JCTC Journal of Chemical Theory and Computation

An Integrated Path Integral and Free-Energy Perturbation−Umbrella Sampling Method for Computing Kinetic Isotope Effects of Chemical Reactions in Solution and in Enzymes

Dan Thomas Major[†] and Jiali Gao*[,†,‡]

Department of Chemistry and Supercomputing Institute, Digital Technology Center, University of Minnesota, Minneapolis, Minnesota 55455, and Centro Nacional de Supercomputación, Programa Biología Computacional, C/ Jordi Girona 29, 08034 Barcelona, Spain

Received December 19, 2006

Abstract: An integrated centroid path integral and free-energy perturbation−umbrella sampling (PI-FEP/UM) method for computing kinetic isotope effects (KIEs) for chemical reactions in solution and in enzymes is presented. The method is based on the bisection sampling in centroid path integral simulations to include nuclear quantum effects to the classical potential of mean force. The required accuracy for computing kinetic isotope effects is achieved by coupled free-energy perturbation and umbrella sampling for reactions involving different isotopes. The use of FEP with respect to different masses results in relatively small statistical uncertainties, whereas if KIEs are computed directly by the difference in free energies obtained from the quantum mechanical potentials of mean force for different isotopes, the statistical errors are significantly greater. The PI-FEP/UM method is illustrated in two applications. The first reaction is the decarboxylation of N-methyl picolinate in water, for which the primary ^{13}C and secondary ^{15}N KIEs have been determined. The second reaction is the proton-transfer reaction between nitroethane and an acetate ion in water. In both cases, the computational results are in accord with experimental data, and the findings provide further insight into the mechanism of these reactions in water.

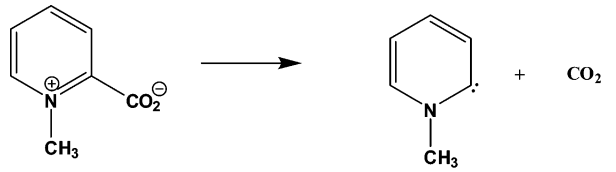
1. Introduction

Proton and hydride transfer reactions are ubiquitous in chemical and biological processes. Because of their relatively small mass, zero-point energy and tunneling contributions are significant in determining free-energy barriers.[1] Consequently, it is necessary to include nuclear quantum mechanical effects in computation of the rate constant.[2,3] The incorporation of nuclear quantum effects is also important in reactions involving heavy atom transfers since one of the most direct, experimental assessments of the transition state of a chemical reaction is through measurements of kinetic isotope effects (KIEs),[4] which are of quantum mechanical origin. This, however, is a great challenge to computation because the measured heavy atom KIEs are typically less than a few percent, which requires an accuracy of a fraction of 0.1 kcal/mol in the computed free-energy difference between isotope substitutions. The difficult task is further exacerbated by the need to sample configurational space of the macromolecular system to achieve statistical convergence.[5,6] Thus, it is highly desirable to develop accurate and practical methods for estimating kinetic isotope effects for chemical reactions in solution and in biological environments.[7] In this paper, we describe an integrated path integral and free-energy perturbation/umbrella sampling (PI-FEP/UM) approach in molecular dynamics simulations using a combined quantum mechanical and molecular mechanical

* Corresponding author e-mail: gao@chem.umn.edu.
[†] University of Minnesota.
[‡] Centro Nacional de Supercomputación.

10.1021/ct600371k CCC: $37.00
Published on Web 04/17/2007

Scheme 1

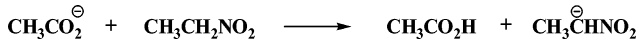


(QM/MM) potential for determining KIEs. In the integrated PI-FEP/UM method coupled with QM/MM potentials, both electronic and nuclear degrees of freedom in the reactive region are treated explicitly by quantum mechanical methods.

Of course, a variety of methods have been developed to treat nuclear quantum effects for gas-phase reactions.[8] In principle, these techniques can be directly extended to condensed-phase systems; however, the size and complexity of these systems unfortunately make it intractable computationally. Thus, a main goal is to develop new methods, or to extend gas-phase techniques to condensed phases or biomolecular systems. One method that has been successfully introduced to computational enzymology is the ensemble-averaged variational transition-state theory with QM/MM sampling (EA-VTST-QM/MM), which has been applied to a number of enzyme systems.[2,3,9−13] Both primary and secondary KIEs can be computed using the EA-VTST-QM/MM method, and the method includes contributions of multidimensional tunneling. In another work, a grid-based hybrid approach was used to model quantum effects in hydrogen transfer reactions by numerically solving the vibrational wavefunction of the transferring hydrogen nucleus.[14,15] The method yielded good results for primary KIEs in several hydride transfer reactions,[16] although the authors noted that a major limitation is its complexity, preventing it from extending to quantizing more than one particle.[17]

The third technique, which was among the first applications to incorporate nuclear quantum effects in enzyme reactions, is the quantized classical path (QCP) method developed by Hwang et al.[18−20] The discrete Feynman path integral method[21] has been used in a variety of applications since it offers an efficient and general approach for treating nuclear quantum effects in condensed-phase simulations.[20,22−36] Here, we focus on applications to biomolecular systems. The centroid path integral molecular dynamics method provides a procedure that can be used to directly estimate the quantum mechanical activation free energy.[28,37] However, these methods are computationally expensive for modeling large systems such as enzymes, especially if combined QM/MM potentials are used to represent the potential surface. The QCP method is also based on centroid path integral sampling, but it is formulated as a correction to the classical potential of mean force (PMF).[18,19,23] Thus, the classical simulations and quantum corrections are fully separated, making it particularly attractive and efficient for modeling enzymatic reactions. Unfortunately, the QCP method has not been widely used, and it is very difficult to obtain converged results using "standard" sampling schemes. Recently, we developed a practical procedure, called BQCP,[5,6] by extending the bisection sampling method developed by Ceperley and Pollock[38] for free-particle sampling to centroid path integral simulations. This has enabled us to obtain converged results through a series of validation studies.[5,6] The remarkable free-particle sampling efficiency of the bisection scheme in centroid path integral simulations stems from the fact that every "Monte Carlo" move is accepted and is independent of the previous beads' distribution, and it has been thoroughly discussed by Ceperley.[39] The BQCP method[5,6] has been applied to a number of proton and hydride transfer reactions in solution and in enzymes.[6,34−36] The computed KIEs are in good accord with experiments and with results from the previous EA-VTST-QM/MM method. The BQCP method has now been adopted in another QCP application.[17]

Scheme 2



Although the methods discussed above are based on very different theories, a common strategy is to estimate approximately the quantum mechanical rate constant by introducing a quantum correction factor to bridge the classical transition state theory. We define

$$k \equiv k_{\mathrm{qm}} = \gamma k_{\mathrm{TST}} \tag{1}$$

where k_{TST} is the transition-state theory rate constant and γ is the generalized transmission coefficient,[2] which includes the classical dynamic recrossing factor, Γ, and the quantum correction factor κ, which is defined as follows:

$$\kappa = \frac{k_{\mathrm{qm}}}{k_{\mathrm{TST}}} = \mathrm{e}^{-\beta(\Delta F^{\neq}_{\mathrm{qm}} - \Delta F^{\neq}_{\mathrm{TST}})} \tag{2}$$

In eq 2, $\beta = 1/k_{\mathrm{B}}T$ with k_{B} being Boltzmann's constant and T the temperature, and $\Delta F^{\neq}_{\mathrm{qm}}$ and $\Delta F^{\neq}_{\mathrm{TST}}$ are, respectively, the quantum and classical free energy of activation. Here, we have implicitly assumed that the classical and quantum Γ factors are identical. The different methods applied to enzymatic reactions to incorporate nuclear quantum effects differ in the specific approximations to estimate the free-energy difference in eq 2.

In this article, we develop an efficient sampling strategy for accurate computation of KIEs for chemical reactions in solutions and in enzymes. The method and computational details are illustrated by two systems, while the computational study also provides insights into the interpretation of the observed KIEs. The first reaction is the decarboxylation of N-methyl picolinate in water, a model for the reaction catalyzed by orotidine 5′-monophosphate decarboxylase (Scheme 1). The second is the deprotonation of nitroethane by an acetate ion, a model for the nitroalkane oxidase reaction in the first step in the oxidation of nitroalkanes to aldehydes and ketones (Scheme 2). In the former reaction, heavy-atom primary ^{13}C and secondary ^{15}N KIEs are computed, while for the latter, ^{2}H KIE is determined.

2. Theory and Method

2.1. Centroid Path Integral Method. In the discrete path integral method, each quantized nucleus is represented by a ring of P quasi-particles called beads, whose coordinates are

denoted as $\mathbf{r} = \{\mathbf{r}_i;\ i = 1, ..., P\}$.[21] The discrete paths are circular with $\mathbf{r}_{P+1} = \mathbf{r}_1$. For convenience, we limit the discussion to a single quantized atom embedded in a classical solvent while extension to many-quantized particles is obvious. Each bead is connected to its two neighbors via harmonic springs and is subjected to a fraction, $1/P$, of the full classical potential, $U(\mathbf{r}_i, \mathbf{S})$, where $\mathbf{S}$ represents all classical solvent coordinates. In the centroid path integral, the centroid position, $\bar{\mathbf{r}}$, is used as the principle variable, and the canonical QM partition function of the hybrid system can be written as follows:[21]

$$Q_P^{\mathrm{qm}} = \int \mathrm{d}\mathbf{S} \int \mathrm{d}\mathbf{s} \left(\frac{P}{2\pi\lambda^2}\right)^{3P/2} \int \mathrm{d}\mathbf{R}\, \mathrm{e}^{-\beta V^{\mathrm{qm}}(\mathbf{r},\mathbf{S})} \tag{3}$$

where $\int \mathrm{d}\mathbf{R} = \int \mathrm{d}\mathbf{r}_1 ... \int \mathrm{d}\mathbf{r}_P \delta(\bar{\mathbf{r}} = \mathbf{s})$, P is the number of quasi-particles of the discrete path, the delta function $\delta(\bar{\mathbf{r}} = \mathbf{s})$ is introduced for use in later discussion, and the centroid coordinate, $\bar{\mathbf{r}}$, of the quasi-particles, $\mathbf{r} = \{\mathbf{r}_i;\ i = 1, ..., P\}$, is defined as

$$\bar{\mathbf{r}} = \frac{1}{P}\sum_{i=1}^{P} \mathbf{r}_i \tag{4}$$

In eq 3, the effective quantum mechanical potential $V^{\mathrm{qm}}(\mathbf{r},\mathbf{S})$ is given by

$$V^{\mathrm{qm}}(\mathbf{r},\mathbf{S}) = \frac{P}{2\beta\lambda^2}\sum_{i}^{P} (\mathbf{r}_i - \mathbf{r}_{i+1})^2 + \frac{1}{P}\sum_{i}^{P} U(\mathbf{r}_i,\mathbf{S}) \tag{5}$$

and λ^2 is the square of the de Broglie thermal wavelength of a particle of mass M:

$$\lambda^2 = \frac{\beta\hbar^2}{M} \tag{6}$$

where $\hbar$ is Planck's constant and M is the mass of the particle.

We introduce the effective semiclassical potential $U^{\mathrm{eff}}(\mathbf{s};\ \mathbf{r},\mathbf{S})$

$$U^{\mathrm{eff}}(\mathbf{s} = \bar{\mathbf{r}};\ \mathbf{r},\mathbf{S}) = \frac{P}{2\beta\lambda^2}\sum_{i}^{P} (\mathbf{r}_i - \mathbf{r}_{i+1})^2 + U(\bar{\mathbf{r}},\mathbf{S}) \tag{7}$$

where the centroid of the quantized particle is constrained at its classical position, $\mathbf{s}$, and they are used interchangeably below. The classical partition function Q_P^{cm} (cm is classical mechanics) is

$$Q_P^{\mathrm{cm}} = \int \mathrm{d}\mathbf{S} \int \mathrm{d}\mathbf{s}\, \mathrm{e}^{-\beta U(\mathbf{s},\mathbf{S})} \left(\frac{P}{2\pi\lambda^2}\right)^{3P/2} \times \int \mathrm{d}\mathbf{R} \exp\left[-\frac{P}{2\lambda^2}\sum_{i}^{P} (\Delta\mathbf{r}_i)^2\right]$$
$$= \mathrm{e}^{-\beta F^{\mathrm{cm}}}\, \mathrm{e}^{-\beta F_{\mathrm{FP}}^{\circ}} \tag{8}$$

where $\mathbf{s}$ is the classical position vector of the quantized particle, F^{cm} and F_{FP}° are the free energies of the classical system [without the quantized free particle (FP)] and the free particle, respectively, and $\Delta\mathbf{r}_i = \mathbf{r}_i - \mathbf{r}_{i+1}$. Then, eq 3 can be rewritten as follows

$$Q_P^{\mathrm{qm}} = Q_P^{\mathrm{cm}} \times \frac{\int \mathrm{d}\mathbf{S} \int \mathrm{d}\mathbf{s}\, \mathrm{e}^{-\beta U(\mathbf{s},\mathbf{S})} \int \mathrm{d}\mathbf{R} \exp\left[-\frac{P}{2\lambda^2}\sum_{i}^{P} (\Delta\mathbf{r}_i)^2\right] \mathrm{e}^{-\beta\Delta\bar{U}(\bar{\mathbf{r}}=\mathbf{s},\mathbf{S})}}{\int \mathrm{d}\mathbf{S} \int \mathrm{d}\mathbf{s}\, \mathrm{e}^{-\beta U(\mathbf{s},\mathbf{S})} \int \mathrm{d}\mathbf{R} \exp\left[-\frac{P}{2\lambda^2}\sum_{i}^{P} (\Delta\mathbf{r}_i)^2\right]}$$
$$= \mathrm{e}^{-\beta[F^{\mathrm{cm}} + F_{\mathrm{FP}}^{\circ}]} \langle \mathrm{e}^{-\beta[F(\bar{\mathbf{r}}=\mathbf{s},\mathbf{S}) - F_{\mathrm{FP}}^{\circ}]} \rangle_U \tag{9}$$

In eq 9, the average $\langle ... \rangle_U$ is obtained according the potential $U(\bar{\mathbf{r}},\mathbf{S})$ over classical coordinates, and

$$\Delta\bar{U}(\bar{\mathbf{r}},\mathbf{S}) = \frac{1}{P}\sum_{i}^{P} \{U(\mathbf{r}_i,\mathbf{S}) - U(\bar{\mathbf{r}},\mathbf{S})\} \tag{10}$$

Equation 9 also defines that, for a fixed "classical" configuration $(\bar{\mathbf{r}},\mathbf{S})$, the free-particle sampling carried out without the external potential $U(\bar{\mathbf{r}},\mathbf{S})$ yields the partial partition function

$$Q_P^{\mathrm{qm}}(\bar{\mathbf{r}},\mathbf{S}) = \mathrm{e}^{-\beta[F(\bar{\mathbf{r}},\mathbf{S}) - F_{\mathrm{FP}}^{\circ}]} = \langle \mathrm{e}^{-\beta\Delta\bar{U}(\bar{\mathbf{r}},\mathbf{S})} \rangle_{FP,\bar{\mathbf{r}}} \tag{11}$$

where $F(\bar{\mathbf{r}},\mathbf{S})$ is the free energy of a quantized particle whose centroid position is constrained to its classical coordinate, $\bar{\mathbf{r}} = \mathbf{s}$, in the presence of the rest of the classical particles, and it is related to the free-particle averaging which is defined as

$$\langle ... \rangle_{\mathrm{FP},\bar{\mathbf{r}}} = \frac{\int \mathrm{d}\mathbf{R}\, \{\cdots\}\, \delta(\bar{\mathbf{r}}) \exp\left[-\frac{P}{2\lambda^2}\sum_{i}^{P} (\Delta\mathbf{r}_i)^2\right]}{\int \mathrm{d}\mathbf{R}\, \delta(\bar{\mathbf{r}}) \exp\left[-\frac{P}{2\lambda^2}\sum_{i}^{P} (\Delta\mathbf{r}_i)^2\right]} \tag{12}$$

Thus, the free-particle sampling of eqs 11 and 12 yields the quantum free-energy difference relative to a reference free particle for a fixed classical configuration. The idea of using classical Monte Carlo simulations to generate particle distributions and then to use eq 12 to make quantum corrections was described by Sprik et al.[23] Later, Warshel et al. pointed out that the expression of eq 9 is particularly useful since the quantum free energy of the system can be obtained by first carrying out classical trajectories for averaging classical configurations $(\bar{\mathbf{r}},\mathbf{S})$, then by determining the quantum contributions through free-particle sampling (eq 11).

The quantum mechanical average of the ground-state property A in the centroid path integral can be expressed as

$$\langle A \rangle = \frac{\int \mathrm{d}\mathbf{S} \int \mathrm{d}\mathbf{s} \int \mathrm{d}\boldsymbol{R} A(\mathbf{r},\mathbf{S})\, \mathrm{e}^{-\beta V^{\mathrm{qm}}(\mathbf{r},\mathbf{S})}}{\int \mathrm{d}\mathbf{S} \int \mathrm{d}\mathbf{s} \int \mathrm{d}\mathbf{R}\, \mathrm{e}^{-\beta V^{\mathrm{qm}}(\mathbf{r},\mathbf{S})}} \tag{13}$$

where the integral over $\mathrm{d}\mathbf{R}$ is under the constraint that $\bar{\mathbf{r}} = \mathbf{s}$. In analogy with the derivation of eq 9, eq 13 can be simplified as follows:

$$\langle A \rangle = \frac{\langle \langle A(\mathbf{r},\mathbf{S})\, \mathrm{e}^{-\beta\Delta\bar{U}(\bar{\mathbf{r}},\mathbf{S})} \rangle_{\mathrm{FP},\bar{\mathbf{r}}} \rangle_U}{\langle \langle \mathrm{e}^{-\beta\Delta\bar{U}(\bar{\mathbf{r}},\mathbf{S})} \rangle_{\mathrm{FP},\bar{\mathbf{r}}} \rangle_U} \tag{14}$$

Equation 14 represents a weighted Boltzmann average of all path integrals and classical configurations. If one makes the approximation

$$\langle A(\mathbf{r},\mathbf{S})\, e^{-\beta\Delta\bar{U}(\bar{\mathbf{r}},\mathbf{S})}\rangle_{\mathrm{FP},\bar{\mathbf{r}}} \approx e^{-\beta[F(\bar{\mathbf{r}},\mathbf{S})-F^{\circ}_{\mathrm{FP}}]}\langle A(\mathbf{r},\mathbf{S})\rangle_{\mathrm{FP},\bar{\mathbf{r}}} \quad (15)$$

the average can be estimated as a Boltzmann weighted average by the individual quantum free energies of classical configurations:

$$\langle A\rangle \approx \sum_{\bar{\mathbf{r}},\mathbf{S}} \frac{e^{-\beta[F(\bar{\mathbf{r}},\mathbf{S})-F^{\circ}_{\mathrm{FP}}]}}{Q_P^{\mathrm{qm}}}\langle A(\mathbf{r},\mathbf{S})\rangle_{\mathrm{FP},\bar{\mathbf{r}}} \quad (16)$$

Note that $\bar{\mathbf{r}} = \mathbf{s}$. Equation 16 has been shown to yield reasonable results for a system consisting of one electron embedded in random hard spheres.[23] If a sampling procedure can be formulated such that the classical configurations are chosen with the probability of $e^{-\beta[F(\bar{\mathbf{r}},\mathbf{S})-F^{\circ}_{\mathrm{FP}}]}/Q_P^{\mathrm{qm}}$, eq 16 may be enumerated by a simple numerical average,[23] but it is not clear how such a sampling scheme can be easily obtained.

$$\langle A\rangle \approx \sum_{\bar{\mathbf{r}},\mathbf{S}} \langle A(\mathbf{r},\mathbf{S})\rangle_{\mathrm{FP},\bar{\mathbf{r}}} \quad (16a)$$

2.2. The Quantized Classical Path (QCP) Method. Warshel et al. pointed out that the most significant result of the QCP method is the use of double averaging[23] in centroid path integral calculations over all configurations generated by classical trajectories.[18,19] From eq 9, the quantum mechanical potential of mean force, defined as a function of the centroid reaction coordinate, $\bar{z}$, can readily be expressed as follows:

$$\begin{aligned} W^{\mathrm{qm}}(\bar{z}) &= W^{\mathrm{cm}}(\bar{z}) - \frac{1}{\beta}\ln\frac{Q_P^{\mathrm{qm}}(\bar{z})}{Q^{\mathrm{cm}}(\bar{z})} \\ &= W^{\mathrm{cm}}(\bar{z}) - \frac{1}{\beta}\ln\langle\langle e^{-\beta\Delta\bar{U}(\bar{z})}\rangle_{\mathrm{FP},\bar{z}}\rangle_U \end{aligned} \quad (17)$$

where Q^{cm} is the cm partition function (without quantized free particles), and $W^{\mathrm{qm}}(\bar{z})$ and $W^{\mathrm{cm}}(z)$ are the centroid quantum mechanical and classical mechanical potentials of mean force, respectively. The advantage of this formulation is that one can sample the FP distribution separately at each classical configuration (i.e., centroid position) and then average over classical configurations obtained from molecular dynamics simulations.

The free-energy difference in eq 2 can then be conveniently obtained from the formula

$$\Delta F^{\neq}_{\mathrm{qm}} - \Delta F^{\neq}_{\mathrm{TST}} = [W^{\mathrm{qm}}(\bar{z}^{\neq}_{\mathrm{qm}}) - W^{\mathrm{qm}}(\bar{z}^{\mathrm{R}}_{\mathrm{qm}})] - [W^{\mathrm{cm}}(z^{\neq}_{\mathrm{cm}}) - W^{\mathrm{cm}}(z^{\mathrm{R}}_{\mathrm{cm}})] \quad (18)$$

where the symbol $\bar{z}^{\neq}_{\mathrm{qm}}$ specifies the value of the centroid reaction coordinate, at which $W^{\mathrm{qm}}(\bar{z})$ has the maximum value, and $\bar{z}^{\mathrm{R}}_{\mathrm{qm}}$ is the coordinate at the reactant state. Notice that the free energy of the free particle cancels out in eq 18; thus, it does not contribute to the calculation of the free-energy difference in eqs 2 and 18. Analogously, $z^{\neq}_{\mathrm{cm}}$ and $z^{\mathrm{R}}_{\mathrm{cm}}$ are the corresponding values of the classical reaction coordinate. Of course, the locations of the transition state in the quantum and classical potential of mean force are not necessarily identical.

2.3. The BQCP Method. A central issue with all path-integral formulations is the sampling method employed. A number of efficient sampling schemes have been proposed and reviewed.[23,24,38−40] However, we had considerable difficulties achieving convergence in QCP calculations using "standard" Monte Carlo or molecular dynamics methods. Recently, we implemented the bisection method of Ceperley and Pollock to sample the free-particle distribution,[38,39] which is based on multilevel sampling and the Lévy Brownian bridge construction,[41] and we found that the convergence in eq 11 can be easily obtained.[5,6] Since the free-particle distribution is known exactly at a given temperature, each ring-bead distribution is generated directly according to this distribution and thus 100% accepted. Furthermore, each new configuration is created independently, starting from a single initial bead position, allowing the new configuration to move into a completely different region of configurational space.

The original bisection sampling method was not developed for centroid path integral simulations, and there is no simple way of constraining the centroid position in the bisection sampling. In our implementation, we decided to first make the bisection sampling as originally proposed by Ceperley and Pollock[38,39] and enforced the first and last beads to be identical to enclose the polymer ring.[5,6] Then, we made a rigid-body translation of the centroid position of the new beads' configuration to coincide with the target (classical) coordinate. Since the free-particle distribution is known exactly at a given temperature, each ring-bead distribution is generated according to this distribution and thus 100% accepted.[39] Furthermore, in this construction, each new configuration is created independently, starting from a single initial bead position, allowing the new configuration to move into a completely different region of configurational space. We call this sampling scheme over classical configurations, in connection with the ideas of Sprik et al.[23] and Warshel et al.'s quantized classical path,[18,19] the BQCP method. The BQCP convergence has been thoroughly tested[5,6] and applied to several condensed-phase and enzyme systems.[6,34−36] Of course, the present sampling approach is not restricted to fixing the centroid position to correct the classical reaction path, and it can be used as an independent variable to obtain a "quantum" reaction path. Although this is a straightforward extension of Ceperley's bisection sampling method,[39] to our knowledge, it has not been utilized previously in centroid path integral simulations, and more importantly, it allows us to achieve fast convergence.

To make convenient the discussion of the kinetic isotope effect calculations, we first describe the bisection sampling procedure and its scalability to an isotopic substitution of the quantized particle. The goal is to develop an efficient sampling scheme that facilitates an accurate prediction of KIEs. Specifically, since we're dealing with free-particle sampling, we would like to generate entirely

uncorrelated distributions at each sampling step, to obtain maximum efficiency, while minimizing sampling noise.

Let the initial distribution of beads be $\{\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_P, \mathbf{r}_{P+1}\}$ where $P = 2^L$ is the number of beads and L is the number of bisection levels, and $\mathbf{r}_1$ and $\mathbf{r}_{P+1}$ are the initial and final positions of the polymer chain. Note that the ends of the sequence in the present discussion are enclosed for a ring of beads and $\mathbf{r}_1 = \mathbf{r}_{P+1}$. To obtain a new distribution, the positions for $P - 1$ beads, $\{\mathbf{r}_2, ..., \mathbf{r}_P\}$, will be randomly generated, starting from $\mathbf{r}_1$ and $\mathbf{r}_{P+1}$. Specifically, the bisection sampling procedure begins at the coarsest level, level 1, and the coordinate of the bead in the midpoint of the sequence is first sampled, which is placed at the geometrical center of the two ends plus a random displacement according to its (Gaussian) free-particle distribution of width $\sigma_1 = 2^{L-1}(\lambda_M^2/2P)$, where λ_M^2 is the square of de Broglie wavelength of a particle of mass M. Next, at level 2, the two new intervals $(\mathbf{r}_1, \mathbf{r}_{P/2})$ and $(\mathbf{r}_{P/2}, \mathbf{r}_{P+1})$ will be used to generate the coordinates of their two midpoints, with a distribution width of $\sigma_2 = 2^{L-2}(\lambda_M^2/2P)$. This process continues until all $P - 1$ positions are generated at level L. For convenience, which is obvious below, we set the local coordinate origin at $\mathbf{r}_1 = (0,0,0)$, and we summarize the procedure below.

Level 1:

$$\mathbf{r}_{1+P/2} = \frac{1}{2}(\mathbf{r}_1 + \mathbf{r}_{P+1}) + \lambda_M \sqrt{\frac{2^{L-2}}{P}}\eta_1 = \lambda_M \theta_{1+P/2} \tag{19}$$

where η_1 is a vector of normal distribution with zero mean and unit variance (which is obtained with the Box-Muller method[42]), and $\theta_{1+P/2} = \sqrt{2^{L-2}/P}\eta_1$.

Level 2:

$$\begin{aligned}\mathbf{r}_{1+P/4} &= \frac{1}{2}(\mathbf{r}_1 + \mathbf{r}_{1+P/2}) + \lambda_M \sqrt{\frac{2^{L-3}}{P}}\eta_2(1) \\ &= \lambda_M\left\{\frac{1}{2}\theta_{1+P/2} + \sqrt{\frac{2^{L-3}}{P}}\eta_2(1)\right\} = \lambda_M \theta_{1+P/4}\end{aligned} \tag{20}$$

$$\begin{aligned}\mathbf{r}_{1+3P/4} &= \frac{1}{2}(\mathbf{r}_{1+P/2} + \mathbf{r}_{P+1}) + \lambda_M \sqrt{\frac{2^{L-3}}{P}}\eta_2(2) \\ &= \lambda_M\left\{\frac{1}{2}\theta_{1+P/2} + \sqrt{\frac{2^{L-3}}{P}}\eta_2(2)\right\} = \lambda_M \theta_{1+3P/4}\end{aligned} \tag{21}$$

where $\eta_2(1)$ and $\eta_2(2)$ are two random vectors of normal distribution.

Level k:

$$\begin{aligned}&\mathbf{r}_{1+(2^m-1)P/2^k} \\ &= \frac{1}{2}(\mathbf{r}_{1+(2^m-2)P/2^k} + \mathbf{r}_{1+2^mP/2^k}) + \lambda_M\sqrt{\frac{2^{L-1-k}}{P}}\eta_k(m) \\ &= \lambda_M\left\{\frac{1}{2}[\theta_{1+(2^m-2)P/2^k} + \theta_{1+2^mP/2^k}] + \sqrt{\frac{2^{L-1-k}}{P}}\eta_k(m)\right\} \\ &= \lambda_M \theta_{1+(2^m-1)P/2^k}\end{aligned} \tag{22}$$

where $m = 1, ..., 2^{k-1}$.

In general, any particle position can be expressed as

$$\mathbf{r}_i = \lambda_M \theta_i \qquad i = 1, 2, ..., P \tag{23}$$

where the vector θ_i is a random vector, properly scaled and generated according to the free-particle distribution, and associated with earlier levels of bisection sampling, as has been explicitly shown for levels 1, 2, and k in eqs 19−22. In eq 23, we have defined $\theta_1 = 0$. Finally, all beads' positions are translated such that the geometric center (centroid) of the polymer ring $\mathbf{r}_{CP} = \lambda_M(1/P \sum_i^P \theta_i)$ is made to coincide with the coordinate of the classical particle.

A key point of interest is that each of the random bead positions generated by the bisection method is proportional to the de Broglie wavelength, which is dependent on the mass of the quantized particle. We also see that the positions of the $P - 1$ sampled beads are independent of the initial position of bead number 1. Moreover, the new position of the "seed" bead, bead number 1, is also independent of the position at the previous step. This would not be the case if we perform the bisection sampling with two different endpoints. In essence, one can simply generate a library of random beads' configurations and randomly choose them in BQCP calculations, but the time needed to generate new bead positions is minimal compared to the computer time required to compute their potential energies. So, there is little advantage for creating such a library. These features make the present procedure especially efficient in centroid path integral simulations. Note that, if external potentials are applied in a Monte Carlo sampling scheme for free-particle configurations, the efficiency of the bisection sampling scheme will be significantly hampered because the beads' distribution can easily get trapped in a local minimum, making it difficult to move out.

2.4. The Integrated Path Integral−Free-Energy Perturbation/Umbrella Sampling (PI-FEP/UM) Method for Computing KIEs. In principle, one can carry out separate centroid path integral simulations to make quantum mechanical corrections to the classical potential of mean force for different isotopes. Then, one can use the activation free energies for different isotopic reactions to compute the corresponding kinetic isotope effects. However, the statistical errors associated with these separate calculations are at least 1 order of magnitude greater than the free-energy difference for different isotopic reactions, too large to be useful for computing KIEs.

We have developed a free-energy perturbation (FEP) method, in the framework of centroid path integral simulations, to obtain directly the free-energy difference from isotope substitutions along a reaction path. Two algorithms are examined below, termed the PI-UM method and the PI-FEP/UM method. Although the two methods yield similar computed KIEs, the statistical uncertainties are significantly different, with the latter PI-FEP/UM approach giving smaller uncertainties comparable to experimental results due to the use of FEP with respect to bead distribution, which is proportional to the mass of the quantized particle. Thus, the method is an FEP over mass.

Considering an atom-transfer reaction where the light atom of mass M_L is replaced by a heavier isotope of mass M_H, we use exactly the same sequence of random numbers, that is, displacement vectors $\{\theta_i\}$, to generate the bisection path integral distribution for both isotopes. Thus, the resulting coordinates of these two bead distributions differ only by the ratio of the corresponding masses:

$$\frac{\mathbf{r}_{i,L}}{\mathbf{r}_{i,H}} = \frac{\lambda_{M_L}\theta_i}{\lambda_{M_H}\theta_i} = \sqrt{\frac{M_H}{M_L}} \qquad i = 1, 2, ..., P \tag{24}$$

where $\mathbf{r}_{i,L}$ and $\mathbf{r}_{i,H}$ are the coordinates for bead i of the corresponding light and heavy isotopes. The free-energy difference, or equivalently, the ratio of the partition functions, between the heavy and light isotopes can then be determined by the formula

$$\frac{Q_L^{qm}(\bar{z})}{Q_H^{qm}(\bar{z})} = \left\{\frac{Q_L^{qm}(\bar{z})}{Q^{cm}(\bar{z})}\right\}\left\{\frac{Q^{cm}(\bar{z})}{Q_H^{qm}(\bar{z})}\right\} = \frac{\langle\delta(z-\bar{z})\langle e^{-\beta\Delta\bar{U}_L}\rangle_{FP}\rangle_U}{\langle\delta(z-\bar{z})\langle e^{-\beta\Delta\bar{U}_H}\rangle_{FP}\rangle_U} \tag{25}$$

where $\Delta\bar{U}_L = (1/P)\sum_i\{U(\mathbf{r}_{i,L}) - U(\bar{\mathbf{r}})\}$ and $\Delta\bar{U}_H = (1/P)\sum_i\{U(\mathbf{r}_{i,H}) - U(\bar{\mathbf{r}})\}$.

Equation 25 offers a convenient approach to determine KIEs. We first carry out classical umbrella sampling simulations to generate the classical potential of mean force (outer average). Then, for each configuration saved in the classical simulation, we carry out centroid path integral calculations by free-particle sampling to obtain the quantum mechanical correction along the classical path (inner average). The bisection samplings are performed in such a way that the same set of displacement vectors, properly scaled by the mass factor of eq 24, are used to obtain the distributions for the two different isotopes. Thus, we obtain two separate quantized potentials of mean force for the two isotopic reactions, and we call this algorithm the path integral−umbrella sampling (PI-UM) method, in which umbrella sampling is used to compute the classical mechanical PMF and the centroid path integral sampling is used to make quantum corrections. The required accuracy is achieved because the isotopic substitution represents only a slight perturbation to the original isotopic bead distribution, proportional to the square root of the ratio between the two masses. This is different than separate PI samplings for different isotopes, which is very difficult to converge to the required precision.

The second algorithm, which we refer to here as the PI-FEP/UM method, is to use the result of eq 14 to obtained the free-energy difference using free-energy perturbation between light and heavy isotopic masses. First, we note that the sampling scheme presented in eqs 23 and 24 yields the following identity for the kinetic energy term of the discrete beads:

$$\frac{P}{2\beta\lambda_{M_L}^2}\sum_i^P(\mathbf{r}_{i,L} - \mathbf{r}_{i+1,L})^2 = \frac{P}{2\beta\lambda_{M_H}^2}\sum_i^P(\mathbf{r}_{i,H} - \mathbf{r}_{i+1,H})^2 \tag{26}$$

Therefore,

$$\frac{Q_H^{qm}(\bar{z})}{Q_L^{qm}(\bar{z})} = \left(\frac{M_H}{M_L}\right)^{3P/2}\frac{\int d\bar{\mathbf{r}}\int d\mathbf{R}_H\, e^{-\beta V^{qm}(\mathbf{r}_H)}}{\int d\bar{\mathbf{r}}\int d\mathbf{R}_L\, e^{-\beta V^{qm}(\mathbf{r}_L)}} = \frac{\int d\bar{\mathbf{r}}\int d\mathbf{R}_L\, e^{-\beta V^{qm}(\mathbf{r}_H)}}{\int d\bar{\mathbf{r}}\int d\mathbf{R}_L\, e^{-\beta V^{qm}(\mathbf{r}_L)}} \tag{27}$$

where we have substituted the integration variable, $d\mathbf{R}_H = (M_L/M_H)^{3P/2}\, d\mathbf{R}_L$. The significance of this coordinate change is that the ratio of the partition functions for the heavy and light isotopes at reaction coordinate z can be obtained exactly by free-energy perturbation (see also eq 14) via sampling only the light (or heavy) isotopic particles:

$$\frac{Q_H^{qm}(\bar{z})}{Q_L^{qm}(\bar{z})} = \frac{\left\langle\delta(z-\bar{z})\left\langle\exp\left[-\frac{\beta}{P}\sum_i\Delta U_i^{L\to H}\right]e^{-\beta\Delta\bar{U}_L}\right\rangle_{FP,L}\right\rangle_U}{\langle\delta(z-\bar{z})\, e^{-\beta[F_L(\bar{z},\mathbf{S})-F_{FP}^o]}\rangle_U} \tag{28}$$

where the superscripts or subscripts L and H specify computations done using light or heavy isotopes, and $\Delta U_i^{L\to H} = U(\mathbf{r}_{i,H}) - U(\mathbf{r}_{i,L})$ represents the difference in potential energy at the heavy and light bead positions $\mathbf{r}_{i,H}$ and $\mathbf{r}_{i,L}$. In eq 28, we obtain the free-energy (inner average) difference between the heavy and light isotopes by carrying out the bisection path integral sampling with the light atom and then perturbing the heavy isotope masses, i.e., positions according to eq 24. Then, the free-energy difference between the light and heavy isotope ensembles is weighted by the Boltzmann factor due to quantization of classical configurations (outer average).

An approximate approach is to employ eq 16, such that the free-energy difference between heavy and light isotopes is Boltzmann-weighted over all classical configurations:

$$\frac{Q_H^{qm}(\bar{z})}{Q_L^{qm}(\bar{z})} \approx \sum_{\mathbf{s},\mathbf{S}} P_{P,L}^{qm}(z{:}\bar{\mathbf{r}} = \mathbf{s},\mathbf{S})\left\langle\exp\left[-\frac{\beta}{P}\sum_i\Delta U_i^{L\to H}\right]\right\rangle_{(\bar{z})FP,L} \tag{29}$$

where

$$P_{P,L}^{qm}(z{:}\bar{\mathbf{r}} = \mathbf{s},\mathbf{S}) = \frac{Q_{P,L}^{qm}(z{:}\bar{\mathbf{r}},\mathbf{S})}{Q_{P,L}^{qm}} = \frac{e^{-\beta[F_{P,L}(z:\bar{\mathbf{r}},\mathbf{S})-F_{FP,L}^o]}}{e^{-\beta[F_{P,L}^{qm}-F_{FP,L}^o]}} \tag{30}$$

The kinetic isotope effects are computed as follows:

$$KIE = \frac{k^L}{k^H} = \left[\frac{Q_H^{qm}(\bar{z}_H^{\neq})}{Q_L^{qm}(\bar{z}_L^{\neq})}\right]\left[\frac{Q_L^{qm}(\bar{z}_L^R)}{Q_H^{qm}(\bar{z}_H^R)}\right] e^{-\beta[F_{cm}^R(\bar{z}_L^R)-F_{cm}^R(\bar{r}_H^R)]} \tag{31}$$

where $F_{cm}^R(\bar{z}_R^L)$ and $F_{cm}^R(\bar{z}_R^H)$ are the free energies of the mode in the reactant (R) state that correlates with the progress coordinate z for the light and heavy isotopes. A method for estimating their values has been described.[1]

3. Computational Details

3.1. Potential Energy Function. In the present study, we use a combined quantum mechanical and molecular me-

chanical potential in molecular dynamics simulations, in which the solute is represented explicitly by an electronic structure method and the solvent is approximated by the three-point charge TIP3P model for water.[43] The details have been described in a number of articles.[44,45] We used the standard semiempirical Austin Model 1 (AM1) method to treat the decarboxylation reaction of N-methyl picolinate.[46] As in numerous other decarboxylation reactions, the AM1 method yielded good energetic results in comparison with high-level electronic structural results.[47−51] In the case of the deprotonation of nitroethane by an acetate ion, the AM1 model failed to yield adequate energetic results, and a set of specific reaction parameters (SRP) was developed within the AM1 formalism to fit density functional theory results. The performance of the SRP-AM1 model has been reported previously,[6,34] and we focus here on testing the performance of the nuclear quantum effects using the PI-FEP method.

3.2. Simulation Details. All simulations are performed using periodic boundary conditions in the isothermal−isobaric (NPT) ensemble at 25 °C and 1 atm. In the decarboxylation reaction of N-methyl picolinate, a total of 888 water molecules were included in a cubic box of about $30 \times 30 \times 30$ Å^3. For the deprotonation of nitroethane by an acetate ion, 898 water molecules were included in the simulations. Nonbonded electrostatics were treated by the particle-mesh Ewald summation method for QM/MM simulations.[52] van der Waals interactions were smoothed to zero at 9.5 Å on the basis of group−group separations. The bond lengths and angles of solvent water molecules were constrained by the SHAKE algorithm,[53] and an integration step of 1 fs was used for all calculations.

The PMF profiles were obtained using the umbrella sampling technique.[54] In this approach, the reaction was divided into a series of "windows", in which a biasing potential was applied to allow sufficient sampling of high-energy regions along the reaction pathway. The effect of the biasing potential was subsequently removed when the separate simulation windows are combined to produce the overall PMF by using the weighted histogram analysis method.[55] In the current simulations, 12 windows were used for the decarboxylation reaction, while 32 windows were employed in the deprotonation reaction. In both reactions, the system was slowly heated to the target temperature over the course of 25 ps and thereafter equilibrated for 100−200 ps. Subsequently, each window was further equilibrated for 25−50 ps before data collection commenced. Each window was sampled for ca. 100−150 ps, totaling at least 2 ns of sampling for the decarboxylation and ca. 4 ns for the deprotonation reaction. The reaction coordinate was defined as the distance of the cleaving C−C bond for the decarboxylation reaction and as the difference between the breaking Cα−H and forming O−H bonds in the deprotonation reaction.

The BQCP simulations employed 97 800 classical configurations for each isotope (^{12}C, ^{13}C, ^{14}N, and ^{15}N) for the decarboxylation reaction, combined with 10 path-integral steps per classical step. For the deprotonation reaction, 112 807 classical configurations were used for each isotope (^{1}H and ^{2}H), combined with 10 path-integral steps per classical step. For the decarboxylation reaction, the isotope labeled atom was quantized in addition to the three neighboring atoms. For the deprotonation reaction, the nitroethane Cα atom, the abstracting acetate oxygen, the transferring proton, and the secondary hydrogen were quantized. Each quantized atom was described by 32 beads. Thus, 32−36 million QM/MM energy calculations are needed in each reaction. To estimate the standard errors in the computed kinetic isotope effects, the entire path integral simulations were divided into 10 separated blocks, each treated independently. The standard uncertainties ($\pm 1\sigma$) were determined from these 10 blocks and the total averages for both the PI-UM (eq 25) and the PI-FEP/UM (eq 31) methods.

All simulations employed the CHARMM program,[56] in which the methods described here have been implemented, and all path-integral simulations used a parallel version that efficiently distributes integral calculations for the quantized beads.[6]

4. Results and Discussions

4.1. Decarboxylation Reaction of N-Methyl Picolinate in Water. The primary and secondary heavy atom kinetic isotope effects for the decarboxylation of N-methyl picolinate have been studied by Rishavy and Cleland, and this was considered as a model to probe the mechanism of orotidine 5′-monophosphate decarboxylation catalyzed by orotidine monophosphate decarboxylase (Scheme 1).[47,57−59] Due to the small magnitude of heavy atom KIEs combined with the inherent statistical noise in condensed-phase simulations, it is very difficult to compute these quantities through computer simulations.

The reaction profile for the decarboxylation of N-methyl picolinate monotonically increases as a function of the cleaving C−C bond distance in the gas phase, reaching a plateau of 18.1 kcal/mol from MPW1PW91/6-311++(3df,2p)//MPW1PW91/6-31+G(d) calculations. The energy of reaction is 17.7 and 11.6 kcal/mol using MP2/6-311++-(3df,2p) and the semiempirical AM1 model, respectively. Thus, AM1 underestimates the reaction energy by about 6 kcal/mol. At all computational levels, there is no barrier separating N-methyl picolinate and the decarboxylation product, a characteristic feature in this class of decarboxylation reactions (refs 48−51, see above). The relatively smaller energy of reaction for N-methyl picolinate decarboxylation, compared to that of orotate ($\Delta E = 36.4$ kcal/mol),[59] is due to the positive charge at the N1 position of N-methyl picolinate. Since the main goal here is to illustrate the performance of the PI-FEP/UM method, the small difference between the AM1 and high-level electronic structure results is deemed acceptable.

The potential of mean force for the decarboxylation of N-methyl picolinate was obtained as a function of the C_2-C_{O2} separation in aqueous solution at 25 °C and 1 atm (see Figure 1). Solvent effects are significant, increasing the free-energy barrier by 15.2 kcal/mol to a value of 26.8 kcal/mol, which is accompanied by a net free energy of reaction of 24.7 kcal/mol. If the error introduced by the AM1 method for the gas-phase reaction is corrected, our best estimate of the free-energy barrier is 32.9 kcal/mol for the decarboxyla-

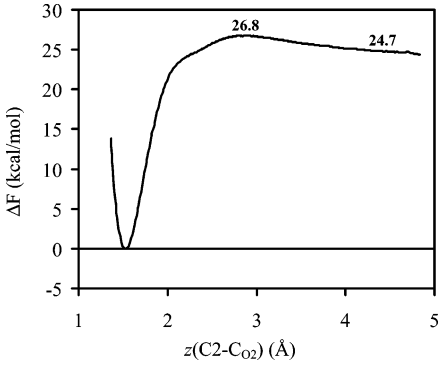


Figure 1. Classical potential of mean force as a function of the C_2-C_{O2} distance of the cleaving bond for the decarboxylation reaction of N-methyl picolinate in water.

tion reaction of N-methyl picolinate in water. This is in contrast to the decarboxylation of N-methyl orotate in water, which has little solvent effect.[47] The difference is again due to the presence of the positive charge in the pyridine ring, which is annihilated in the decarboxylation reaction. However, the similarity is that there is a small solvent reorganization barrier for CO_2 recombination, which is about 2 kcal/mol.[47]

The free-energy contributions to the potential of mean force by quantizing the carboxylate group and the C2 carbon in the pyridine ring are displayed in Figure 2A both for ^{12}C and ^{13}C at the carboxyl position, whereas the ratio of the two quantum partition functions from the PI-FEP/UM simulations is shown in Figure 2B. The nuclear quantum effects are non-negligible even for bond cleavage involving two carbon atoms, which reduce the free energy barrier by 0.45 kcal/mol. The intrinsic ^{13}C primary KIE has been determined by the formula of eq 25, along with a factor correcting for the reduced mass of the reaction coordinate, which is treated as a harmonic oscillator of masses 12 (C2) and 44 or 45 (CO_2). The PI-UM method yields a computed KIE of 1.0346 ± 0.8773 at 25 °C for the decarboxylation of N-methyl picolinate in water (Table 1). For comparison, the use of the PI-FEP/UM method resulted in a similar KIE (1.0345 ± 0.0028), but the associated statistical error is much smaller. To emphasize the sensitivity of the computational result, the computed KIE is equivalent to a free-energy difference of merely 0.0187 kcal/mol (Figure 2). Thus, even with the use of exactly the same configuration distributions, scaled by the corresponding masses of different isotopes, in path integral sampling, the fluctuation of the average quantum correction to the classical mechanical PMF is of similar magnitude as that of the computed KIE, if separated quantum mechanical PMFs are used (Figure 2A). On the other hand, the average over the exponential factor in free-energy perturbation simulations with the PI-FEP/UM method, which yields the free-energy difference between different isotopic substitutions, greatly reduces the statistical uncertainty. Experimentally, the primary ^{13}C KIE for the decarboxylation of N-methyl picolinate was determined to be 1.0212 ± 0.0002 in ethylene glycol at 120 °C, which was extrapolated to a value of 1.0281 ± 0.0003 at 25 °C.[57] The agreement between the computational result and experimental value is reasonable, suggesting that the decarboxylation

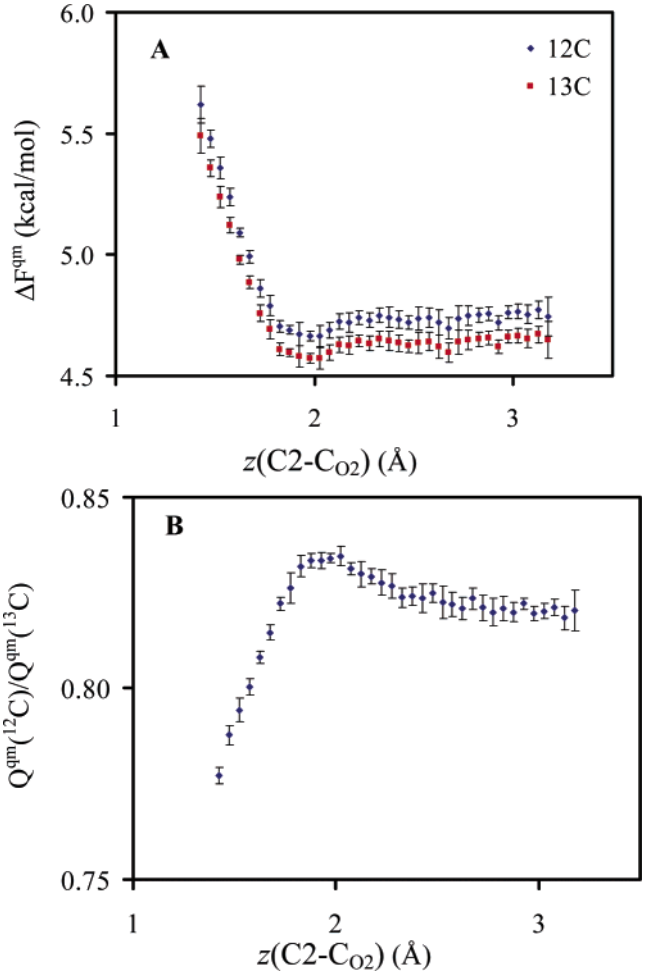


Figure 2. Nuclear quantum-mechanical free-energy corrections for the decarboxylation reaction of N-methyl picolinate in aqueous solution along the centroid reaction coordinate as defined in Figure 1. The C_2 atom of pyridinium and the carboxylate carbon and oxygen atoms are quantized. (A) Total free-energy contributions are displayed in blue for ^{12}C and in red for ^{13}C substitution at the carboxyl carbon position using the PI-UM algorithm. (B) The ratio of the quantum mechanical partition functions (equivalent to the free-energy difference) between ^{12}C and ^{13}C isotopic substitutions at the carboxyl carbon position from the PI-FEP/UM method.

Table 1. Computed and Experimental Primary $^{12}C/^{13}C$ and Secondary $^{14}N/^{15}N$ Kinetic Isotope Effects for the Decarboxylation of N-Methyl Picolinate at 25 °C in Water

	$^{12}k/^{13}k$	$^{14}k/^{15}k$
exptl (120 °C)[a]	1.0212 ± 0.0002	1.0053 ± 0.0002
exptl (25 °C)[a]	1.0281 ± 0.0003	1.0070 ± 0.0003
PI-UM	1.0346 ± 0.8773	1.0067 ± 0.8862
PI-FEP/UM	1.0345 ± 0.0028	1.0083 ± 0.0016

[a] In ethylene glycol with quinoline, ref 57.

reaction itself is the rate-limiting step. This is further supported by the secondary KIE discussed below.

Figure 3 depicts the nuclear quantum correction to the classical PMF by quantizing the N1 nitrogen and C2 and C6 carbons in the pyridine ring as well as the methyl carbon atom. Both ^{14}N and ^{15}N isotopes are used, and the ratio between their quantum mechanical distribution functions

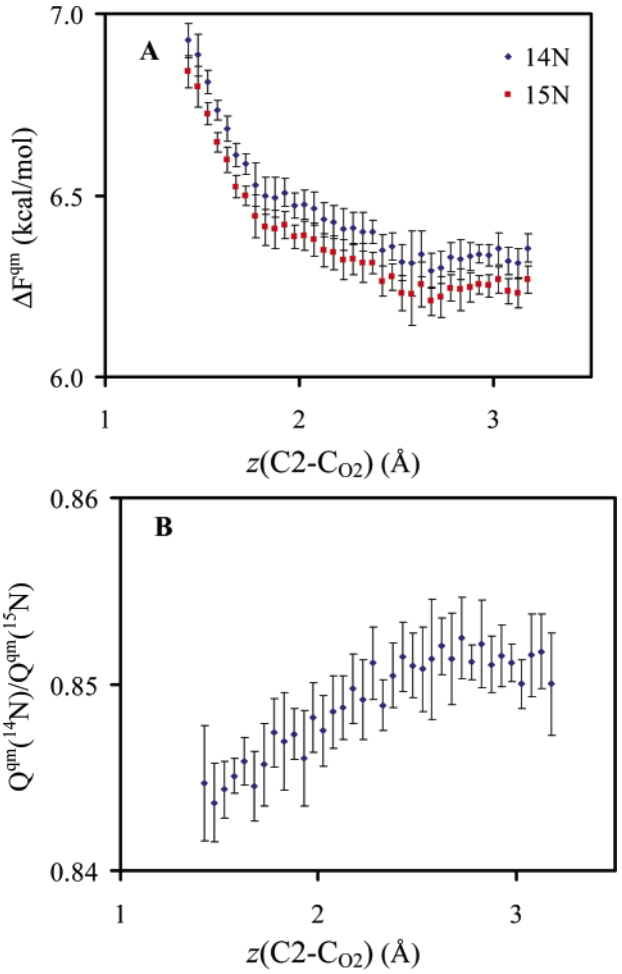


Figure 3. Nuclear quantum-mechanical free-energy corrections for the N-methyl picolinate decarboxylation reaction in aqueous solution along the centroid reaction coordinate as defined in Figure 1. Quantized atoms include the primary N_1 atom in pyridinium and the three adjacent (secondary) carbon atoms (the methyl and C_2 and C_6 carbon atoms). (A) Total free-energy contributions are shown in blue for ^{14}N and in red for ^{15}N substitution for the nitrogen atom from PI-UM calculations. (B) The ratio of the quantum partition functions (equivalent to the free-energy difference) between ^{14}N and ^{15}N isotopic substitutions for the nitrogen atom from PI-FEP/UM simulations.

along the centroid reaction coordinate is shown in Figure 3B. These calculations yield the secondary ^{15}N KIE, stemming from the loss of bending and torsional modes involving the carboxylate moiety at the transition state. Obviously, this is a very small secondary effect, and the standard fluctuations are too large when quantum corrections to the PMF for the two isotopes are used directly (Figure 3A). Fortunately, the present PI-FEP/UM method gives sufficiently stable and well-converged results as depicted in Figure 3B, and the computed secondary nitrogen KIE is computed as 1.0083 ± 0.0016 (Table 1). (The ratio of the reduced mass for the reaction coordinate for the sceondary isotope effects is assumed to be one). The corresponding experimental value is 1.0070 ± 0.0003 extrapolated to 25 °C.[57] The agreement between theory and experiment is also good, further supporting a unimolecular decarboxylation mechanism in this model reaction.

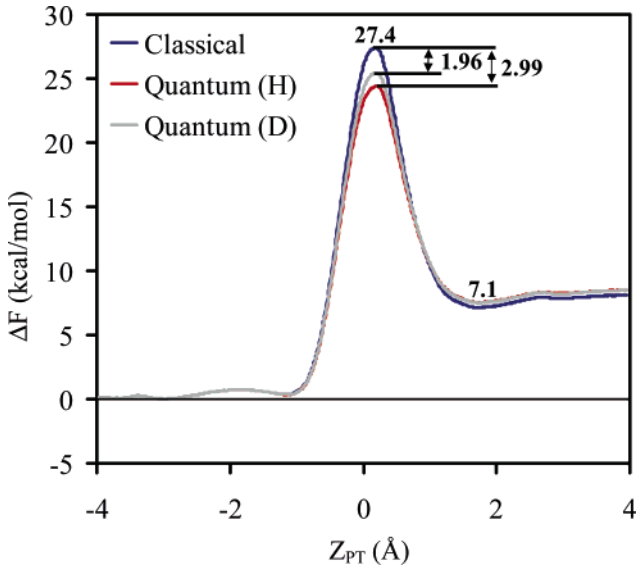


Figure 4. Classical and quantum potentials of mean force for the deprotonation of nitroethane by an acetate ion in aqueous solution. The reaction coordinate is defined as the difference of distances of the transferring proton between the donor atom (C_1) and acceptor atom (O), $Z_{PT} = r(C_1-H) - r(H-O)$.

4.2. Deprotonation of Nitroethane by Acetate Ion in Water. The initial step in the oxidation of nitroalkanes by nitroalkane oxidase (NAO) is the abstraction of a Cα proton of nitroalkane.[60,61] When the substrate is nitroethane, the deprotonation step is rate-limiting.[62] Analyses of the crystallographic structure and mutagenesis experiments suggest that Asp402 is the base responsible for the proton abstraction.[63,64] Recently, Valley and Fitzpatrick determined the KIEs for the dideuterated substrate [1,1-$^2H_2$]nitroethane in NAO, and its model reaction of proton/deuteron abstraction by an acetate ion.[65] In the NAO-catalyzed reaction, the KIE was found to be 9.2 ± 0.4, while in aqueous solution, the KIE was found to be 7.8 ± 0.1. Here, we report the KIE results for the proton abstraction of nitroethane by an acetate ion in aqueous solution.

In the present study, we employ a reparameterized semiempirical model using the AM1 formalism, and this SRP model was fitted to the Gaussian theory 3 (G3) results, and the details of the SRP parametrization procedure and its application to the NAO enzymatic reaction have been described previously.[6] The aim here is to examine our PI-FEP/UM method for KIE calculations, but we emphasize that the SRP model is a good one, and the computed energy of reaction for the proton abstraction of nitroethane by acetate is 8.7 kcal/mol, in adequate accord with the G3 result of 10.3 kcal/mol.[6,34] The classical PMF for the deprotonation of nitroethane by acetate is presented in Figure 4. The computed barrier is 27.4 kcal/mol, while the computed free energy of reaction is 7.1 kcal/mol, slightly higher than the experimental value of 5.8 kcal/mol obtained from the relevant experimental pK_a.[34]

Since light atoms are involved in the proton abstraction reaction, quantization of the "primary" and "secondary"

hydrogen atoms as well as the donor carbon atom and acceptor oxygen atom has a major impact on the computed free energy of activation, by lowering the barrier height by 3.0 kcal/mol. Thus, the estimated free energy of activation is 24.4 kcal/mol, in excellent agreement with the experimental value of 24.8 kcal/mol. This result underscores the importance of including nuclear quantum mechanical effects to accurately predict the activation free energy in reactions involving hydrogen transfers. We note that the overall barrier reduction is identical to that estimated previously using the EA-VTST-QM/MM method.[34] In addition, to dissect the specific contributing factors of the nuclear quantum effects, we have used the multidimensional tunneling algorithms developed by Truhlar and co-workers,[8] extended to enzyme applications,[1−3] to determine the average tunneling transmission factor, yielding a value of $\langle\kappa\rangle = 1.3$.[34] This suggests that that tunneling only makes minor contributions in the present case for the aqueous reaction. The PI-FEP/UM simulations include both zero-point and tunneling contributions that are not separable.

To directly compare the computed KIE with experimental results, which employed the dideuterated compound, 1,1-$^2H_2$[nitroethane], we also made the same substitutions in our computation at the Cα position. The experimental KIE of 7.8 corresponds to a free-energy difference between the two isotopic reactions of 1.21 kcal/mol. Employing the PI-FEP/UM method (Figure 5A), we obtain a computed value of 6.13 ± 0.68 (Table 2), somewhat smaller than the experimental result. However, this corresponds to a free-energy error of only 0.15 kcal/mol, or 12%. Using the more efficient algorithm, the PI-FEP/UM method (Figure 5B), we obtained a computed deuterium KIE of 8.50 ± 0.01 (Table 2). It is interesting to note that there is noticeable difference in the computed deuterium primary KIE between the two computational algorithms because the effects are much greater than the heavy isotope effects. Previously, we reported a study of the same reaction in water using the EA-VTST-QM/MM method, and a total KIE of 6.0 was obtained.[6,34]

5. Conclusions

We have described an integrated path integral and free-energy perturbation−umbrella sampling method for computing kinetic isotope effects for chemical reactions in solution and in enzymes. The approach is based on the BQCP sampling method, a combination of the bisection sampling scheme and centroid path integral simulations. The required accuracy for KIE calculations is achieved by the combined use of free-energy perturbation and umbrella sampling between different isotopic reactions. This is made possible because the bisection sampling scheme leads to a linear coordinate scaling relationship proportional to the ratio of the masses of different isotopes. Thus, the centroid path integral sampling for different quantized isotopic atoms can be similarly scaled, giving rise to small errors in free-energy perturbation/umbrella sampling simulations. The idea of making quantum mechanical corrections using the centroid path integral over configurations sampled by classical mechanical simulations has been explored previously, notably by Sprik et al. and the quantized classical path approach by

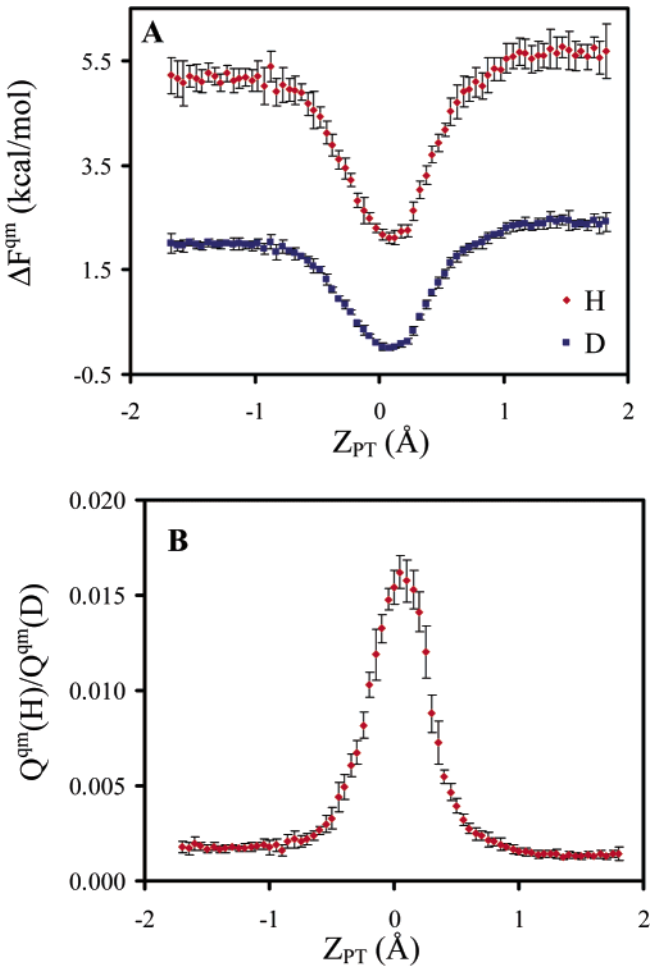


Figure 5. Nuclear quantum-mechanical free-energy correction for the deprotonation reaction of nitroethane by acetate in aqueous solution. The primary and secondary hydrogen atoms on C_1 carbon and the donor (C_1) and acceptor (O) atoms have been quantized. The reaction coordinate is defined in the legend of Figure 4. (A) The total free-energy contributions are shown in red for ^1H and in blue for ^2H substitutions, and (B) the ratio of the quantum-mechanical partition functions is displayed between hydrogen and deuterium substitutions at the C_1 carbon.

Table 2. Computed and Experimental $^1H/^2H$ Kinetic Isotope Effects for the Deprotonation of Nitroethane by Acetate at 25 °C in Water

	$^1k/^2k$
exptl[a]	7.8 ± 0.1
PI-UM	6.13 ± 0.68
PI-FEP/UM	8.50 ± 0.01

[a] 25 °C in aqueous solution, ref 63.

Warshel et al., which has been applied to several enzyme systems.[18−20,66] Here, we formulated two algorithms for computing KIEs for chemical reactions in condensed phases and in enzymes. We found that the PI-FEP/UM method yields relatively small statistical uncertainties in the computed KIEs, enabling the computed results to be compared with experimental results with reasonable confidence. However, when KIEs are obtained by taking the free-energy difference obtained from the quantum mechanical potentials of mean

force for the two isotopes, the statistical noise becomes too large to be comfortable; even the estimated KIEs are similar to those from the PI-FEP/UM calculations.

The PI-FEP/UM method is illustrated by computing the kinetic isotope effects for two chemical reactions in aqueous solution. In the first reaction, the decarboxylation of N-methyl picolinate involves heavy atom transfer reactions, and both the primary $^{12}C/^{13}C$ and the secondary $^{14}N/^{15}N$ KIEs are computed, and the results are found in good accord with experimental results. In particular, the computed primary and secondary KIEs are 1.0345 ± 0.0028 and 1.0083 ± 0.0016, and the corresponding experimental values are 1.0281 ± 0.0003 and 1.0070 ± 0.0003 at 25 °C. In the second application, the proton-transfer reaction between nitroethane and by an acetate ion, a prototypical model for the enzymatic process by nitroalkane oxidase, was found to lower the free-energy barrier from the classical simulations by 3 kcal/mol when nuclear quantum effects are included using the PI-FEP/UM method. The computed primary deuterium isotope effect is 8.50 ± 0.01, also in reasonable accord with experimental results (7.8 ± 0.1). The present study shows that the PI-FEP/UM method in connection with BQCP sampling allows computation of both primary and secondary kinetic isotope effects involving light as well as heavy atom transfers in solution. The method has also been applied to enzymatic reactions, providing a useful tool for the study of enzyme reaction mechanisms and for interpreting experimental KIEs.

Acknowledgment. This work has been supported in part by the National Institutes of Health and by the Army Research Laboratory through the Army High-Performance Computing Research Center (AHPCRC) under the auspices of Army Research Laboratory DAAD 19-01-2-0014. J.G. thanks the Ministerio de Educación y Ciencia, España, for partial support.

References

(1) Garcia-Viloca, M.; Alhambra, C.; Truhlar, D. G.; Gao, J. *J. Chem. Phys.* **2001**, *114*, 9953−9958.

(2) Garcia-Viloca, M.; Gao, J.; Karplus, M.; Truhlar, D. G. *Science* **2004**, *303*, 186−195.

(3) Pu, J.; Gao, J.; Truhlar, D. G. *Chem. Rev.* **2006**, *106*, 3140−3169.

(4) Cleland, W. W. *Arch. Biochem. Biophys.* **2005**, *433*, 2−12.

(5) Major, D. T.; Gao, J. *J. Mol. Graphics Modell.* **2005**, *24*, 121−127.

(6) Major, D. T.; Garcia-Viloca, M.; Gao, J. *J. Chem. Theory Comput.* **2006**, *2*, 236−245.

(7) Kohen, A.; Limbach, H.-H. *Isotope Effects in Chemistry and Biology*; Taylor and Francis: Boca Raton, FL, 2006.

(8) Fernández-Ramos, A.; Millar, J. A.; Klippenstein, S. J.; Truhlar, D. G. *Chem. Rev.* **2006**, *106*, 4518−4584.

(9) Alhambra, C.; Gao, J.; Corchado, J. C.; Villà, J.; Truhlar, D. G. *J. Am. Chem. Soc.* **1999**, *121*, 2253−2258.

(10) Alhambra, C.; Corchado, J. C.; Sánchez, M. L.; Gao, J.; Truhlar, D. G. *J. Am. Chem. Soc.* **2000**, *122*, 8197−8203.

(11) Alhambra, C.; Corchado, J.; Sanchez, M. L.; Garcia-Viloca, M.; Gao, J.; Truhlar, D. G. *J. Phys. Chem. B* **2001**, *105*, 11326−11340.

(12) Gao, J.; Truhlar, D. G. *Annu. Rev. Phys. Chem.* **2002**, *53*, 467−505.

(13) Gao, J.; Ma, S.; Major, D. T.; Pu, J.; Truhlar, D. G. *Chem. Rev.* **2006**, *106*, 3188−3209.

(14) Billeter, S. R.; Webb, S. P.; Iordanov, T.; Agarwal, P. K.; Hammes-Schiffer, S. *J. Chem. Phys.* **2000**, *114*, 6925−6936.

(15) Billeter, S. R.; Webb, S. P.; Agarwal, P. K.; Iordanov, T.; Hammes-Schiffer, S. *J. Am. Chem. Soc.* **2001**, *123*, 11262−11272.

(16) Hammes-Schiffer, S. *Acc. Chem. Res.* **2006**, *39*, 93−100.

(17) Wang, Q.; Hammes-Schiffer, S. *J. Chem. Phys.* **2006**, *125*, 184102.

(18) Hwang, J. K.; Chu, Z. T.; Yadav, A.; Warshel, A. *J. Phys. Chem.* **1991**, *95*, 8445−8448.

(19) Hwang, J. K.; Warshel, A. *J. Phys. Chem.* **1993**, *97*, 10053−10058.

(20) Hwang, J.-K.; Warshel, A. *J. Am. Chem. Soc.* **1996**, *118*, 11745−11751.

(21) Feynman, R. P.; Hibbs, A. R. In *Quantum Mechanics and Path Integrals*; McGraw-Hill: New York, 1965.

(22) Chandler, D.; Wolynes, P. G. *J. Chem. Phys.* **1981**, *74*, 4078−4095.

(23) Sprik, M.; Klein, M. L.; Chandler, D. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1985**, *31*, 4234−4244.

(24) Berne, B. J.; Thirumalai, D. *Annu. Rev. Phys. Chem.* **1986**, *37*, 401−424.

(25) Voth, G. A.; Chandler, D.; Miller, W. H. *J. Chem. Phys.* **1989**, *91*, 7749−7759.

(26) Cao, J.; Berne, B. J. *J. Chem. Phys.* **1993**, *99*, 2902−2916.

(27) Chakrabarti, N.; Carrington, T., Jr.; Roux, B. *Chem. Phys. Lett.* **1998**, *293*, 209−220.

(28) Jang, J.; Voth, G. A. *J. Chem. Phys.* **1999**, *111*, 2357−2370.

(29) Marx, D.; Tuckerman, M. E.; Martyna, G. J. *Comput. Phys. Commun.* **1999**, *118*, 166−184.

(30) Tuckerman, M. E.; Marx, D. *Phys. Rev. Lett.* **2001**, *86*, 4946−4949.

(31) Iftimie, R.; Schofield, J. *J. Chem. Phys.* **2001**, *114*, 6763−6773.

(32) Mielke, S. L.; Truhlar, D. G. *Chem. Phys. Lett.* **2003**, *378*, 317−322.

(33) Olsson, M. H. M.; Siegbahn, P. E. M.; Warshel, A. *J. Am. Chem. Soc.* **2004**, *126*, 2820−2828.

(34) Major, D. T.; York, D. M.; Gao, J. L. *J. Am. Chem. Soc.* **2005**, *127*, 16374−16375.

(35) Major, D. T.; Nam, K.; Gao, J. L. *J. Am. Chem. Soc.* **2006**, *128*, 8114−8115.

(36) Major, D. T.; Gao, J. L. *J. Am. Chem. Soc.* **2006**, *128*, 16345−16357.

(37) Cao, J.; Voth, G. A. *J. Chem. Phys.* **1994**, *100*, 5093−5105.

(38) Pollock, E. L.; Ceperley, D. M. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1984**, *30*, 2555−2568.

(39) Ceperley, D. M. *Rev. Mod. Phys.* **1995**, *67*, 279−355.

(40) Tuckerman, M. E.; Berne, B. J.; Martyna, G. J.; Klein, M. L. *J. Chem. Phys.* **1993**, *99*, 2796−2808.

(41) Levy, P. *Compos. Math.* **1939**, *7*, 283.

(42) Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling, W. T. In *Numerical Recipes. The Art of Scientific Computing*, 2nd ed.; Cambridge University Press: New York, 1992; pp 279−281.

(43) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.

(44) Gao, J. *Rev. Comput. Chem.* **1995**, *7*, 119−185.

(45) Gao, J.; Thompson, M. A. In *Combined Quantum Mechanical and Molecular Mechanical Methods*; American Chemical Society: Washington, DC, 1998; American Chemical Society Symp. Ser. #712.

(46) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902−3909.

(47) Wu, N.; Mo, Y.; Gao, J.; Pai, E. F. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 2017−2022.

(48) Gao, J. *J. Am. Chem. Soc.* **1995**, *117*, 8600−8607.

(49) Gao, D.; Pan, Y.-K. *J. Org. Chem.* **1999**, *64*, 1151−1159.

(50) Gao, D.; Pan, Y.-K. *J. Org. Chem.* **1999**, *64*, 4492−4501.

(51) Acevedo, O.; Jorgensen, W. L. *J. Am. Chem. Soc.* **2005**, *127*, 8829−8834.

(52) Nam, K.; Gao, J.; York, D. M. *J. Chem. Theory Comput.* **2005**, *1*, 2−13.

(53) Allen, M. P.; Tildesley, D. J. In *Computer Simulation of Liquids*; Oxford University Press: Oxford, U. K., 1987.

(54) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187−199.

(55) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Comput. Chem.* **1992**, *13*, 1011−1021.

(56) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.

(57) Rishavy, M. A.; Cleland, W. W. *Biochemistry* **2000**, *39*, 4569−4574.

(58) Radzicka, A.; Wolfenden, R. *Science* **1995**, *267*, 90−93.

(59) Lee, J. K.; Houk, K. N. *Science* **1997**, *276*, 942−945.

(60) Fitzpatrick, P. F. *Acc. Chem. Res.* **2001**, *34*, 299−307.

(61) Fitzpatrick, P. F.; Orville, A. M.; Nagpal, A.; Valley, M. P. *Arch. Biochem. Biophys.* **2005**, *433*, 157−165.

(62) Gadda, G.; Fitzpatrick, P. F. *Biochemistry* **2000**, *39*, 1406−1410.

(63) Valley, M. P.; Fitzpatrick, P. F. *J. Am. Chem. Soc.* **2003**, *125*, 8738−8739.

(64) Nagpal, A.; Valley, M. P.; Fitzpatrick, P. F.; Orville, A. M. *Biochemistry* **2006**, *45*, 1138−1150.

(65) Valley, M. P.; Fitzpatrick, P. F. *J. Am. Chem. Soc.* **2004**, *126*, 6244−6245.

(66) Feierberg, I.; Luzhkov, V.; Aqvist, J. *J. Biol. Chem.* **2000**, *275*, 22657−22662.

CT600371K