S.; Walker, R. C. Routine microsecond molecular dynamics simulations with AMBER on GPUs. 1. Generalized born. *J. Chem. Theory Comput.* **2012**, *8* (5), 1542.

(3) Salomon-Ferrer, R.; Götz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *J. Chem. Theory Comput.* **2013**, *9* (9), 3878−3888.

(4) Henriksen, N. M.; Roe, D. R.; Cheatham, T. E., III. Reliable Oligonucleotide Conformational Ensemble Generation in Explicit Solvent for Force Field Assessment Using Reservoir Replica Exchange Molecular Dynamics Simulations. *J. Phys. Chem. B* **2013**, *117* (15), 4014−4027.

(5) Bergonzo, C.; Henriksen, N. M.; Roe, D. R.; Swails, J. M.; Roitberg, A. E.; Cheatham, T. E. Multi-dimensional Replica Exchange Molecular Dynamics Yields a Converged Ensemble of an RNA Tetranucleotide. *J. Chem. Theory Comput.* **2014**, *10*, 492−499.

(6) Sugita, Y.; Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* **1999**, *314* (1), 141−151.

(7) Wu, X.; Brooks, B. R. Self-guided Langevin dynamics simulation method. *Chem. Phys. Lett.* **2003**, *381* (3), 512−518.

(8) Babin, V.; Roland, C.; Sagui, C. Adaptively biased molecular dynamics for free energy calculations. *J. Chem. Phys.* **2008**, *128* (13), 134101−134101.

(9) Hamelberg, D.; Mongan, J.; McCammon, J. A. Accelerated molecular dynamics: a promising and efficient simulation method for biomolecules. *J. Chem. Phys.* **2004**, *120*, 11919.

(10) Lednev, I. K.; Karnoup, A. S.; Sparrow, M. C.; Asher, S. A. α-Helix peptide folding and unfolding activation barriers: a nanosecond UV resonance Raman study. *J. Am. Chem. Soc.* **1999**, *121* (35), 8074−8086.

(11) Oliveberg, M.; Tan, Y.-J.; Fersht, A. R. Negative activation enthalpies in the kinetics of protein folding. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92* (19), 8926−8929.

(12) Cavalli, A.; Ferrara, P.; Caflisch, A. Weak temperature dependence of the free energy surface and folding pathways of structured peptides. *Proteins: Struct., Funct., Bioinf.* **2002**, *47* (3), 305−314.

(13) Okur, A.; Roe, D. R.; Cui, G.; Hornak, V.; Simmerling, C. Improving Convergence of Replica-Exchange Simulations through Coupling to a High-Temperature Structure Reservoir. *J. Chem. Theory Comput.* **2007**, *3* (2), 557−568.

(14) Li, H.; Li, G.; Berg, B. A.; Yang, W. Finite reservoir replica exchange to enhance canonical sampling in rugged energy surfaces. *J. Chem. Phys.* **2006**, *125* (14), 144902−144902.

(15) Lyman, E.; Ytreberg, F. M.; Zuckerman, D. M. Resolution Exchange Simulation. *Phys. Rev. Lett.* **2006**, *96* (2), 028105.

(16) Affentranger, R.; Tavernelli, I.; Di Iorio, E. E. A Novel Hamiltonian Replica Exchange MD Protocol to Enhance Protein Conformational Space Sampling. *J. Chem. Theory Comput.* **2006**, *2* (2), 217−228.

(17) Christen, M.; van Gunsteren, W. F. Multigraining: An algorithm for simultaneous fine-grained and coarse-grained simulation of molecular systems. *J. Chem. Phys.* **2006**, *124* (15), 154106.

(18) Liu, P.; Voth, G. A. Smart resolution replica exchange: An efficient algorithm for exploring complex energy landscapes. *J. Chem. Phys.* **2007**, *126* (4), 045106.

(19) Fukunishi, H.; Watanabe, O.; Takada, S. On the Hamiltonian replica exchange method for efficient sampling of biomolecular systems: Application to protein structure prediction. *J. Chem. Phys.* **2002**, *116* (20), 9058−9067.

(20) Jang, S.; Shin, S.; Pak, Y. Replica-Exchange Method Using the Generalized Effective Potential. *Phys. Rev. Lett.* **2003**, *91* (5), 058305.

(21) Cheng, X.; Cui, G.; Hornak, V.; Simmerling, C. Modified Replica Exchange Simulation Methods for Local Structure Refinement. *J. Phys. Chem. B* **2005**, *109* (16), 8220−8230.

(22) Ostermeir, K.; Zacharias, M. Hamiltonian replica-exchange simulations with adaptive biasing of peptide backbone and side chain dihedral angles. *J. Comput. Chem.* **2014**, *35* (2), 150−158.

(23) de Oliveira, C. A. F.; Hamelberg, D.; McCammon, J. A. Coupling accelerated molecular dynamics methods with thermodynamic integration simulations. *J. Chem. Theory Comput.* **2008**, *4* (9), 1516−1525.

(24) Sinko, W.; de Oliveira, C. A. F.; Pierce, L. C.; McCammon, J. A. Protecting high energy barriers: A new equation to regulate boost energy in accelerated molecular dynamics simulations. *J. Chem. Theory Comput.* **2011**, *8* (1), 17−23.

(25) Yang, L.; Grubb, M. P.; Gao, Y. Q. Application of the accelerated molecular dynamics simulations to the folding of a small protein. *J. Chem. Phys.* **2007**, *126*, 125102.

(26) Shen, T.; Hamelberg, D. A statistical analysis of the precision of reweighting-based simulations. *J. Chem. Phys.* **2008**, *129* (3), 034103.

(27) Ceriotti, M.; Brain, G. A.; Riordan, O.; Manolopoulos, D. E. The inefficiency of re-weighted sampling and the curse of system size in high-order path integration. *Proc. R. Soc. London, Ser. A* **2012**, *468* (2137), 2−17.

(28) Markwick, P. R.; McCammon, J. A. Studying functional dynamics in bio-molecules using accelerated molecular dynamics. *Phys. Chem. Chem. Phys.* **2011**, *13* (45), 20053−20065.

(29) Wereszczynski, J.; McCammon, J. A. Using Selectively Applied Accelerated Molecular Dynamics to Enhance Free Energy Calculations. *J. Chem. Theory Comput.* **2010**, *6* (11), 3285−3292.

(30) Doshi, U.; Hamelberg, D. Improved Statistical Sampling and Accuracy with Accelerated Molecular Dynamics on Rotatable Torsions. *J. Chem. Theory Comput.* **2012**, *8* (11), 4004−4012.

(31) Trebst, S.; Troyer, M.; Hansmann, U. H. Optimized parallel tempering simulations of proteins. *J. Chem. Phys.* **2006**, *124*, 174903.

(32) Patriksson, A.; van der Spoel, D. A temperature predictor for parallel tempering simulations. *Phys. Chem. Chem. Phys.* **2008**, *10* (15), 2073−2077.

(33) Mitsutake, A.; Sugita, Y.; Okamoto, Y. Generalized-ensemble algorithms for molecular simulations of biopolymers. *Pept. Sci.* **2001**, *60* (2), 96−123.

(34) Fajer, M.; Hamelberg, D.; McCammon, J. A. Replica-exchange accelerated molecular dynamics (REXAMD) applied to thermodynamic integration. *J. Chem. Theory Comput.* **2008**, *4* (10), 1565−1569.

(35) Jiang, W.; Roux, B. Free energy perturbation hamiltonian replica-exchange molecular dynamics (FEP/H-REMD) for absolute ligand binding free energy calculations. *J. Chem. Theory Comput.* **2010**, *6* (9), 2559−2565.

(36) Arrar, M.; de Oliveira, C. A. F.; Fajer, M.; Sinko, W.; McCammon, J. A. w-REXAMD: A Hamiltonian Replica Exchange Approach to Improve Free Energy Calculations for Systems with

Kinetically Trapped Conformations. *J. Chem. Theory Comput.* **2012**, *9* (1), 18−23.

(37) Sugita, Y.; Kitao, A.; Okamoto, Y. Multidimensional replica-exchange method for free-energy calculations. *J. Chem. Phys.* **2000**, *113*, 6042−6051.

(38) Yildirim, I.; Stern, H. A.; Tubbs, J. D.; Kennedy, S. D.; Turner, D. H. Benchmarking AMBER force fields for RNA: comparisons to NMR spectra for single-stranded r (GACC) are improved by revised χ torsions. *J. Phys. Chem. B* **2011**, *115* (29), 9261−9270.

(39) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79* (2), 926−935.

(40) Joung, I. S.; Cheatham, T. E., III. Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J. Phys. Chem. B* **2008**, *112* (30), 9020−9041.

(41) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Struct., Funct., Bioinf.* **2006**, *65* (3), 712−725.

(42) Pérez, A.; Marchán, I.; Svozil, D.; Sponer, J.; Cheatham, T. E., III; Laughton, C. A.; Orozco, M. Refinement of the AMBER Force Field for Nucleic Acids: Improving the Description of alpha/gamma Conformers. *Biophys. J.* **2007**, *92* (11), 3817−3829.

(43) Banáš, P.; Hollas, D.; Zgarbová, M.; Jurečka, P.; Orozco, M.; Cheatham, T. E., III; Šponer, J. i.; Otyepka, M. Performance of molecular mechanics force fields for RNA simulations: Stability of UUCG and GNRA hairpins. *J. Chem. Theory Comput.* **2010**, *6* (12), 3836−3849.

(44) Zgarbová, M.; Otyepka, M.; Šponer, J. i.; Mládek, A. t.; Banáš, P.; Cheatham, T. E., III; Jurecka, P. Refinement of the Cornell et al. nucleic acids force field based on reference quantum chemical calculations of glycosidic torsion profiles. *J. Chem. Theory Comput.* **2011**, *7* (9), 2886−2902.

(45) Case, D. A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, C.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Swails, J. M.; Goetz, A. W.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wolf, R. M.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M. J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Salomon-Ferrer, R.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. *Amber 12*; University of California: San Francisco, CA, 2012.

(46) Case, D. A.; Cheatham, T. E., III; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber Biomolecular Simulation Programs. *J. Comput. Chem.* **2005**, *26* (16), 1668−1688.

(47) Darden, T.; York, D.; Pedersen, L. Partical mesh Ewald - an Nlog(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 8577−8593.

(48) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of cartesian equations of motion of a system with constraints - molecular dynamics of N-alkanes. *J. Comput. Phys.* **1977**, *23* (3), 327−341.

(49) Roe, D. R.; Cheatham, T. E., III. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9*, 3084−3095.

(50) Ester, M.; Kriegel, H. P.; Sander, J.; Xu, X. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96)*; Simoudis, E., Han, J., Fayyad, U., Eds.; AAAI Press: Palo Alto, CA, 1996; pp 226−231.

(51) Kullback, S.; Leibler, R. A. On information and sufficiency. *Ann. Math. Stat.* **1951**, *22* (1), 79−86.

(52) Abraham, M. J.; Gready, J. E. Ensuring mixing efficiency of replica-exchange molecular dynamics simulations. *J. Chem. Theory Comput.* **2008**, *4* (7), 1119−1128.