Characterization of Conformational Equilibria Through Hamiltonian and Temperature Replica-Exchange Simulations: Assessing Entropic and Environmental Effects

JOSÉ D. FARALDO-GÓMEZ,[1] BENOÎT ROUX[1,2]
[1]*Department of Pediatrics, Institute for Molecular Pediatric Sciences, Gordon Center for Integrative Science, University of Chicago, Chicago, Illinois 60637*
[2]*Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, Illinois 60637*

Received 14 April 2006; Revised 16 June 2006; Accepted 2 July 2006
DOI 10.1002/jcc.20652
Published online 6 March 2007 in Wiley InterScience (www.interscience.wiley.com).

Abstract: Molecular dynamics simulations based on the replica-exchange framework (REMD) are emerging as a useful tool to characterize the conformational variability that is intrinsic to most chemical and biological systems. In this work, it is shown that a simple extension of the replica-exchange method, known as Hamiltonian REMD, greatly facilitates the characterization of conformational equilibria across large energetic barriers, or in the presence of substantial entropic effects, overcoming some of the difficulties of REMD based on temperature alone. In particular, a comparative assessment of the HREMD and TREMD approaches was made, through computation of the gas-phase free-energy difference between the so-called D_{2d} and S_4 states of tetrabutylammonium (TBA), an ionic compound of frequently used in biophysical studies of ion channels. Taking advantage of the greater efficiency of the HREMD scheme, the conformational equilibrium of TBA was characterized in a variety of conditions. Simulation of the gas-phase equilibrium in the 100–300 K range allowed us to compute the entropy difference between these states as well as to describe its temperature dependence. Through HREMD simulations of TBA in a water droplet, the effect of solvation on the conformational equilibrium was determined. Finally, the equilibrium of TBA in the context of a simplified model of the binding cavity of the KcsA potassium channel was simulated, and density maps for D_{2d} and S_4 states analogous to those derived from X-ray crystallography were constructed. Overall, this work illustrates the potential of the HREMD approach in the context of computational drug design, ligand-receptor structural prediction and more generally, molecular recognition, where one of the most challenging issues remains to account for conformational flexibility as well for the solvation and entropic effects thereon.

 J Comput Chem 28: 1634–1647, 2007

Key words: replica-exchange; conformational change; entropy; solvation effects; drug design

Introduction

In trying to characterize the microscopic origins of chemical and biological processes, one of the greatest challenges is to account for the conformational variability that in most cases is intrinsic to the underlying molecular system, and may even be indispensable for the process to occur. Computer simulation methods such as molecular dynamics (MD) have great potential for such investigations, since their framework relies on a physically realistic description of the energetic and dynamic properties of the system under study. Nevertheless, the scope of standard MD simulations is limited in practice; for example, the simulation of a chemical or biological process during which the system explores two or more conformational states separated by substantial free-energy barriers ($\geq 10\, k_B T$) may require an enormous computational cost, simply because the sampling of configurational space accomplished by this method is such that regions of high free energy are rarely visited. In such cases, the conformational transitions that are characteristic of the process considered will certainly be poorly described.

To overcome sampling limitations because of large energetic barriers, advanced simulation methods, such as umbrella sampling,[1,2] meta-dynamics,[3] or accelerated dynamics,[4] typically rely on the introduction of biasing potentials that modify the free-energy surface along specific multidimensional coordinates, as well as on suitable unbiasing schemes to recover the appropriate statistical

Correspondence to: B. Roux; e-mail: roux@uchicago.edu

weight of the resulting ensemble of configurations. Similarly, the exploration of complex and rugged free-energy landscapes such as those underlying protein folding can be enhanced greatly by generating a large number of unbiased dynamical trajectories by means of multiple standard MD simulations using world-wide computing resources,[5] or through algorithms based on generalized Monte Carlo schemes, such as transition path sampling.[6]

Replica-exchange molecular dynamics (REMD) is an alternative approach that combines biasing methods that alter the free-energy landscape and a Monte Carlo-based, multiple-simulation strategy.[7] In its most common implementation to date, the free-energy hypersurface of the molecular system is modified through temperature, which gradually increases across the ensemble of simulations (or replicas), in the hope that regions of the configurational space that are inaccessible at room temperature, for a given computational time, will be more frequently visited at high temperatures.[8,9] However, while temperature-REMD may be useful in specific cases, its general applicability is clearly limited by the possibility that the free-energy landscape at high temperature is drastically different from that at the temperature of interest, to the point that one or more of the conformational states involved in the equilibrium under study is no longer statistically significant. Moreover, a large number of replicas is required to simulate equilibria involving large free-energy barriers, or systems with many degrees of freedom,[10] which makes this method impractical for many systems of interest (see also Zuckerman and Lyman[11] for further discussion).

In this work, we show that a simple extension of the replica-exchange method, known as Hamiltonian REMD,[12,13] greatly facilitates the characterization of conformational equilibria across large energetic barriers, and in a variety of environments, overcoming some of the difficulties of REMD based on temperature alone. The ability to carry out such analyses is of great interest not only from the methodological viewpoint, for example, as a means to improve the quality of atomistic or coarse-grained forcefields; it is also of considerable importance in a variety of applications, notably computational drug design, ligand-receptor structural prediction, and more generally, molecular recognition, where one of the most challenging issues remain to account for conformational flexibility, as well for the solvation and entropic effects thereon.

To illustrate the performance of the HREMD method, which makes use of a gradual perturbation of the potential energy function across the system of replicas, we simulate the conformational equilibrium of tetrabutylammonium (TBA), an ionic compound with particular relevance in the fields of zeolites and biological ion channels, and illustrate the influence of entropic and environmental effects on this equilibrium. More specifically, we compute the room-temperature free-energy difference between the D_{2d} and S_4 conformers in several environments, namely in the gas phase, in water, and in association with a reduced model of the potassium channel KcsA. In addition, we also characterize the gas-phase equilibrium in the 100–300 K temperature range, which allows us to quantify the entropic contribution to the equilibrium energetics. On the basis of our observations, we anticipate that this or similar approaches will greatly assist future computational analyses of chemical and biological processes where the conformational variability of the underlying molecular systems is a determinant factor.

Methods

Hamiltonian and Temperature Replica-Exchange Simulation Framework

Molecular dynamics (MD) simulations typically aim to quantify some structural or dynamical property of a molecular system at equilibrium, for a given thermodynamic ensemble. For example, in the NVT or canonical ensemble, the equilibrium average value of a property A that may be dependent on the coordinates $\mathbf{X}$ and momenta $\mathbf{P}$ of the particles in the sytem is

$$\langle A \rangle = \int d\mathbf{X} d\mathbf{P}\, A(\mathbf{X}, \mathbf{P})\, \rho(\mathbf{X}, \mathbf{P}) \tag{1}$$

$$\rho(\mathbf{X}, \mathbf{P}) \propto \exp\left\{\frac{-\mathcal{H}_0(\mathbf{X}, \mathbf{P})}{k_B T_0}\right\} \tag{2}$$

where $\rho(\mathbf{X}, \mathbf{P})$ denotes the probability density of a given configuration of the phase space, and $\mathcal{H}_0$ and T_0 are the Hamiltonian and temperature of the system. In practice, however, this average is computed simply as the mean value of A throughout the simulation

$$\langle A \rangle \approx \langle A \rangle_{\text{sim}} = \frac{1}{N_s} \sum_{s=1}^{N_s} A_{\text{sim}}(s) \tag{3}$$

where N_s is the number of configurations generated, and $A(s)$ is the value of A for each of them. The extent to which the approximation in eq. (3) is reasonable depends on two factors: first, the simulation algorithm must be such that the ensemble of configurations generated is statistically consistent with the probability density $\rho(\mathbf{X}, \mathbf{P})$ in eq. (2); second, the ensemble of configurations must include those regions of the phase space $(\mathbf{X}, \mathbf{P})$ that contribute the most to the integral in eq. (1).

Modern MD algorithms are designed to satisfy the first of these conditions; however, insufficient computational power severely limits the extent of the configurational sampling that MD simulations can accomplish for many systems of interest. Typically, regions of the configurational space that may contribute significantly to the integral in eq. (1) are separated from the starting configuration of the simulation by substantial free-energy barriers, which by construction the MD algorithm rarely overcomes; thus, the approximation in eq. (3) is compromised.

To tackle this problem, methodologies such as umbrella sampling implement biasing potentials that force the system to explore relevant regions of the configurational space, and subsequently make use of unbiasing schemes to recover the consistency of the sampling with eq. (2). In replica-exchange molecular dynamics (REMD), by contrast, the simulation is coupled to an ensemble of simulations of the same molecular system, which are designed to more easily access relevant regions of the configurational space. In particular, these additional simulations, or replicas, may have a different Hamiltonian $\mathcal{H}$, a different temperature T, or both, which are chosen so as to reduce the free-energy barriers characteristic of the system with the Hamiltonian and temperature of interest, $\mathcal{H}_0$ and T_0. By allowing replicas to exchange their configurations every so often, the latter will have a much better chance of exploring all relevant

regions of the configurational space contributing to the integral in eq. (1), even though in actuality it has not climbed the free-energy barriers separating those regions.

As in umbrella sampling, however, it is crucial in REMD to maintain consistency of the configurational sampling with the probability density characteristic of the system under study. For a system of N replicas, and assuming the Hamiltonian can be separated into potential and kinetic energy functions, i.e., $\mathcal{H}(\mathbf{X}, \mathbf{P}) = \mathcal{U}(\mathbf{X}) + \mathcal{K}(\mathbf{P})$, this overall probability density function can be written as

$$\rho(\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_N, \mathbf{P}_1, \mathbf{P}_2, \ldots, \mathbf{P}_N) = \chi(\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_N)\, \xi(\mathbf{P}_1, \mathbf{P}_2, \ldots, \mathbf{P}_N) \quad (4)$$

$$\chi(\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_N) \propto \prod_{i=1}^{N} \exp\{-\beta_i\, \mathcal{U}_i(\mathbf{X}_i)\} \quad (5)$$

$$\xi(\mathbf{P}_1, \mathbf{P}_2, \ldots, \mathbf{P}_N) \propto \prod_{i=1}^{N} \exp\{-\beta_i\, \mathcal{K}_i(\mathbf{P}_i)\} \quad (6)$$

where β_i is $1/k_B T_i$. Although the MD algorithm is designed to sample the phase space according to the χ and ξ distributions in the time periods between exchanges, one needs to ensure that the exchange of configurations between replicas is consistent with these distributions as well. With respect to the spatial part χ, this can be achieved by allowing exchanges between replicas only with a certain probabilty, p_{ex}, which is configuration-dependent. To derive this exchange probability, we can think of the replica-exchange process as a Markov chain of states of the replica ensemble. That is, two states m and n can be understood as the state of the ensemble before and after a pair of replicas (i,j) have exchanged their respective configurations. Thus, the transition probability between them, π_{mn}, will be related to the exchange probability between replicas by the expressions

$$\begin{aligned} \pi_{mn} &= p_{ex}(i,j) \\ \pi_{nm} &= 1 - p_{ex}(i,j) \end{aligned} \quad (7)$$

For this Markov process to have a limiting distribution equal to χ, the transition probability π_{mn} must satisfy the condition of microscopic reversibility $\chi_m \pi_{mn} = \chi_n \pi_{nm}$. Following Metropolis,[14] a computationally efficient choice of π_{mn} would be

$$\begin{aligned} \pi_{mn} &= 1 \quad \chi_n \geq \chi_m \\ \pi_{mn} &= \chi_n / \chi_m \quad \chi_n < \chi_m \end{aligned} \quad (8)$$

Combining eq. (5) and eqs. (7 and 8) lead to a suitable definition of the exchange probability $p_{ex}(i,j)$

$$\begin{aligned} p_{ex}(i,j) &= 1 \quad \Delta(i,j) \leq 0 \\ p_{ex}(i,j) &= \exp(-\Delta(i,j)) \quad \Delta(i,j) > 0 \end{aligned} \quad (9)$$

$$\Delta(i,j) = \beta_i[\mathcal{U}_i(\mathbf{X}_j) - \mathcal{U}_i(\mathbf{X}_i)] + \beta_j[\mathcal{U}_j(\mathbf{X}_i) - \mathcal{U}_j(\mathbf{X}_j)] \quad (10)$$

which will guarantee that the sampling of the coordinate space in every replica will be consistent with χ, as if no exchanges were taking place.

Concerning the kinetic part of the canonical distribution, ξ, an efficient and convenient strategy is to reinitialize the velocities of all particles in the system upon every exchange of configurations, extracting their new values from a Maxwell-Boltzmann distribution at temperature T_i, which is by definition consistent with eq. (6). It can be shown that this approach is as accurate or more in reproducing the canonical ensemble (e.g., for a system of harmonic oscillators) as other schemes where velocity scaling factors of the form β_i/β_j are introduced after each exchange, or where the full Hamiltonian function is included in the Metropolis test in eqs. (9 and 10).

All simulations in this report implement the scheme outlined above or variations thereof. For example, in the constant-temperature Hamiltonian REMD simulations (HREMD) presented hereafter, only some of the terms of the potential energy function $\mathcal{U}(\mathbf{X})$ are modified across replicas (see the following section for further details), specifically through a scaling parameter, i.e., $\mathcal{U}_i = \mathcal{U}_A + \lambda_i \mathcal{U}_B$ (a similar scheme has been recently reported by Affentranger et al.[15]); thus eq. (10) reduces to

$$\Delta(i,j) = \beta\,(\lambda_i - \lambda_j)[\mathcal{U}_B(\mathbf{X}_j) - \mathcal{U}_B(\mathbf{X}_i)] \quad (11)$$

In temperature REMD (TREMD), by contrast, the Hamiltonian is the same for all replicas and therefore

$$\Delta(i,j) = (\beta_i - \beta_j)[\mathcal{U}(\mathbf{X}_j) - \mathcal{U}(\mathbf{X}_i)] \quad (12)$$

From eqs. (10–12), it is also apparent that the replica-exchange algorithm may in practice involve either the exchange of configurations between replicas, i.e., the exchange of atomic coordinates for all particles in the system, while each replica retains its temperature and/or Hamiltonian, or the exchange of temperatures and/or Hamiltonian functions, while each replica preserves the configuration of the simulation system. In this work, the latter approach was adopted, since it was found to be more convenient computationally.

In all cases, it is important to realize that the efficiency of the REMD scheme relies on $\Delta(i,\, j)$ being a sufficiently small number, on average, to yield a large enough acceptance probability of exchange, and thus ensure that the replicas are actually coupled. For this reason, exchanges are typically attempted only between replicas whose Hamiltonian or temperature are adjacent in the ensemble, i.e., $\mathcal{H}_j = \mathcal{H}_{i+1}$ and $T_j = T_{i+1}$, and the temperatures T_i and coupling parameters λ_i are generally exponentially distributed. Similarly, the frequency at which exchanges are attempted should be as large as possible, but not so large that it precludes the replicas from sampling significantly diverse regions of the configurational space between exchanges.

Description of the Simulation System

We have analyzed the conformational equilibrium of tetrabutylammonium (TBA) in a variety of conditions, including the interior of a simplified model of the KcsA K^+ channel protein. TBA is an interesting example because it primarily exists in either of two conformations, known as D_{2d} and S_4, which are separated by a large energy barrier and have different degeneracy (six and three, respectively);[16] additionally, it is expected that entropic and solvation

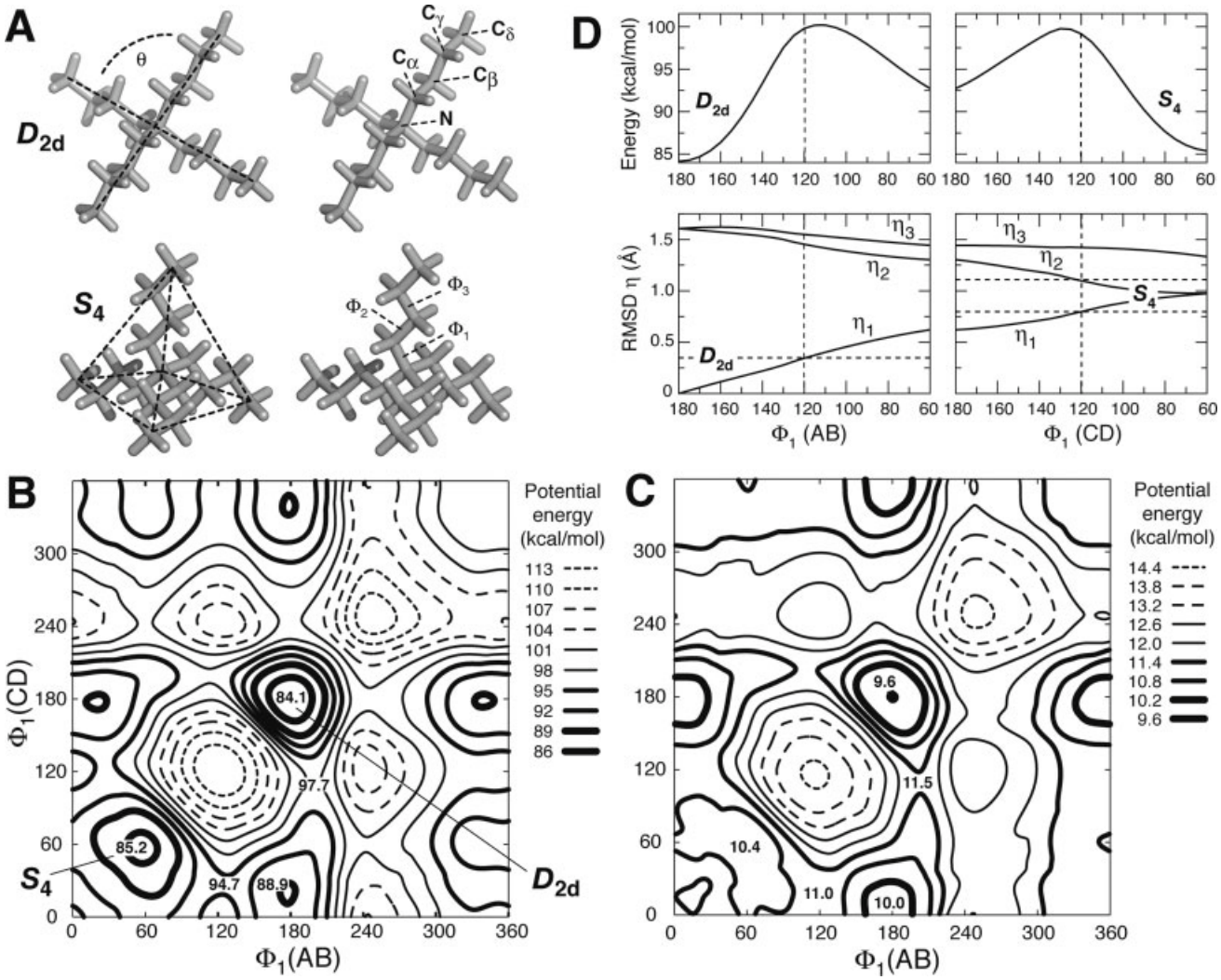

Figure 1. A. Molecular representation of tetrabutylammonium (TBA) in its most probable conformations, known as D_{2d} and S_4. The transition between these two states, which are three- and six-fold degenerate, respectively, involves a concerted 120° rotation of the central torsions Φ_1 (C'_α, N, C_α, C_β). B and C. Potential-energy surface along two of the torsions Φ_1 (C'_α, N, C_α, C_β), for the energy function given in eq. (14) with $\lambda = 1.0$ and $\lambda = 0.10$, respectively. D. Top panel: one-dimensional potential-energy profile connecting the D_{2d} and S_4 states, first moving along the x axis in (B), and then along the y axis. Bottom panel, for the same pathway, evolution in the root-mean-squared deviation, η_i, with respect to each of the three degenerate D_{2d} ideal conformations; as can be seen, the minimum of the three curves (in this example, η_1) discriminates unambiguously the D_{2d} and S_4 states, as well as their intermediate state.

effects will also influence its conformational equilibrium, reflecting the alternate packing of the hydrocarbon chains and different exposure of the core of the molecule (Fig. 1A).

Structurally, the ideal geometries of the D_{2d} and S_4 conformers differ in the arrangement of the 12 central torsions Φ_1 (C'_α, N, C_α, C_β), of which only four are independent, and thus these two conformations are also characteristic of TBA analogues such as tetraethylammonium (TEA) and tetrapentylammonium (TPA). To model the conformational equilibrium of this molecule, we employed the CHARMM27 forcefield,[17] corrected so as to exactly reproduce the quantum-mechanical potential-energy difference between the D_{2d} and S_4 conformations of TEA, namely 0.9 kcal/mol in favor of D_{2d}.[18] Because all bonded terms of the classical potential-energy function are identical in the ideal D_{2d} and S_4 geometries, this correction is applied to the four pseudoangles θ (C'_β, N, C_β), via the Saxon-Wood flat-bottom potential:

$$\mathcal{U}_{\text{corr}} = h\left[1 + \exp\left\{\frac{p_2 - \|\theta - \theta_{\text{ref}}\|}{p_1}\right\}\right]^{-1} \qquad (13)$$

where $h = 1.4$ kcal/mol, $\theta_{\text{ref}} = 90.0$, $p_1 = 0.03$, and $p_2 = 0.3$. The resulting potential-energy surface connecting the D_{2d} and S_4 conformations of TBA, shown in Figure 1B, yields an energy difference

of 1.1 kcal/mol in favor of D_{2d}, with energy barriers on the order of 10 kcal/mol separating these and their intermediate states.

Simulation Details and Statistical Analysis

All simulations were carried out with a modified version of the CHARMM c32a2 molecular simulation software,[19] using its Langevin-equation integrator with a collision frequency $\beta = 5\,\mathrm{ps}^{-1}$, and full electrostatics and van der Waals interactions. In the Hamiltonian REMD simulations, the CHARMM potential energy function was modified across replicas through a single scaling factor λ_i ranging from 0.10 to 1.00, that is,

$$\mathcal{U}_i = \mathcal{U}_{\mathrm{bonds}} + \mathcal{U}_{\mathrm{angles}}^{(\mathrm{core})} + \lambda_i \left[\mathcal{U}_{\mathrm{elec}} + \mathcal{U}_{\mathrm{vdw}} + \mathcal{U}_{\mathrm{torsions}} + \mathcal{U}_{\mathrm{angles}}^{(\mathrm{other})}\right] + (1 - 0.75\,(1-\lambda_i))\,\mathcal{U}_{\mathrm{corr}} \equiv \mathcal{U}_{\mathrm{A}} + \lambda_i \mathcal{U}_{\mathrm{B}} \quad (14)$$

The modified potential energy surface of TBA, according to eq. (14) and $\lambda_i = 0.10$, is shown in Figure 1C; while the relative energy of the D_{2d} and S_4 conformations is changed only slightly with respect to that with $\lambda_i = 1.00$ (0.8 kcal/mol in favor of D_{2d}), the energy barriers separating these and the intermediate state are dramatically reduced (from more that 10 to less than 2 kcal/mol). In the replica simulation of TBA in water, the additional contributions to the potential energy, namely water–water bonded and nonbonded terms as well as TBA-water interactions, were not scaled and thus were identical across replicas; by contrast the TBA-protein interactions in the simulation of the simplified model of KcsA were scaled down to facilitate the conformational search within the cavity.

To assess the convergence of the REMD method as a function of simulation time, our strategy has been to set up all replicas in each simulation in either the D_{2d} or the S_4 conformation, and simulate the replica system until approximately the same populations of each state, $P(D_{2d})$ and $P(S_4)$, were achieved in the ensemble corresponding to $T = 300\,\mathrm{K}$ and $\lambda = 1.00$. To compute these populations, we recorded the value of a conformational coordinate η that unambiguously distinguishes these two conformations as well as the transition state between them (Fig. 1D). Specifically, this coordinate is defined as

$$\eta = \min\{\eta_1, \eta_2, \eta_3\} \quad (15)$$

where η_i denotes the root-mean-square deviation of a given conformation of TBA during the simulation with respect to the ideal geometries of each of the degenerate D_{2d} conformations. From the probability density $\rho(\eta)$, we can derive the populations of each state, and thus their relative free energy, via the expression

$$\Delta F = -k_{\mathrm{B}}T\,\ln\left\{\frac{P(D_{2d})}{P(S_4)}\right\} = -k_{\mathrm{B}}T\,\ln\left\{\frac{\int_{\eta\in D_{2d}} d\eta\,\rho(\eta)}{\int_{\eta\in S_4} d\eta\,\rho(\eta)}\right\} \quad (16)$$

where the integration range for the D_{2d} and S_4 states are $\eta \in [0, 0.3]$ and $\eta \in [0.85, 1.20]$, respectively. In practice, the reported values of the free-energy difference derive from a block average comprising the latest 80% of the sampling obtained at $T = 300\,\mathrm{K}$ and with $\lambda = 1.00$ from each simulation, that is

$$\overline{\Delta F} = \frac{1}{m}\sum_{i=1}^{m} \Delta F_i \quad (17)$$

with $m = 4$ blocks for all time-scales. The error of this block average is defined as the standard error of the mean

$$e = \frac{\sigma(\Delta F)}{\sqrt{m}} = \sqrt{\frac{\sum_{i=1}^{m}(\Delta F_i - \overline{\Delta F})^2}{m^2}} \quad (18)$$

Lastly, the vibrational analysis used for the calculation of conformational entropies of TBA in the D_{2d} and S_4 conformations was carried out with the VIBRAN module of CHARMM.[20]

Results and Discussion

Hamiltonian Vs. Temperature REMD

We first compare the efficiency of the temperature and Hamiltonian REMD (TREMD and HREMD) approaches in reproducing the conformational equilibrium of TBA, for an equivalent computational cost. To this end, we carried out two independent 13-replica simulations for each method, where the starting conformation of TBA in all replicas, $\mathbf{X}_{\mathrm{init}}$, was set to be either D_{2d} or S_4; the sampling time per replica in all four simulations was 100 ns. The temperature range used in the TREMD simulations was 300–600 K, while the potential-energy scaling factor λ in the HREMD simulations ranged between 1.00 and 0.10, and the temperature was 300 K for all replicas; in both cases, the values of temperature and/or λ assigned to each replica followed an exponential distribution. Exchanges were attempted only between randomly chosen replicas that were adjacent in temperature or scaling parameter λ, every 1 ps of simulation.

The proportion of replica exchanges that were accepted during these simulations are shown in Table 1, for those pairs within which exchange attempts were made. By virtue of the exponential distribution of the values of temperature and λ, the acceptance ratios were found to be uniform across replica pairs in the TREMD simulations, and slightly less so for the HREMD, because of the fact that the scaling parameter is applied to only a subset of the potential energy function. At any rate, these acceptance ratios, around 74% in TREMD and in the 50–65% range for HREMD, are expected to be more than acceptable. To validate this expectation, we analyzed the fraction of the simulation time during which each replica had a given temperature, or a given scaling parameter. The range of these residence times across all replicas are shown in Table 2. As can be seen, in all four simulations, all replicas were found to run at all temperatures or with all scaling parameters, for a fraction of the simulation time that is close to the ideal value, i.e., 7.7% or 1/13 of the simulation time.

Although the analysis above demonstrates that the REMD scheme is coupling all replicas efficiently and uniformly, it remains to be determined whether the ensemble of configurations obtained at each temperature, or with a given Hamiltonian, are correct from the statistical viewpoint, i.e., whether they remain within the

Table 1. Probabilities of Replica-Exchange Acceptance From the TREMD (Left) and HREMD (Right) Gas-Phase Simulations, for the Pairs of Replicas Whose Temperature T_i or Hamiltonian-Scaling Parameter λ_i Are Adjacent in the Replica Ensemble (No Exchanges Were Attempted Otherwise).

	Exchange acceptance ratio (% attempts)			Exchange acceptance ratio (% attempts)	
(T_i, T_j)	$X_{init} = D_{2d}$	$X_{init} = S_4$	(λ_i, λ_j)	$X_{init} = D_{2d}$	$X_{init} = S_4$
(300, 318)	73.6	73.5	(1.00, 0.83)	50.7	50.9
(318, 337)	73.3	73.8	(0.83, 0.68)	49.1	48.6
(337, 357)	73.8	74.1	(0.68, 0.56)	50.8	50.6
(357, 378)	74.1	74.7	(0.56, 0.46)	51.3	51.4
(378, 400)	74.3	74.1	(0.46, 0.38)	53.7	53.8
(400, 424)	73.6	74.1	(0.38, 0.32)	59.2	59.6
(424, 449)	73.9	74.1	(0.32, 0.26)	54.6	53.9
(449, 476)	73.2	73.3	(0.26, 0.22)	64.1	64.4
(476, 504)	73.8	74.2	(0.22, 0.18)	59.5	59.7
(504, 534)	74.1	74.0	(0.18, 0.15)	64.3	64.5
(534, 566)	73.8	73.8	(0.15, 0.12)	58.9	58.9
(566, 600)	73.6	73.6	(0.12, 0.10)	67.1	67.2

The overall acceptance ratios in the TREMD simulations were 73.8% and 73.9%, depending on whether the starting conformation of the replicas was D_{2d} or S_4; the overall acceptance ratio in the HREMD simulations was 56.9% in both cases.

corresponding canonical distribution. In TREMD, a simple test relies on deriving the potential-energy probability density characteristic of each ensemble of configurations, $\rho_\beta(U)$, and computing pairwise ratios between them. As shown in eqs. (19 and 20) below, the natural logarithm of these ratios must have a linear relationship with the potential energy, which depends on the temperatures of the ensembles that are compared. That is,

$$\rho_\beta(U) \equiv \langle \delta[\mathcal{U}(\mathbf{X}) - U] \rangle = \frac{\int d\mathbf{X} \delta[\mathcal{U}(\mathbf{X}) - U] \mathrm{e}^{-\beta\,\mathcal{U}(\mathbf{X})}}{\int d\mathbf{X}\, \mathrm{e}^{-\beta\,\mathcal{U}(\mathbf{X})}} = C_\beta\, \mathrm{e}^{-\beta\,U} \quad (19)$$

$$\ln\left\{\frac{\rho_\beta(U)}{\rho_{\beta'}(U)}\right\} \propto (\beta' - \beta)U \quad (20)$$

An analogous test can be carried out for HREMD, except that now the potential-energy probability density is derived only for the part of the Hamiltonian that is modified across replicas, $\mathcal{U}_\mathrm{B}$; the linear dependence of the pair-wise ratios with the value of $\mathcal{U}_\mathrm{B}$ is determined by the respective scaling parameters, that is,

$$\rho_\lambda(U_\mathrm{B}) \equiv \langle \delta[\mathcal{U}_\mathrm{B}(\mathbf{X}) - U_\mathrm{B}] \rangle = \frac{\int d\mathbf{X} \delta[\mathcal{U}_\mathrm{B}(\mathbf{X}) - U_\mathrm{B}]\, \mathrm{e}^{-\beta\,\mathcal{U}_\lambda(\mathbf{X})}}{\int d\mathbf{X} \mathrm{e}^{-\beta\,\mathcal{U}_\lambda(\mathbf{X})}} = C_\lambda\, \mathrm{e}^{-\beta\,\lambda U_\mathrm{B}} \int d\mathbf{X}\, e^{-\beta\,\mathcal{U}_\mathrm{A}(\mathbf{X})} \quad (21)$$

$$\ln\left\{\frac{\rho_\lambda(U_\mathrm{B})}{\rho_{\lambda'}(U_\mathrm{B})}\right\} \propto \beta\,(\lambda' - \lambda)\,U_\mathrm{B} \quad (22)$$

As shown in Figure 2, the result of such an analysis for the current TREMD and HREMD simulations is satisfactory, i.e., in spite of the coupling of replicas through the exchange process, the potential-energy distributions of the ensemble of configurations obtained in these simulations are statistically consistent with the assumed temperatures and Hamiltonian functions. As discussed earlier, this simply reflects that the exchange probability derived in eqs. (4–10) was used correctly.

Having established the correctness of our implementation of the REMD algorithm from a methodological viewpoint, we can now assess the perfomance of the TREMD and HREMD simulations comparatively. Specifically, we have analyzed the computed free-energy difference between the D_{2d} and S_4 states as a function of simulation time, following the procedure outlined in eqs. (15–18). As shown in Figure 3 and Table 3, while both the TREMD and HREMD methods roughly converge to similar values of the free-energy difference after 100 ns of simulation, the HREMD is clearly more efficient in the current case. For example, comparing the consensus value from the two replica simulations that start in either the D_{2d} or S_4 conformation, the HREMD simulation yields the correct sign of ΔF about five times faster than the TREMD simulation ($\sim$10 vs. 50 ns, respectively). In fact, it is unclear from the TREMD data alone whether a significant change in the free-energy difference may be expected if, say, the simulation time is extended 5- or 10-fold, since the profiles derived from the two TREMD simulations cross only once, after around 60 ns. By contrast, the HREMD simulations decorrelate from their starting conditions within a few nanoseconds, and yield consistent values of ΔF throughout the second half of the simulation, thus providing much greater precision for an equivalent computational cost.

Table 2. Minimum and Maximum Residence Times (in % of the Total Simulation Time) During Which Any Given Replica in the Ensemble Was Running at Temperature T_i or With Hamiltonian Scaling Parameter λ_i, in Either the TREMD (Left) or HREMD (Right) Gas-Phase Simulations.

	Replica residence times (min–max in % simulation time)			Replica residence times (min–max in % simulation time)	
T_i	$X_{init} = D_{2d}$	$X_{init} = S_4$	λ_i	$X_{init} = D_{2d}$	$X_{init} = S_4$
300	6.2–11.2	5.2–10.5	1.00	6.2–8.8	6.5–9.2
318	6.2–10.6	5.8–10.2	0.83	6.3–8.6	6.8–9.0
337	6.5–10.1	5.9–9.8	0.68	6.1–8.4	6.9–8.2
357	6.7–9.4	6.4–9.2	0.56	6.5–8.1	7.3–8.1
378	6.6–8.7	7.0–9.0	0.46	7.0–8.2	7.2–8.5
400	6.8–8.2	7.3–8.4	0.38	7.3–8.3	7.3–8.5
424	7.3–8.2	7.3–8.2	0.32	7.2–8.0	7.2–8.2
449	7.0–8.3	6.8–8.3	0.26	7.2–8.3	7.3–8.0
476	6.3–8.4	6.5–8.8	0.22	7.0–8.4	7.3–8.1
504	5.8–8.7	6.0–9.0	0.18	7.0–8.6	7.2–8.4
534	5.4–9.1	5.6–9.1	0.15	6.7–9.0	7.3–8.4
566	5.0–9.3	5.1–9.3	0.12	6.7–8.8	7.0–8.4
600	4.6–9.6	4.9–9.5	0.10	6.9–8.7	6.9–8.5

For a simulation of infinite length, the residence time would be equal to 1 over the number of replicas, i.e. 7.7%.

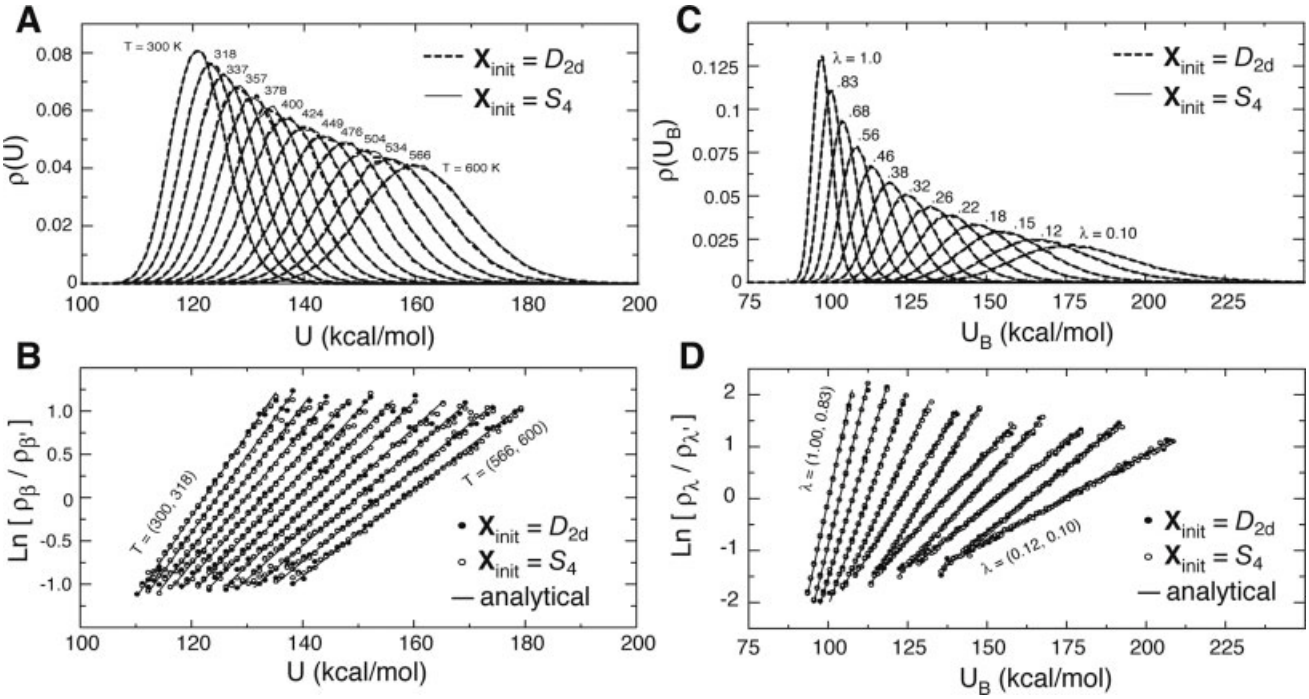

Figure 2. A. Distributions of the probability density of the potential energy, derived from the TREMD gas-phase simulations. Each distribution corresponds to the ensemble of conformations sampled at temperature T_i by the replica system, either when the starting conformation was D_{2d} or S_4. B. Pairwise ratios of the probability density distributions from the TREMD simulations, compared with the analytical curve (eqs. (19 and 20)). C. Distributions of the probability density of the potential-energy contribution U_B (defined in eq. (14)), derived from the HREMD gas-phase simulations. Each distribution corresponds to the ensemble of conformations sampled with a given scaling parameter λ_i. D. Pairwise ratios of the probability density distributions from the HREMD simulations, compared with the analytical curve (eqs. (21 and 22)).

To understand the reasons underlying the poorer performance of the TREMD approach relative to HREMD, we have analyzed the extent to which the replica ensemble enhances the probability of observing conformational transitions at $T = 300$ K and $\lambda = 1.00$ in each case. Specifically, we identified the number of trajectories in which, by virtue of the exchange between replicas, the molecule first exists at $T = 300$ K and $\lambda = 1.00$, transiently visits other temperatures $T_i > 300$ K or Hamiltonian functions with $\lambda_i < 1.00$, and finally returns to $T = 300$ K and $\lambda = 1.00$; then, we computed the number of these excursions in which the conformational state of TBA differs at the endpoints. As shown in Table 4, the likelihood of a conformational transition in the TREMD simulations was found to be about 15 times smaller than in HREMD; in other words, the extent to which the replica ensemble facilitates the exploration of the conformational equilibrium, in this case, is much more limited when temperature is used to reduce the free-energy barriers that when the Hamiltonian function is modified. Consistently, the TREMD simulations take a substantially longer time to decorrelate from their starting configuration.

A seemingly obvious enhancement of the TREMD approach, in cases such as the current one where large energy barriers are present, would be to increase the temperature range to a sufficiently high value. Leaving aside the fact that this choice would entail a larger number of replicas for an equivalent exchange acceptance probability—and thus a greater computational cost—a larger problem relates to the impact of the temperature itself on the free-energy surface that governs the conformational equilibrium. To illustrate this point, in Figure 4, we plot the probability density of the conformational coordinate defined in eq. (15) from an additional 100-ns, 13-replica HREMD simulation at 600 K analogous to that reported above at 300 K. As can be seen, at this temperature, the S_4 conformation becomes predominant ($\Delta F \sim +1.5$ kcal/mol), while the free-energy barriers between this and the D_{2d} conformation are not reduced significantly; though these barriers can be expected to diminish further at even higher temperatures, it is clear, in this case, that the shift in the conformational equilibrium of TBA will also be more pronounced. From a methodological viewpoint, this poses a problem in that the computation of the relative free energy of the D_{2d} and S_4 conformations can hardly be carried out reversibly via TREMD. In other words, while a TREMD simulation in which all replicas start in the D_{2d} conformation will gradually yield an increasing population of S_4, the opposite cannot be expected, since at low temperatures the free-energy barriers are too large to enable transitions, and at high temperature, the equilibrium is so shifted in favor of the S_4 conformation that the D_{2d} state will be rarely visited. Thus, in problems such as this where the conformational equilibrium is characterized by large energy barriers, and where entropy appears to contribute importantly (see next section), it can be expected that REMD approaches based on the perturbation of the Hamiltonian function will be more efficient and result in a greater precision than

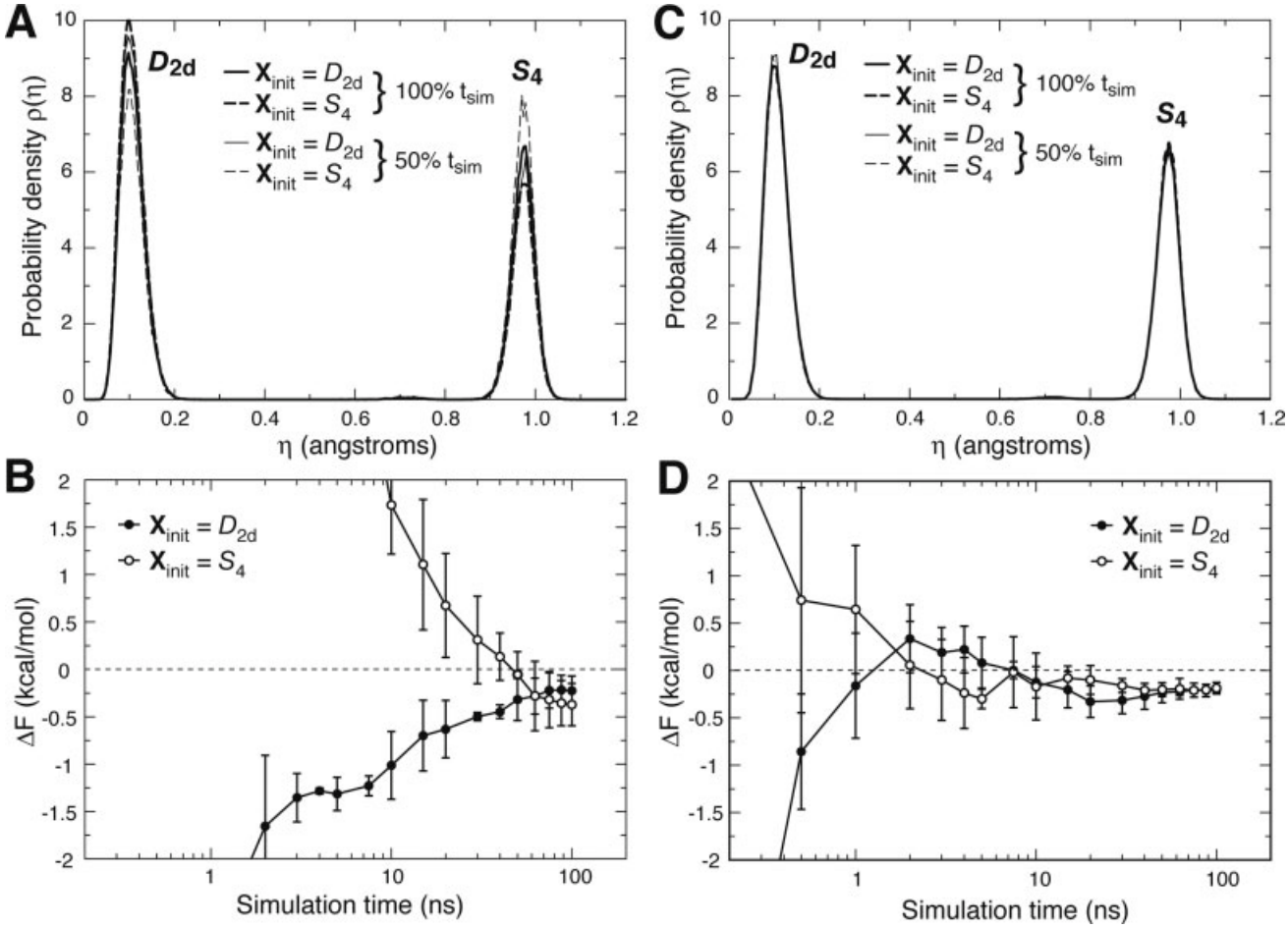

Figure 3. A. Distribution of the probability density of the reaction coordinate η (defined in eq. (15)), computed from the ensemble of conformations at $T = 300\,\mathrm{K}$ generated by the TREMD gas-phase simulations. Several distributions are shown to illustrate the dependence on the starting configuration of the replica system as well as on the duration of the simulation. B. Convergence of the room-temperature free-energy difference between the D_{2d} and S_4 states, computed via eqs. (16–18) from the TREMD simulations. C. Distributions of the probability density of η, calculated through the HREMD gas-phase simulations, for the ensemble of conformations at $T = 300\,\mathrm{K}$ and $\lambda = 1.0$. D. Convergence of the free-energy difference between the D_{2d} and S_4 states, as derived from the HREMD simulations.

those based on increasing the temperature. Nonetheless, it is clear that the applicability of the HREMD methodology relies critically on whether a suitable perturbative scheme can be identified and implemented, in contrast to TREMD, where no insight into the nature of the conformational equilibrium is required a priori.

More broadly speaking, it is worth noting that similar considerations can plausibly be applicable to related problems such as protein folding, where the conformational free-energy barriers may be relatively lower but where entropy also plays a key role. In this context, while a large number of TREMD simulations of protein unfolding can be found in the literature, it is not evident whether the TREMD approach would yield the same conformational free-energy surface regardless of the starting configuration of the replica ensemble. In particular, given the greater entropy of the unfolded state, it is likely that the TREMD method will be in general inefficient at simulating folding transitions (that is, a

Table 3. Free-Energy Difference Between the D_{2d} and S_4 Conformations of TBA, as Derived from the TREMD and HREMD Gas-Phase Simulations, and Dependence of the Computed Values on Simulation Time, as Well as on the Starting Conformation of the Replicas (in kilocalories per mole).

		$\Delta F = F(D_{2d}) - F(S_4)$		
		$t_{sim} = 30$ ns	$t_{sim} = 50$ ns	$t_{sim} = 100$ ns
TREMD	$\mathbf{X}_{init} = D_{2d}$	-0.50 ± 0.04	-0.32 ± 0.22	-0.23 ± 0.16
	$\mathbf{X}_{init} = S_4$	0.31 ± 0.46	-0.05 ± 0.04	-0.37 ± 0.22
HREMD	$\mathbf{X}_{init} = D_{2d}$	-0.31 ± 0.14	-0.23 ± 0.10	-0.21 ± 0.02
	$\mathbf{X}_{init} = S_4$	-0.16 ± 0.07	-0.20 ± 0.07	-0.19 ± 0.06

The values of ΔF shown correspond to block averages over the last 80% of the simulated time in each case, with the corresponding error estimate (see Methods section). For $t_{sim} = 100\,\mathrm{ns}$, the difference in the average internal energy of TBA is $\Delta U_{TBA}^{(gas)} \simeq -1.5\,\mathrm{kcal/mol}$.

Table 4. Efficiency of the Replica Ensemble in Enhancing Conformational Transitions Between the D_{2d} and S_4 States of TBA, in Either the TREMD and or HREMD Gas-Phase Simulations.

		Number of excursions	Percentage of transitions
TREMD	$\mathbf{X}_{\text{init}} = D_{2d}$	46,159	0.24
	$\mathbf{X}_{\text{init}} = S_4$	46,343	0.24
HREMD	$\mathbf{X}_{\text{init}} = D_{2d}$	31,082	3.19
	$\mathbf{X}_{\text{init}} = S_4$	31,109	3.01

To compute this efficiency, we identified all trajectories where TBA starts at $T = 300$ K or $\lambda = 1.00$, subsequently exchanges T or λ with one or more replicas, and ultimately returns to $T = 300$ K or $\lambda = 1.00$. We then compute the proportion of these "excursions" where the conformation of TBA differs at the endpoints of the trajectory.

simulation where the initial configuration of all replicas corresponds to an unfolded protein) since the temperature bias that in principle is meant to accelerate the rate of conformational changes also diminishes the likelihood of the folded state. Nevertheless, whether Hamiltonian REMD can perform better than TREMD in protein folding has not been established either; given the complexity and cooperativity of these processes, it may not be trivial in many cases to define a potential-energy perturbative scheme that enhances the sampling of folding/unfolding transitions. At any rate, it is clear that the extent to which any REMD method represents a substantial improvement over less computationally costly approaches, such as high-temperature unfolding simulations, depends on whether the free-energy surfaces thus determined have a greater precision, which in turn relies on the appropriate, reversible sampling of the conformational equilibrium.

Computation of Conformational Entropy Through Combined Hamiltonian and Temperature REMD

As seen earlier, entropic effects contribute significantly to the conformational equilibrium of TBA, in particular, favoring the S_4 state. Evidently, the greater degeneracy of this state, namely two-fold relative to the D_{2d} conformation, is consistent with this observation, and can be expected to shift the equilibrium towards S_4 by $-k_B \ln(1/2) = 1.32 \times 10^{-3}$ kcal/mol per temperature degree. However, this contribution alone cannot account for the changes in ΔF upon temperature increases from $T \sim 0$ (-1.1 kcal/mol) to 300 and 600 K (-0.2 and $+1.5$ kcal/mol, respectively), and therefore an additional entropic contribution to the conformational equilibrium must exist.

A possible approach to estimate this additional contribution, $\Delta S_{\text{conf}}(T)$, across a given temperature range, relies on the computation of the corresponding temperature-dependent free-energy differences, which can then be related to the conformational entropy through the expresssions

$$\Delta F(T) = \Delta U(T) - T\Delta S(T) \quad (23)$$

$$-T\Delta S(T) = -k_B T \ln(1/2) - T\Delta S_{\text{conf}}(T) \quad (24)$$

where $\Delta U(T)$ denotes the average potential energy of the system, or internal energy, at a given temperature. Taking advantage of the fact that the conformational equilibrium of TBA can be efficiently described through the HREMD approach described earlier, even at room temperature, it is possible to readily estimate $\Delta F(T)$, $\Delta U(T)$, and thus $\Delta S(T)$, across any temperature range, by coupling the Hamiltonian replicas to an additional ensemble of temperature replicas. Specifically, we used this combined Hamiltonian and temperature REMD scheme (HTREMD) to simulate the conformational equilibrium of TBA in the 100–300 K temperature range, since this is particularly challenging from the perspective of sampling in the context of traditional MD simulations. Half of the 26 replicas employed in this case spanned this temperature range, with $\lambda = 1.00$, while the other half spanned the range in λ, from 1.00 to 0.10, at $T = 300$ K. As before, two independent simulations were carried out where all 26 replicas start in either the D_{2d} and S_4 conformation.

The results from these simulations are summarized in Figure 5. As in the HREMD simulation reported previously, after $\sim$100 ns simulation per replica, the free-energy difference at room temperature is well converged to around -0.20 kcal/mol, which corresponds approximately to a population ratio of 59:41 (Fig. 5A). As the temperature decreases and the potential energy term in eq. (23) becomes dominant, the equilibrium population of S_4 diminishes (Fig. 5B), down to less than 1% at 100 K, and $\Delta F \simeq -0.92$ kcal/mol. Although the convergence of the low-temperature free-energy differences is not as satisfactory as for 300 K, it is encouraging, from the methodological standpoint, that this subtle balance can be reproduced regardless of the starting configuration of the simulation.

On the basis of the values of $\Delta F(T)$ and $\Delta U(T)$, we can now evaluate the entropic term $-T\Delta S(T)$ in eq. (24). As shown in Figure 5C, the temperature dependence of this term is roughly linear in the 100–300 K range, i.e., in this temperature range the entropy

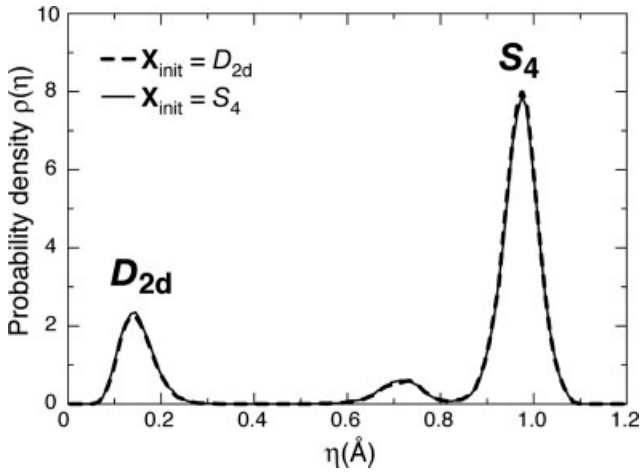

Figure 4. Distribution of the probability density of the reaction coordinate η (defined in eq. (15)), calculated from the HREMD gas-phase simulations at $T = 600$ K. The corresponding free-energy differences between the D_{2d} and S_4 states are $+1.53 \pm 0.17$ kcal/mol for the simulation starting in D_{2d} and $+1.47 \pm 0.18$ kcal/mol for the simulation starting in S_4. The corresponding difference in the average potential energy of TBA is $\Delta U \simeq -2.3$ kcal/mol in both cases; thus, at this temperature, $-T\Delta S(T) \simeq -3.8$ kcal/mol (eq. (23)).

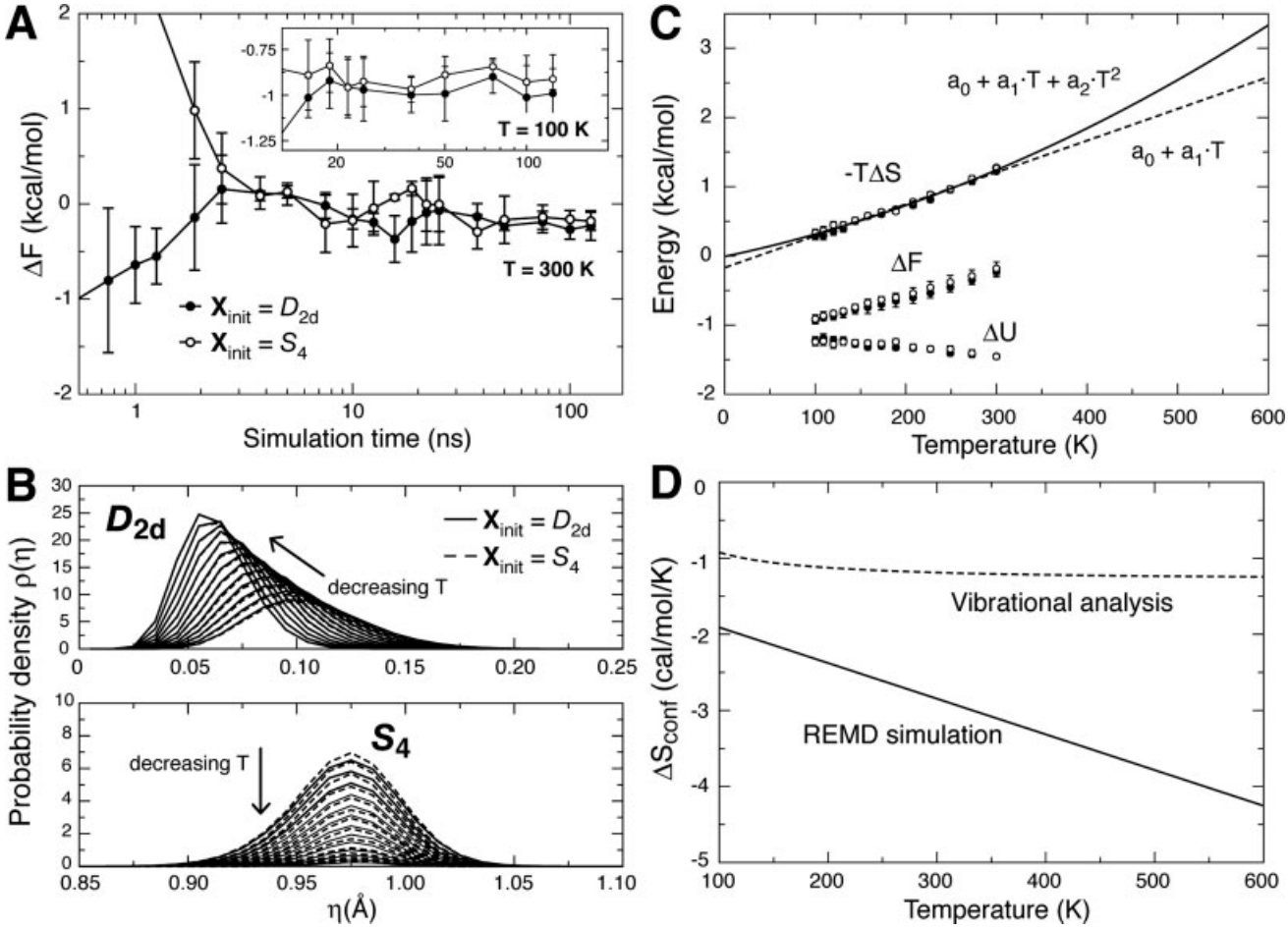

Figure 5. A. Convergence of the free-energy difference between the D_{2d} and S_4 states at 300 K and 100 K, computed via eqs. (16–18) from the combined Hamiltonian and temperature replica-exchange MD simulations of TBA in the gas phase. B. Distribution of the probability density of the conformational coordinate η (defined in eq. (15)), as function of the temperature (all curves correspond to $\lambda = 1.0$). C. Temperature-dependence of the free-energy difference between the D_{2d} and S_4 states, $\Delta F(T)$, and of the internal energy (average potential energy), $\Delta U(T)$. From these data, we derive the $-T\Delta S(T)$ using eq. (23), which we then fit to either a linear or quadratic function; the correlation coefficients for the linear and quadratic fits are 99.3 and 99.4%, respectively, but only the quadratic fit is in agreement with the low- and high-temperature expected values, namely 0 and 3.8 kcal/mol (Fig. 4), respectively (neither limit is used in the fit). D. Comparison of the conformational entropy difference, $\Delta S_{conf}(T)$, derived from vibrational analysis of the D_{2d} and S_4 conformations, and from the HTREMD simulations through eqs. (23 and 24).

difference between the D_{2d} or S_4 is approximately constant. However, extrapolation to $T \sim 0$ or $T = 600$ K fails to agree with the expected values, namely 0 and 3.8 kcal/mol, respectively (estimated from the data in Fig. 4). By contrast, a quadratic fit of the form $-T\Delta S(T) = a_0 + a_1 T + a_2 T^2$ does yield much better agreement for both the low and high temperature limits, in particular, with $a_0 = -0.0009$ kcal/mol, $a_1 = 2.76 \times 10^{-3}$ kcal/mol/K, and $a_2 = 4.69 \times 10^{-6}$ kcal/mol/K^2. Thus, from eq. (24), our resulting estimate of the conformational entropy difference is $\Delta S_{conf}(T) \simeq -(1.44 \times 10^{-3} + 4.69 \cdot 10^{-6}\, T)$ kcal/mol/K. Note that at room temperature, ΔS_{conf} outweighs the effect of the different degeneracy of the two states by a factor of two, and is thus the main contribution to the entropic effect on the conformational equilibium of TBA. Interestingly, calculation of ΔS_{conf} through vibrational analysis of TBA in either the D_{2d} or S_4 conformation yields values between -0.9×10^{-3} and -1.25×10^{-3} kcal/mol/K in the 100–600 K range (Fig. 5D), which approximately accounts for the temperature-independent component of ΔS_{conf}, as derived from the HTREMD simulations.

Broadly speaking, differences in entropy between two states of a system at a given temperature reflect that the range of conformational space that is accessible along one or more of the system's degrees of freedom is wider in one state than in the other. Upon heating of the system, these differences may or may not be amplified, depending on the extent to which the free-energy barriers along these entropic degrees of freedom are also distinct between

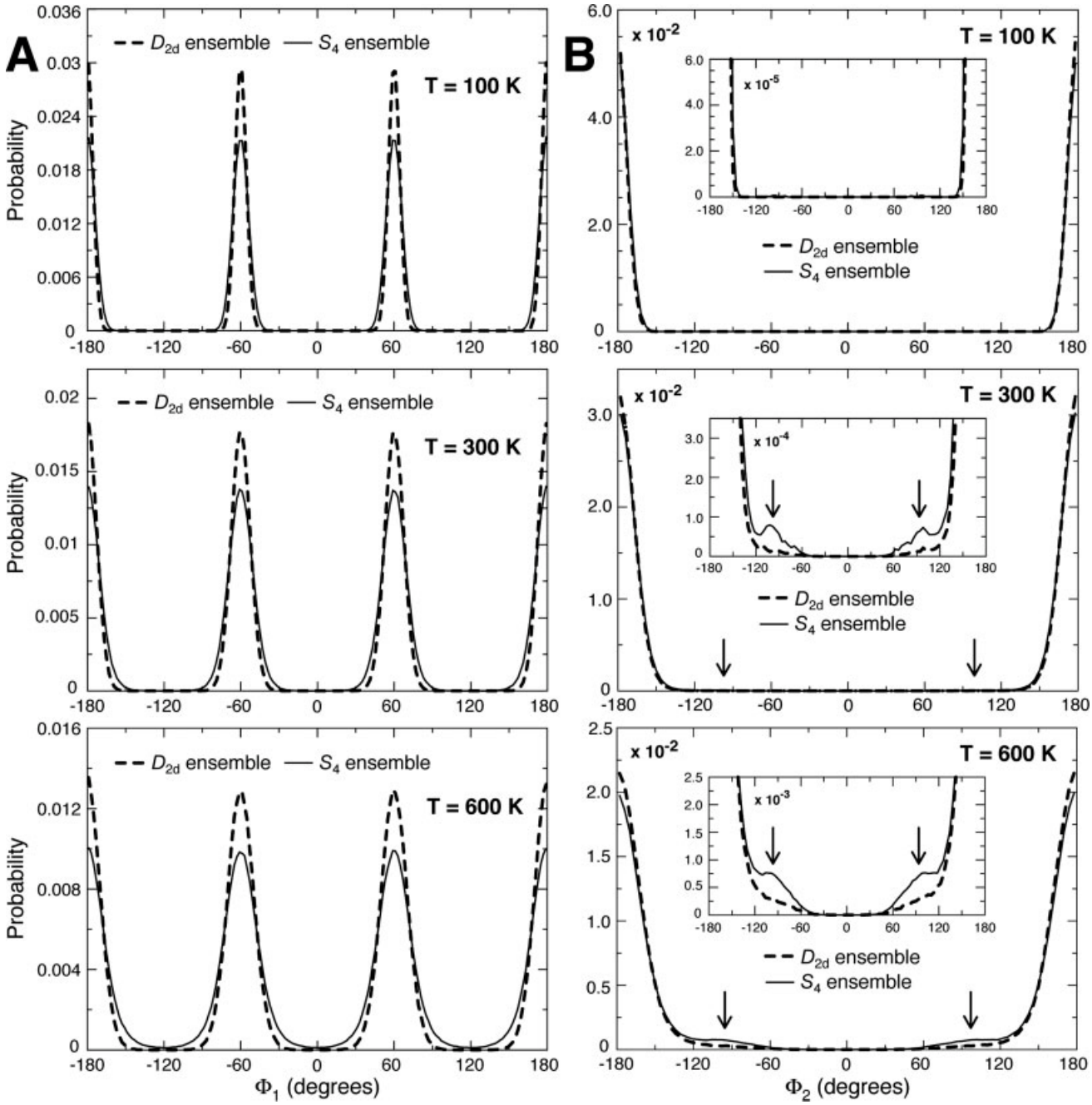

Figure 6. Temperature-dependence of the probability distributions of A. torsion Φ_1 ($C'_\alpha, N, C_\alpha, C_\beta$), and B. torsion Φ_2 ($N, C_\alpha, C_\beta, C_\gamma$), from the ensemble of TBA conformations in either D_{2d} or S_4 state obtained through the HTREMD gas-phase simulations.

the two states. As shown earlier, our results indicate that the conformational entropy difference between the D_{2d} or S_4 states of TBA, ΔS_{conf}, appears to have both temperature-dependent and temperature-independent components (the degeneracy effect itself is temperature independent). To identify the degrees of freedom that contribute to ΔS_{conf}, we analyzed the populations of all torsions and angles throughout the molecule, for the ensemble of either D_{2d} or S_4 conformations, from the HREMD simulations at 100, 300 and 600 K. Only two of the torsions and none of the angles were found to have distinct probability distributions when comparing the two states, namely the central torsion Φ_1 ($C'_\alpha, N, C_\alpha, C_\beta$), and the adjacent one along the alkane chain, Φ_2 ($N, C_\alpha, C_\beta, C_\gamma$). As shown in Figure 6, for both torsions the spread of distributions characteristic of the S_4 state is larger, consistent with its greater entropy. However, an interesting distinction between the two torsions is that, while the difference in the Φ_1 distributions between the D_{2d} or S_4 states is

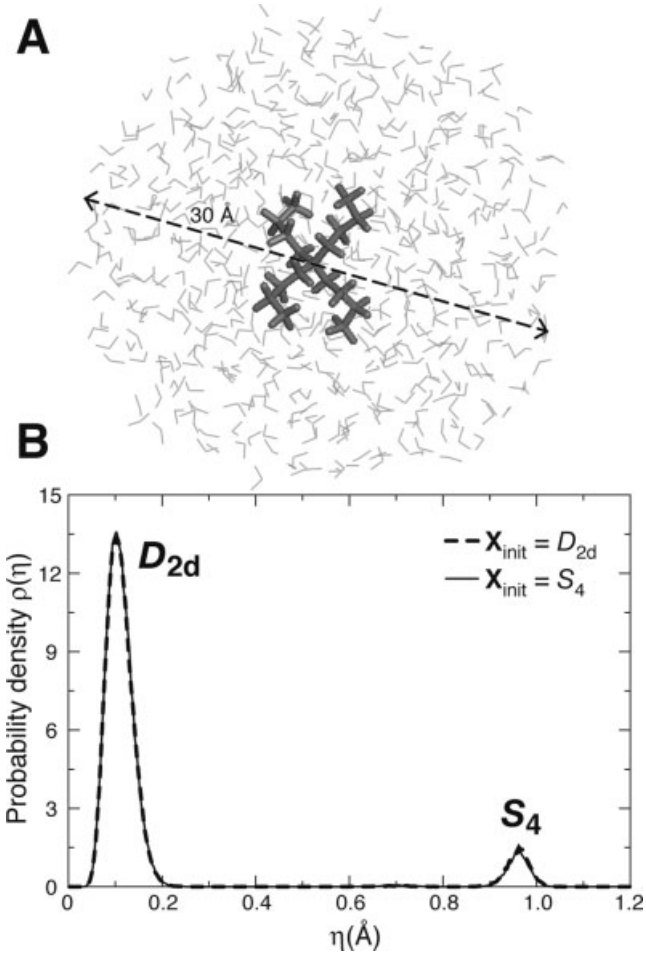

Figure 7. A. Molecular representation of the TBA-water spherical droplet used in the HREMD solvent simulations at $T = 300$ K. B. Distributions of the probability density of the reaction coordinate η, after 30 ns of simulation. The corresponding free-energy differences between the D_{2d} and S_4 states are given in Table 5.

roughly constant throughout the 100–600 K range, the difference in the Φ_2 distributions is activated by temperature, so that the energetic barriers that constrain this torsion in the trans configuration are reduced upon heating. Thus, from this analysis, we can conclude that the temperature-independent component of $\Delta S_{conf}(T)$ is likely to be because of a greater conformational freedom along Φ_1 ($C'_\alpha, N, C_\alpha, C_\beta$) in the S_4 state, while its temperature dependency originates from an increased range in the torsional space accessible along of Φ_2 ($N, C_\alpha, C_\beta, C_\gamma$), the extent of which is enhanced by temperature.

Simulation of Room-Temperature Conformational Equilibria in Solvent and Within a Protein Binding Site

The availability of a framework such as HREMD that enables a realistic simulation of the conformational equilibrium of a molecule in the gas phase, at a small cost and with sufficient precision, provides a stepping stone for the characterization of equilibria in more complex environments. In the case of many small molecules and ligands, for example, it is desirable to be able to describe their conformational equilibria as they transfer from bulk water into the active sites of binding proteins or enzymes, as a means to further our atomic-level understanding of the structural and energetic aspects of these processes. In this final section, we illustrate the potential of the HREMD approach for such problems through simulations of TBA in a homogeneous water environment, as well as within a simplified model of the KcsA channel. To characterize the room-temperature conformational equilibrium in bulk water, we used a spherical droplet of 543 TIP3P water molecules (ca. 30 Å in diameter), alongside the SSBP boundary potential, which provides a constant pressure environment as well as the appropiate reaction field energy (Fig. 7A). The HREMD simulation procedure was analogous to that used above for the ideal gas, with two independent 30-ns simulations with alternate starting conformations, each comprising 13 replicas that span a range in λ between 1.0 to 0.1, according to eq. (14).

The resulting profiles of the probability density of the reaction coordinate η are plotted in Figure 7B, showing very good agreement between both simulations, and a marked shift in the equilibrium towards the D_{2d} conformation, with respect to the ideal gas equilibrium analyzed earlier. In particular, the free-energy difference between the D_{2d} and S_4 states is now around -1.5 kcal/mol (Table 5), which corresponds to a solvation free-energy difference $\Delta\Delta G \simeq -1.3$ kcal/mol. This shift in the conformational equilibrium is correlated with a more favorable electrostatic interaction energy between TBA and its water environment for the D_{2d} conformation, namely $\Delta U_{elec} \simeq -2.6$ kcal/mol relative to S_4.

Table 5. Free-Energy Difference Between the D_{2d} and S_4 Conformations of TBA.

	$\Delta G = G(D_{2d}) - G(S_4)$[a]			
	TBA-bulk water		TBA-KcsA model	
	$t_{sim} = 15$ ns	$t_{sim} = 30$ ns	$t_{sim} = 50$ ns	$t_{sim} = 75$ ns
$\mathbf{X}_{init} = D_{2d}$	-1.59 ± 0.51	-1.42 ± 0.10	-1.73 ± 0.21	-1.68 ± 0.33
$\mathbf{X}_{init} = S_4$	-1.71 ± 0.56	-1.62 ± 0.36	-1.65 ± 0.22	-1.58 ± 0.36
	Ensemble − averaged energies[b]			
	TBA-bulk water		TBA-KcsA model	
	$\langle U_{elec} \rangle$	$\langle U_{vdW} \rangle$	$\langle U_{TBA} \rangle$	
D_{2d} ensemble	-80.4 ± 7.4	-28.1 ± 2.4	-120.0 ± 4.8	
S_4 ensemble	-77.8 ± 7.2	-29.1 ± 2.3	-123.3 ± 5.0	

[a]Free-energy difference between the D_{2d} and S_4 conformations of TBA, as derived from the HREMD simulations in bulk water or within the simplified model of KcsA's binding cavity (in kilocalories per mole). As the dependence of the block-averaged values on simulation time, as well as on the starting conformation of the replicas, is shown.

[b]Most relevant contributions to the equilibrium energetics of TBA in each system. In the water droplet simulation, the equilibrium shift towards D_{2d}, with respect to the gas phase (Table 3), is primarily a result of the electrostatic interaction energy between TBA and its water environment. By constrast, the shift in the context of the protein binding site is mostly due to an increased internal energy of TBA while in the S_4 conformation, which appears to be accomodated within the channel only under strain.

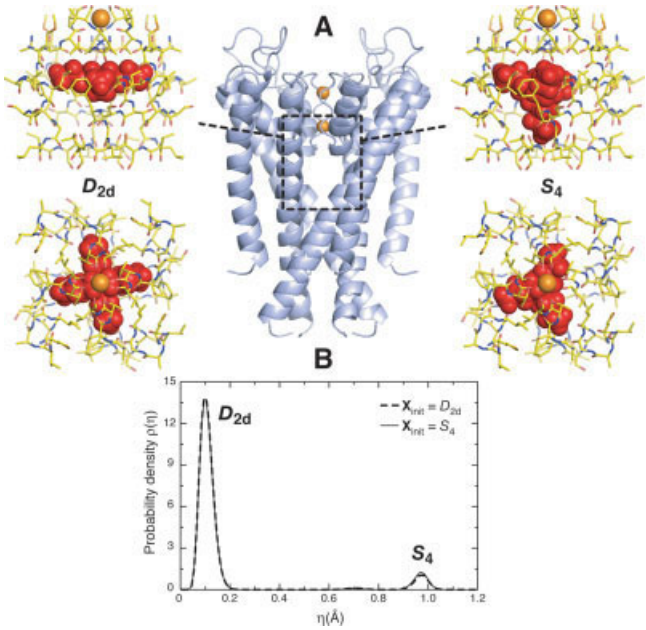

Figure 8. A. Molecular representation of TBA (red) within the binding cavity of KcsA (blue, yellow), as used in the HREMD simulations of the simplified ligand-protein system at $T = 300$ K. The top panels represent the view along the plane of the membrane, while the bottom panels depict the view along the four-fold symmetry axis of the channel. The so-called selectivity filter above the binding cavity includes a potassium ion (orange). B. Distributions of the probability density of the conformational coordinate η, after 75 ns of simulation. The corresponding free-energy differences between the D_{2d} and S_4 states are given in Table 5. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

A similar shift in the equilibrium towards the D_{2d} state was observed in the context of the model of the binding cavity of the KcsA channel. These simulations used a reduced model of the protein comprising a four-fold symmetric tetramer of residues 73–76 and 96–107 (PDB code 1K4C), alongside a single potassium ion in the so-called selectivity filter, located in the site known as S3 (Fig. 8A). All atoms in the channel were fixed in space and uncharged, and thus these simulations are expected to reveal how the equilibrium is affected by the shape of the binding site. As before, we prepared two independent HREMD simulations starting in either the D_{2d} or S_4 conformations, which, in this case, included 13 replicas and were carried out for 75 ns.

In Figure 8B, we plot the probability density profiles along the conformational coordinate η resulting from both simulations, which again show very good convergence. The corresponding free-energy differences are shown in Table 5; these reflect a population ratio of approximately 94:6. In contrast to the bulk-water equilibrium, the shift towards the D_{2d} state, of about -1.5 kcal/mol relative to that in the gas phase, is now primarily because of a change in the internal energy difference between the D_{2d} and S_4 conformations, from $\Delta U_{\mathrm{TBA}}^{(\mathrm{gas})} \simeq -1.5$ kcal/mol (Table 3) to $\Delta U_{\mathrm{TBA}}^{(\mathrm{channel})} \simeq -3.3$ kcal/mol (Table 5), while the interaction energy with the channel is comparable in both states. That is to say, while in the S_4 state TBA must adopt a strained conformation to optimize its interaction energy with the channel; by constrast, such strain is not required in the D_{2d} state.

In addition, two observations from these simulations are worth noting, which may be relevant for the interpretation of experimental data. The simulated D_{2d} and S_4 ensembles, when transformed into three-dimensional density maps (Fig. 9), such as those derived from crystallographic experiments, display a four-fold symmetry along the axis of the channel. Nonetheless, TBA is not four-fold symmetric in either the D_{2d} and S_4 conformations. For example, as shown in Figure 8A for the S_4 state, at any given time three of the alkane chains of the TBA typically lie in the plane perpendicular to the channel's axis (i.e., in the plane of the membrane), while the fourth chain points downwards; however, the average density from the S_4 ensemble is four-fold symmetric, since this instantaneous arrangement can be attained in four different orientations within the cavity. Therefore, it is unclear the extent to which the available electron density map of the TBA-KcsA complex,[21] which was derived with imposed noncrystallographic symmetry and is, obviously, an ensemble average across the crystal lattice, can in fact discriminate between the two states of TBA.

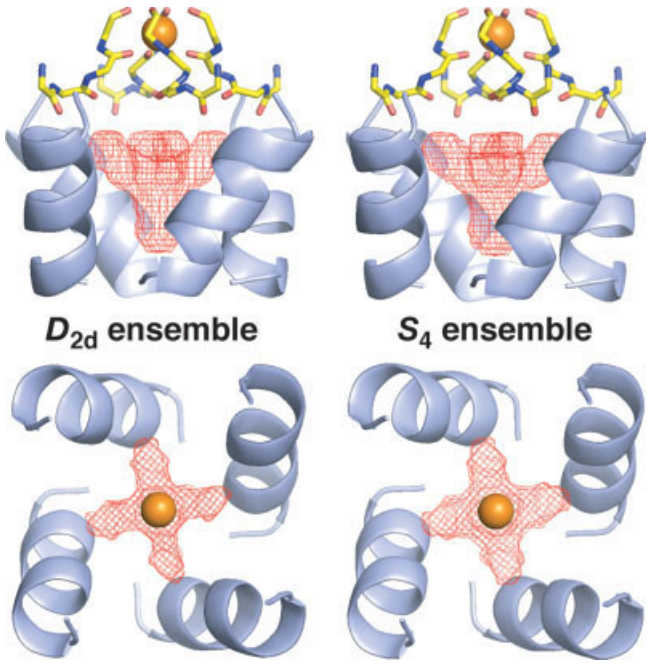

Figure 9. Three-dimensional density maps of TBA (red mesh) within the binding cavity of KcsA, derived from the HREMD simulations. The views of the system are like those in Figure 8. The density maps were derived through a mass-weighted, normalized mapping of all configurations of TBA, in either the D_{2d} or S_4 states, onto a three-dimensional grid. The mesh surfaces in the figure were constructed by rolling a 1.4-Å sphere over the subset of grid-points whose density exceeded 25% of the maximum density across the grid. The density maps shown were not explicitly symmetrized; the four-fold symmetry arises from the extensive conformational sampling (the instantaneous configurations of TBA, whether D_{2d} or S_4, are not four-fold symmetric either). [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

It is also apparent from the density map constructed from the D_{2d} ensemble (Fig. 9) that, while in this conformation, TBA may adopt an orientation relative to the channel in which not all four alkane chains are in the plane of the membrane, but rather, one that resembles the arrangement characteristic of S_4, with one chain pointing downwards. Given the extreme simplicity of the protein model used here, it remains to be determined whether this configuration is relevant physiologically. To this end, further HREMD simulations including a more realistic model of the channel are currently underway, through which we hope to elucidate the implications of these observations for our current understanding of TBA-mediated blockage of potassium channels, and which we hope will also provide a more meaningful comparison with experimental data.

Conclusions

The performance of two replica-exchange simulation methods has been compared through the computation of the gas-phase free-energy difference between the so-called D_{2d} and S_4 states of TBA, an ionic compound of some importance in the field of ion-channel physiology. These simulations are also of general interest, in that they illustrate important features of replica-exchange sampling techniques. The HREMD approach was found to be much more efficient than TREMD for an equivalent computational cost. The poorer performance of TREMD is primarily because of the substantial energy barrier that separates the D_{2d} and S_4 states, which is difficult to overcome even at high temperatures. In addition, the fact the entropy of the S_4 state is significantly lower than that of the D_{2d} state, and is thus increasingly favored upon heating, precludes the TREMD simulation from undergoing fast transitions between these states in a reversible fashion. This lack of reversibility leads to a prolonged dependence of the computed free-energy on the initial conditions of the simulation, i.e., a slow convergence. By contrast, the HREMD scheme, designed to reduce the magnitude of the free-energy barrier without drastically shifting the equilibrium of the two stable states, yields well converged free-energy differences within an affordable simulation time.

The relevance of these observations transcend our test system; for example, in the context of protein folding, the significantly higher entropy of the unfolded state, especially in large proteins, will probably hinder the ability of TREMD simulations to explore folding pathways reversibly. However, whether Hamiltonian REMD can perform better than TREMD in folding problems remains an open question; given the complexity and cooperativity of these processes, it may not be trivial, in many cases, to identify and/or implement a potential-energy perturbative scheme that enhances the sampling of folding/unfolding transitions. Nonetheless, we expect the HREMD approach to be generally more useful in tackling a variety of problems that feature large free-energy barriers, such as in conformational changes driven by isomerization events.

Taking advantage of the efficiency of the HREMD scheme, it was straighforward to characterize the conformational equilibrium of TBA in a variety of conditions. In particular, we were able to quantitatively assess the effect of entropy and solvation on its conformational equilibrium as well as to gain insights into its mode of association with a binding protein. Overall, this work illustrates the usefulness of the HREMD method to enhance conformational sampling in the context of computational drug design, analysis of ligand-receptor interactions and, more generally, molecular recognition.

Acknowledgments

J. D. F.-G. would like to thank D. Sezer and Drs. D. M. Zuckerman and A. C. Pan for estimulating discussions on the replica-exchange framework and statistical mechanics, as well as Dr. L. R. Forrest for her advice on the design of this project and her critical assessment of this manuscript. This work was supported by a grant GM-62342 from the NIH.

References

1. Torrie, G.; Valleau, J. J Comp Phys 1977, 23, 187.
2. Souaille, M.; Roux, B. Comp Phys Comm 2001, 135, 40.
3. Ensing, B.; De Vivo, M.; Liu, Z.; Moore, P.; Klein, M. Acc Chem Res 2006, 39, 73.
4. Hamelberg, D.; Mongan, J.; McCammon, J. J Chem Phys 2004, 120, 11919.
5. Pande, V.; Baker, I.; Chapman, J.; Elmer, S.; Khaliq, S.; Larson, S.; Rhee, Y.; Shirts, M.; Snow, C.; Sorin, E.; Zagrovic, B. Biopolymers 2003, 68, 91.
6. Bolhuis, P.; Chandler, D.; Dellago, C.; Geissler, P. Annu Rev Phys Chem 2002, 53, 291.
7. Sugita, Y.; Okamoto, Y. Chem Phys Lett 2000, 329, 261.
8. Sugita, Y.; Okamoto, Y. Chem Phys Lett 1999, 314, 141.
9. Earl, M.; Deem, D. J. Phys Chem Chem Phys 2005, 7, 3910.
10. Liu, P.; Kim, B.; Friesner, R.; Berne, B. Proc Nat Acad Sci USA 2005, 102, 13749.
11. Zuckerman, D.; Lyman, E. J Chem Theor Comp 2006, 2, 1200.
12. Fukunishi, H.; Watanabe, O.; Takada, S. J Chem Phys 2002, 116, 9058.
13. Sugita, Y.; Kitao, A.; Okamoto, Y. J Chem Phys 2000, 113, 6042.
14. Allen, M. P.; Tildesley, D. J. Computer Simulation of Liquids; Oxford University Press: Oxford (UK). 1987.
15. Affentranger, R.; Tavernelli, I.; Di Iorio, E. J Chem Theor Comp 2006, 2, 217.
16. Luzhkov, V.; Osterberg, F.; Acharya, P.; Chattopadhyaya, J.; Aqvist, J. Phys Chem Chem Phys 2002, 4, 4640.
17. MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. J Phys Chem B 1998, 102, 3586.
18. Crouzy, S.; Berneche, S.; Roux, B. J Gen Physiol 2001, 118, 207.
19. Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. J Comp Chem 1983, 4, 187.
20. Brooks, B.; Janezic, D.; Karplus, M. J Comp Chem 1995, 16, 1522.
21. Lenaeus, M.; Vamvouka, M.; Focia, P.; Gross, A. Nat Struct Mol Biol 2005, 12, 454.