The Journal of Chemical Physics

RESEARCH ARTICLE | MAY 01 2023

Estimation of frequency factors for the calculation of kinetic isotope effects from classical and path integral free energy simulations

Timothy J. Giese ; Darrin M. York

J. Chem. Phys. 158, 174105 (2023)
https://doi.org/10.1063/5.0147218

The Journal of Chemical Physics ARTICLE scitation.org/journal/jcp

Estimation of frequency factors for the calculation of kinetic isotope effects from classical and path integral free energy simulations

Cite as: J. Chem. Phys. 158, 174105 (2023); doi: 10.1063/5.0147218
Submitted: 20 February 2023 • Accepted: 17 April 2023 •
Published Online: 1 May 2023

Timothy J. Giese and Darrin M. York[a)]

AFFILIATIONS

Laboratory for Biomolecular Simulation Research, Institute for Quantitative Biomedicine and Department of Chemistry and Chemical Biology, Rutgers University, Piscataway, New Jersey 08854, USA

[a)]Author to whom correspondence should be addressed: Darrin.York@rutgers.edu

ABSTRACT

We use the modified Bigeleisen–Mayer equation to compute kinetic isotope effect values for non-enzymatic phosphoryl transfer reactions from classical and path integral molecular dynamics umbrella sampling. The modified form of the Bigeleisen–Mayer equation consists of a ratio of imaginary mode vibrational frequencies and a contribution arising from the isotopic substitution's effect on the activation free energy, which can be computed from path integral simulation. In the present study, we describe a practical method for estimating the frequency ratio correction directly from umbrella sampling in a manner that does not require normal mode analysis of many geometry optimized structures. Instead, the method relates the frequency ratio to the change in the mass weighted coordinate representation of the minimum free energy path at the transition state induced by isotopic substitution. The method is applied to the calculation of $^{16/18}$O and $^{32/34}$S primary kinetic isotope effect values for six non-enzymatic phosphoryl transfer reactions. We demonstrate that the results are consistent with the analysis of geometry optimized transition state ensembles using the traditional Bigeleisen–Mayer equation. The method thus presents a new practical tool to enable facile calculation of kinetic isotope effect values for complex chemical reactions in the condensed phase.

I. INTRODUCTION

The understanding of reaction mechanisms in the condensed phase is of broad interest in biology, medicine, and catalyst design.[1] Mechanistic studies of biocatalysis are particularly challenging, often requiring coordinated experimental and computational efforts.[2] A key goal of computational studies is to characterize the chemical mechanism by determining the minimum free energy paths passing through distinct rate-controlling transition states. By examining the sensitivity of the minimum free energy path in response to environmental perturbations, one gains insight into the key factors that control reactivity and thereby facilitate enzyme design.[3] The validation of computational predictions can be challenging, however, due to the inherent difficulty of directly observing transition state ensembles from the experiment. Kinetic isotope effect (KIE) measurements are one of the most sensitive experimental techniques to probe the changes in structure and bonding as a reaction proceeds to the transition state.[4–6] The primary and secondary KIE measurements are often able to distinguish between competing mechanistic pathways;[4,7] however, these measurements frequently require recourse to computational modeling to fully interpret. Hence, a vital component of computational mechanistic studies is the ability to calculate KIE values that are consistent with the predicted pathways and transition state ensembles.

KIE values for reactions of small molecules in the gas phase or implicit solvent environments are often computed from the Bigeleisen–Mayer equation (BME).[8] The BME introduces nuclear quantum effects from a harmonic approximation of the uncoupled vibrations determined from normal mode analysis, and it ignores quantum tunneling effects. To apply the BME, one performs geometry optimizations on a single (or small number of) reactant state (RS) and transition state (TS) structure(s) to obtain the vibrational frequencies of the light and heavy isotopologues from normal mode analysis. Similar approaches based on generalized transition state theory have also been developed to include multi-dimensional tunneling and anharmonicity effects.[9–11]

In explicitly modeled condensed phase environments, the reactant and transition states are described by an expanded ensemble of structures identified by the stationary points along a minimum free energy path traversing a multi-dimensional free energy surface.[12] The expanded ensemble requires one to introduce an averaging procedure to account for the structural variations of both the solute and environment[13] because the KIE estimated from individual samples may span a relatively wide range.[14] Rather than performing an exhaustive set of geometry optimizations for each reactant and transition state structure in the ensemble, one can leverage the configurational averaging provided by centroid path integral molecular dynamics (PIMD) sampling[15–19] and calculate the KIE from a modified form of the BME that introduces anharmonicity and quantum tunneling within Feynman's path integral framework.[20–24] The modified BME has two components: a ratio of imaginary mode vibrational frequencies and a component resulting from the change in activation free energy due to isotopic substitution. The frequency ratio correction (FRC) is the imaginary mode's contribution to the BME originating from the Redlich–Teller product rule.[25–27] The change in activation free energy can be computed from PIMD umbrella sampling,[15,16,28–31] but the calculation of an average FRC might appear, at first, to require geometry optimization and normal mode analysis of a transition state ensemble.

We describe a simple method for estimating the FRC directly from umbrella sampling that does not require normal mode analysis of geometry optimized structures. The method relates the FRC to the change in the mass weighted coordinate representation of the minimum free energy path at the transition state induced by isotopic substitution. We apply the method to the calculation of FRC and KIE values for a series of six non-enzymatic phosphoryl transfer reactions (Fig. 1) simulated under periodic boundary conditions with explicit solvent.[14] We validate the approach by generating distributions of reference FRC values from normal mode analysis of geometry optimized umbrella sampling configurations. The reference normal mode analysis distributions are compared to the FRC values produced by the new method, and the two approaches are shown to yield consistent results. We compute the $^{16/18}$O and $^{32/34}$S primary KIE values from the modified BME using the new FRC values and the change in activation free energy obtained from PIMD simulations.[14] In this manner, the modified BME is evaluated purely from the analysis of classical and PIMD umbrella sampling without recourse to geometry optimization. The modified BME KIE values are validated against distributions of reference BME results calculated from geometry optimized umbrella sampling configurations, and the two approaches are shown to be in excellent agreement.

The remainder of the paper is outlined as follows. Sec. II provides the relevant background for the BME and modified BME approaches and gives an overview of a strategy to compute KIE values using the new method. Section III formulates the new approach for calculating the FRC, and we detail our simulation protocol. Section IV presents and discusses our FRC and KIE results for six nonenzymatic phosphoryl transfer reactions. Section V summarizes the main conclusions of the paper.

II. BACKGROUND

A KIE is defined as the ratio $\eta = k_L/k_H$ of light (L) and heavy (H) isotopologue rate constants, k. The KIE is a phenomenon

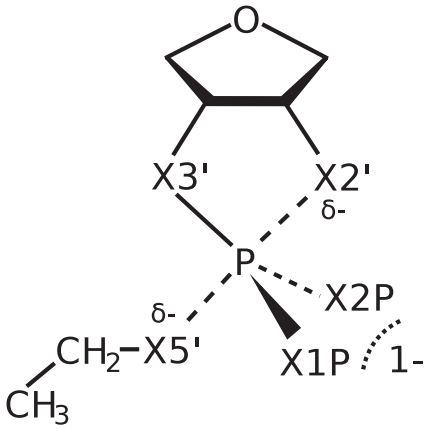

FIG. 1. Illustration of the nonenzymatic phosphoryl transfer models explored in this study. The X2′, X3′, and X5′ refer to the RNA naming scheme for these positions, and X1P and X2P are the nonbridge positions. The schematic depicts the transition state, and the reactant state is characterized by a fully formed P–X5′ bond and a cleaved P–X2′ bond. The native model has oxygen at all positions. The remaining five models involve sulfur substitutions at key positions: S12 (X1P and X2P), S1P (X1P), S2′ (X2′), S3′ (X3′), and S5′ (X5′). These substitutions are commonly used as mechanistic probes in experimental studies of RNA catalysis. X2′ and X5′ KIE values refer to either ^{16}O/^{18}O or ^{32}S/^{34}S isotopic substitutions as appropriate for each model.

that primarily arises from the mass influence on the nuclear wavefunction ψ_N (within the Born–Oppenheimer approximation) and is exhibited, in part, by a relative change in the reactant and transition state zero point energies. The primary $^{16/18}$O KIE values in RNA 2′-O-transphosphorylation reactions require precise experimental measurements and careful computational protocols because their values are typically within 3% of unity and the reactions occur in complex biological environments.[4,23,32] For these systems, KIE values should be estimated to within 1% error to categorize them as being "large inverse" (<0.97), "inverse" (0.97–0.99), "near unity" (0.99–1.01), "normal" (1.01–1.03), or "large normal" (>1.03).[4,32] Reported uncertainties in the experimental $^{16/18}$O KIE measurements of transphosphorylation reactions have been in the range of 0.3%–0.4%;[23] however, the reliability of calculated values heavily depends on the approach undertaken.[33]

A. Bigeleisen–Mayer equation and related ensemble approaches

The BME uses a rigid rotor, harmonic oscillator approximation of ψ_{nuc}, and the Redlich–Teller product rule[25–27] to derive an expression for the KIE that depends on the vibrational frequencies of the reactant state (ω_i) and transition state $(\omega_i^\ddagger)$ determined from normal mode analysis.[8,34] When tunneling effects are ignored, the BME is given by Eq. (1),

$$\eta_{BME} = \frac{\omega_{L,1}^\ddagger}{\omega_{H,1}^\ddagger}\left(\frac{\prod_{i=2}^{3N-6}\frac{\omega_{L,i}^\ddagger}{\omega_{H,i}^\ddagger}\frac{\sinh(\hbar\beta\omega_{H,i}^\ddagger/2)}{\sinh(\hbar\beta\omega_{L,i}^\ddagger/2)}}{\prod_{i=1}^{3N-6}\frac{\omega_{L,i}}{\omega_{H,i}}\frac{\sinh(\hbar\beta\omega_{H,i}/2)}{\sinh(\hbar\beta\omega_{L,i}/2)}}\right). \quad (1)$$

A derivation of Eq. (1) is provided in the supplementary material. The ratios of vibrational frequencies are the contributions of the

translational and rotational partition functions to the KIE upon enforcing the Redlich–Teller product rule, and the hyperbolic sine functions are contributions from the vibrational partition functions in the harmonic oscillator approximation. The term in brackets is the exponentiated change in activation free energy induced by isotopic substitution when the motion along the imaginary mode is constrained to be located at the transition state. The leading ratio of imaginary mode frequencies is the contribution of the missing degree of freedom to the translational and rotational partition functions.

When the BME is applied to explicitly-modeled condensed-phase environments, one can observe large variations in the predicted KIE values because many plausible reactant and transition state structures can be found.[14,35] For example, when nonenzymatic phosphoryl transfer reactions were optimized from different initial configurations in explicit solvent, the BME led to 2′ and 5′ KIE values distributed over a wide 3% range with a 1% standard deviation.[14] This observation is not a criticism; a range of values is expected, but it does imply that the confidence in any single result is low, so the result of many optimizations should be averaged.

Ensemble-averaged variational transition state theory (EA-VTST) is another approach that similarly initiates a series of transition state searches departing from umbrella sampling configurations.[36–39] The EA-VTST method performs canonical variational theory analysis on each structure to produce a distribution of transmission coefficients.[40] The transmission coefficients can include tunneling and recrossing corrections,[41,42] thus making EA-VTST an attractive alternative to the BME. Although sophisticated modeling of transmission coefficients may be necessary for some situations,[43,44] the BME continues to see widespread use and success.[45,46]

B. Path integral molecular dynamics with thermodynamic free energy perturbation

The need for configurational averaging is one of the main appeals for using centroid PIMD to compute KIE values.[15–19] PIMD directly includes a model[47] for ψ_N that can be used to evaluate the free energy of changing isotopic masses (light–to–heavy) with a free energy perturbation approach pioneered by Gao *et al.*[15,16,29–31] This was followed by the development of the thermodynamic free energy perturbation (TD-FEP) method,[28] which employs a conceptually similar strategy. The TD-FEP method has been implemented into the i-PI software[48] to take advantage of new generalized Langevin thermostats[49,50] intended to make PIMD simulations more practical by requiring fewer ring polymer beads to reach convergence. By performing TD-FEP analysis of the PIMD sampling at the reactant and transition states (producing isotopic substitution free energies ΔG_{RS} and ΔG_{TS}), one obtains the change in activation free energy induced by isotopic substitution ($\Delta\Delta G = \Delta G_{RS} - \Delta G_{TS}$), and the KIE can be estimated from the modified BME [Eq. (2)],[20–24]

$$\eta_{mBME} = \frac{\omega_L^{\ddagger}}{\omega_H^{\ddagger}}\,\eta_{PIMD}, \tag{2}$$

$$\eta_{PIMD} = \exp(-\beta\Delta\Delta G). \tag{3}$$

The isotopic substitution free energy at the transition state, ΔG_{TS}, is estimated from constrained sampling performed at a first-order saddle point on the free energy surface. In practice, one often approximates the use of a constraint by subjecting the system to a sufficiently large biasing potential that maintains the transition state structure along the minimum free energy path in a reduced-dimensional space of reaction coordinates. The leading factor of imaginary mode frequencies, referred to as the frequency ratio correction (FRC), is the contribution of the rotational and translational partition functions to the KIE from the unsampled degree of freedom in accordance with the Redlich–Teller product rule. By evaluating $\Delta\Delta G$ from PIMD/TD-FEP umbrella sampling, one extends the BME to include anharmonicity and tunneling effects within the path integral framework.[21] The modified BME first appeared in Appendix B of Ref. 51; it has been successfully applied in previous studies,[20,22–24] and an extended discussion of the relationship between the BME and the modified BME has been the subject of review.[21] Previous applications have frequently used Feynman path integrals to evaluate the modified BME within the context of Kleinert's variational perturbation theory.[52] Recently, we have used the modified BME to calculate KIE values from explicit PIMD umbrella sampling performed with the i-PI software.[14] The supplementary material includes a comparison of KIE values calculated from the BME and modified BME in the harmonic approximation.

C. Overview of the strategy to compute KIE values from umbrella sampling

The presence of the FRC in Eq. (2) gives the impression that one might still need to optimize transition state structures and perform normal mode analysis to obtain a distribution of its value,[14,20,35] or by approximating it from an optimized transition state in a dielectric continuum[22] or microsolvated environment.[53] In the present study, we describe a simple method for estimating the FRC directly from umbrella sampling; that is, we explore an approximation for the FRC that does not require normal mode analysis of geometry optimized structures. Instead, the method relates the FRC to the change in the mass weighted coordinate representation of the minimum free energy path at the transition state induced by isotopic substitution.

The reader may question how the transition state and imaginary mode are defined if geometry optimization and normal mode analysis on the potential energy surface are not first performed. In the context of simulation, the transition state is the location of a first-order saddle point on a reduced dimensional free energy surface described by one or more reaction coordinates. The reactant state is similarly defined as a minimum on the free energy surface. In other words, the reactant and transition states are ensembles characterized by their reaction coordinate values as opposed to being specific structures corresponding to potential energy stationary points. The sampling of the free energy surface is often enhanced by including umbrella biasing potentials, and the unbiased free energy surface is calculated from either umbrella integration,[54] weighted histogram analysis,[55] the multistate Bennett acceptance ratio method,[56] or the variational free energy profile method.[57,58] The minimum free energy path is a parametric curve of reaction coordinates connecting the reactant and product states by traversing a first-order saddle point on the reduced dimensional surface. The "imaginary mode" is the minimum free energy path as it passes through the transition state.

With the approach described here, one could estimate a KIE using the following strategies:

1. Perform the umbrella sampling finite temperature string method to obtain a minimum free energy path.[59–62] This is a classical minimum free energy path without nuclear quantum effects. As such, both the light and heavy isotopologues share the same pathway.
2. Sample the reactant state and transition state (as identified by the minimum free energy path) with PIMD/TD-FEP to compute η_{PIMD}.[28]
3. Analyze the umbrella sampling performed in step 1 to obtain the mass weighted coordinate atomic displacements along the minimum free energy path. The FRC estimate is a harmonic analysis made from these displacements.
4. Calculate the KIE with Eq. (2).

Section III describes the method for approximating the FRC in more detail.

III. METHODS

A. Estimation of the FRC from umbrella sampling

The method for estimating the FRC is founded upon a solution of the Euler–Lagrange equation for the harmonic motion of a system of N classical particles with masses m_a and Cartesian positions $r_{ak}(q)$ constrained to move along a generalized coordinate q near a stationary point. In other words, $r_{ak}(q)$ is a $3N$ parametric curve describing the minimum energy path, and q is the curve's progress variable. Let $s_{ak}(q) = m_a^{1/2} r_{ak}(q)$ denote the path in mass weighted coordinates, then one can show,

$$\frac{\omega_{\mathrm{L}}^{\ddagger}}{\omega_{\mathrm{H}}^{\ddagger}} = \frac{\|\mathbf{s}'_{\mathrm{H}}(q^{\ddagger})\|}{\|\mathbf{s}'_{\mathrm{L}}(q^{\ddagger})\|}, \tag{4}$$

where $\|\mathbf{s}'(q^{\ddagger})\|$ is the magnitude of the mass weighted coordinate displacement vector along the path at the transition state, $q^{\ddagger}$,

$$\|\mathbf{s}'(q^{\ddagger})\| = \left[\sum_{a=1}^{N} m_a \sum_{k=1}^{3} (\partial r_{ak}(q^{\ddagger})/\partial q)^2\right]^{1/2}. \tag{5}$$

A derivation of Eq. (4) is provided in the supplementary material. The squared magnitude of the displacement is proportional to the reduced mass of the mode, $\mu = \alpha \|\mathbf{s}'(q^{\ddagger})\|^2$, in the curvilinear coordinate system parametrized by q.[63,64] Equation (4) effectively reads: $\omega_{\mathrm{L}}^{\ddagger}/\omega_{\mathrm{H}}^{\ddagger} = \sqrt{\mu_{\mathrm{H}}/\mu_{\mathrm{L}}}$ because the normalization factors $\alpha = \|\mathbf{r}'(q^{\ddagger})\|^{-2}$ cancel in the ratio.

In the context of simulation, q is the progress of the reaction along a minimum free energy path, which we regard as a parametric curve in a reduced set of reaction coordinates, $\boldsymbol{\xi}(q)$, that define the umbrella potential locations used to perform umbrella sampling along the path. The reaction coordinates are often chosen to be internal coordinates rather than Cartesian positions, and they are assigned by chemical intuition so that the minimum free energy path is characterized by the relative displacement of a select few atoms. By construction, the average relative displacement between the selected atoms has a dependence on q. Umbrella sampling performed at $\boldsymbol{\xi}(q)$ yields a trajectory of atomic positions from which we extract the $3N_{\mathrm{sele}}$ coordinates of the selected atoms. Because $\boldsymbol{\xi}(q)$ only describes the internal displacements of the atoms, the extracted coordinates are translated to a common center of mass and aligned to reduce the coordinate root mean square deviation. The average aligned coordinates will be denoted $\langle r_{ak}\rangle_q$. This procedure can be applied to the umbrella sampling performed on either side of $q^{\ddagger}$ to approximate the derivative $\partial\langle r_{ak}\rangle_{q^{\ddagger}}/\partial q$ from finite difference, leading to Eq. (6),

$$\left\langle \frac{\omega_{\mathrm{L}}^{\ddagger}}{\omega_{\mathrm{H}}^{\ddagger}} \right\rangle \approx \left[\frac{\sum_{a=1}^{N_{\mathrm{sele}}} m_{\mathrm{H},a} \sum_{k=1}^{3} \left(\langle r_{ak}\rangle_{q^{\ddagger}+\Delta q} - \langle r_{ak}\rangle_{q^{\ddagger}-\Delta q} \right)^2}{\sum_{a=1}^{N_{\mathrm{sele}}} m_{\mathrm{L},a} \sum_{k=1}^{3} \left(\langle r_{ak}\rangle_{q^{\ddagger}+\Delta q} - \langle r_{ak}\rangle_{q^{\ddagger}-\Delta q} \right)^2} \right]^{1/2}. \tag{6}$$

The coordinates $\langle r_{ak}\rangle_{q^{\ddagger}+\Delta q}$ and $\langle r_{ak}\rangle_{q^{\ddagger}-\Delta q}$ are translated and aligned before taking their difference. In the classical approximation, the difference $\langle r_{ak}\rangle_{q^{\ddagger}+\Delta q} - \langle r_{ak}\rangle_{q^{\ddagger}-\Delta q}$ is mass-independent (because the minimum free energy path is mass-independent); however, the calculation of the averages required us to center and align the atoms. For this purpose, we use light isotopes. The use of the heavy isotopes does not produce a meaningful difference in our results.

The "selected atoms" are those atoms used to describe the motion of the imaginary frequency at the transition state $[\partial r_{ak}(q^{\ddagger})/\partial q \neq 0]$, whereas all other atoms are ignored $[\partial r_{ak}(q^{\ddagger})/\partial q = 0]$. The decision of which atoms to "select" has been the topic of discussion for decades.[46,65–72] These discussions have typically been in the context of computing the nuclear quantum effects from a truncated Hessian to approximate the relevant vibrations of a larger system. Stern and Wolfsberg introduced the "two bond rule," which suggests the KIE can be reasonably approximated by excluding all atoms located more than two bonds from the isotopic substitution so long as the estimate is made around room temperature and the force constants of the truncated modes do not significantly differ from the full system.[65,66] The present study does not diagonalize a Hessian because the vibrational motion is constrained to a single coordinate (the minimum free energy path), and the nuclear quantum effects are modeled through PIMD. Instead, the selected atoms are exclusively used in Eq. (6), so the selection should include those atoms that significantly contribute to the mass weighted coordinate displacement. Motivated by Stern and Wolfsberg's rule, we examine an analogous set of rules in the supplementary material (Tables S1 and S2) for a series of 20 nonenzymatic phosphoryl transfer models whose transition states were found in the gas phase with B3LYP/6-31++G(3df,2p).[73] The 2′ and 5′ FRCs were computed from the normal mode analysis of the full Hessian and compared to Eq. (4) by approximating $\mathbf{s}'(q^{\ddagger})$ from finite differentiation of the imaginary mode using only those atoms within B bonds of P. In summary, the mean relative percent error of the FRCs was 0.3% when all atoms within two bonds of P were included, which is similar in magnitude to the uncertainty in the experimental KIE measurements.[23]

In order to apply Eq. (6), one must be mindful of whether the reaction coordinates are a sufficient description of the atom's relative motion with consideration for those atoms that are located several bonds from the chemical event. Equation (6) tacitly assumes that the Cartesian distributions of the selected atoms are monomodal. If this assumption is not true, then it may be necessary to redefine the minimum free energy path to include additional reaction coordinates that prevent rotations around a bond, for example. Furthermore, a

meaningful result will only be obtained if the statistical uncertainty in the averages is less than the displacement used to approximate $\partial\langle r_{ak}\rangle_{q^\ddagger}/\partial q$. In contrast, excessively large displacements will also be a poor approximation of the derivative. We will explore the sensitivity of the results to displacement size and sampling.

B. Computational details

We previously examined the six nonenzymatic phosphoryl transfer models in Ref. 14 to train a deep potential range corrected (DPRc) machine learning potential (Δ-MLP)[74] that supplements the second order density functional tight binding[75–77] (DFTB2) quantum mechanical (QM)/molecular mechanical (MM) Hamiltonian with a nonelectronic neural network correction parametrized to reproduce PBE0/6-31G* QM/MM energies and forces. The DFTB2 model is evaluated with the MIO parameter set and referred to as DFTB2/MIO. Because the primary concern of the present study is the estimation of the FRC, we reuse some of our previous results reported in Ref. 14. Specifically, the reactant and transition state locations along the minimum free energy path defined by the reaction coordinate $\xi_{PT} = R_{P-X5'} - R_{P-X2'}$ were previously identified by umbrella sampling; η_{PIMD} was previously calculated from PIMD/TD-FEP sampling performed at the reactant state and transition state; and many structures departing from reactant and transition state umbrella sampling configurations were optimized to produce distributions of η_{BME} and FRC values from normal mode analysis. In the present study, we perform new umbrella sampling in the vicinity of the transition state to evaluate Eq. (6) from simulation, and we compare these FRC distributions to our previously reported normal mode analysis values. We calculate KIE values from Eq. (2) using the new FRC estimates and our previously reported η_{PIMD} values. These KIE values are then compared to our previously reported distribution of η_{BME} results obtained from normal mode analysis of optimized structures.

In this paragraph, we briefly summarize some of the relevant computational details in Ref. 14 to facilitate the comparison with our new results. We implemented the DPRc correction into the DeePMD-kit software package[78] and interfaced it to a development version of AMBER's SANDER program,[79] which was used to compute free energy surfaces of the reactions from DFTB2/MIO QM/MM+DPRc umbrella sampling in a periodic box of 1510 TIP4P/Ew waters.[14] These surfaces were used to identify the reactant state and transition state. PIMD/TD-FEP sampling was performed at the reactant and transition states to calculate η_{PIMD}, and distributions of FRC and η_{BME} values were obtained from normal mode analysis of optimized reactant and transition state geometries. The optimizations were performed with DFTB2/MIO QM/MM+DPRc with explicit solvent in periodic boundary conditions; however, only the coordinates of the solute and 6 Å of nearby waters were allowed to move. In this manner, reactant state optimizations were readily performed to produce Hessians whose dimensions were ~750 × 750. The transition state structures were found by partial optimization of the solute and 6 Å of nearby waters with a constrained value of ξ_{PT}. This relaxation was followed by an unrestrained transition state search that included 3 Å of nearby waters (producing Hessians whose dimensions were ~300 × 300). This process was repeated from different starting configurations, leading to a normal mode analysis of 26, 19, 62, 30, 19, and 111 transition state structures for the native, S12, S1P, S2′, S3′, and S5′ model reactions. The distributions of normal mode analysis FRC values are characterized by their mean, standard deviation, and maximum and minimum observed values.

The normal mode analysis FRC distributions are compared to estimates made from umbrella sampling using Eq. (6). For each nonenzymatic phosphoryl transfer reaction, we ran new DFTB2/MIO QM/MM+DPRc simulations in the vicinity of the transition state. Specifically, we sampled the reaction with classical molecular dynamics at six ξ_{PT} locations corresponding to three displacements from the transition state: ±0.1, ±0.2, and ±0.3 Å. The umbrella sampling was performed with SANDER in the canonical ensemble using a 1 fs time step and the Langevin thermostat to maintain a temperature of 298 K.[80] The ξ_{PT} reaction coordinate was harmonically restrained with a 200 kcal mol^{-1} Å^{-2} force constant. Electrostatics were modeled with the QM/MM particle mesh Ewald method using a 1 Å^3 grid spacing, an 8 Å real-space cutoff, and tinfoil boundary conditions.[81–84] The Lennard-Jones interactions were truncated at 8 Å, and a long-range tail correction was used to model the long-range interactions.[85] The initial configurations were restarts departing from the production sampling reported in Ref. 14, whose densities were previously equilibrated for 100 ps at a pressure of 1 atm with a Berendsen barostat[86] at 298 K. In the present study, we carried out each simulation for 100 ps, producing a trajectory of 1000 saved configurations, and we repeated the simulations four times using different thermostat random number seeds and each of the four neural network parameter sets developed in Ref. 14. When only the ξ_{PT} coordinate is restrained, the ethoxide leaving group is free to rotate around the P–X5′ bond, which causes a diffusion of the C2′ and C5′ distributions; therefore, we introduce additional C5′–X5′–P–XP1 and C2′–X2′–P–XP2 torsion restraints to concentrate the distributions of these atoms. These torsions were restrained with a force constant of 20 kcal mol^{-1} radian^{-2}, and the equilibrium positions were chosen by partitioning the unrestrained angular distributions observed in the windows near the transition state into two clusters and calculating the mean of the most probable cluster. All six of the displaced simulations use the same torsion restraints, and the equilibrium values for each nonenzymatic phosphoryl transfer model are listed in the supplementary material (Table S3). Although the torsion restraints were introduced out of an abundance of caution, they had only a minor effect on our results; the KIE values analyzed from sampling performed without the torsion restraints reproduce our reported results with a mean absolute deviation (MAD) of 0.1%, which is an order of magnitude smaller than their standard deviations. A comparison of KIE values with and without torsion restraints is shown in Fig. S1 within the supplementary material.

IV. RESULTS AND DISCUSSION

We examine the nucleophile (X2′) and leaving group (X5′), $^{16/18}$O and $^{32/34}$S FRC, and primary KIE values for the series of nonenzymatic phosphoryl transfer reactions illustrated in Fig. 1. The X2′ and X5′ FRC distributions produced from Eq. (6) and normal mode analysis are compared in Fig. 2. To generate a distribution of values using Eq. (6) with a displacement of $\Delta\xi_{PT}$, we read the 4000 saved frames from the simulations performed at $\xi_{PT}^{\ddagger} + \Delta\xi_{PT}$ and split them into 20 equal segments, each containing 200 frames, to yield 20 sets of coordinate averages. The procedure is

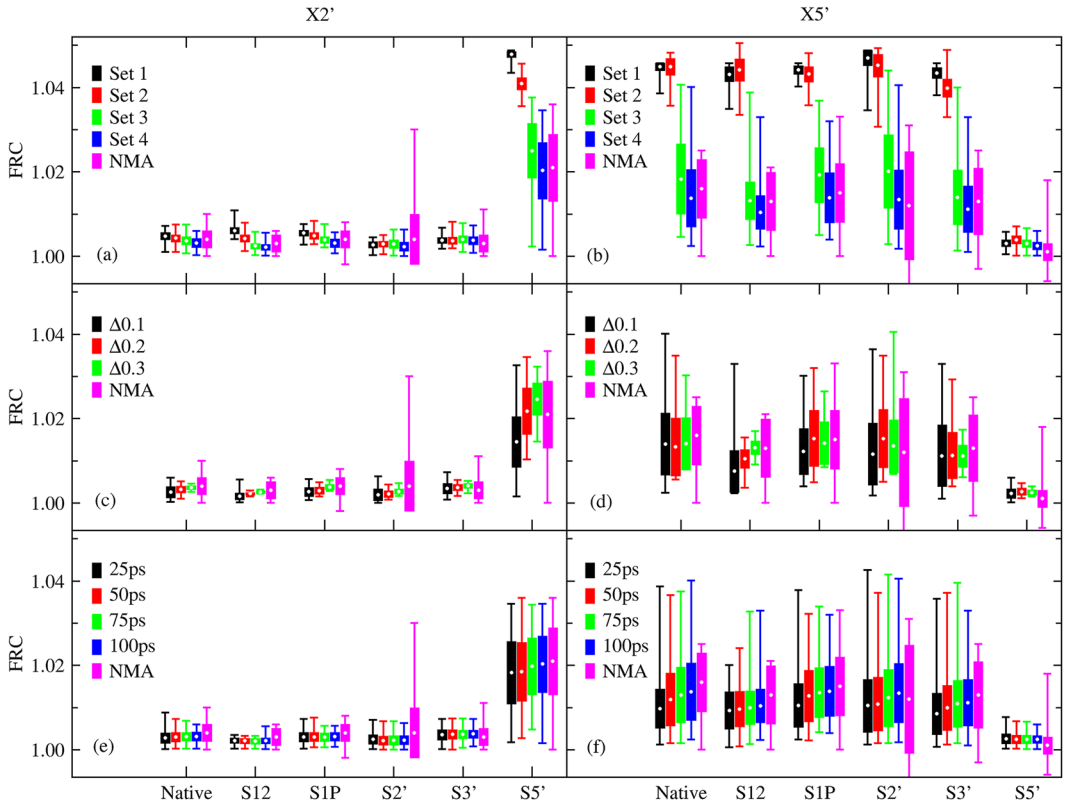

FIG. 2. Sensitivity of the X2′ and X5′ FRC with respect to atom selection (a) and (b), finite difference step length (c) and (d), and sampling (e) and (f). Parts (a)–(d) use 100 ps/window of sampling. Parts (a), (b), (e), and (f) use the aggregate distribution of results from the 0.1, 0.2, and 0.3 Å displacements. Parts (c)–(f) use atom selection set 4. NMA refers to the distribution of values obtained from the normal mode analysis of optimized structures in an explicit MM solvent (taken from Ref. 14). The box plots describe the distribution of FRC values. The white circle is the mean, the colored box spans 1 standard deviation on either side of the mean, and the bars reach the minimum and maximum values.

repeated to obtain 20 sets of coordinate averages from the $\xi^{\ddagger}_{\mathrm{PT}} - \Delta\xi_{\mathrm{PT}}$ simulations, and Eq. (6) is applied 400 times corresponding to each combination of averages. The mean, standard deviation, and minimum and maximum values of these distributions are illustrated in Fig. 2. Figures 2(c) and 2(d) compare the distributions obtained using different displacements, whereas Figs. 2(a), 2(b), 2(e), and 2(f) illustrate the distribution of 1200 aggregate estimates made from the ±0.1, ±0.2, and ±0.3 Å displacements. Figures 2(a) and 2(b) compare the FRC distributions from 100 ps/window of sampling with different atom selections. We refer to the selections as follows. Set 1 consists of three atoms: P, X2′, and X5′. Set 2 selects all atoms within one bond of P. Set 3 selects all atoms within two bonds of P. Set 4 selects all atoms within two bonds of P and any hydrogens within three bonds of P. Figures 2(e) and 2(f) compare the FRC distributions using selection Set 4, but different sampling lengths per simulation.

The results shown in Figs. 2(a) and 2(b) suggest the normal mode analysis FRC is not well-reproduced by Eq. (6) when the selection is limited to those atoms directly bonded to P (1.5% MAD); however, the FRC values agree to within 1 standard deviation of the normal mode analysis values when Eq. (6) is used with selection Sets 3 (0.2% MAD) and 4 (0.1% MAD). This observation is consistent with the atom selection tests performed on the 20 nonenzymatic phosphoryl transfer models provided in the supplementary material. Figures 2(c) and 2(d) show that the predicted FRC values are stable with respect to the size of the displacement. Displacements of 0.1, 0.2, and 0.3 Å with Set 4 produce MAD values of 0.2%, 0.2%, and 0.1%, respectively, when compared to the normal mode analysis distributions. Finally, Figs. 2(e) and 2(f) demonstrate that the FRC distributions produced from different amounts of sampling are similar.

The η_{PIMD} and η_{mBME} estimates of the X2′ and X5′ KIE values are compared to structurally averaged η_{BME} results in Figs. 3(a) and 3(b), respectively. The η_{mBME} values shown in Fig. 3(b) were evaluated with Eq. (2) using the 100 ps FRC distributions shown in Figs. 2(e) and 2(f) and the η_{PIMD} values taken from our previous study.[14] The error bars represent 1 standard deviation. Although we use the η_{BME} values as a reference to make comparisons, we describe any discrepancies as "differences" rather than "errors" because the BME ignores quantum tunneling and anharmonicity, whereas η_{PIMD} and η_{mBME} include a treatment for these by virtue of PIMD sampling. In the current application, quantum tunneling is expected

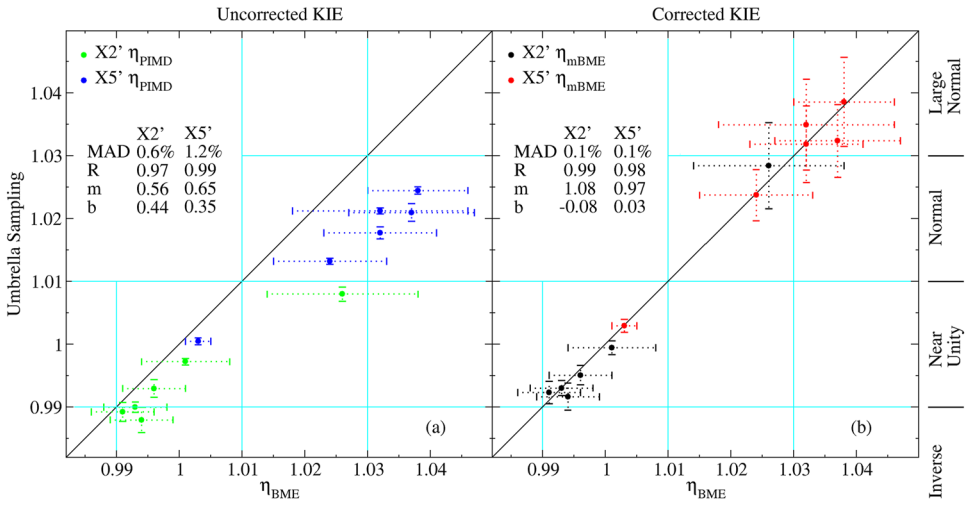

FIG. 3. Comparison of X2′ and X5′ η_{BME}, η_{PIMD}, and η_{mBME} KIE values. (a) The uncorrected PIMD/TD-FEP KIE value, η_{PIMD} [Eq. (3)]. (b) The modified BME KIE value, η_{mBME} [Eq. (2)], using the FRC estimated from Eq. (6). MAD is the mean absolute deviation with respect to the average η_{BME} values. "R," "m," and "b" are the Pearson correlation coefficient, slope, and intercept of a linear regression, respectively. Cyan grid lines indicate heuristic boundaries distinguishing large normal (>1.03), normal (1.01–1.03), near unity (0.99–1.01), and inverse (0.97–0.99) KIE values.[4,32]

to be small, so the comparisons are meant to quantify a level of consistency between the methods. Obtaining consistent results between the methods is of particular interest because they are computed using very different approaches: the η_{BME} values are produced solely from normal mode analysis of geometry optimized structures, and the η_{mBME} values are produced solely from classical and path integral sampling. To place a better perspective on the role of the FRC, Fig. 3(a) shows an analogous comparison of the uncorrected η_{PIMD} values to the distribution of η_{BME} results. The η_{PIMD} error bars are the standard deviation of four independent sets of PIMD simulations. Without the FRC correction, the η_{PIMD} values deviate from η_{BME} by 0.6% (X2′) and 1.1% (X5′) MAD. These differences are reduced to 0.1% MAD when η_{PIMD} is corrected by the FRC calculated by Eq. (6).

We previously mentioned that the KIE values of phosphoryl transfer reactions are conveniently categorized as being "large normal," "normal," "near unity," "inverse," or "large inverse."[4,32] The cyan grid lines in Fig. 3 mark the relevant boundaries between these categorizations. Data points that reside within a diagonal block indicate agreement between the umbrella sampling and BME categorizations. The uncorrected η_{PIMD} values [Fig. 3(a)] categorize 7 out of the 12 values differently than the BME averages. In contrast, the categorization of all KIE values agrees with the BME estimates when the FRC is included [Fig. 3(b)]. This is related to the observation that linear regression of the η_{PIMD} values produces slopes (0.56 and 0.65 for X2′ and X5′, respectively) and intercepts (0.44 and 0.35 for X2′ and X5′, respectively) that deviate from their ideal values of 1 and 0, respectively, even though the correlation coefficients are high. The η_{mBME} values are also strongly correlated, but the regressions produce slopes near unity and intercepts close to zero. Taken together, this demonstrates support for the use of the FRC estimation procedure as a practical method for calculating KIE values from the classical and PIMD simulations.

The native nonenzymatic reaction explored in this study is a truncated model of the uridylyl-3′-guanosine 2′-O-transphosphorylation reaction catalyzed in basic conditions. Experimental $^{16/18}$O KIE values of uridylyl-3′-guanosine at the O2′ and O5′ positions have been reported in pH 14 conditions.[23] Table I compares the BME and modified BME calculations to the experimental values. The experimental uncertainties are the reported standard deviations produced by a numerical fit to high performance liquid chromatography chromatograms. The BME and modified BME values are in very good agreement with each other, and their O5′ KIE values are in good agreement with the experiment. There is a greater discrepancy between the calculated and experimental O2′ KIE values; the differences between experiment and computation are about the same as the propagated uncertainty. All of the values suggest the O2′ KIE is inverse, however. The neural network parameters were not trained to reproduce KIE values; they were trained to reproduce PBE0/6-31G* QM/MM energies and forces, so some discrepancy with the experimental values is not unexpected. It was previously shown that the DFTB2/MIO QM/MM+DPRc model well-reproduced the KIE values predicted from the PBE0/6-31G* QM/MM simulation.[14] Further improvement with respect to experiment may be obtained by training the machine learning potential to a higher level of theory.

The computational cost of estimating KIE values from either the BME or modified BME is dominated by the classical sampling used to determine the minimum free energy path. The method we've described for estimating the FRC does not formally require any further calculation; it is an analysis of that sampling. The classical free energy surface of the native model reaction, for example, was produced from umbrella sampling at 86 windows. To validate the convergence of the free energy surface, it was necessary to sample each window for 100 ps using a 1 fs time step and repeat each simulation three times with different random number seeds.[87] This

TABLE I. Comparison of the native model reaction KIE values with the experiment. Experimental values are uridylyl-3′-guanosine at pH 14 taken from Ref. 23.

Method	O2′	O5′
Expt.	0.984 ± 0.003	1.034 ± 0.004
η_{BME}	0.991 ± 0.005	1.038 ± 0.008
η_{mBME}	0.992 ± 0.005	1.039 ± 0.009

protocol was performed with each of the four neural network parameter sets, resulting in 103×10^6 aggregate energy evaluations. To calculate the BME, samples from the reactant and transition states are extracted to initiate geometry optimizations. To calculate the modified BME, the reactant and transition state reaction coordinate values are used to initiate PIMD umbrella sampling.

The PIMD simulations for the calculation of ΔG_{RS} and ΔG_{TS} were run for 10 ps with a 0.25 fs time step and six ring polymer beads using the PIGLET thermostat.[49] This protocol corresponds to 0.5×10^6 energy evaluations to estimate η_{PIMD}, which is less than 1% of the computational effort used to converge the classical free energy surface. To estimate the native model reaction KIE values from the BME, we geometry optimized 40 samples from both the reactant and transition states. The reactant state minimizations readily succeed without difficulty, requiring 461 evaluations to reach convergence on average and ~1500 evaluations to estimate the Hessian from finite differentiation of the atomic forces. To increase the likelihood of finding a transition state with the partitioned rational function optimization method,[88,89] we precondition the structures by performing partial minimizations that constrain the reaction coordinate value. Each transition state required 5026 energy evaluations on average and ~600 additional evaluations to calculate the Hessian. If all of the optimizations succeed, the BME estimate of the native model reaction could be evaluated with only 0.3×10^6 energy evaluations.

In our opinion, both the BME and modified BME methods are practical; neither of them requires excessive computational effort. The discussion earlier does not adequately convey the amount of human effort one must devote to each method, however. Our practical experience is that transition state searches often fail to locate a stationary point when many degrees of freedom are involved. Most numerical optimization algorithms have an inherent radius of convergence, which the extracted samples are not guaranteed to lie within.[90,91] Even when a transition state search does converge, it may converge to an unintended saddle point, so one must manually examine each structure and verify that the imaginary mode vibration is qualitatively correct. As described earlier, we performed 40 transition state searches for the native model reaction; however, we were only successful in locating 26 of the structures. One approach to overcome this difficulty would be to develop new optimization strategies that improve the reliability of locating transition states.[13] In the present study, we have taken a different approach by exploring a method for calculating KIE values directly from umbrella sampling. In doing so, we extend the BME to include anharmonicity and tunneling effects within the path integral framework.[20–24] The calculation of isotopic substitution free energies from PIMD has previously been described;[28] however, to apply the approach to the calculation of KIE values, one must sample a transition state, which necessarily requires the imposition of a constraint. The contribution of the unsampled motion to the KIE appears as a ratio of imaginary mode vibrational frequencies in accordance with the Redlich–Teller product rule.[25–27] The present study describes a method for estimating this ratio from a harmonic analysis of the umbrella sampling used to calculate the minimum free energy profile.

V. CONCLUSIONS

We described a practical method for estimating the light-to-heavy ratio of imaginary mode vibrational frequencies from umbrella sampling performed along a minimum free energy path in condensed phase environments. These ratios are used in the modified BME to calculate KIE values from classical and PIMD/TD-FEP simulations[28] to incorporate anharmonicity and tunneling effects within the Feynman path integral framework.[20–24] By calculating the FRC from simulation, one obtains an estimate of its ensemble average without needing to geometry optimize many starting configurations.[14,35] We applied the new method to condensed phase simulations of six nonenzymatic phosphoryl transfer reactions. It was shown that the distribution of $^{16/18}$O and $^{32/34}$S FRC values estimated from umbrella sampling is in good agreement (0.2% MAD) with geometry optimized normal mode analysis results as long as the average position of those atoms within two bonds of P is considered. The inclusion of all H's located within three bonds of P further improved the comparison to 0.1% MAD. The FRC estimates were used in the modified BME to compute the 2′ and 5′ primary KIE values. The modified BME (computed solely from classical and PIMD umbrella sampling) was shown to be consistent with average BME values (computed solely from normal mode analysis of geometry optimized structures) with a MAD of 0.1%.

SUPPLEMENTARY MATERIAL

See the supplementary material for a derivation of Eq. (4), a comparison of approximate FRC values using Eq. (6) for 20 nonenzymatic phosphoryl transfer reactions in the gas phase, a comparison of condensed phase KIE values computed with and without torsion restraints, and a comparison of the BME and modified BME in the harmonic approximation.

ACKNOWLEDGMENTS

The authors are grateful for the financial support provided by the National Institutes of Health (Grant No. GM107485 to D.M.Y.). Computational resources were provided by the Office of Advanced Research Computing (OARC) at Rutgers, The State University of New Jersey; the Advanced Cyberinfrastructure Coordination Ecosystem: Services and Support (ACCESS) program, which is supported by the National Science Foundation under Grant Nos. 2138259, 2138286, 2138307, 2137603, and 2138296 (supercomputer Expanse at SDSC through allocation under Grant No. CHE190067); and the Texas Advanced Computing Center (TACC) at the University of Texas at Austin, URL: http://www.tacc.utexas.edu (supercomputer Frontera through allocation under Grant No. CHE20002).

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

T.J.G. and D.M.Y. contributed equally to this work.

Timothy J. Giese: Conceptualization (equal); Funding acquisition (equal); Methodology (equal); Visualization (equal); Writing – original draft (equal); Writing – review & editing (equal). **Darrin**

M. York: Conceptualization (equal); Funding acquisition (lead); Methodology (equal); Visualization (equal); Writing – original draft (equal); Writing – review & editing (equal).

DATA AVAILABILITY

A development version of SANDER interfaced to the DeepMD-Kit library is openly available in the GitLab repository at https://gitlab.com/RutgersLBSR/AmberDPRc. The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1] M. Garcia-Viloca, J. Gao, M. Karplus, and D. G. Truhlar, "How enzymes work: Analysis by modern rate theory and computer simulations," Science **303**, 186–195 (2004).

[2] A. Ganguly, B. P. Weissman, T. J. Giese, N.-S. Li, S. Hoshika, S. Rao, S. A. Benner, J. A. Piccirilli, and D. M. York, "Confluence of theory and experiment reveals the catalytic mechanism of the Varkud satellite ribozyme," Nat. Chem. **12**, 193–201 (2020).

[3] C. S. Gaines, J. A. Piccirilli, and D. M. York, "The L-platform/L-scaffold framework: A blueprint for RNA-cleaving nucleic acid enzyme design," RNA **26**, 111–125 (2020).

[4] A. C. Hengge, "Isotope effects in the study of phosphoryl and sulfuryl transfer reactions," Acc. Chem. Res. **35**, 105–112 (2002).

[5] P. Paneth, "Chlorine kinetic isotope effects on enzymatic dehalogenations," Acc. Chem. Res. **36**, 120–126 (2003).

[6] W. W. Cleland, "The use of isotope effects to determine enzyme mechanisms," Arch. Biochem. Biophys. **433**, 2–12 (2005).

[7] B. Weissman, Ş. Ekesan, H.-C. Lin, S. Gardezi, N.-S. Li, T. J. Giese, E. McCarthy, M. E. Harris, D. M. York, and J. A. Piccirilli, "Dissociative transition state in hepatitis delta virus ribozyme catalysis," J. Am. Chem. Soc. **145**, 2830–2839 (2023).

[8] J. Bigeleisen and M. Wolfsberg, "Theoretical and experimental aspects of isotope effects in chemical kinetics," Adv. Chem. Phys. **1**, 15–76 (1958).

[9] W.-P. Hu and D. G. Truhlar, "Deuterium kinetic isotope effects and their temperature dependence in the gas-phase S_N2 reactions $X^- + CH_3Y \rightarrow CH_3 + Y^-$ (X, Y = Cl, Br, I)," J. Am. Chem. Soc. **117**, 10726–10734 (1995).

[10] A. Fernández-Ramos, J. A. Miller, S. J. Klippenstein, and D. G. Truhlar, "Modeling the kinetics of biomolecular reactions," Chem. Rev. **106**, 4518–4584 (2006).

[11] T. Yu, J. Zheng, and D. G. Truhlar, "Multi-structural variational transition state theory. Kinetics of the 1,4-hydrogen shift isomerization of the pentyl radical with torsional anharmonicity," Chem. Sci. **2**, 2199 (2011).

[12] T. J. Giese, Ş. Ekesan, and D. M. York, "Extension of the variational free energy profile and multistate Bennett acceptance ratio methods for high-dimensional potential of mean force profile analysis," J. Phys. Chem. A **125**, 4216–4232 (2021).

[13] T. D. Poulsen, M. Garcia-Viloca, J. Gao, and D. G. Truhlar, "Free energy surface, reaction paths, and kinetic isotope effect of short-chain acyl-CoA dehydrogenase," J. Phys. Chem. B **107**, 9567–9578 (2003).

[14] T. J. Giese, J. Zeng, Ş. Ekesan, and D. M. York, "Combined QM/MM, machine learning path integral approach to compute free energy profiles and kinetic isotope effects in RNA cleavage reactions," J. Chem. Theory Comput. **18**, 4304–4317 (2022).

[15] J. Gao, "Chapter fourteen—Enzymatic kinetic isotope effects from path-integral free energy perturbation theory," Methods Enzymol. **577**, 359–388 (2016).

[16] A. Vardi-Kilshtain, N. Nitoker, and D. T. Major, "Nuclear quantum effects and kinetic isotope effects in enzyme reactions," Arch. Biochem. Biophys. **582**, 18–27 (2015).

[17] J. K. Hwang, Z. T. Chu, A. Yadav, and A. Warshel, "Simulations of quantum mechanical corrections for rate constants of hydride-transfer reactions in enzymes and solutions," J. Phys. Chem. **95**, 8445–8448 (1991).

[18] J. K. Hwang and A. Warshel, "A quantized classical path approach for calculations of quantum mechanical rate constants," J. Phys. Chem. **97**, 10053–10058 (1993).

[19] J.-K. Hwang and A. Warshel, "How important are quantum mechanical nuclear motions in enzyme catalysis?," J. Am. Chem. Soc. **118**, 11745–11751 (1996).

[20] K.-Y. Wong, J. P. Richard, and J. Gao, "Theoretical analysis of kinetic isotope effects on proton transfer reactions between substituted α-methoxystyrenes and substituted acetic acids," J. Am. Chem. Soc. **131**, 13963–13971 (2009).

[21] K.-Y. Wong, Y. Xu, and L. Xu, "Review of computer simulations of isotope effects on biochemical reactions: From the Bigeleisen equation to Feynman's path integral," Biochim. Biophys. Acta **1854**, 1782–1794 (2015).

[22] K.-Y. Wong, Y. Xu, and D. M. York, "*Ab initio* path-integral calculations of kinetic and equilibrium isotope effects on base-catalyzed RNA transphosphorylation models," J. Comput. Chem. **35**, 1302–1316 (2014).

[23] H. Gu, S. Zhang, K.-Y. Wong, B. K. Radak, T. Dissanayake, D. L. Kellerman, Q. Dai, M. Miyagi, V. E. Anderson, D. M. York, J. A. Piccirilli, and M. E. Harris, "Experimental and computational analysis of the transition state for ribonuclease A-catalyzed RNA 2′-*O*-transphosphorylation," Proc. Natl. Acad. Sci. U. S. A. **110**, 13002–13007 (2013).

[24] À. González-Lafont and J. M. Lluch, "Kinetic isotope effects in chemical and biochemical reactions: Physical basis and theoretical methods of calculation," Wiley Interdiscip. Rev.: Comput. Mol. Sci. **6**, 584–603 (2016).

[25] O. Redlich, "Eine allgemeine beziehung zwischen den schwingungsfrequenzen isotoper molekeln," Z. Phys. Chem. **28B**, 371–382 (1935).

[26] W. R. Angus, C. R. Bailey, C. K. Ingold, A. H. Leckie, C. G. Raisin, J. W. Thompson, and C. L. Wilson, "Infra-red spectrum of hexadeuterobenzene and the structure of benzene," Nature **136**, 680 (1935).

[27] W. R. Angus, A. H. Leckie, C. R. Bailey, C. G. Raisin, J. L. Gleave, C. L. Wilson, and C. K. Ingold, "Raman spectra of deuterobenzenes and the structure of benzene," Nature **135**, 1033–1034 (1935).

[28] M. Ceriotti and T. E. Markland, "Efficient methods and practical guidelines for simulating isotope effects," J. Chem. Phys. **138**, 014112 (2013).

[29] D. T. Major and J. Gao, "An integrated path integral and free-energy perturbation-umbrella sampling method for computing kinetic isotope effects of chemical reactions in solution and in enzymes," J. Chem. Theory Comput. **3**, 949–960 (2007).

[30] J. Gao, "Computation of kinetic isotope effects for enzymatic reactions," Sci. China. Chem. **54**, 1841–1850 (2011).

[31] Y. Fan, A. Cembran, S. Ma, and J. Gao, "Connecting protein conformational dynamics with catalytic function as illustrated in dihydrofolate reductase," Biochemistry **52**, 2036–2049 (2013).

[32] K.-Y. Wong, H. Gu, S. Zhang, J. A. Piccirilli, M. E. Harris, and D. M. York, "Characterization of the reaction path and transition states for RNA transphosphorylation models from theory and experiment," Angew. Chem., Int. Ed. **51**, 647–651 (2012).

[33] P. Paneth and A. Dybala-Defratyka, "Isotope effects as analytical probes: Applications of computational theory," in *Computational Techniques for Analytical Chemistry and Bioanalysis* (The Royal Society of Chemistry, 2020).

[34] J. Bigeleisen and M. G. Mayer, "Calculation of equilibrium constants for isotopic exchange reactions," J. Chem. Phys. **15**, 261–267 (1947).

[35] V. López-Canut, J. Ruiz-Pernía, I. Tuñón, S. Ferrer, and V. Moliner, "Theoretical modeling on the reaction mechanism of *p*-nitrophenylmethylphosphate alkaline hydrolysis and its kinetic isotope effects," J. Chem. Theory Comput. **5**, 439–442 (2009).

[36] D. G. Truhlar, J. Gao, M. Garcia-Viloca, C. Alhambra, J. Corchado, M. Luz Sanchez, and T. D. Poulsen, "Ensemble-averaged variational transition state theory with optimized multidimensional tunneling for enzyme kinetics and other condensed-phase reactions," Int. J. Quantum Chem. **100**, 1136–1152 (2004).

[37] C. Alhambra, J. Corchado, M. L. Sánchez, M. Garcia-Viloca, J. Gao, and D. G. Truhlar, "Canonical variational theory for enzyme kinetics with the protein mean force and multidimensional quantum mechanical tunneling dynamics. Theory and application to liver alcohol dehydrogenase," J. Phys. Chem. B **105**, 11326–11340 (2001).

[38] D. G. Truhlar, J. Gao, C. Alhambra, M. Garcia-Viloca, J. Corchado, M. L. Sánchez, and J. Villà, "The incorporation of quantum effects in enzyme kinetics modeling," Acc. Chem. Res. **35**, 341–349 (2002).

[39]N. Kanaan, S. Ferrer, S. Martí, M. Garcia-Viloca, A. Kohen, and V. Moliner, "Temperature dependence of the kinetic isotope effects in thymidylate synthase. A theoretical study," J. Am. Chem. Soc. **133**, 6692–6702 (2011).
[40]J. Pu, J. Gao, and D. G. Truhlar, "Multidimensional tunneling, recrossing, and the transmission coefficient for enzymatic reactions," Chem. Rev. **106**, 3140–3169 (2006).
[41]A. Kuppermann and D. G. Truhlar, "Exact tunneling calculations," J. Am. Chem. Soc. **93**, 1840–1851 (1971).
[42]B. C. Garrett and D. G. Truhlar, "Semiclassical tunneling calculations," J. Phys. Chem. **83**, 2921–2926 (1979).
[43]M. Garcia-Viloca, C. Alhambra, D. G. Truhlar, and J. Gao, "Hydride transfer catalyzed by xylose isomerase: Mechanism and quantum effects," J. Comput. Chem. **24**, 177–190 (2003).
[44]J. Pang, J. Pu, J. Gao, D. G. Truhlar, and R. K. Allemann, "Hydride transfer reaction catalyzed by hyperthermophilic dihydrofolate reductase is dominated by quantum mechanical tunneling and is promoted by both inter- and intramonomeric correlated motions," J. Am. Chem. Soc. **128**, 8015–8023 (2006).
[45]Q. Liu, J. A. Tossell, and Y. Liu, "On the proper use of the Bigeleisen–Mayer equation and corrections to it in the calculation of isotopic fractionation equilibrium constants," Geochim. Cosmochim. Acta **74**, 6965–6983 (2010).
[46]I. H. Williams, "Kinetic isotope effects from QM/MM subset Hessians: 'Cut-off' analysis for S_N2 methyl transfer in solution," J. Chem. Theory Comput. **8**, 542–553 (2012).
[47]J. Cao and G. A. Voth, "The formulation of quantum statistical mechanics based on the Feynman path centroid density. IV. Algorithms for centroid molecular dynamics," J. Chem. Phys. **101**, 6168–6183 (1994).
[48]M. Ceriotti, J. More, and D. E. Manolopoulos, "i-PI: A Python interface for ab initio path integral molecular dynamics simulations," Comput. Phys. Commun. **185**, 1019–1026 (2014).
[49]M. Ceriotti and D. E. Manolopoulos, "Efficient first-principles calculation of the quantum kinetic energy and momentum distribution of nuclei," Phys. Rev. Lett. **109**, 100604 (2012).
[50]M. Ceriotti, D. E. Manolopoulos, and M. Parrinello, "Accelerating the convergence of path integral dynamics with a generalized Langevin equation," J. Chem. Phys. **134**, 084104 (2011).
[51]K.-Y. Wong, "Simulating biochemical physics with computers," Ph.D. thesis, University of Minnesota, 2008.
[52]K.-Y. Wong, "Review of Feynman's path integral in quantum statistics: From the molecular Schrödinger equation to Kleinert's variational perturbation theory," Commun. Comput. Phys. **15**, 853–894 (2014).
[53]M. Rostkowski, H. K. V. Schürner, A. Sowińska, L. Vasquez, M. Przydacz, M. Elsner, and A. Dybala-Defratyka, "Isotope effects on the vaporization of organic compounds from an aqueous solution–insight from experiment and computations," J. Phys. Chem. B **125**, 13868–13885 (2021).
[54]J. Kästner and W. Thiel, "Bridging the gap between thermodynamic integration and umbrella sampling provides a novel analysis method: 'Umbrella integration,'" J. Chem. Phys. **123**, 144104 (2005).
[55]B. W. Zhang, J. Xia, Z. Tan, and R. M. Levy, "A stochastic solution to the unbinned WHAM equations," J. Phys. Chem. Lett. **6**, 3834–3840 (2015).
[56]M. R. Shirts and J. D. Chodera, "Statistically optimal analysis of samples from multiple equilibrium states," J. Chem. Phys. **129**, 124105 (2008).
[57]T.-S. Lee, B. K. Radak, A. Pabis, and D. M. York, "A new maximum likelihood approach for free energy profile construction from molecular simulations," J. Chem. Theory Comput. **9**, 153–164 (2013).
[58]T.-S. Lee, B. K. Radak, M. Huang, K.-Y. Wong, and D. M. York, "Roadmaps through free energy landscapes calculated using the multidimensional vFEP approach," J. Chem. Theory Comput. **10**, 24–34 (2014).
[59]E. Vanden-Eijnden and M. Venturoli, "Revisiting the finite temperature string method for the calculation of reaction tubes and free energies," J. Chem. Phys. **130**, 194103 (2009).
[60]W. E, W. Ren, and E. Vanden-Eijnden, "Finite temperature string method for the study of rare events," J. Phys. Chem. B **109**, 6688–6693 (2005).
[61]L. Maragliano, A. Fischer, E. Vanden-Eijnden, and G. Ciccotti, "String method in collective variables: Minimum free energy paths and isocommittor surfaces," J. Chem. Phys. **125**, 024106 (2006).
[62]Y. Zhou, P. Ojeda-May, M. Nagaraju, B. Kim, and J. Pu, "Mapping free energy pathways for ATP hydrolysis in the *E. coli* ABC transporter HlyB by the string method," Molecules **23**, 2652 (2018).
[63]C. A. Gonzalez, T. C. Allison, and F. Louis, "General expression for the effective mass in the one-dimensional treatment of tunneling corrections," J. Phys. Chem. A **105**, 11034–11040 (2001).
[64]D. G. Truhlar and B. C. Garrett, "Reduced mass in the one-dimensional treatment of tunneling," J. Phys. Chem. A **107**, 4006–4007 (2003).
[65]M. Wolfsberg and M. J. Stern, "Validity of some approximation procedures used in the theoretical calculation of isotope effects," Pure Appl. Chem. **8**, 225–242 (1964).
[66]M. J. Stern and M. Wolfsberg, "Simplified procedure for the theoretical calculation of isotope effects involving large molecules," J. Chem. Phys. **45**, 4105–4124 (1966).
[67]L. Melander and W. H. Saunders, Jr., *Reaction Rates of Isotopic Molecules* (John Wiley and Sons, New York, 1980).
[68]E. K. Thornton and E. R. Thornton, *Isotope Effects in Chemical Reactions* (Van Nostrand Reinhold, New York, 1970), Chap. 4, pp. 213–285.
[69]G. D. Ruggiero, S. J. Guy, S. Martí, V. Moliner, and I. H. Williams, "Vibrational analysis of the chorismate rearrangement: Relaxed force constants, isotope effects and activation entropies calculated for reaction in vacuum, water and the active site of chorismate mutase," J. Phys. Org. Chem. **17**, 592–601 (2004).
[70]H. Li and J. H. Jensen, "Partial Hessian vibrational analysis: The localization of the molecular vibrational energy and entropy," Theor. Chim. Acta **107**, 211–219 (2002).
[71]Y. Tao, W. Zou, S. Nanayakkara, M. Freindorf, and E. Kraka, "A revised formulation of the generalized subsystem vibrational analysis (GSVA)," Theor. Chem. Acc. **140**, 31 (2021).
[72]K. Swiderek, A. Dybala-Defratyka, and D. R. Rohr, "A new scheme to calculate isotope effects," J. Mol. Model. **17**, 2175–2182 (2011).
[73]T. J. Giese, B. A. Gregersen, Y. Liu, K. Nam, E. Mayaan, A. Moser, K. Range, O. N. Faza, C. S. Lopez, A. R. de Lera, G. Schaftenaar, X. Lopez, T.-S. Lee, G. Karypis, and D. M. York, "QCRNA 1.0: A database of quantum calculations of RNA catalysis," J. Mol. Graphics Modell. **25**, 423–433 (2006).
[74]J. Zeng, T. J. Giese, Ş. Ekesan, and D. M. York, "Development of range-corrected deep learning potentials for fast, accurate quantum mechanical/molecular mechanical simulations of chemical reactions in solution," J. Chem. Theory Comput. **17**, 6993–7009 (2021).
[75]Q. Cui, M. Elstner, E. Kaxiras, T. Frauenheim, and M. Karplus, "A QM/MM implementation of the self-consistent charge density functional tight binding (SCC-DFTB) method," J. Phys. Chem. B **105**, 569–585 (2001).
[76]M. Elstner, T. Frauenheim, E. Kaxiras, G. Seifert, and S. Suhai, "A self-consistent charge density-functional based tight-binding scheme for large biomolecules," Phys. Status Solidi B **217**, 357–376 (2000).
[77]T. A. Niehaus, M. Elstner, T. Frauenheim, and S. Suhai, "Application of an approximate density-functional method to sulfur containing compounds," J. Mol. Struct.: THEOCHEM **541**, 185–194 (2001).
[78]H. Wang, L. Zhang, J. Han, and W. E, "DeePMD-kit: A deep learning package for many-body potential energy representation and molecular dynamics," Comput. Phys. Commun. **228**, 178–184 (2018).
[79]D. A. Case, K. Belfon, I. Y. Ben-Shalom, S. R. Brozell, D. S. Cerutti, T. E. Cheatham III, V. W. D. Cruzeiro, T. A. Darden, R. E. Duke, G. Giambasu, M. K. Gilson, H. Gohlke, A. W. Goetz, R. Harris, S. Izadi, S. A. Izmailov, K. Kasavajhala, K. Kovalenko, R. Krasny, T. Kurtzman, T. Lee, S. Le-Grand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, V. Man, K. Merz, Y. Miao, O. Mikhailovskii, G. Monard, H. Nguyen, A. Onufriev, F. Pan, S. Pantano, R. Qi, D. R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C. L. Simmerling, N. Skrynnikov, J. Smith, J. Swails, R. C. Walker, J. Wang, R. M. Wilson, R. M. Wolf, X. Wu, Y. Xiong, Y. Xue, D. M. York, and P. A. Kollman, AMBER 20, University of California, San Francisco, CA, 2020.
[80]R. Zwanzig, "Nonlinear generalized Langevin equations," J. Stat. Phys. **9**, 215–220 (1973).
[81]K. Nam, J. Gao, and D. M. York, "An efficient linear-scaling Ewald method for long-range electrostatic interactions in combined QM/MM calculations," J. Chem. Theory Comput. **1**, 2–13 (2005).

[82]T. Kubař, K. Welke, and G. Groenhof, "New QM/MM implementation of the DFTB3 method in the gromacs package," J. Comput. Chem. **36**, 1978–1989 (2015).
[83]T. Darden, D. York, and L. Pedersen, "Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems," J. Chem. Phys. **98**, 10089–10092 (1993).
[84]U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen, "A smooth particle mesh Ewald method," J. Chem. Phys. **103**, 8577–8593 (1995).
[85]T. J. Giese and D. M. York, "A GPU-accelerated parameter interpolation thermodynamic integration free energy method," J. Chem. Theory Comput. **14**, 1564–1582 (2018).
[86]H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. Dinola, and J. R. Haak, "Molecular dynamics with coupling to an external bath," J. Chem. Phys. **81**, 3684–3690 (1984).
[87]T. J. Giese, J. Zeng, and D. M. York, "Multireference generalization of the weighted thermodynamic perturbation method," J. Phys. Chem. A **126**, 8519–8533 (2022).
[88]J. Kästner, J. M. Carr, T. W. Keal, W. Thiel, A. Wander, and P. Sherwood, "DL-FIND: An open-source geometry optimizer for atomistic simulations," J. Phys. Chem. A **113**, 11856–11865 (2009).
[89]J. Baker, "An algorithm for the location of transition states," J. Comput. Chem. **7**, 385–395 (1986).
[90]P. Y. Ayala and H. B. Schlegel, "A combined method for determining reaction paths minima and transition state geometries," J. Chem. Phys. **107**, 375–384 (1997).
[91]C. Choi and R. Elber, "Reaction path study of helix formation in tetrapeptides: Effect of side chains," J. Chem. Phys. **94**, 751–760 (1991).