JCTC Journal of Chemical Theory and Computation

# Computing Relative Free Energies of Solvation Using Single Reference Thermodynamic Integration Augmented with Hamiltonian Replica Exchange

Ilja V. Khavrutskii* and Anders Wallqvist

*Biotechnology HPC Software Applications Institute, Telemedicine and Advanced Technology Research Center, U.S. Army Medical Research and Materiel Command, Fort Detrick, Maryland 21702, United States*

Received June 15, 2010

**Abstract:** This paper introduces an efficient single-topology variant of Thermodynamic Integration (TI) for computing relative transformation free energies in a series of molecules with respect to a single reference state. The presented TI variant that we refer to as Single-Reference TI (SR-TI) combines well-established molecular simulation methodologies into a practical computational tool. Augmented with Hamiltonian Replica Exchange (HREX), the SR-TI variant can deliver enhanced sampling in select degrees of freedom. The utility of the SR-TI variant is demonstrated in calculations of relative solvation free energies for a series of benzene derivatives with increasing complexity. Of note, the SR-TI variant with the HREX option provides converged results in a challenging case of an amide molecule with a high (13−15 kcal/mol) barrier for internal cis/trans interconversion using simulation times of only 1 to 4 ns.

## A. Introduction

The ability to predict solvation free energies for a series of small molecules is important in drug design.[1,2] It allows for an assessment of solvation and, in general, partition properties of novel molecules before their syntheses. Furthermore, solvation free energy is an essential ingredient for predicting proton affinities of small molecules and their binding affinities to biomolecular drug targets in water.[2−6] Therefore, computational tools that provide accurate solvation free energies are indispensable in drug design, particularly in the lead optimization stages where molecules of interest are not immediately available for experimental evaluation.

Sustainable high quality predictions of solvation free energies require rigorous computational methods that can properly account for explicit ligand−solvent interactions and ligand flexibility.[4,7−12] Although various empirical methods exist that have been tuned to reproduce solvation free energies of available compounds, the quality of their predictions decays quickly for novel, flexible molecules.[10,11] Therefore, more rigorous methods are required in such cases.[4,12−20] Currently, the most rigorous methods are Thermodynamic Integration (TI) and the closely related Free Energy Perturbation (FEP).[2,5,6,21−30] Because of the computational equivalence of TI and FEP methods, we limit ourselves to TI in this paper. While the TI method incorporates ligand flexibility via Molecular Dynamics (MD) or Monte Carlo (MC) simulations, it inherits sampling issues associated with these simulation methods.[23,31−33] These sampling issues become particularly severe in systems with hindered conformational transitions, often rendering the results of TI calculations unreliable.[20,23,31−35]

Proper accounting for ligand flexibility requires an enhanced sampling technique. Previous attempts to achieve enhanced sampling have produced a multitude of sophisticated computational methods, such as conformational flooding,[36,37] hyper dynamics,[38−40] accelerated molecular dynamics,[41−45] simulated scaling,[46−48] and generalized ensemble methods such as temperature and Hamiltonian replica exchange methods.[20,49−60] In principle, all of these methods could be coupled with TI to compute solvation free energies. However, it is not clear which combination would give the most accurate results most efficiently.

* To whom correspondence should be addressed. E-mail: ikhavrutskii@bioanalysis.org.

10.1021/ct1003302 
Published on Web 10/21/2010

A combination of TI with Hamiltonian Replica Exchange (HREX) has been demonstrated to enhance sampling and improve convergence in solvation free energy calculations.[58,59] This combination improves upon standard TI at the smallest expense and is, therefore, the most promising first step. The improvements are attributed to the so-called "Hamiltonian tunnel" that enriches otherwise isolated MD simulations with configurations from all of the other TI windows. Recently, the HREX-TI method has been further enhanced by introducing the double-tunneling scheme.[35] This latter work highlights the importance of the simulation setup and provides a way to accelerate specific hindered degrees of freedom. In addition, it demonstrates that the cost of the simulations could be further reduced via an optimal selection of TI windows. Enhanced sampling of the hindered degrees of freedom is achieved through the "Hamiltonian tunnel" that connects the molecules on the original, hindered potentials to their unhindered counterparts.[20,35,56,57]

Computing relative as opposed to absolute solvation free energies could be more advantageous for enhanced sampling. To rank a series of molecules according to their solvation free energies, one can use either an absolute or a relative scale. For a ranking based on relative solvation free energies, one needs to relate all of the molecules from the series to a single reference state.[61,62] This seeming disadvantage has prompted many researchers to turn to the absolute scale. We note, however, that relative TI simulations provide opportunities for incorporating procedures to select specific degrees of freedom for enhanced sampling.[35] For example, hindered internal rotations can greatly benefit from enhanced sampling.[20,35,56,57] Such opportunities appear to be absent in absolute TI calculations. Therefore, while TI can yield both absolute and relative solvation free energies, the latter can provide additional sampling benefits.

Finally, the choice of topology appears critical to achieving the most efficient sampling with HREX-TI. Thus, Yang et al. demonstrated that dual topology TI simulations benefited greatly from HREX, whereas in the single topology case, the effect of replica exchange was small.[35] Hence, the authors concluded that dual topology TI would be the method of choice to combine with HREX for enhanced sampling.

In the present paper, we developed an approach to enhance sampling of hindered degrees of freedom in the single topology TI framework that was inspired by the single reference state extrapolation method.[15,19] Thus, we remove the roadblock identified by Yang et al.[35] This new strategy allows for enhanced sampling of hindered degrees of freedom in a controlled way by choosing a reference state in which the desired degrees of freedom are sampled freely. The added flexibility of this approach comes from the fact that the reference state does not have to correspond to a real molecule as in standard TI.

This paper is organized as follows. We first briefly recap the conventional TI methodology and point out its drawbacks. We then present the Single Reference-TI (SR-TI) approach and highlight its features that help overcome the referred drawbacks. Following the computational details, we describe results of SR-TI simulations for a series of benzene derivatives. Specifically, we focused on *para*-substituted phenols, hydroxylated benzenes, and aryl alcohols. Finally, we take on a real challenge—an amide molecule with a hindered internal rotation that separates its *cis* and *trans* isomers.[63] This difficult test case presents a sampling problem that can only be solved using SR-TI with the HREX option. This test case also validates the SR-TI approach by comparing the results of the regular and HREX SR-TI simulations against each other and against the amide bond torsional Potential of Mean Force (PMF). To the best of our knowledge, this is the first direct calculation of the amide hydration free energy that automatically accounts for *cis*/*trans* interconversion.

## B. Methodology

**B.1. General TI.** TI methods are simulation methods concerned with computing free energy or reversible work for an alchemical transformation of a molecule A to a different molecule B using rigorous statistical mechanical principles.[21,23,28,64] Depending on the topology of the alchemical system, TI methods can be divided into single and dual topology methods.[65,66] Dual topology TI methods simultaneously propagate Hamiltonians ($H_A$ and $H_B$) of both molecules coupled through their shared environment, whereas single topology methods create a hybrid molecule, a union of the molecules A and B, and propagate the resulting single Hamiltonian ($H_{AB}$). While both approaches have their advantages and disadvantages, in this paper, we will focus on single topology TI methods.

To better understand the chemistry behind TI simulations, it is useful to consider the potential energy function $V$ that along with kinetic energy $K$ comprises the Hamiltonian of the system ($H = K + V$). For a system of $N$ atoms described by a configuration $R$ with $3N$ coordinates, its potential energy depends on the identity of the atoms and the way they are connected by covalent bonds. Typically, TI methods use Molecular Mechanical (MM) potentials that have assigned atom types and do not break covalent bonds. Therefore, an MM potential can be expressed as a simple sum of bonded and nonbonded terms. The general form of all of these latter terms is well-known,[67–69] whereas the fine details of each term depend on the types of atoms involved.[70,71]

In brief, bonded terms include harmonic terms for stretching, $V^{bond}$ (two atoms with one bond), and bending, $V^{angle}$ (three atoms with two bonds), as well as an anharmonic, torsional term for twisting, $V^{torsion}$ (four atoms with three bonds in a chain). In some cases additional harmonic, improper torsional terms, $V^{improper}$ (four atoms with three bonds sharing one of the atoms), are added. These latter terms enforce a particular configuration of the shared atom with respect to the plane formed by the other atoms. Hence, improper torsions can be used to enforce planarity or a particular chirality of the four-atom group. On the other hand, the anharmonic, torsional terms are essential in representing different conformations of the molecule separated by corresponding barriers. Because the bonded terms cannot act through space, the MM potential is completed by adding nonbonded terms. The essential nonbonded terms include pairwise, long-range electrostatic, $V^{Coulomb}$, and short-range

Lennard-Jones (LJ), $V^{LJ}$, potentials. The LJ terms among other things ensure that atoms of the opposite charge do not collapse on top of each other. This level of description is more than sufficient to understand the alchemical transformations behind the TI method and their important consequences.

In single topology TI, most common alchemical transformations can be represented by two simple operations, namely, "atom substitution" and "atom creation/annihilation." Both transformations modify the MM potential terms, but to a different extent. Atom substitution transforms one atom type into another, by changing all bonded and nonbonded parameters correspondingly. This transformation proportionally modifies all of the associated bonded terms, including the anharmonic torsional terms. In contrast, atom annihilation completely voids the anharmonic torsional terms associated with the annihilated atom but does not modify the harmonic terms for stretching and bending.[65,66] The improper torsional terms that involve the annihilated atoms can be handled either way depending on the situation. In addition, atom annihilation "switches off" all of the corresponding nonbonded terms. Therefore, the atoms that are "switched off" remain attached to the molecule via their original harmonic interactions but no longer interact with the other atoms through space. Consequently, atom creation requires that the atom be present in the hybrid molecule in the "switched off" state. The "switched off" atoms are often referred to as *dummy* atoms, as opposed to *real* atoms. Needless to say, when dummy atoms are involved with the hybrid molecule, standard rules of valency do not have to apply.

The reversible work for the alchemical transformation is derived by following the changes in the MM potential of the hybrid system. During the alchemical transformation from molecule A to molecule B, the total number of atoms in the system, including the dummy atoms, does not change. However, their contributions to the MM potential energy do change depending on whether the atoms are being "switched on/off" or substituted. Throughout the alchemical transformation, the hybrid potential $V_{AB}$ remains well-defined. Therefore, the associated reversible work can be computed. This reversible work formally corresponds to changing the potential of the system $V_{AB}$ from the state that represents the original molecule A to the state that corresponds to molecule B.

The efficiency of computing the reversible work associated with the alchemical transformation depends on the way the hybrid Hamiltonian is constructed. Typically, the hybrid Hamiltonian of the system ($H_{AB}$) that includes the potential $V_{AB}$ is linearly interpolated between the end points, molecules A and B, using a coupling parameter $\lambda$:

$$H_{AB}(\lambda) = (1 - \lambda)H_A + \lambda H_B \quad (1)$$

The coupling parameter in TI spans an interval from $\lambda = 0.0$ (molecule A) to $\lambda=1.0$ (molecule B) and is an analogue of the reaction coordinate. However, linear interpolation of Coulomb electrostatic and LJ nonbonded interactions presents a sampling problem in cases with dummy atoms, known as the end-point catastrophe.[72,73] To alleviate this problem, several approaches have been proposed.[16,72–75] Among those, we find the use of soft-core potentials to be the most suited for our purposes. Instead of going to infinity at zero interaction distances (when atoms collapse on top of each other), soft-core potentials take on finite values due to built-in distance offsets. It is worth mentioning that although the approach introduced by Yang et al. employed the soft-core potentials with the HREX option, it could also work without them.[35] Hence, this approach could be regarded as an alternative solution to the end-point catastrophe problem.

In the present work, we employ a free energy integration procedure that requires a finite potential and its derivatives and hence necessitates the use of soft-core potentials. The available soft-core potentials also differ in styles. In the present work, we employ the original GROMOS style soft-core (SC) potentials as implemented in GROMACS:[37,73,76]

$$V^{sc}_{AB}(r) = (1 - \lambda)V_A(R_A(r,\lambda)) + \lambda V_B(R_B(r,\lambda)) \quad (2)$$

$$R_A(r,\lambda) = (\alpha\sigma_A^6\lambda^p + r^6)^{1/6} \quad (3)$$

$$R_B(r,\lambda) = (\alpha\sigma_B^6(1 - \lambda)^p + r^6)^{1/6} \quad (4)$$

where $p = 2$, $r$ is the distance between a given pair of atoms, $\alpha$ is the soft core parameter, and $\sigma$ is the radius of interaction computed from LJ parameters. Although different strategies combine electrostatic and LJ soft-core transformations differently, sometimes completely decoupling the two,[14,75,77–80] a simultaneous change of both potentials should reduce the computational burden.

Given all of the Hamiltonian interpolation rules underlying a particular alchemical transformation, we can begin collecting necessary statistics for computing the corresponding reversible work. One of the most efficient procedures to do that is the multiconfigurational TI.[23,64] In this method, the Hamiltonian derivatives with respect to $\lambda$ are collected in $M$ independent MD or MC simulation windows. The $\lambda$ values of these windows span the interval [0; 1]. Thus, each $i$th window corresponds to the hybrid Hamiltonian with a fixed value of $\lambda_i$, and its simulation provides the ensemble averaged Hamiltonian derivatives, $\langle(\partial H_{AB})/(\partial\lambda)|_{\lambda_i}\rangle$. Integration of the mean Hamiltonian derivatives from all the windows yields the reversible work.

**B.2. Fourier Bead Integration.** The reversible work can be computed from the multiconfigurational TI data in many different ways. Methods such as the Weighted Histogram Analysis Method (WHAM),[81–84] Bennett's Acceptance Ratio (BAR),[16,85–87] or simple line integration have been used most frequently. Recently, overlap histogramming[35] was used to integrate the results of TI simulations where the distributions of the Hamiltonian derivatives in two adjacent windows overlap. Here, we employ a Fourier beads[88–90] variant of the line integration procedure suggested by Bennett.[85] Single topology TI simulations have been noted to provide faster convergence for the mean Hamiltonian derivatives than dual topology ones,[35] thus justifying the use of the line integration approach. In this method, we do not need to have the histograms of the adjacent windows overlap. However, we require that changes in the mean Hamiltonian derivatives are sufficiently smooth that they could be interpolated between all of the windows. We caution that in the absence

of the histogram overlap, TI simulations with the HREX option will become inefficient.[35,57–59,91]

The adaptation of the Fourier bead integration procedure to TI is straightforward. The Fourier interpolated values of the Hamiltonian derivatives and of the corresponding $\lambda$ values are represented as continuous functions of a parameter $\xi$ defined on the interval [0; 1]:

$$\tilde{X}(\xi) = \langle X|_{\lambda=0}\rangle + (\langle X|_{\lambda=1}\rangle - \langle X|_{\lambda=0}\rangle)\xi + \sum_{k=1}^{M} a_k \sin(k\pi\xi) \quad (5)$$

Here, X refers to either the derivative of the Hamiltonian or the corresponding $\lambda$ value and $a_k$'s are the Fourier amplitudes. The brackets indicate ensemble averaging, and "$\sim$" indicates the continuous function representing the interpolated ensemble-averaged values. Note that we only need to interpolate the values of the coupling parameter $\lambda$ if the TI windows are not distributed evenly on the interval [0; 1]. The corresponding Fourier interpolation amplitudes are derived by the discrete Fourier transform procedure using the simulation data from all $M$ windows:

$$a_k = 2\sum_{l=1}^{M} [\langle X|_{\lambda=\lambda_l}\rangle - \langle X|_{\lambda=0}\rangle - (\langle X|_{\lambda=1}\rangle - \langle X|_{\lambda=0}\rangle)\xi_l] \sin(l\pi\xi_l)\Delta\xi \quad (6)$$

where the set $\{\xi_l | l = \overline{1, M}\}$ forms a uniform grid by construction:

$$\xi_l = \frac{l-1}{M-1} \quad (7)$$

$$\Delta\xi = \frac{1}{M-1} \quad (8)$$

Using the interpolated values, we can take all of the required derivatives analytically and then compute the work via the line integral, also, in principle, analytically. In practice, we use a fine grid to evaluate the integral using a simple trapezoidal rule.

$$W(\xi) = \int_0^{\xi} \left\langle \frac{\partial \tilde{H}_{AB}}{\partial\lambda}\bigg|_{\lambda(\xi')} \right\rangle \frac{d\lambda(\xi')}{d\xi'} d\xi' \quad (9)$$

We stress that all of the Fourier beads methodology can be applied to any multiconfigurational TI calculations that collect the mean derivatives of the hybrid Hamiltonian with respect to the coupling parameter and can be trivially generalized to multiple coupling parameters.[88–90]

**B.3. Drawbacks of Conventional TI.** Multiconfigurational TI with MD simulations, like other methods employing independent windows, suffers an inherent drawback that can deteriorate the quality of the computed free energy estimates. Because of the random nature of the thermal fluctuations, certain rare transitions can occur in some windows but not others, creating incoherence across the windows over time. With different windows exploring nonoverlapping regions of configurational space, the final free energy estimates might get skewed.[23,31–34,64,77,92–95] Another related issue that affects the results of the multiconfigurational TI is the bias from the initial configuration that causes the simulation to explore only a limited region of the configurational space near the starting configuration. Both of these issues are due to sampling limitations that stem from the inability of MD simulations to consistently overcome large barriers ($\gg k_B T$) that correspond to rare events. Because their effect on the quality of the final free energy estimates can be significant, they need to be addressed. Although different solutions have been proposed to tackle some of these issues,[31,41,42,55–60,94–96] we believe that HREX-TI is one of the most cost-effective improvements of TI methodologies in general.

**B.4. Single Reference-TI Approach.** The SR-TI approach presented here computes relative solvation free energies for a series of related molecules. It has been developed to aid lead optimization efforts in drug design. The SR-TI approach is a variant of single topology TI methods and has been inspired by the Single Reference State Extrapolation (SRSE) method.[15,18,19] The SRSE method attempts to compute free energies for an alchemical transformation of a series of related molecules A to a common reference state B by using only a single simulation of the reference state B. The reference state B, by construction, includes all of the atoms necessary to form any molecule A in the series. Hence, the alchemical transformation free energies are derived by evaluating the differences in the MM potential of the real and reference states over the configurations sampled by the reference state only. Computationally, this is roughly equivalent to running a single TI window for the series of molecules and is, therefore, very economical. Unfortunately, the accuracy of the SRSE method depends on the overlap between the configurations of the reference simulation and those of the real state. This makes the choice of the reference state complicated and the results approximate. The SR-TI approach naturally avoids these issues albeit at a higher cost.

In a nutshell, SR-TI computes free energies or reversible works for the alchemical transformations of each molecule A in a series down to a common reference molecule B. This is similar to previous TI calculations using a common reference state.[18,61,62] The fact that the reference molecule B is a single substructure of all of the molecules A in the series greatly simplifies the chemistry in SR-TI. Thus, SR-TI needs to annihilate atoms of molecule A that extend beyond the common reference structure B and substitute their shared atoms to match atom types to those within the reference. This is illustrated in Figure 1 with the example of pyridine and toluene as a series of two molecules. Thus, the SR-TI approach never needs to create new atoms. The SR-TI approach makes comparing different molecules A from a given series simple. Importantly, unlike in earlier TI simulations using a common reference,[18,61,62] in SR-TI, reference molecule B does not need to represent a real molecule. For example, in Figure 1, reference molecule B comprises only the heavy atoms of the benzene. Once the common core structure for the series has been determined, the hybrid topology and parameter files for each molecule A can be set up independently. Nevertheless, because the real atoms (but not necessarily the dummy atoms) in the common reference B are the same across the series, any

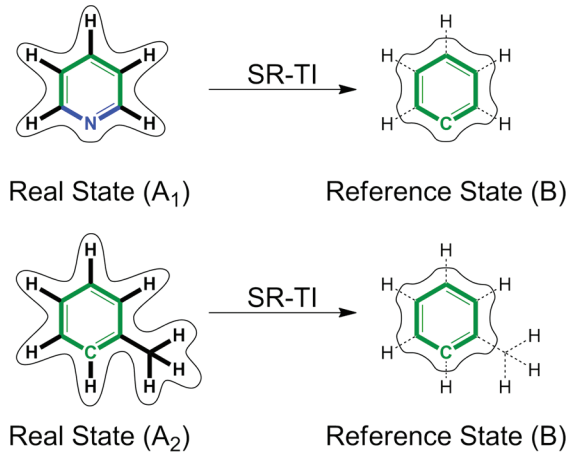


***Figure 1.*** Two kinds of alchemical operations and a single reference state in SR-TI. The top pane represents the transformation of pyridine ($A_1$) to a benzene core reference state (B), by "atom substitution" of the nitrogen atom and "atom annihilation" of the hydrogen atoms. The bottom pane shows another transformation of toluene ($A_2$) to the same reference state (B) with only "atom annihilations." In this latter case, the internal rotation of the methyl group becomes enhanced in the reference state. The common reference substructure is highlighted in bold green. The atoms and bonds that are shown in bold correspond to the real atoms, whereas regular labels with thin dashed bonds represent the dummy atoms. The contours surrounding molecules represent molecular volume, which is always reduced in the reference state. We emphasize that the reference states on the right are considered identical, despite the differences in the dummy atoms.

molecule A can be conveniently compared to all of the other molecules from the series. Therefore, the SR-TI approach can be used to study differences between stereoisomers, including *cis*/*trans* isomers and enantiomers.

Besides the computational convenience, SR-TI can overcome the bias of the initial configuration inherent in conventional TI methods. In SR-TI, the volume of molecule A always reduces down to that of the common reference structure B (see Figure 1). At the same time, it is possible to also reduce the complexity of the reference state and its overall polarity. As a result, the reference end state can sample confined spaces, such as solvent cages, more efficiently than the real end state. In addition, if reference molecule B were less polar than molecule A, it could escape traps due to specific interactions with the solvent. Last but not least, the TI windows that are closer to reference state B enjoy enhanced sampling of the torsional degrees of freedom, which involve the dummy atoms (recall that these terms become void in the reference molecule). Therefore, by appropriately choosing the reference substructure, one can enhance sampling not only of the orientational, but also of the specific torsional degrees of freedom. Although, these sampling benefits can remove the initial configuration bias and ease the overlap issues at or near the reference end state, they do not apply to the real end state or molecule A automatically.

**B.5. Single Reference−TI with Hamiltonian Replica Exchange.** With an appropriate technology, SR-TI can achieve enhanced sampling across all TI windows. We have noted that SR-TI naturally achieves enhanced sampling in the windows at and near the reference state. To enhance sampling in all TI windows, we can invoke exchanges of configurations between the SR-TI windows by employing Hamiltonian replica exchange (HREX).[20,53,55–60,96] Note that unphysical states have been used to enhance sampling of hindered degrees of freedom with the help of HREX before.[56,57] Although HREX has been proposed[53] and applied[35,57–59,96] in combination with conventional TI previously, the benefits are most pronounced in dual topology HREX-TI.[35]

In the case of the NPT ensemble, for example, where all of the TI windows are run at the same temperature $T$ and pressure $P$, the HREX option can be implemented as follows. Consider a vertical excitation for an $i$th window with configuration $R_i$ from the Hamiltonian $H_{AB}$ on the alchemical energy surface $V_{AB}$ at $\lambda_i$ to that at $\lambda_j$:

$$\Delta_{ij}(R_i) = \beta(V_{AB}(R_i, \lambda_j) - V_{AB}(R_i, \lambda_i)) \tag{10}$$

where $\beta = 1/k_B T$ is the inverse temperature and $k_B$ is the Boltzmann constant. Now consider the energy change upon exchange of the two adjacent windows $i$ and $j$. The total change in the energy of the generalized ensemble upon the Hamiltonian exchange between two configurations $R_i$ and $R_j$ is as follows:

$$\Delta\Delta = \Delta_{ij}(R_i) + \Delta_{ji}(R_j) \tag{11}$$

and the final Metropolis acceptance criterion is

$$W(R_i \leftrightarrow R_j) = \begin{cases} 1, & \text{for } \Delta\Delta \le 0 \\ \exp(-\Delta\Delta), & \text{for } \Delta\Delta > 0 \end{cases} \tag{12}$$

Here, $W$ is the probability of the exchange. Although different exchange protocols exist that vary in efficiency,[53,54,91,97–100] we follow the standard exchange protocol developed previously for temperature replica exchange (TREX).[53,54,97] Specifically, we perform HREX by alternating exchange attempts for all odd and all even pairs of replicas. Odd pairs are replicas $i = 2n - 1$ and $j = 2n$, and even pairs are replicas $i = 2n$ and $j = 2n + 1$, where $n$ starts at 1 with the largest index not to exceed the total number of TI windows $M$. The exchanges are attempted at regular, predetermined time intervals. Best of all, because HREX maintains the original ensembles within each window, the methodology for computing the reversible work remains unchanged. All we need to do is follow $\lambda$ values through all of the exchanges. Thus, the only additional cost associated with evaluation of the Metropolis exchange criteria is well compensated for by the benefits of the enhanced sampling.

The combination of HREX with single-topology SR-TI addresses both the initial bias and the overlap drawbacks of conventional TI methods in much the same way as the previously developed dual topology HREX-TI variant.[35] Conveniently, SR-TI provides flexibility in partitioning a given molecule between reference and dummy atoms that grants control over hindered torsional degrees of freedom, molecular volume, and polarity of the system. The HREX option creates a flow of replicas across the Hamiltonian space of the whole alchemical transformation. Such a flow of

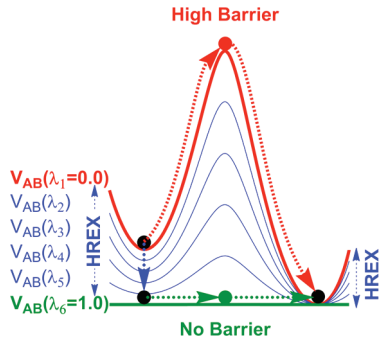


***Figure 2.*** An illustration of the "Hamiltonian tunnel" opened by the HREX option during SR-TI calculations. The potential energy surface at $\lambda = 0$ has an insurmountable barrier, which disappears completely in the potential at $\lambda = 1.0$. The red and green dashed arrows represent regular transitions on the same surface through corresponding transition states. The blue arrows represent the Hamiltonian replica exchanges that allow the system to hop between the surfaces. Instead of going over a high barrier on the $\lambda = 0$ surface, the system arrives at the surface with the lower barrier, say at $\lambda = 1$ through HREX, crosses over to the other minimum, and is then brought back to the $\lambda = 0$ surface on the other side of the barrier.

replicas has two main benefits. First, it opens a "Hamiltonian tunnel" (see Figure 2) that allows consistent crossing of high energy barriers over a short period of time that would have been impossible otherwise.[46,48,56] Second, the flow of replicas mixes configurations from different TI windows and thus provides superb overlap in configurational space. Ultimately, HREX SR-TI is an improvement over regular SR-TI and permits calculations of high quality solvation free energies. In tandem, regular and HREX SR-TI calculations can be used as a tool to identify sampling issues.

The HREX option should be applied with caution to SR-TI calculations involving chiral atoms outside the common reference structure, as the chiral atoms can invert their configurations in the dummy state. However, this issue may be avoided if the chirality is maintained through harmonic, improper torsional terms.

## C. Computational Details

**C.1. Small Molecule Parameters.** We obtained initial coordinates of the small molecules from their SMILES[101,102] using a program called CORINA.[103] An exhaustive conformational search was attempted using a companion program ROTATE,[104] followed by structural refinement with the semiempirical AM1 method,[105] as implemented in open source MOPAC7, version 1.11.[106] Where multiple conformations existed, the AM1 partial charges for each conformation were symmetrized across the equivalent atoms and combined into a conformation-independent set of charges using Boltzmann weighting by the AM1 energies at the target temperature of 300 K with appropriate degeneracies. Finally, the AM1 charges were augmented through the BCC procedure[107,108] as implemented in the Antechamber program[109,110] from Amber Tools, version 1.2. The resulting conformation-independent, properly symmetrized set of AM1BCC charges is expected to reproduce HF/6-31G* RESP charges to a good approximation. The remainder of the parameters, including vdW and bonded terms, were generated in an automated fashion by the Antechamber program[109,110] to comply with the GAFF force field.[71]

**C.2. Single Reference State.** The choice of the reference state for SR-TI simulations is flexible. Recall that the single reference state does not have to correspond to a real molecule. The size of the reference state cannot exceed that of the largest common substructure and cannot be less than one atom from that substructure. For computational convenience, the total charge of the reference state should be an integer. To derive parameters for the reference state, we can employ parameters of its nearest real molecule. We do that by turning the atoms used to complete the reference structure to the nearest molecule (preferably hydrogens) into dummy atoms and adding their partial charges to those of the nearest real atoms of the common substructure. This procedure is similar to the one used in AutoDock4.0 to derive united-atom parameters from all-atom ones.[111–113] For the molecules studied in the present work, we chose the benzene substructure, comprising only six heavy atoms, as the reference state (see Figure 1).

**C.3. SR-TI Setup.** The alchemical transformation turns all the atoms of the original molecules that are outside the benzene core reference state into dummy atoms by altering their partial charges and vdW parameters simultaneously. To avoid the end-point catastrophe, we employ GROMOS style soft-core electrostatic and LJ potentials[73] as implemented in GROMACS.[76,114–116] Specifically, we employ $p = 2$, the soft-core parameter $\alpha = 1.5$, and the radius of interaction $\sigma = 0.3$ (see eqs 2−4 and the GROMACS 4.0 manual for a description[117]). Thus, the SR-TI simulations in this paper assess the reversible works needed to alchemically change each molecule to the benzene core reference state.

To automate the alchemical transformation setup using molecular mechanical Amber 99SB[70] and GAFF force fields[71] in GROMACS, we developed an in-house PERL script by augmenting an existing script used to convert Amber parameter and topology files to GROMACS format.[95]

**C.4. MD Simulations.** All of the simulations were run using a single precision version of GROMACS. Water solution was modeled with an explicit TIP3P water model using a cubic simulation box that extended at least 10 Å beyond the solute molecule. The box was prepared using the LEAP module from Amber Tools, version 1.2.[118] For water simulations, prior to production runs in the NPT ensemble at $T = 300$ K and $P = 1$ atm, each system was first minimized and then equilibrated in a series of runs with gradually vanishing harmonic restraints on all of the atoms of the solute. The equilibration protocol involved 500 steps of steepest descent minimization with a force constant, fc = 100 kJ/mol/nm$^2$, followed by 5000 steps of NVT run with fc = 100 kJ/mol/nm$^2$, followed by a series of NPT runs of 10 000 steps with fc progressing as 100, 10, 1, 0.1, 0.01, and 0.0 kJ/mol/nm$^2$. The whole equilibration procedure totaled 65 000 steps or 130 ps. The production run employed a Langevin thermostat and Berendsen barostat,[114–117] with identical collision frequencies of 2 ps$^{-1}$. For gas phase

simulations, the equilibration procedure in the canonical ensemble at $T = 300$ K was performed in a way similar to that in water, but without any harmonic restraints.

Throughout the simulations, all of the bonds containing hydrogen atoms were constrained using LINCS,[119] and the integration time step was set to 2 fs. To compute nonbonded interactions in water, we employed periodic boundary conditions (PBC) with particle mesh Ewald for electrostatics[114–117] using a 1 nm real space cutoff, while switching van der Waals interactions off over the range between 0.8 and 0.9 nm. In the gas phase, no PBCs were used, and all of the interactions were computed explicitly without any cutoffs. Unless stated otherwise, all of the simulations have been repeated two times with different random seeds to generate initial velocities. Typically, production runs were 1-ns-long, but in some cases the runs were extended to 4 ns per window. The coordinates of the system were recorded every 1000 steps.

**C.5. Regular SR-TI Simulations.** To obtain the alchemical free energies or reversible works, the TI procedure split the interval from the real state of the molecule at $\lambda = 0$ to the reference state at $\lambda = 1$ into $M$ windows, separated by equal $\lambda$ intervals. The majority of work was done with $M = 12$ windows, but in some cases additional simulations were performed with $M = 23$ windows. For each molecule, all TI windows had identical starting configurations. For each window, we recorded $(\partial V)/(\partial\lambda)$ values at every time step. The mean values $\langle(\partial V)/(\partial\lambda)\rangle$ for all of the windows were assembled into the final work using the Fourier beads integration procedure described in the Methodology section. The standard deviations were calculated from two independent runs. The differences between the work values in gas and water phases yielded the relative hydration free energies with respect to the reference state.

**C.6. HREX SR-TI Simulations.** To use the HREX option, we have developed a PERL script interfaced with GROMACS. Unless otherwise stated, replica exchanges were attempted every 1000 MD steps or 2 ps. For the majority of simulations, we employed a total of 500 exchange attempts over 1 ns simulation time. In special cases, the number of exchanges was increased to 2000, extending the simulation time to 4 ns. The exchanges were attempted using a standard procedure described in the Methodology section. Upon acceptance, only $\lambda$ values were exchanged between windows, while keeping the coordinates and velocities in place to expedite restarting of the simulations. Following the exchange attempts, simulations in all of the windows were restarted with different random seeds provided by the PERL random number generator to reinitialize the Langevin dynamics and avoid possible random seed artifacts.[120] All of the other simulation details were the same as with regular SR-TI simulations.

**C.7. Torsional PMF.** To get the PMF for the hindered amide bond rotation, we performed umbrella sampling simulations[121] using 72 equally spaced windows to span the range from $-180$ to $+180°$. We employed the harmonic biasing potential with a force constant of 2000 KJ/mol/rad$^2$. All of the equilibration and production protocols were identical to those in the regular SR-TI approach. The coordinates were saved every 50 steps during the 4 ns simulation time. The results of the simulations were unbiased and reconstructed into the final PMF using two independent methods. Specifically, we applied WHAM[83,84,122] and the HFB[88–90] method to the same data set, without regard to periodicity. Both methods gave nearly identical PMFs.

**C.8. *cis/trans* Ratio from the HREX SR-TI Simulations.** To compute the *cis*/*trans* ratios for the amide molecule from the HREX SR-TI simulations, we developed an additional PERL script that assembled a full length simulation trajectory from the individual short pieces produced by HREX SR-TI between exchange attempts for the specified $\lambda$ value. In the present paper, we only used the trajectory for the real state of the molecule ($\lambda = 0.0$). The values of the dihedral angle were then extracted using the TRJCONV and G_ANGLE tools from GROMACS.[114–116] The configurations with the dihedral angles in the range between $-100$ and $+100°$ were considered *cis*, and all the other configurations were considered *trans*.

## D. Results and Discussion

**D.1. *para*-Phenols.** To test the SR-TI approach, we first computed hydration free energies for a series of *para*-phenols *p*-$C_6H_4(OH)(X)$, where X = H, F, Cl, Br, I, $CH_3$, $C_2H_5$, CN, and $OCH_3$, along with benzene, for a total of 10 molecules. For this series, we did not employ the HREX option and used the benzene molecule without the hydrogen atoms as the common reference. Thus, for the *para*-phenols, the alchemical transformation annihilates all of the hydrogen atoms of the benzene ring, along with the OH and the X groups. The results are summarized in Table 1, along with experimental as well as computational data from other research groups for 7 out of the 10 molecules.

Table 1 demonstrates that SR-TI results are in good agreement with previous benchmark TI calculations.[79] The largest difference between the results of the two calculations is 0.37 kcal/mol. Our uncertainties (based on two independent simulations) are in general slightly higher than those from previous benchmark calculations. This is to be expected, as the bootstrap method used in the latter case is known to underestimate the uncertainties by a factor of 3.[14,80] The exception is provided by *p*-ethylphenol (X = $C_2H_5$), which shows the largest uncertainty of 0.21 kcal/mol for the reversible work in water. Overall, SR-TI systematically overestimated hydration free energies compared to the previous TI benchmarks. Note that the referred benchmark calculations were done using NVT simulations, while annihilating each molecule as a whole (absolute scale). In contrast, the SR-TI simulations were performed in the NPT ensemble and annihilated only those parts of each molecule that extended beyond the common reference substructure. Given these differences, we find the agreement between our and previous benchmark simulations particularly satisfying.

Like previous benchmark calculations, SR-TI overestimates the experimental relative hydration free energies for *para*-phenols. The SR-TI predictions for the phenols with simple aliphatic substituents, namely, *p*-cresol (X = $CH_3$) and *p*-ethylphenol (X = $CH_2CH_3$), are the closest to the

***Table 1.*** Alchemical and Relative Hydration Free Energies for a Series of *para*-Substituted Phenols[a]

| compound | SR-TI results | | | | previous results | |
|---|---|---|---|---|---|---|
| X,Y | $\Delta G_G$ (SD) | $\Delta G_W$ (SD) | $\Delta\Delta G$ (SD) | $\Delta\Delta G^{SR\text{-}TI}$ (SD) | $\Delta\Delta G^E$ | $\Delta\Delta G^{TI}$ (SD) |
| *H,H* | *−8.04 (0.02)* | *−5.24 (0.01)* | *−2.80 (0.02)* | *0.00 (0.03)* | *0.00* | *0.00 (0.03)* |
| H,OH | 9.27 (0.04) | 16.64 (0.08) | −7.37 (0.09) | −4.57 (0.09) | −5.75 | −4.97 (0.04) |
| F,OH | 7.52 (0.02) | 14.29 (0.04) | −6.77 (0.05) | −3.97 (0.05) | −5.33 | −4.29 (0.04) |
| Cl,OH | 9.72 (0.02) | 16.81 (0.01) | −7.09 (0.02) | −4.29 (0.03) | −6.17 | −4.66 (0.03) |
| Br,OH | 11.56 (0.01) | 19.06 (0.01) | −7.51 (0.01) | −4.70 (0.03) | −6.27 | −4.77 (0.03) |
| I,OH | 11.46 (0.00) | 18.50 (0.07) | −7.04 (0.07) | −4.24 (0.07) | N/A | N/A |
| $CH_3$,OH | 11.14 (0.01) | 18.39 (0.06) | −7.25 (0.06) | −4.45 (0.07) | −5.27 | −4.66 (0.03) |
| $CH_2CH_3$,OH | 9.90 (0.02) | 17.06 (0.21) | −7.15 (0.21) | −4.35 (0.21) | −5.27 | −4.38 (0.04) |
| CN,OH | 7.99 (0.03) | 17.33 (0.02) | −9.35 (0.04) | −6.55 (0.04) | −9.31 | −6.91 (0.04) |
| $OCH_3$,OH | 8.44 (0.00) | 17.41 (0.04) | −8.97 (0.04) | −6.17 (0.04) | N/A | N/A |

[a] Compound labels X,Y refer to the benzene substituent groups in the *para* position to each other. The free energy values are given in kcal/mol. The standard deviations (SDs) are computed on the basis of two independent simulations. The SR-TI simulations employed 12 windows run for 1 ns each in canonical (gas phase) and in NPT (water) ensembles at $T = 300$ K and $P = 1$ atm. $\Delta G_G$ and $\Delta G_W$ are the SR-TI work values for the alchemical transformation to the benzene core in the gas phase and water, respectively, and $\Delta\Delta G = \Delta G_G - \Delta G_W$ is the corresponding relative hydration free energy. The values $\Delta\Delta G^{SR\text{-}TI}$, $\Delta\Delta G^{TI}$, and $\Delta\Delta G^E$ are the hydration free energies from SR-TI, earlier TI calculations, and the experiment with respect to benzene. The corresponding absolute numbers for the reference benzene are −2.80, −0.70, and −0.86 kcal/mol.[79]

***Table 2.*** Alchemical and Relative Hydration Free Energies for Benzene and its Hydroxylated Derivatives[a]

| compound | | *M* | $EX_G$ [$EX_W$], % | $\Delta G_G$ (SD) | $\Delta G_W$ (SD) | $\Delta\Delta G$ (SD) | $\Delta\Delta G^{SR\text{-}TI}$ (SD) | $\Delta\Delta G^{SRSE}$ |
|---|---|---|---|---|---|---|---|---|
| *benzene* | | *12* | | *−8.04 (0.02)* | *−5.24 (0.01)* | *−2.80 (0.02)* | *0.00 (0.03)* | *0.00* |
| | HREX | 12 | 75 [34] | −8.04 (0.05) | −5.23 (0.01) | −2.82 (0.05) | −0.02 (0.05) | |
| | | 23 | | −8.07 (0.03) | −5.20 (0.03) | −2.87 (0.04) | −0.07 (0.05) | |
| phenol | | 12 | | 9.27 (0.04) | 16.64 (0.08) | −7.37 (0.09) | −4.57 (0.09) | −5.28 |
| | HREX | 12 | 74 [25] | 9.25 (0.01) | 16.68 (0.03) | −7.42 (0.03) | −4.62 (0.04) | |
| | | 23 | | 9.26 (0.03) | 16.90 (0.07) | −7.64 (0.08) | −4.84 (0.08) | |
| catechol | | 12 | | 1.64 (0.07) | 11.87 (0.11) | −10.23 (0.13) | −7.43 (0.13) | −7.00 |
| | HREX | 12 | 64 [21] | 1.51 (0.14) | 11.80 (0.07) | −10.28 (0.15) | −7.48 (0.15) | |
| | | 23 | | 1.58 (0.15) | 12.15 (0.08) | −10.57 (0.17) | −7.77 (0.17) | |
| pyrogallol | | 12 | | 9.34 (0.05) | 21.46 (0.09) | −12.12 (0.11) | −9.32 (0.11) | −9.61 |
| | HREX | 12 | 58 [19] | 9.31 (0.07) | 21.54 (0.03) | −12.23 (0.08) | −9.43 (0.08) | |
| | | 23 | | 9.27 (0.06) | 21.98 (0.03) | −12.71 (0.07) | −9.91 (0.07) | |

[a] The free energy values are given in kcal/mol. The standard deviations (SDs) are computed on the basis of two independent simulations. The SR-TI simulation employed *M* windows run for 1 ns each in canonical (gas phase) and in NPT (water) ensembles at $T = 300$ K and $P = 1$ atm. The HREX label indicates that the option was turned on during the simulations, in which case exchanges were attempted every 2 ps. $EX_G$ and $EX_W$ indicate the average acceptance ratio in the gas phase and water, respectively. $\Delta G_G$ and $\Delta G_W$ are the SR-TI work values for the alchemical transformation to the reference benzene core in the gas phase and water, respectively, and $\Delta\Delta G = \Delta G_G - \Delta G_W$ is the corresponding relative hydration free energy. $\Delta\Delta G^{SR\text{-}TI}$ and $\Delta\Delta G^{SRSE}$ are hydration free energies relative to benzene from the SR-TI approach and single reference state extrapolation (SRSE)[18] method, respectively.

experimental values and overestimate the hydration free energy by 0.82 and 0.92 kcal/mol, respectively. The largest disagreement (of 2.76 kcal/mol) is found for *p*-cyanophenol.

**D.2. Hydroxylated Benzenes.** To validate the HREX SR-TI approach, we computed hydration free energies for a series of hydroxylated benzenes with and without the HREX option. We selected benzene, phenol, catechol (benzene-1,2-diol), and pyrogallol (benzene-1,2,3-triol) for this study following previously published work.[18] For these four molecules, we used the same benzene core reference as above. We anticipated that for the benzene and phenol molecules, both of which have been studied in the previous series, the SR-TI approach with and without the HREX option should give identical results. The actual results are summarized in Table 2. For this series, we compare the results with the previous calculations employing the more approximate SRSE method that inspired the present work.[18]

Table 2 reveals that the regular SR-TI results correspond very well with the SRSE results. As expected, the discrepancies between SR-TI and the SRSE method are much larger than between the two TI approaches in the previous section. The largest discrepancy of 0.72 kcal/mol is found for the phenol. Nevertheless, given the computational savings that the more approximate SRSE method provides over the SR-TI, this level of agreement can be considered satisfactory.

Table 2 further demonstrates good agreement of the results from the regular SR-TI approach with those from the HREX SR-TI approach. Thus, we find that the relative free energies computed with and without the HREX option are identical within the specified uncertainties for all four molecules.

Here, we assess the dependence of the computed work estimates on the number of windows used in SR-TI. For simplicity, we employ the regular SR-TI approach in this test. Up to this point, the results discussed were obtained using 12 TI windows. With the regular SR-TI, we can simply insert and simulate 11 new windows precisely between the 12 old windows. In this way, we obtain a TI simulation with a total of 23 equally spaced windows (see Table 2). As is seen from Table 2, while benzene has nearly identical hydration free energies, the other three molecules show significant differences with 12 and 23 windows. In particular, pyrogallol shows the largest difference of 0.59 kcal/mol. Generally, the hydration free energies computed with 23 windows are lower than those with 12 windows. Interest-

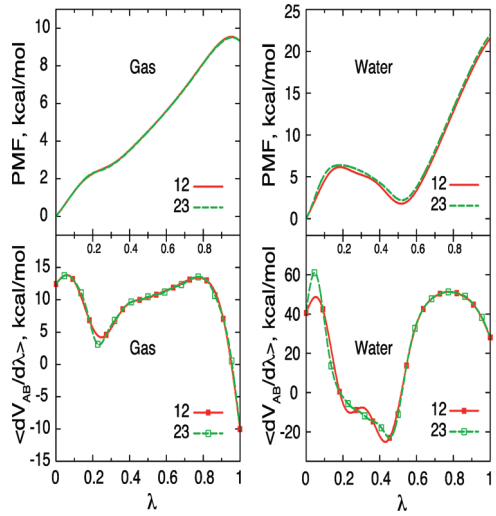


***Figure 3.*** Integration of the TI data for pyrogallol using the Fourier beads method. The bottom panels show the actual simulation data as squares and the Fourier beads fit as lines for 12 and 23 windows for a regular SR-TI simulation in the gas phase and in water. The top panels show the corresponding potentials of mean force (PMFs) from the Fourier fitted curves with respect to the coupling parameter $\lambda$. The values of these PMF curves at $\lambda = 1.0$ give the corresponding reversible works for alchemical transformation of pyrogallol to the benzene core.

ingly, the gas phase work values are independent of the number of windows. Therefore, the water phase contributions create the observed disparity.

To understand the dependence of the computed work in water on the number of windows, we plotted the mean force $\langle(\partial V_{AB})/(\partial\lambda)|_{\lambda_i}\rangle$ profiles, along with their integrals (see Supporting Information). For brevity, Figure 3 shows only the results for pyrogallol. As seen from Figure 3, the mean force peaks sharply near $\lambda = 0.045$. Going from pyrogallol down to catechol and then to phenol gradually lowers the peak, which vanishes completely at benzene (see Supporting Information). Therefore, this relatively sharp peak causes the 12-window interpolation to overestimate the relative hydration free energies.

Appearance of the sharp peak in the mean force profile near $\lambda = 0$ follows from the properties of the GROMOS soft-core potential used in this work.[73] This particular soft-core potential (see eqs 2−4) is known to create peaks in the $\langle(\partial V_{AB})/(\partial\lambda)|_{\lambda_i}\rangle$ profile near $\lambda = 0.0$ due to the so-called "soft-core effect of hydrogens without the LJ interactions."[76,115] In the present case, the peak comes from the real hydrogen atoms of the hydroxyl groups. Clearly, this problem can be alleviated by introducing additional windows, as was done here. In addition, the peak size can be manipulated by reducing the value of the SC-$\sigma$ parameter. Finally, the issue can be addressed by using a different form of the soft-core potential that peaks precisely at $\lambda = 0.0$.[16] Although trivial to implement, this latter option requires further study and has not been pursued in this work.

With HREX SR-TI, as with any replica exchange method, it is useful to know the exchange rate. Table 2 provides the exchange acceptance ratios for both gas and water phase transformations. As expected, Table 2 demonstrates that the acceptance ratio is significantly larger in the gas phase than in explicit water. However, we also notice a trend that the more atoms of molecule A need to be "switched off" to get to the reference state B, the lower the acceptance ratio becomes. Nevertheless, in the present case, the acceptance ratio does not affect the results significantly in either phase. The differences in free energies computed with the SR-TI approach with and without the HREX option suggest that regular MD achieves sufficient sampling for the molecules studied here.

**D.3. Aryl-Alcohols.** To further assess the HREX SR-TI approach for computing hydration free energies, we turned to a more complex set of molecules. Specifically, we looked at a series of aryl-alcohols that proved challenging for both SRSE and regular TI approaches.[18] For completeness, we have considered terminal aryl-alcohols of the form $C_6H_5-(CH_2)_n-OH$ and their dimethylated analogues $C_6H_5-C(CH_3)_2-(CH_2)_{n-1}-OH$ where $n = 1$, 2, and 3. These molecules have additional torsional degrees of freedom that might benefit from the enhanced sampling of HREX SR-TI. To enhance sampling of these degrees of freedom, we once again used the benzene core as the reference. Armed with the results of the previous sections, we can use the differences between the values from SR-TI with and without the HREX option as a measure of nonergodicity stored in these torsional degrees of freedom. The results are summarized in Table 3 along with comparisons to the earlier calculations and experimental data where available.

We find that hydration free energies computed with SR-TI always fall within the error bars from the earlier SRSE extrapolated values. However, the error bars on the SRSE results are rather large in this case, reducing their predictive power compared to that of SR-TI. In addition, for the three linear molecules $C_6H_5-(CH_2)_n-OH$, our results compare favorably with previous TI calculations and experimental measurements.[79] In particular, we find that our calculations overestimate the experimental relative solvation free energies for the linear aryl-alcohols by at most 1.57 kcal/mol. Compared to previous TI benchmarks, SR-TI free energies are overestimated by at most 0.44 kcal/mol. In both cases, the discrepancy is systematic.

Similarly to hydroxylated benzenes, the hydration free energies for the aryl-alcohols computed with and without the HREX option are nearly identical. Analysis of the data for this series of molecules reinforces our previous observation that the exchange rate is inversely proportional to the number of atoms that is "switched off" within the series. The largest molecule with 22 out of 28 atoms "switched off" in the reference state exhibits the lowest acceptance ratio of 16% in water (see Table 3). Recall that we enhance sampling in all of the torsional degrees of freedom outside the benzene core of the molecules when using the HREX option. The lack of sizable differences between the results with and without the replica exchange suggests that regular sampling of the torsional degrees of freedom is sufficiently ergodic. Therefore, to clearly demonstrate the utility of the HREX option, we examined a molecule with a hindered torsional degree of freedom.

**D.4. N-(2-Hydroxy-phenyl)formamide.** To demonstrate the utility of the HREX option, we have investigated

***Table 3.*** Alchemical and Relative Hydration Free Energies for a Series of Aryl Alcohols[a]

| | | SR-TI results | | | | previous results | | |
|---|---|---|---|---|---|---|---|---|
| compound | $EX_G[EX_W]$,% | $\Delta G_G$ (SD) | $\Delta G_W$ (SD) | $\Delta\Delta G$ (SD) | $\Delta\Delta G^{SR\text{-}TI}$ (SD) | $\Delta\Delta G^{SRSE}$ (SD) | $\Delta\Delta G^{E}$ | $\Delta\Delta G^{TI}$ (SD) |
| L1 | | −8.36 (0.01) | −1.29 (0.08) | −7.07 (0.08) | −4.27 (0.09) | | −5.76 | −4.71 (0.04) |
| HREX | 71 [22] | −8.37 (0.06) | −1.26 (0.05) | −7.11 (0.08) | −4.31 (0.08) | | | |
| L2 | | −7.70 (0.04) | −0.46 (0.13) | −7.24 (0.14) | −4.44 (0.14) | −5.90 (1.84) | −5.93 | −4.63 (0.04) |
| HREX | 68 [21] | −7.70 (0.05) | −0.61 (0.11) | −7.09 (0.12) | −4.29 (0.12) | | | |
| L3 | | −6.52 (0.02) | 0.77 (0.01) | −7.30 (0.02) | −4.49 (0.03) | −7.01 (3.09) | −6.06 | −4.80 (0.04) |
| HREX | 65 [19] | −6.47 (0.04) | 0.83 (0.03) | −7.31 (0.05) | −4.50 (0.05) | | | |
| B1 | | 5.35 (0.02) | 12.23 (0.11) | −6.88 (0.11) | −4.08 (0.11) | −5.45 (5.40) | | |
| HREX | 63 [18] | 5.35 (0.01) | 12.17 (0.00) | −6.82 (0.01) | −4.02 (0.03) | | | |
| B2 | | −15.67 (0.12) | −8.86 (0.06) | −6.81 (0.14) | −4.01 (0.14) | −8.64 (6.04) | | |
| HREX | 57 [17] | −15.68 (0.03) | −8.97 (0.16) | −6.72 (0.16) | −3.91 (0.16) | | | |
| B3 | | −13.58 (0.11) | −6.39 (0.01) | −7.19 (0.11) | −4.39 (0.11) | −2.12 (5.61) | | |
| HREX | 52 [16] | −13.59 (0.01) | −6.33 (0.12) | −7.26 (0.12) | −4.46 (0.12) | | | |

[a] The free energy values are given in kcal/mol. In the compound column, Ln refers to linear $C_6H_5-(CH_2)_n-OH$, and Bn refers to their branched, dimethylated analogues $C_6H_5-C(CH_3)_2(CH_2)_{n-1}OH$. The HREX label indicates that the option was turned on during the simulations, in which case exchanges were attempted every 2 ps. $EX_G$ and $EX_W$ indicate the average acceptance ratio in the gas phase and water, respectively. The standard deviations (SDs) are computed on the basis of two independent simulations. The SR-TI simulations employed 12 windows run for 1 ns each in canonical (gas phase) and in NPT (water) ensembles at $T = 300$ K and $P = 1$ atm. $\Delta G_G$ and $\Delta G_W$ are the SR-TI work values for the alchemical transformations to the reference benzene core in the gas phase and water, respectively, and $\Delta\Delta G = \Delta G_G - \Delta G_W$ is the corresponding relative hydration free energy. $\Delta\Delta G^{SR\text{-}TI}$, $\Delta\Delta G^{SRSE}$, $\Delta\Delta G^{E}$, and $\Delta\Delta G^{TI}$ are relative hydration free energies with respect to benzene from the SR-TI, single reference state extrapolation (SRSE),[18] experiment and earlier TI calculations, correspondingly.[79]

molecules with an N−C amide bond. The rotation about the amide bond is so strongly hindered that amides are typically considered locked in one of the two conformations at $T = 300$ K, namely *cis* or *trans*.[13,28,123–128] Of the two isomers, the *trans* isomer is considered the most favorable, and the *cis* isomer is often completely ignored.[63,128–130] Earlier attempts to compute hydration free energies of some amides produced results conflicting with experimental data.[13,123,125,127] Interestingly, for simple amides, the hydration free energies of the *cis* and *trans* isomers have been found experimentally to be nearly identical.[128] This argued that the preference for the *trans* isomer, which has important implications for peptide and protein structure in general, is not due to hydration. For the purposes of our study, we wanted to examine an amide with substantially different hydration free energies in the *cis* and *trans* conformations. Consequently, we placed a formamide group on the benzene ring and added a hydroxyl group in the *ortho* position allowing for intramolecular interactions. In what follows, we refer to the resulting N-(2-hydroxy-phenyl)formamide simply as "amide" for brevity.

An exhaustive conformational search for the amide identified five groups of isomers in the gas phase (see Supporting Information). Specifically, employing the AM1 semiempirical potential,[105] we found two groups for *trans* and three groups for *cis* conformations of the amide. All of the groups, except the first *trans* group, contained two iso-energetic conformations that were mirror reflections of each other. The *trans* isomer that did not have such a degeneracy was completely planar. Thus, we identified nine distinct conformations in total. Interestingly, the AM1 potential[105] ranked the first *cis* group to be the most favorable in the gas phase, while the second *trans* group that had a possibility for intramolecular hydrogen bonding was ranked the least favorable.

We used the information from all *cis* and *trans* conformations to derive a conformation-independent set of AM1BCC atomic charges[107,108] as described in the Methodology section. Reoptimizing geometries of these conformations using the conformation-independent charges with the GAFF molecular mechanical (MM) potential[71,110] changed the ranking only slightly (see the Supporting Information). Most importantly, the MM potential strongly favored (by 3.0 kcal/mol) the second *trans* group of conformations with an intramolecular hydrogen bond. This group became the new ground state and was 1.3 kcal/mol lower than the lowest energy *cis* group. We did not attempt to make any empirical adjustments to correct for this behavior in the MM potential.

Using the MM potential derived above, we computed the reversible works or potentials of mean force (PMFs) for the amide bond torsion. The PMFs were computed from MD simulations with a relatively stiff harmonic biasing potential on the amide bond dihedral angle. To unbias the results, we employed two independent approaches, namely Weighted Histogram Analysis Method (WHAM)[83,84,122] and umbrella integration with Harmonic Fourier Beads (HFB).[88–90] Figure 4 depicts the final PMFs in the gas and water phases for 360° rotation about the amide bond. In computing the PMFs, we ignored the periodicity and treated the data near −180 and near +180 independently. In what follows, we use the PMF extrema, i.e., the local minima and maxima (see the Supporting Information), instead of introducing *cis* and *trans* indicator functions,[94] to make quantitative comparisons. The resulting asymmetry in the PMFs provides a measure of uncertainty, which is between 0.3 and 0.5 kcal/mol. Importantly, the PMFs confirm that the rotation about the amide bond is strongly hindered with barriers ranging between 13.3 and 15.0 kcal/mol. Hence, *cis* and *trans* isomers will not interconvert during regular MD simulations on a nanosecond time scale. Furthermore, in the gas phase, the *cis* isomer is lower than the *trans* by 0.3 kcal/mol, despite the bias in the force field toward the *trans* isomer. In contrast, in water, the *trans* isomer is lower than *cis* by 1.5 kcal/mol. Such a significant change in the PMF upon hydration suggests that the selected molecule is well suited for the ultimate HREX SR-TI validation.

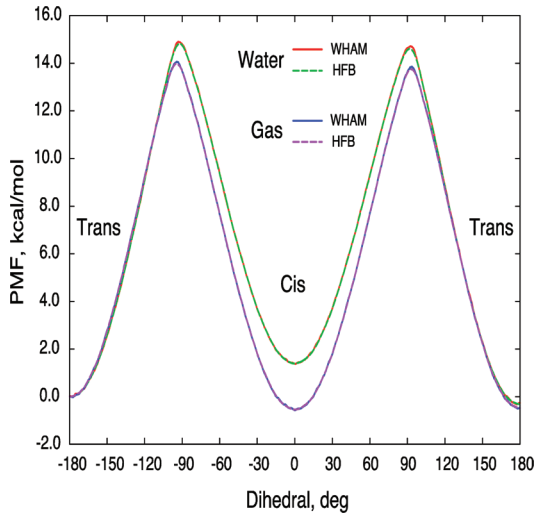


***Figure 4.*** Potential of mean force (PMF) for rotation about the amide bond in the gas phase and in water. The PMFs were computed from umbrella sampling simulations with harmonic dihedral restraints using two independent methods, namely, the weighted histogram analysis method (WHAM) and umbrella integration with harmonic Fourier beads method.

To validate the SR-TI approach, its results must satisfy certain criteria. Specifically, in the regular SR-TI simulations, the amide molecule should not cross the barrier separating the *cis* and *trans* conformations and, therefore, should maintain the conformation of the initial configuration near $\lambda = 0$. Thus, the regular SR-TI simulations in either phase should provide corresponding relative free energies of *cis* and *trans* isomers directly. Furthermore, regardless of the initial configuration, each HREX SR-TI simulation should yield the *cis*/*trans* ratio matching the free energy differences computed with the regular SR-TI approach. Similarly, the averaged hydration free energy computed with HREX SR-TI, along with its gas and water phase components, should be bounded by the corresponding values for *cis* and *trans* conformations from regular SR-TI. Finally, the computed results should be consistent with the amide torsional PMFs.

Table 4 summarizes the results of multiple SR-TI simulations with and without the HREX option. To ensure that the amide dihedral is activated in the HREX simulations, we employed the benzene core as the reference structure. All nine conformations of the amide identified during the exhaustive search were used to initiate SR-TI simulations, resulting in nine independent SR-TI simulations per option. In each case, we ran two simulations per isomer, one with 12 and another with 23 windows. All of the results can be found in the Supporting Information. Table 4 shows only the averaged results from these simulations, for all isomers together, and for *cis* and *trans* isomers individually.

The regular SR-TI simulations demonstrate that *cis* and *trans* conformations have similar free energies in the gas phase, and that hydration significantly favors the *trans* conformation over the *cis*. These results are in good agreement with the torsional PMFs.

Comparing the results with 12 and 23 windows, we find a discrepancy of about 0.4 kcal/mol between relative free energies of *cis* and *trans* conformations both in the gas and water phases. Similar discrepancy was observed in the torsional PMFs. Thus, the regular SR-TI simulations with 12 windows underestimate the free energy of the *trans* conformation, which is the opposite of the reversible work for the alchemical transformation reported in Table 4. This discrepancy is due to the appearance of a sharp peak in the mean force profiles near $\lambda = 0$ (see the Supporting Information) as was seen for the hydroxylated benzenes in water. However, because in the amide case, this discrepancy is present in both the gas and water phases, it fortuitously cancels out in the final free energy difference result. The fact that the *trans* and not the *cis* conformation is affected strongly argues that this peak is related to the intramolecular hydrogen bond formed between the hydroxyl group and the amide carbonyl.

Using the results of the regular SR-TI simulations, we can construct the expectation values for the *cis*/*trans* ratios (*C*/*T*) in the gas and water phases as follows:

$$C/T = \exp\left[\frac{\Delta G(cis) - \Delta G(trans)}{k_B T}\right] \qquad (13)$$

where $k_B$ is the Boltzmann constant and $\Delta G(cis)$ and $\Delta G(trans)$ are the reversible works to alchemically transform the amide in the *cis* and *trans* configurations to the reference state, respectively. Because these expectation values employ the relative free energies of the *cis* and *trans* conformations, they are particularly sensitive to small variations and hence depend on the number of windows. With 12 windows, the *C*/*T* ratios are 1.91 and 0.10, in the gas phase and water, respectively, whereas with 23 windows, these numbers change to 0.98 and 0.06, respectively.

Now we can examine the results of the HREX SR-TI simulations. Table 4 shows that with 12 windows the rate of exchanges in water is only 11%. Note that merely 11 atoms out of a total 17 are "switched off" here. This is the lowest exchange rate we have seen. However, even at such a low rate, the HREX option cuts the standard deviation of the computed hydration free energy in half (from 0.87 in regular SR-TI to 0.42 kcal/mol in HREX SR-TI). Nevertheless, over 4 ns of simulation time, the *cis*/*trans* ratios in the gas and water phases are far from the expected values (1.91 and 0.10, respectively). Attesting to the validity of HREX SR-TI, the computed hydration free energies stay well within the bounds set by regular SR-TI.

Increasing the number of windows can significantly improve the HREX SR-TI results. We have seen that the number of windows can affect the results of regular SR-TI simulations and their derived expectation values. Here, we demonstrate that increasing the number of windows about two times greatly affects the exchange rate and consequently improves the overall HREX SR-TI performance. In particular, going from 12 to 23 windows raises the acceptance ratio in water from 11 to 40%. This, in turn, reduces the standard deviation in the computed hydration free energy by about a factor of 4 (from 0.86 to 0.20 kcal/mol). Furthermore, the averaged hydration free energies for simulations initiated from *cis* and *trans* conformations get within 0.36 kcal/mol of each other. With 23 windows, the *C*/*T* ratios in the gas and water phases begin to approach the expected values (of

***Table 4.*** Results of Regular and HREX SR-TI Simulations for N-(2-Hydroxy-phenyl)formamide[a]

| | regular SR-TI, 4 ns | | | HREX SR-TI, 4 ns | | | | HREX SR-TI, last 3 ns | |
|---|---|---|---|---|---|---|---|---|---|
| selection | $\Delta G_G$ | $\Delta G_W$ | $\Delta\Delta G$ | $EX_G$ [$EX_W$], % | $\Delta G_G$ | $\Delta G_W$ | $\Delta\Delta G$ | $C/T_G$ [$C/T_W$] | $\Delta\Delta G$ | $C/T_G$ [$C/T_W$] |
| | | | | | 12 windows | | | | | |
| all-avg | 16.35 | 30.92 | **−14.57** | 56 [11] | 16.32 | 31.06 | **−14.74** | 0.99 [0.88] | | |
| all-SD | 0.20 | 0.68 | **0.87** | 1 [0] | 0.05 | 0.43 | **0.42** | 0.38 [0.95] | | |
| cis-avg | 16.48 | 30.47 | **−13.99** | 56 [11] | 16.31 | 30.84 | **−14.53** | 0.99 [1.31] | | |
| cis-SD | 0.04 | 0.05 | **0.07** | 1 [0] | 0.05 | 0.34 | **0.36** | 0.36 [0.89] | | |
| trans-avg | 16.09 | 31.82 | **−15.73** | 56 [11] | 16.34 | 31.49 | **−15.15** | 1.00 [0.03] | | |
| trans-SD | 0.07 | 0.04 | **0.08** | 2 [0] | 0.06 | 0.13 | **0.09** | 0.50 [0.06] | | |
| | | | | | 23 windows | | | | | |
| all-avg | 16.49 | 31.05 | **−14.56** | 77 [40] | 16.51 | 31.58 | **−15.08** | 1.18 [0.24] | **−15.24** | 1.20 [0.08] |
| all-SD | 0.05 | 0.86 | **0.86** | 1 [0] | 0.08 | 0.19 | **0.20** | 0.52 [0.16] | **0.12** | 0.60 [0.03] |
| cis-avg | 16.49 | 30.48 | **−13.99** | 77 [40] | 16.51 | 31.46 | **−14.96** | 1.27 [0.33] | **−15.20** | 1.12 [0.09] |
| cis-SD | 0.02 | 0.05 | **0.06** | 1 [0] | 0.10 | 0.07 | **0.10** | 0.61 [0.08] | **0.10** | 0.60 [0.02] |
| trans-avg | 16.50 | 32.20 | **−15.71** | 77 [40] | 16.50 | 31.82 | **−15.32** | 0.99 [0.04] | **−15.33** | 1.34 [0.05] |
| trans-SD | 0.08 | 0.04 | **0.10** | 0 [0] | 0.04 | 0.11 | **0.08** | 0.26 [0.01] | **0.13** | 0.70 [0.00] |

[a] The free energy values are given in kcal/mol. In the selection column, all, cis, and trans labels refer to average (avg) and standard deviation (SD) values computed over all nine, just six *cis*, and just three *trans* conformations of the amide, respectively. The number of windows used in the simulations is indicated on a separate line. The HREX SR-TI simulations attempted exchanges every 2 ps. $EX_G$ and $EX_W$ indicate the average acceptance ratio in the gas phase and water, respectively. Similarly, $C/T_G$ and $C/T_W$ refer to cis/trans ratios in gas and water phases. All simulations employed the specified number of windows run for 4 ns each in canonical (gas phase) and in NPT (water) ensembles at $T = 300$ K and $P = 1$ atm. $\Delta G_G$ and $\Delta G_W$ are the SR-TI work values for the alchemical transformations to the reference benzene core in the gas phase and water, respectively, and $\Delta\Delta G = \Delta G_G - \Delta G_W$ is the corresponding relative hydration free energy. The expectation values for $C/T$ in either phase were computed using alchemical free energies from regular SR-TI as follows: $C/T = \exp[(\Delta G(\text{cis}) - \Delta G(\text{trans}))/(k_B T)]$, where $k_B$ is the Boltzmann constant. With 12 beads, these values were 1.91 and 0.10 for the gas and water phases, respectively, whereas with 23 windows they changed to 0.98 and 0.06.

0.98 and 0.06, respectively). Hence, in difficult cases like the amide, increasing the number of windows can greatly improve the results of the HREX simulations.

To assess the effect of exchange frequency on the results, we repeated SR-TI simulations with the HREX option attempting simulations every 500 steps (every 1 ps) as opposed to every 1000 steps. While doubled the number of exchange attempts over the 4 ns run, it did not significantly affect the final solvation free energy (see the Supporting Information). This confirms that our previous simulations have reached convergence. We note that increasing the frequency of exchanges could provide additional computational savings as long as the time it takes to evaluate the Metropolis acceptance criteria is much less than the time to run MD simualtions between the exchanges. However, we did not attempt to identify the corresponding limit on exchange frequency. A previous work employing HREX with FEP suggested that attempting exchanges every 100 steps (0.2 ps) is close to the limit.[96]

Equilibrating the generalized ensemble improves the HREX SR-TI predictions. Typically, before we begin the Hamiltonian exchanges, each window is equilibrated the same way as in regular SR-TI (see the Methodology section). However, this equilibration does not involve any exchanges, and hence at the beginning, the simulations are still biased by their starting configurations. Therefore, additional equilibration is desired for the generalized HREX ensemble itself. To demonstrate this, we use the simulation with 23 windows. Specifically, we divide the 4 ns of simulation time, which has 2000 exchange attempts, into two periods $t_e + t_p = 4$ ns, where $t_e$ and $t_p$ are the equilibration and production periods. We use only the $t_p$ portion to recalculate the hydration free energies and the *cis*/*trans* ratios. In particular, we start with $t_p = 3$ and proceed in decrements of 1 ns or 500 exchange attempts. Note that we have already discussed the results with $t_p = 4$ ns, which corresponds to using all of the HREX simulation time, in previous paragraphs. In Table 4, we only show additional results for $t_p = 3$ ns, which has the lowest hydration free energy of −15.24 kcal/mol with the lowest standard deviation of 0.12 kcal/mol. Clearly, the recomputed hydration free energy remains bounded by the corresponding regular SR-TI values for *cis* and *trans* conformations. Finally, for $t_p = 3$ ns, the *C*/*T* ratio in water is 0.08 and practically matches the expected value of 0.06. Thus, introducing an equilibration time for the generalized ensemble can greatly improve the quality of predictions.

This work has demonstrated that the HREX option is absolutely essential to get high quality results for solutes with multiple configurations that have distinct solvation properties and are separated by high energy barriers. In the case of the amide studied here, HREX SR-TI consistently circumvents barriers as high as 15 kcal/mol over the course of 4 ns. Importantly, HREX SR-TI achieves this outstanding result using a modest number of simulation windows. Such a dramatic enhancement in sampling is simply impossible with either regular SR-TI or conventional TI.

We feel that the difficult amide test case has helped us validate the SR-TI approach and establish best protocols for its use. The *C*/*T* ratios predicted with regular SR-TI and HREX SR-TI are in excellent agreement. Furthermore, the final hydration free energy from HREX SR-TI is independent of whether the simulation starts from the *cis* or *trans* isomer. The final value stays within the clear bounds defined by regular SR-TI for *cis* and *trans* isomers separately and, independently, by the dihedral PMF results. All of these results provide confidence in the approach necessary for future applications.

## E. Conclusion

To conclude, we have presented a single-topology TI variant, called SR-TI augmented with HREX, that provides reliable estimates of relative solvation free energies for series of related molecules even in the presence of hindered conformational transitions. The key difference from conventional TI methods is that SR-TI transforms all of the molecules from a particular series down to a single reference state that does not have to correspond to a physical state. The choice of the reference state is flexible and allows rational selection of torsional degrees of freedom for enhanced sampling. Furthermore, a reduction in molecular volume in the reference state allows for better mobility in confined spaces. The benefits of the enhanced sampling and mobility in the reference state can be passed on to the real state using the HREX option. In addition, the HREX option improves overlap in configuration space between the TI simulation windows. Therefore, combining the SR-TI approach with Hamiltonian replica exchange brings considerable improvements over current single-topology TI methods. Thus, we feel that the SR-TI approach with and without the HREX option is a useful addition to the family of rigorous, high quality TI methods. Application of this methodology to more complex problems, including ligand binding, is currently in progress in our laboratory and will be described in a subsequent paper.

**Acknowledgment.** We would like to thank Dr. In-Chul Yeh, Dr. Michael S. Lee, and Dr. Hyung-June Woo for helpful discussions. Also, we acknowledge the National Cancer Institute (NCI) for allocation of computing time and staff support at the Advanced Biomedical Computing Center (ABCC) at NCI−Frederick. This work was sponsored by the U.S. Department of Defense High Performance Computing Modernization Program (HPCMP), under the High Performance Computing Software Applications Institutes (HSAI) initiative. The opinions and assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the U.S. Army or the U.S. Department of Defense.

**Supporting Information Available:** Figures 1S−4S show the results of integration of the simulation data for the series of hydroxybenzenes that demonstrate the appearance of a sharp peak near $\lambda = 0.0$ due to hydrogen atoms of the hydroxyl groups. Figures 5S and 6S show the same results for the *cis* and *trans* isomers of the amide. For the amide, all of the isomers identified during the exhaustive conformation search and their AM1 and AM1BCC/GAFF energies are presented. A summary of the torsional PMF extrema is also presented. Additional simulation and analysis results are provided in several tables. This material is available free of charge via the Internet at http://pubs.acs.org.

### References

(1) Gunthrie, J. P. *J. Phys. Chem. B* **2009**, *113*, 4501.

(2) Lybrand, T. P.; Ghosh, I.; McCammon, J. A. *J. Am. Chem. Soc.* **1985**, *107*, 7793.

(3) Jorgensen, W. L.; Nguyen, T. B. *J. Comput. Chem.* **1992**, *14*, 195.

(4) Ewing, T. J. A.; Lybrand, T. P. *J. Phys. Chem.* **1994**, *98*, 1748.

(5) Lybrand, T. P.; McCammon, J. A.; Wipff, G. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 833.

(6) Tembe, B. L.; McCammon, J. A. *Comp. Chem.* **1984**, *8*, 281.

(7) Severance, D. L.; Essex, J. W.; Jorgensen, W. L. *J. Comput. Chem.* **1995**, *16*, 311.

(8) Bordner, A. J.; Cavasotto, C. N.; Abagyan, R. A. *J. Phys. Chem. B* **2002**, *106*, 11009.

(9) Woolf, T. B.; Roux, B. *J. Am. Chem. Soc.* **1994**, *116*, 5916.

(10) Richards, N. G. J.; Williams, P. B.; Tute, M. S. *Int. J. Quantum Chem.* **1992**, *44*, 219.

(11) Mannhold, R.; Waterbeemd, H. v. d. *J. Comput.-Aided Mol. Des.* **2001**, *15*, 337.

(12) Garrido, N. M.; Queimada, A. J.; Jorge, M.; Macedo, E. A.; Economou, I. G. *J. Chem. Theory Comput.* **2009**, *5*, 2436.

(13) Morgantini, P.-Y.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 6057.

(14) Nicholls, A.; Mobley, D. L.; Guthrie, J. P.; Chodera, J. D.; Bayly, C. I.; Cooper, M. D.; Pande, V. S. *J. Med. Chem.* **2008**, *51*, 769.

(15) Pitera, J. W.; Gunsteren, W. F. v. *J. Phys. Chem. B* **2001**, *105*, 11264.

(16) Shirts, M. R.; Pande, V. S. *J. Chem. Phys.* **2005**, *122*, 134508(13).

(17) Shivakumar, D.; Deng, Y.; Roux, B. *J. Chem. Theory Comput.* **2009**, *5*, 919.

(18) Mordasini, T. Z.; McCammon, J. A. *J. Phys. Chem. B* **2000**, *104*, 360.

(19) Liu, H.; Mark, A. E.; Gunsteren, W. F. v. *J. Phys. Chem.* **1996**, *100*, 9485.

(20) Hritz, J.; Oostenbrink, C. *J. Phys. Chem. B* **2009**, *113*, 12711.

(21) Kirkwood, J. G. *J. Chem. Phys.* **1935**, *3*, 300.

(22) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420.

(23) Mitchel, M. J.; McCammon, J. A. *J. Comput. Chem.* **1991**, *12*, 271.

(24) Jorgensen, W. L.; Ravimohan, C. *J. Chem. Phys.* **1985**, *83*, 3050.

(25) Okazaki, S.; Nakanishi, K.; Touhara, H. *J. Chem. Phys.* **1979**, *71*, 2421.

(26) Postma, J. P. M.; Berendsen, H. J. C.; Haak, J. R. *Faraday Symp. Chem. Soc.* **1982**, *17*, 55.

(27) Singh, U. C.; Brown, F. K.; Bash, P. A.; Kollman, P. A. *J. Am. Chem. Soc.* **1987**, *109*, 1509.

(28) Bash, P. A.; Singh, U. C.; Langridge, R.; Kollman, P. A. *Science* **1987**, *236*, 564.

(29) Mezei, M.; Beveridge, D. L. *Ann. Acad. Sci. N. Y.* **1986**, *482*, 1.

(30) Warshel, A. *J. Phys. Chem.* **1982**, *86*, 2218.

(31) Straatsma, T. P.; McCammon, J. A. *J. Chem. Phys.* **1994**, *101*, 5032.

(32) Straatsma, T. P.; McCammon, J. A. *J. Chem. Phys.* **1989**, *90*, 3300.

(33) Straatsma, T. P.; McCammon, J. A. *J. Chem. Phys.* **1989**, *91*, 3631.

(34) Deng, Y.; Roux, B. *J. Chem. Theory Comput.* **2006**, *2*, 1255.

(35) Min, D.; Li, H.; Li, G.; Bitetti-Putzer, R.; Yang, W. *J. Chem. Phys.* **2007**, *126*, 144109.

(36) Grubmueller, H. *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.* **1995**, *52*, 2793.

(37) Lange, O. F.; Schaefer, L. V.; Grubmueller, H. *J. Comput. Chem.* **2006**, *27*, 1693.

(38) Voter, A. F. *Phys. Rev. Lett.* **1997**, *78*, 3908.

(39) Voter, A. F. *J. Chem. Phys.* **1997**, *106*, 4665.

(40) Perez, D.; Uberuaga, B. P.; Shim, Y.; Amar, J. G.; Voter, A. F. *Annu. Rep. Comput. Chem.* **2009**, *5*, 79.

(41) Fajer, M.; Hamelberg, D.; McCammon, J. A. *J. Chem. Theory Comput.* **2008**, *4*, 1565.

(42) Fajer, M.; Swift, R. V.; McCammon, J. A. *J. Comput. Chem.* **2009**, *30*, 1719.

(43) Hamelberg, D.; McCammon, J. A. *Annu. Rep. Comput. Chem.* **2006**, *2*, 221.

(44) Hamelberg, D.; Mongan, J.; McCammon, J. A. *J. Chem. Phys.* **2004**, *120*, 11919.

(45) Hamelberg, D.; Shen, T.; McCammon, J. A. *J. Am. Chem. Soc.* **2005**, *127*, 1969.

(46) Li, H.; Fajer, M.; Yang, W. *J. Chem. Phys.* **2007**, *126*, 024106/1.

(47) Zheng, L.; Carbone, I. O.; Lugovskoy, A.; Berg, B. A.; Yang, W. *J. Chem. Phys.* **2008**, *129*, 034105/1.

(48) Zheng, L.; Yang, W. *J. Chem. Phys.* **2008**, *129*, 124107/1.

(49) Hansmann, U. H. E. *Phys. A (Amsterdam, Neth.)* **2010**, *389*, 1400.

(50) Hansmann, U. H. E. *Chem. Phys. Lett.* **1997**, *281*, 140.

(51) Hansmann, U. H. E.; Okamoto, Y. *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.* **1996**, *54*, 5863.

(52) Lyubartsev, A. P.; Martsinovskii, A. A.; Shevkunov, S. V.; Vorontsov-Velyaminov, P. N. *J. Chem. Phys.* **1992**, *96*, 1776.

(53) Sugita, Y.; Kitao, A.; Okamoto, Y. *J. Chem. Phys.* **2000**, *113*, 6042.

(54) Sugita, Y.; Okamoto, Y. *Chem. Phys. Lett.* **1999**, *314*, 141.

(55) Itoh, S. G.; Okumura, H.; Okamoto, Y. *J. Chem. Phys.* **2010**, *132*, 134105/1.

(56) Kwak, W.; Hansmann, U. H. E. *Phys. Rev. Lett.* **2005**, *95*, 138102(4).

(57) Hritz, J.; Oostenbrink, C. *J. Chem. Phys.* **2008**, *128*, 144121(10).

(58) Woods, C. J.; Essex, J. W.; King, M. A. *J. Phys. Chem. B* **2003**, *107*, 13703.

(59) Woods, C. J.; King, M. A.; Essex, J. W. *Lect. Notes Comput. Sci. Eng.* **2006**, *49*, 251.

(60) Fukunishi, H.; Watanabe, O.; Takada, S. *J. Chem. Phys.* **2002**, *116*, 9058.

(61) Jorgensen, W. L.; Nguyen, T. B. *J. Comput. Chem.* **1993**, *14*, 195.

(62) Carlson, H. A.; Nguyen, T. B.; Orozco, M.; Jorgensen, W. L. *J. Comput. Chem.* **1993**, *14*, 1240.

(63) Drakenberg, T.; Forsen, S. *J. Chem. Soc. D* **1971**, 1404.

(64) Straatsma, T. P.; McCammon, J. A. *J. Chem. Phys.* **1991**, *95*, 1175.

(65) Boresch, S.; Karplus, M. *J. Phys. Chem. A* **1999**, *103*, 103.

(66) Boresch, S.; Karplus, M. *J. Phys. Chem. A* **1999**, *103*, 119.

(67) Levitt, M.; Warshel, A. *Nature (London)* **1975**, *253*, 694.

(68) Karplus, M.; McCammon, J. A. *Nature (London)* **1979**, *277*, 578.

(69) McCammon, J. A.; Gelin, B. R.; Karplus, M. *Nature (London)* **1977**, *267*, 585.

(70) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins* **2006**, *65*, 712.

(71) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157.

(72) Simonson, T. *Mol. Phys.* **1993**, *80*, 441.

(73) Beutler, T. C.; Mark, A. E.; van Schaik, R. C.; Gerber, P. R.; van Gunsteren, W. F. *Chem. Phys. Lett.* **1994**, *222*, 529.

(74) Cross, A. J. *Chem. Phys. Lett.* **1986**, *128*, 198.

(75) Steinbrecher, T.; Mobley, D. L.; Case, D. A. *J. Chem. Phys.* **2007**, *127*, 214108(13).

(76) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. *J. Comput. Chem.* **2005**, *26*, 1701.

(77) Sitkoff, D.; Sharp, K. A.; Honig, B. *J. Phys. Chem.* **1994**, *98*, 1978.

(78) Mobley, D. L.; Bayly, C. I.; Cooper, M. D.; Dill, K. A. *J. Phys. Chem. B* **2009**, *113*, 4533.

(79) Mobley, D. L.; Bayly, C. I.; Cooper, M. D.; Shirts, M. R.; Dill, K. A. *J. Chem. Theory Comput.* **2009**, *5*, 350.

(80) Mobley, D. L.; Dumont, E.; Chodera, J. D.; Dill, K. A. *J. Phys. Chem. B* **2007**, *111*, 2242.

(81) Ferrenberg, A. M.; Swendsen, R. H. *Phys. Rev. Lett.* **1988**, *61*, 2635.

(82) Ferrenberg, A. M.; Swendsen, R. H. *Phys. Rev. Lett.* **1989**, *63*, 1195.

(83) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Comput. Chem.* **1992**, *13*, 1011.

(84) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. *J. Comput. Chem.* **1995**, *16*, 1339.

(85) Bennett, C. H. *J. Comput. Phys.* **1976**, *22*, 245.

(86) Shirts, M. R.; Chodera, J. D. *J. Chem. Phys.* **2008**, *129*, 124105/1.

(87) Shirts, M. R.; Mobley, D. L.; Chodera, J. D. *Annu. Rep. Comput. Chem.* **2007**, *3*, 41.

(88) Khavrutskii, I. V.; Arora, K.; Brooks, C. L., III. *J. Chem. Phys.* **2006**, *125*, 174108.

(89) Khavrutskii, I. V.; Dzubiella, J.; McCammon, J. A. *J. Chem. Phys.* **2008**, *128*, 044106.

(90) Khavrutskii, I. V.; McCammon, J. A. *J. Chem. Phys.* **2007**, *127*, 124901.

(91) Hritz, J.; Oostenbrink, C. *J. Chem. Phys.* **2007**, *127*, 204104/1.

(92) Mobley, D. L.; Graves, A. P.; Chodera, J. D.; McReynolds, A. C.; Shoichet, B. K.; Dill, K. A. *J. Mol. Biol.* **2007**, *371*, 1118.

(93) Boyce, S. E.; Mobley, D. L.; Rocklin, G. J.; Graves, A. P.; Dill, K. A. *J. Mol. Biol.* **2009**, *394*, 747.

(94) Mobley, D. L.; Chodera, J. D.; Dill, K. A. *J. Chem. Theory Comput.* **2007**, *3*, 1231.

(95) Mobley, D. L.; Chodera, J. D.; Dill, K. A. *J. Chem. Phys.* **2006**, *125*, 084902(16).

(96) Jiang, W.; Hodoscek, M.; Roux, B. *J. Chem. Theory Comput.* **2009**, *5*, 2583.

(97) Okabe, T.; Kawata, M.; Okamoto, Y.; Mikami, M. *Chem. Phys. Lett.* **2001**, *335*, 435.

(98) Nadler, W.; Hansmann, U. H. E. *Los Alamos Natl. Lab., Prepr. Arch., Quant. Biol.* **2007**, 1.

(99) Nadler, W.; Hansmann, U. H. E. *J. Phys. Chem. B* **2008**, *112*, 10386.

(100) Trebst, S.; Troyer, M.; Hansmann, U. H. E. *J. Chem. Phys.* **2006**, *124*, 174903/1.

(101) Weininger, D. *J. Chem. Inf. Comput. Sci.* **1988**, *28*, 31.

(102) Weininger, D.; Weininger, A.; Weininger, J. L. *J. Chem. Inf. Comput. Sci.* **1989**, *29*, 97.

(103) Sadowski, J.; Gasteiger, J.; Klebe, G. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 1000.

(104) Renner, S.; Schwab, C. H.; Gasteiger, J.; Schneider, G. *J. Chem. Inf. Model.* **2006**, *46*, 2324.

(105) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.

(106) Stewart, J. J. P. *MOPAC7*.

(107) Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I. *J. Comput. Chem.* **2000**, *21*, 132.

(108) Jakalian, A.; Jack, D. B.; Bayly, C. I. *J. Comput. Chem.* **2002**, *23*, 1623.

(109) Wang, J. Antechamber. In *Amber Tools*, 1.2 ed.; Case, D. A., Ed.; University of California, San Francisco: San Francisco, CA, 2009.

(110) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. *J. Mol. Graphics Modell.* **2006**, *25*, 247.

(111) Morris, G. M.; Huey, R.; Olson, A. J. *Current Protocols in Bioinformatics*; John Wiley and Sons: New York, 2008; Chapter 8, Unit 8, p 14.

(112) Goodsell, D. S.; Morris, G. M.; Olson, A. J. *J. Mol. Recognit.* **1996**, *9*, 1.

(113) Morris, G. M.; Goodsell, D. S.; Huey, R.; Olson, A. J. *J. Comput.-Aided Mol. Des.* **1996**, *10*, 293.

(114) Berendsen, H. J. C.; van, d. S. D.; van, D. R. *Comput. Phys. Commun.* **1995**, *91*, 43.

(115) Hess, B.; Kutzner, C.; van, d. S. D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435.

(116) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. *J. Comput. Chem.* **2005**, *26*, 1701.

(117) van der Spoel, D.; Lindahl, E.; Hess, B.; Kutzner, C.; van Buuren, A. R.; Apol, E.; Meulenhoff, P. J.; Tieleman, D. P.; Sijbers, A. L. T. M.; Feenstra, K. A.; Drunen, R. v.; Berendsen, H. J. C. *GROMACS User Manual Version 4.0*; The GROMACS development team: Groningen, The Netherlands, 2006.

(118) Zhang, W.; Hou, T.; Schafmeister, C.; Ross, W. S.; Case, D. A. LEaP. In *Amber Tools*, 1.2 ed.; Case, D. A., Ed.; University of California, San Francisco: San Francisco, CA, 2009.

(119) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. *J. Comput. Chem.* **1997**, *18*, 1463.

(120) Cerutti, D. S.; Duke, R.; Freddolino, P. L.; Fan, H.; Lybrand, T. P. *J. Chem. Theory Comput.* **2008**, *4*, 1669.

(121) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187.

(122) Boczko, E. M.; Brooks, C. L., III. *J. Phys. Chem.* **1993**, *97*, 4509.

(123) Cieplak, P.; Caldwell, J.; Kollman, P. *J. Comput. Chem.* **2001**, *22*, 1048.

(124) Mezei, M.; Harrison, S. W.; Ravishanker, G.; Beveridge, D. L. *Isr. J. Chem.* **1986**, *27*, 163.

(125) Spector, T. I.; Kollman, P. A. *J. Phys. Chem. B* **1998**, *102*, 4004.

(126) Rick, S. W.; Berne, B. J. *J. Am. Chem. Soc.* **1996**, *118*, 672.

(127) Jorgensen, W. L.; Gao, J. *J. Am. Chem. Soc.* **1988**, *110*, 4212.

(128) Gerothanassis, I. P.; Demetropoulos, I. N.; Vakka, C. *Biopolymers* **1995**, *36*, 415.

(129) Exarchos, K. P.; Papaloukas, C.; Exarchos, T. P.; Troganis, A. N.; Fotiadis, D. I. *J. Biomed. Inf.* **2009**, *42*, 140.

(130) Troganis, A. N.; Sicilia, E.; Barbarossou, K.; Gerothanassis, I. P.; Russo, N. *J. Phys. Chem. A* **2005**, *109*, 11878.

CT1003302