2014

Approximate Quantum Trajectory Method for Modeling Chemical Reaction Dynamics: Application to Enzymatic Proton Transfer

James William Mazzuca
University of South Carolina - Columbia

Approximate Quantum Trajectory Method for Modeling Chemical Reaction Dynamics: Application to Enzymatic Proton Transfer

by

James William Mazzuca Jr.

Bachelor of Arts
Saint Mary's University of Minnesota, 2009

Submitted in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy in

Chemistry

College of Arts and Sciences

University of South Carolina

2014

Accepted by:

Sophya Garashchuk, Major Professor

Vitaly Rassolov, Committee Member

Ken Shimizu, Committee Member

Andreas Heyden, Committee Member

Lacy Ford, Vice Provost and Dean of Graduate Studies

Dedication

For Kim.

Acknowledgments

I am very grateful to Prof. Sophya Garashchuk for the privilege of doing research in her group for the past five years. Her expertise and professional insight have made me a competent scientist and teacher. She always provided me with the tools and freedom to do my job, and pushed me to excel scientifically and professionally. I continue my career in academia knowing that my experience in her group has prepared me to succeed.

I would like to thank Prof. Vitaly Rassolov for many constructive discussions on theory, and for insight into succeeding in academia. I will be a much better teacher and research advisor as a result of his advice. Thanks to Prof. Jacek Jakowski for his involvement in all stages of the SLO-1 project, and for the opportunity to do collaborative research at Oak Ridge National Lab. Thanks to Prof. Michael Myrick for putting me in a position to educate new teaching assistants and contribute to the USC undergraduate chemistry curriculum. I am grateful to Prof. Linda Shimizu for the opportunity to collaborate on research and bring chemistry to the community. Thanks to theory group members Dr. Tijo Vazhappilly, Dr. Lei Wang, Dr. David Dell'Angelo, Bing Gu, Bryan Nichols, and Brett Cagg for discussing science, providing feedback on my lectures, and for their company during lunch each day.

I would like to thank my parents for encouraging my academic success, and for always offering their support. Finally, a special thanks to my wife, Kim. Her optimism and encouragement were instrumental to the completion of this work, and I look forward to continuing my career in chemistry with her by my side.

Abstract

Molecular dynamics simulations, providing a detailed picture of the reaction mechanism, is an essential tool for theoretical and experimental chemists. In these simulations the nuclei are typically treated as classical particles, but under some conditions (low energies and temperatures, processes involving multiple electronic states) a classical description is inappropriate. Quantum effects of nuclear motion, such as tunneling and zero-point energy can play an important role in determining a reaction mechanism, yet exact quantum dynamics methods are limited to reactive systems of just 3-4 atoms. Central to this work is the development and implementations of an efficient trajectory-based methodology, in which the dominant quantum effects of nuclear motion are included through an approximate "quantum potential" term. A combination of quantum and classical nuclei can be evolved within this approach under the Hamiltonian or Boltzmann operators.

This quantum trajectory (QT) method is applied to the proton transfer in the enzymatic active site of soybean lipoxygenase-1. Experimental evidence suggests that this proton transfer step proceeds by a quantum tunneling mechanism. First, the reaction was examined as occurring within fixed substrate configurations at zero temperature, and the primary H/D kinetic isotope effect was in agreement with exact quantum and experimental results. Next, taking advantage of QT features, the effects of temperature and substrate motion were included into the simulation. Vibrational motion of the linoleic acid substrate was incorporated through on-the-fly density-functional tight-binding (DFTB) electronic structure (ES) calculations. This motion was found to modestly enhance the reaction across the temperatures of 250-

350 K, and in a similar fashion for proton and deuteron. Through application of the quantum-mechanical flux operator and imaginary time evolution, the temperature was incorporated into the proton wavefunction. The experimentally observed weak temperature-dependence of the kinetic isotope effect was reproduced and is understood largely as an effect of the quantum partition function. Linear scaling of the QTES-DFTB code with respect to the number of computing cores (typically run on thousands of cores), makes the developed methodology and code practical to chemical systems of up to 200 atoms.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Chemical reaction dynamics

Dynamics simulations are essential for understanding the time-evolution of chemical systems, and they provide us with valuable insight into reaction mechanisms. In a typical dynamics calculation, we make the Born-Oppenheimer approximation, and represent the total wavefunction of our chemical system as a product of electronic and nuclear pieces. Electrons are fully quantum objects for which we solve the time-independent Schrödinger equation, and nuclei are typically treated as classical, Newtonian particles.

In this framework, the electrons provide a potential and force which act on the nuclei, depending on their configuration. If the potential is calculated for every possible nuclear configuration, we now have a potential energy surface (PES). Local minima represent stable products, and first-order saddle-points are referred to as transition states. We can now view reaction dynamics simply as the evolution of nuclei on a potential energy surface. This framework, with fully quantum electrons and classical nuclei, is adequate for most chemical systems, but there are exceptions.

1.2 Theoretical methods to include quantum effects

In many cases, quantum effects of light nuclei can have a large impact on the reaction. The nuclei may exhibit tunneling, interference, or non-adiabatic behavior, and these effects are ignored in the classical framework. The inclusion of these quantum effects

formally requires solving the time-dependent Schrödinger equation (TDSE) for all nuclei in the system,

$$\hat{H}\psi(\boldsymbol{x},t) = \imath\hbar\frac{\partial\psi(\boldsymbol{x},t)}{\partial t}. \tag{1.1}$$

Traditional methods solve the TDSE on a spatial grid within discrete variable representation, and there is an inherent scaling problem with the system size[1, 2]. Exponentially more effort is required for each quantum degree of freedom, and only very small systems of 4-6 degrees of freedom can be studied using exact quantum dynamics methods. Other methods have been developed which attempt to mitigate this scaling problem, such as the multiconfiguration time-dependent Hartree (MCTDH) method[3, 4, 5]. This method provides a great deal of flexibility based on the importance of quantum effects in each degree of freedom, but it scales exponentially nonetheless.

If the quantum effects of nuclei are viewed as a correction to classical mechanics, the problem can be simplified. In these cases, a semiclassical (SC) approximation would account for these effects in a computationally cheap, and approximate way. Examples include initial value representation methods and ring polymer molecular dynamics[6, 7]. These methods are applicable to high-dimensional systems. A general problem with these methods is the difficulty in evaluating and improving the semiclassical error.

1.3 Approximate quantum trajectory method

We are developing a method which is formally derived from the Schrödinger equation, yet scales like semi-classical methods. Quantum effects are explicitly accounted for, and classical mechanics is a limiting case. We implement a quantum trajectory (**QT**) approach based on the de Broglie-Bohm description of the TDSE[8]. Within this

framework, the wavefunction is written in a polar form[9],

$$\psi(\boldsymbol{x}, t) = A(\boldsymbol{x}, t)e^{\imath S(\boldsymbol{x},t)/\hbar}. \tag{1.2}$$

The wavefunction of a chemical system is represented as an ensemble on Newtonian trajectories (each with index i) with equations of motion

$$\frac{d\boldsymbol{x}_i}{dt} = \frac{\boldsymbol{p}_i}{m}, \qquad \frac{d\boldsymbol{p}_i}{dt} = -\nabla(V+U)|_{\boldsymbol{x}=\boldsymbol{x}_i} \tag{1.3}$$

where $\boldsymbol{x}$ is a vector of the Cartesian coordinates and $\boldsymbol{p}$ is a momentum vector. The quantum potential $U(\boldsymbol{x})$ is defined as

$$U(\boldsymbol{x}) = \frac{-\hbar^2}{2m}\frac{\nabla^2 A(\boldsymbol{x})}{A(\boldsymbol{x})}. \tag{1.4}$$

It is prohibitively difficult to determine the exact form of U in a general case.

We approximate this term through a least-squares fit of the non-classical momentum ∇A, where we define the vector

$$\tilde{\boldsymbol{r}}(\boldsymbol{x}) \approx \frac{\nabla A}{A} \tag{1.5}$$

and the quantum potential can be written as

$$U(\boldsymbol{x}) \approx \frac{-\hbar^2}{2m}(\tilde{\boldsymbol{r}} \cdot \tilde{\boldsymbol{r}} + \nabla \cdot \tilde{\boldsymbol{r}}). \tag{1.6}$$

This term is responsible for all nuclear quantum effects in the dynamics calculations. A similar approach can be used to solve the Boltzmann evolution of a wavefunction through imaginary time.

With this formulation of the quantum potential, a swarm of Newtonian trajectories can be employed to represent a delocalized, time-dependent wavefunction. The

trajectories can be viewed as moving grid points that follow the wavefunction density. The scaling of computational effort required for these calculations is linear with the number of trajectories, and this formulation provides an essentially linear scaling with each additional degree of freedom. The approximate quantum potential calculation is a small addition to the trajectory dynamics, and only one calculation is required to know the quantum potential for the entire trajectory ensemble.

In high-dimensional systems consisting of many atoms, it is desirable to model a mixture of quantum and classical nuclei simultaneously. For example, quantum effects of a carbon nucleus are generally negligible compared to the quantum effects of a hydrogen nucleus, and it would be a very good approximation to treat the carbon nucleus as a totally localized, classical particle throughout the dynamics calculations. For this reason, the method has been generalized to treat classical nuclei through the Ehrenfest approximation.

An obstacle presented by large chemical systems is the availability of a high-dimensional potential energy surface. In general, the calculation of such a surface is prohibitively expensive. We have implemented a density-functional tight-binding (**DFTB**) method to calculate on-the-fly electronic structure (**ES**) for the geometry represented by each trajectory[10, 11]. This way, no analytical surfaces are required for dynamics, and the potential and gradient are calculated only at the required geometries. These electronic structure calculations are performed independently for each trajectory, and for this reason the **QTES-DFTB** method was implemented on supercomputing resources.

1.4 Model systems

The quantum trajectory method was used to study the quantum proton transfer in the active site of soybean lipoxygenase-1 (SLO-1)[12]. The quantum proton was first evolved on analytical three-dimensional potential energy surfaces, representing

particular substrate configurations, as described in Chapter 2. Excited vibrational states of the quantum proton were included in Chapter 3. This model captured the experimental kinetic isotope effect (KIE), but the reaction oscillated indefinitely. In a follow-up study, described in Chapter 4, substrate vibrations were included through the QTES-DFTB method, and reaction rate constants were calculated with the quantum mechanical flux operator. Rate constants were evaluated at several temperatures to determine the effect of substrate vibrations on the proton transfer rate constant and primary kinetic isotope effect.

Chapter 5 details the approximate Boltzmann evolution of quantum trajectories with the momentum-dependent quantum potential. Performance was benchmarked in the quartic well, double well, Morse oscillator, and high-dimensional coupled harmonic oscillators.

1.5 Other projects

The development of a quantum mechanical rate constant operator for a one-dimensional double-well potential is outlined in Chapter 6. We apply this formalism to the proton transfer reaction in the HO–H–CH_3 model system and examine the primary kinetic isotope effect as a function of temperature. Finally, a new thermodynamics experiment was designed for an undergraduate level physical chemistry lab, and this procedure is outlined in Chapter 7. This experiment uses NMR to determine the distribution of N-methylacetamide conformers at several temperatures, and to predict the enthalpy change, ΔH, between the two conformers. Results were verified with straightforward quantum chemistry calculations, and NMR peak assignments were made based on these calculations.

Several other projects were also pursued which are not included as full chapters. Our group investigated the use of a frictional term to simulate energy loss of a quantum system to its environment[13]. Additionally, a collaboration with

Linda Shimizu's experimental organic chemistry group resulted in a publication on the short, strong halogen bonding in co-crystals of pyridyl bis-urea macrocycles and iodoperfluorocarbons[14].

Chapter 2

Ground State Proton Transfer in the SLO-1 Double-Well Potential

2.1 Abstract

An approximate dynamics method, based on donor and acceptor quantum trajectory ensembles, is employed to model hydrogen tunneling and the kinetic isotope effect (KIE) in soybean lipoxygenase-1. The proton is treated as a three-dimensional quantum-mechanical particle moving between the donor and acceptor wells for multiple configurations of the active site. Substitution of the proton with a deuteron reduces the transmission probability, integrated over enzyme configurations, by a factor of 51, which is in reasonable agreement with the experimental value of KIE equal to 81, validating the applicability of the current approach in biological systems. Some of this work appears in Ref. [15], and it is reproduced here with permission from the publisher.

2.2 Introduction

Soybean lipoxygenase-1 (SLO-1) is an enzyme that is chemically responsible for the hydroperoxidation of linoleic acid. The rate-limiting step of this reaction, shown on Fig. 2.1, is the abstraction of a hydrogen atom from the C(11) position of the substrate by the Fe(III)-OH cofactor, and this reaction has been found to exhibit nearly temperature-independent rate constants as well as a very large kinetic isotope effect (KIE) [16, 17, 18]. An experimentally determined KIE of 81 has been measured when

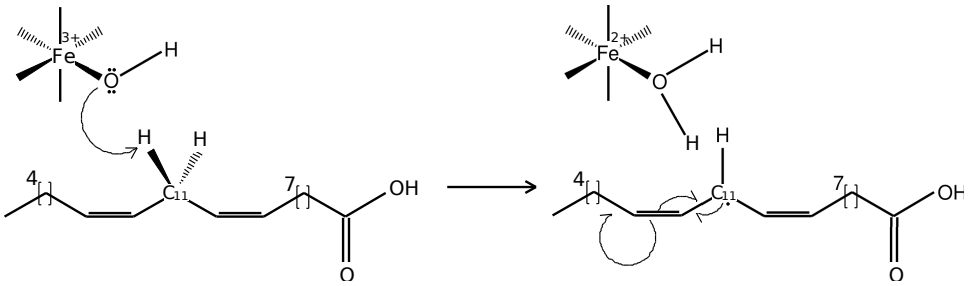

Figure 2.1 Hydrogen abstraction from linoleic acid is the rate-limiting step in the action of SLO-1 .

the transferring hydrogen is replaced with deuterium [16, 17, 19]. Weak dependence of the KIE on temperature points to a fairly rigid active site, while the unusually large value (a typical KIE in CH bond breaking is 5-10) that could not be accounted for by the quantum-mechanical correction to the transition state theory [20] suggested that the abstraction mechanism may occur entirely by quantum tunneling of the transferring proton [17, 21]. This idea sparked theoretical interest in this system and opened doors for the development of methods that must include QM treatment of the transferring proton/deuteron in order to describe this KIE.

Computational studies related to the proton transfer of SLO-1 suggest that enzymes such as this may have evolved to promote tunneling through dynamics in order to facilitate specific reactions [21] or, in other words, that the enzyme specifically acts on the substrate to optimize tunneling, but not necessarily classical transfer. It has been also found in Ref. [19] that the nuclear configuration of the active site of SLO-1 plays a large role in driving the reaction forward by both lowering and narrowing the barrier between donor and acceptor states. Examination of the active site potential energy surfaces while treating the transferring nucleus quantum mechanically has shown that the large KIE is indeed a result of nuclear quantum effects, and the largest effects are seen in the quantum transition state region. The transition state that has been identified through traditional means was not, in terms of the proton transfer reaction, exactly the transition state, but simply a regime in which the classical proton

could overcome the barrier. It has been proposed that a quantum-dynamically corrected transition state may be be shifted toward the reactants relative to the classical transition states.

The cost of traditional time-dependent wavepacket treatment [2] of nuclear motion scales exponentially with the system size, thus we explore the possibility for an alternative dynamics method efficient in high dimensions. We calculate the KIE of the hydrogen-abstraction reaction that occurs in the active site of SLO-1 within the setup of Iyengar *et al* [19], using the approximate quantum trajectory (QT) dynamics with linearized quantum force [22]. This approach is efficient in high dimensions, which will allow us to include motion of the environment in the future. Description of the double-well tunneling dynamics, underlying the proton transfer in SLO-1, is a known challenge for the trajectory-based methods [23]. We overcome this challenge by using two ensembles of QTs, representing the donor and acceptor states, with population transfer between them [24]. This is the first application of the approach, described in Section 2.3, to a multidimensional biochemical system. Section 2.5 contains the details of the numerical implementation, results and discussion. Section 2.7 concludes.

2.3 The approximate quantum trajectory methodology

2.3.1 The quantum trajectory formulation with linearized quantum force

Our dynamics method, based on the quantum trajectory evolution with linearized quantum force of Garashchuk and Rassolov [22], is outlined below. The QT method has its theoretical roots in the hydrodynamic formulation of quantum mechanics [25, 8]. In general, a quantum object is described by a wavefunction which solves the time-dependent Schrödinger equation,

$$\hat{H}\psi(\boldsymbol{x},t) = \imath\hbar\frac{\partial\psi(\boldsymbol{x},t)}{\partial t}. \tag{2.1}$$

A polar form of a wavepacket,

$$\psi(\boldsymbol{x},t) = A(\boldsymbol{x},t)\exp\left(\frac{\imath}{\hbar}S(\boldsymbol{x},t)\right), \tag{2.2}$$

where amplitude A and phase S are real functions and $\boldsymbol{x}$ represents all position coordinates, a three-dimensional vector $\boldsymbol{x} = (x, y, z)$ in the application below, is substituted into Eq. (2.1). Associating ∇S with the trajectory momentum, $\boldsymbol{p} = \nabla S$, Eqs 2.1 and 2.2 are equivalent to the QT evolution under the combined influence of the classical potential V and the quantum potential U,

$$U(\boldsymbol{x},t) = \frac{-\hbar^2}{2m}\frac{\nabla^2 A(\boldsymbol{x},t)}{A(\boldsymbol{x},t)}. \tag{2.3}$$

In the the regime of $m \to \infty$, when a wavepacket describing a heavy particle does not develop interference on the time scale relevant to the dynamics process, U becomes vanishingly small and the trajectory motion (given by Eq. 2.6) becomes classical as expected.

For numerical implementation the wavefunction is represented by an ensemble of trajectories assigned certain weights, w, based on the initial wavefunction amplitude and the volume (dependent on sampling) associated with each trajectory,

$$w_i = \psi^*(\boldsymbol{x}_i,t_0)\psi(\boldsymbol{x}_i,t_0)d\boldsymbol{x}_i(t_0) = A^2(\boldsymbol{x}_i,t_0)d\boldsymbol{x}_i(t_0). \tag{2.4}$$

The space of non-zero amplitude is sufficiently sampled so that for a normalized $\psi(\boldsymbol{x},t_0)$ the sum of trajectory weights is equal to unity,

$$\sum_i^{N_{tr}} w_i \approx \int_{-\infty}^{+\infty} \psi^*\psi d\boldsymbol{x} = 1 \tag{2.5}$$

The trajectories evolve according to

$$\frac{d\boldsymbol{x}_i}{dt} = \frac{\boldsymbol{p}_i}{m}, \quad \frac{d\boldsymbol{p}_i}{dt} = -\nabla\,(V+U)|_{\boldsymbol{x}=\boldsymbol{x}_i}\,, \tag{2.6}$$

and their weights are constant in time [22]. The wavefunction phase at $\boldsymbol{x}_i$ is equal to the action function S_i of each trajectory defined (in units of $\hbar$) by

$$\frac{dS_i}{dt} = \frac{\boldsymbol{p}_i \cdot \boldsymbol{p}_i}{2m} - (V+U)|_{\boldsymbol{x}=\boldsymbol{x}_i} \tag{2.7}$$

Within the trajectory discretization of a wavefunction the position-dependent expectation values are easy to compute,

$$\langle \Omega \rangle = \sum_{i=1}^{N_{tr}} w_i \Omega(\boldsymbol{x}_i). \tag{2.8}$$

The QT formulation of Eqs 2.4 and 2.6 in principle gives the exact QM wavefunction, but accurate calculation of U (responsible for all QM effects!) and its gradient is impractical. For this reason an approximate quantum potential is defined from the global linear least-squares fitting of the nonclassical component of the momentum operator $\boldsymbol{r}$ [22],

$$\boldsymbol{r} \equiv \frac{\nabla A}{A} \approx \tilde{\boldsymbol{r}}(\boldsymbol{x}), \tag{2.9}$$

to become

$$U \approx \frac{-\hbar^2}{2m}(\tilde{\boldsymbol{r}} \cdot \tilde{\boldsymbol{r}} + \nabla \cdot \tilde{\boldsymbol{r}}). \tag{2.10}$$

The least squares fit [26] minimizes $\langle (\boldsymbol{r} - \tilde{\boldsymbol{r}})^2 \rangle$, where $\tilde{\boldsymbol{r}}$ is represented in a linear basis $\boldsymbol{f}$. For a three dimensional system the basis functions can be arranged as a vector $\boldsymbol{f} = (1, x, y, z)$, so the approximate nonclassical momentum components are expressed as

$$\tilde{\boldsymbol{r}} = \mathbf{C}\boldsymbol{f}, \tag{2.11}$$

where $\mathbf{C}$ solves the matrix equation

$$\mathbf{M}\ \mathbf{C} = -\frac{1}{2}\mathbf{F}'. \tag{2.12}$$

The matrices are defined by the outer product of vectors

$$\mathbf{M} = \langle \boldsymbol{f} \otimes \boldsymbol{f} \rangle,\ \mathbf{F}' = \langle \nabla \otimes \boldsymbol{f} \rangle^T \tag{2.13}$$

which, when expanded, are

$$\boldsymbol{M} = \begin{pmatrix} \langle 1 \rangle & \langle x \rangle & \langle y \rangle & \langle z \rangle \\ \langle x \rangle & \langle x^2 \rangle & \langle xy \rangle & \langle xz \rangle \\ \langle y \rangle & \langle xy \rangle & \langle y^2 \rangle & \langle yz \rangle \\ \langle z \rangle & \langle xz \rangle & \langle yz \rangle & \langle z^2 \rangle \end{pmatrix}, \quad \boldsymbol{F'} = \begin{pmatrix} \langle 0 \rangle & \langle 0 \rangle & \langle 0 \rangle \\ \langle 1 \rangle & \langle 0 \rangle & \langle 0 \rangle \\ \langle 0 \rangle & \langle 1 \rangle & \langle 0 \rangle \\ \langle 0 \rangle & \langle 0 \rangle & \langle 1 \rangle \end{pmatrix} \tag{2.14}$$

The approximate quantum potential defined by Eqs 2.10-2.13 is simply a quadratic function of $\boldsymbol{x}$ yielding a linear quantum force for every trajectory. This approximation rigorously conserves energy and is exact for Gaussian wavepackets, but does not presume that $\psi(\boldsymbol{x}, t)$ is necessarily a Gaussian wavefunction. (Some other approaches based on the QT formalism can be found in Refs [27, 28, 29, 30, 31, 32].) This simple approximation captures 'soft' QM effects, such as wavepacket bifurcation, moderate tunneling, zero-point energy, but not the quantum-mechanical (QM) interference [33]. Therefore, to treat the proton tunneling dynamics in SLO-1 which exhibits the double-well character we have used the trajectory dynamics driven by two sets of quantum trajectories representing the donor and acceptor wells.

2.3.2 The two-component description of the donor/acceptor dynamics

In regions of very low probability density, such as near a node, the quantum potential is generally singular because of the vanishingly small denominator $A(\boldsymbol{x}, t)$ in Eq.

2.3, which destroys the simulation [34, 30]. In the case of the approximate quantum potential of Eqs 2.10-2.13, the dynamics are stable, but $U(\boldsymbol{x}, t)$ is inaccurate and the trajectories decohere. One way to deal with this problem is to prevent the node formation by treating the double-well potential as two single-well potentials that overlap one another. Instead of one wavepacket moving across two wells, we can simultaneously evolve separate wavepackets in each of them and allow for the transfer of wavefunction density between them [24]. The total normalized wavefunction of the system becomes a superposition of the wavefunctions in each well defined by the complex "population" coefficients,

$$\psi(\boldsymbol{x}, t) = c_1(t)\phi_1(\boldsymbol{x}, t) + c_2(t)\phi_2(\boldsymbol{x}, t). \tag{2.15}$$

Above, the wavefunction components ϕ_n are normalized, time-dependent functions represented via the QTs experiencing the potential V_n. Functions ϕ_n, represented in the polar form,

$$\phi_n = A_n \exp\left(\frac{\imath}{\hbar} S_n\right), \quad n = 1, 2 \tag{2.16}$$

solve the time-dependent Schrödinger Eq. 2.1 for $V = V_n$ as described in Section 2.3.1.

The time-dependence of population coefficients following from Eq. 2.1 (with the full V) is

$$\imath\hbar \mathbf{S}\dot{\boldsymbol{c}} = \mathbf{V}\boldsymbol{c} \tag{2.17}$$

where we have used the coefficient vector, $\boldsymbol{c} = (c_1, c_2)$ and the matrices,

$$\mathbf{S} = \begin{pmatrix} \langle\phi_1|\phi_1\rangle & \langle\phi_1|\phi_2\rangle \\ \langle\phi_2|\phi_1\rangle & \langle\phi_2|\phi_2\rangle \end{pmatrix} \tag{2.18}$$

$$\mathbf{V} = \begin{pmatrix} \langle\phi_1|V - V_1|\phi_1\rangle & \langle\phi_1|V - V_2|\phi_2\rangle \\ \langle\phi_2|V - V_1|\phi_1\rangle & \langle\phi_2|V - V_2|\phi_2\rangle \end{pmatrix} \tag{2.19}$$

The initial wavefunction defines the initial values of the population coefficients $c_n(t_0)$ evolving in time according to Eq. 2.17. The diagonal elements of $\mathbf{S}$ and $\mathbf{V}$ are easily evaluated as averages. Evaluation of the off-diagonal elements, however, requires approximate steps since we do not know values of ϕ_1 at the positions of the trajectories describing ϕ_2 and vice versa. We evaluate these functions from the least squares fitting of $\boldsymbol{p}_n$ and $\boldsymbol{r}_n$ similar to that of Section 2.3.1. Arranging components of the trajectory momentum of the trajectory set n into a matrix $\boldsymbol{p}_n$, the coefficients of the linear fit to each $x, y, z-$component of $\boldsymbol{p}_n$ form a matrix $\mathbf{B}_n$ using the basis and overlap matrix of Eq. 2.13,

$$\mathbf{M}_n\mathbf{B}_n = \langle \boldsymbol{f}_n \otimes \boldsymbol{p}_n \rangle \,. \tag{2.20}$$

Using $\mathbf{B}_n$ the action function S_n is approximated at any position,

$$\tilde{S}_n(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}\mathbf{B}_n\boldsymbol{x}^\dagger + S_n^{(0)}. \tag{2.21}$$

The vector $\boldsymbol{x}$ defines position in the Cartesian space, $\boldsymbol{x} = (x, y, z)$, and † marks a vector of *extended* dimensionality,

$$\boldsymbol{x}^\dagger = (x, y, z, 2) = (\boldsymbol{x}, 2). \tag{2.22}$$

In Eq. 2.21 $S_n^{(0)}$ is a constant of integration found by computing

$$S_n^{(0)} = \langle S_n \rangle - \frac{1}{2}\langle \boldsymbol{x}\mathbf{B}_n\boldsymbol{x}^\dagger \rangle. \tag{2.23}$$

Likewise, the wavefunction amplitude is approximated using the fitting of nonclassical momentum $\boldsymbol{r}_n$ for the nth trajectory set:

$$\tilde{A}_n(\boldsymbol{x}) = N_n \exp\left(\frac{1}{2}\boldsymbol{x}\mathbf{C}_n\boldsymbol{x}^\dagger\right). \tag{2.24}$$

The constant N_n is the normalization constant for ϕ_n.

With these fitted values, the off-diagonal expectation values dependent on the overlap between the two separate functions ϕ_n are approximated in a symmetrized fashion as:

$$\begin{aligned}\langle\Omega\rangle &= \frac{1}{2}\left(\langle\phi_1|\hat{\Omega}|\phi_2\rangle + \langle\phi_2|\hat{\Omega}|\phi_1\rangle^*\right) \\ &\approx \frac{1}{2}\sum_{i=1}^{N_{tr}} w_{1,i}\Omega(\boldsymbol{x}_{1,i})\frac{\tilde{A}_2(\boldsymbol{x}_{1,i})}{\tilde{A}_1(\boldsymbol{x}_{1,i})}e^{-\frac{\imath}{\hbar}\left(S_{1,i}-\tilde{S}_2(\boldsymbol{x}_{1,i})\right)} \\ &+ \frac{1}{2}\sum_{i=1}^{N_{tr}} w_{2,i}\Omega(\boldsymbol{x}_{2,i})\frac{\tilde{A}_1(\boldsymbol{x}_{2,i})}{\tilde{A}_2(\boldsymbol{x}_{2,i})}e^{-\frac{\imath}{\hbar}\left(S_{2,i}-\tilde{S}_1(\boldsymbol{x}_{2,i})\right)}\end{aligned} \tag{2.25}$$

In Eq. 2.25, the subscript i labels quantities computed along a particular trajectory (position and action) while the quantities with a tilde are fitted functions evaluated at the trajectory positions. The numerical subscripts of 1 or 2 label the set of trajectories from the corresponding state. For example, $\tilde{A}_2(\boldsymbol{x}_{1,i})$ is the fitted amplitude of ϕ_2 at the position $\boldsymbol{x}_{1,i}$ of the ith trajectory describing ϕ_1. The population transfer from donor to acceptor wells can be expressed as changes in wavefunction coefficients over time, the approach we use to study the proton transfer in SLO-1.

2.4 Split-operator exact quantum methodology

Since this work was done to test the effectiveness of the approximate quantum trajectory methodology on a biological system, it was important to verify whether or not the results of the simulations were physically realistic as well as numerically accurate. To this end, all potential energy surfaces examined by the AQP method were also

examined with a grid-bases split-operator method, which we take to be the "exact" solution to the systems that we study. A brief overview of this method can be found in Ref. [35].

We begin by describing the solution to the time-dependent Schrödinger equaton (Eq. 2.1) as

$$\psi(\boldsymbol{x}, t_f) = \hat{U}(t)\psi(\boldsymbol{x}, 0) = \exp\left(-\frac{\imath}{\hbar}\int_0^{t_f} \hat{H}(t)dt\right)\psi(\boldsymbol{x}, 0) \tag{2.26}$$

and for a single, finite time step we can write

$$\psi(\boldsymbol{x}, t + \Delta t) = \exp\left(-\frac{\imath}{\hbar}\hat{H}(t)\Delta t\right)\psi(\boldsymbol{x}, t). \tag{2.27}$$

It is computationally difficult to directly apply this operator in a single time step as a result of the Hamiltonian's dependence on both the kinetic energy (a function of momentum space) and potential energy (a function of position space). It is easiest to solve for these energy contributors if the wavefunction is (fast) Fourier transformed (FFT) between momentum and position space when the respective operators are being evaluated. If we take these steps, we can write our propagation operator as

$$\begin{aligned}\hat{U}(t) &= \exp\left(-\frac{\imath}{\hbar}\hat{H}(t)\Delta t\right) \\ &\approx \xrightarrow{\text{FFT}} \exp\left(-\frac{\imath}{\hbar}\hat{K}(t)\Delta t/2\right) \xrightarrow{\text{FFT}^{-1}} \exp\left(-\frac{\imath}{\hbar}\hat{V}(t)\Delta t\right) \xrightarrow{\text{FFT}} \exp\left(-\frac{\imath}{\hbar}\hat{K}(t)\Delta t/2\right)\end{aligned} \tag{2.28}$$

where each arrow requires that the FFT or inverse FFT of the wavefunction is evaluated before the operator is applied. The separation of the operator as shown in Eq. 2.28 is commonly referred to as the split-operator method, and its error is quadratic with Δt.

As can be clearly seen, this method is an approximation as a result of discretized time steps. The wavefunction itself is also discretized onto a grid with a finite grid spacing. The scaling of the computational effort depends linearly on the number of grid points (N) and the number of grid points typically depend exponentially with the degrees of freedom (D), so the scaling is overall exponential as N^D. This scaling makes calculations for more than two fully three-dimensional particles (6 degrees of freedom) unfeasible. It is only applicable to very low-dimensional systems, thus our method is a desirable alternative for large systems since the AQP method scales linearly with degrees of freedom for a given number of trajectories [22].

The split-operator method is used as a benchmark method with which to compare against our AQP simulations because of its systematically improvable nature. At an infinitely small time step, and an infinite grid, this method would be exact. We need only be sure, for our simulations, that the grid adequately covers the area of interest, that grid points are sufficiently close to one-another, and that the time step is adequately small so that our results are converged with respect to those parameters of the calculation.

2.5 Simulation of the proton transfer in SLO-1

2.5.1 The proton potential energy surfaces

To study QM effects on the proton transfer in SLO-1 we use the potential energy surfaces (PESs) describing the regime where quantum effects are important. Jakowski and coworkers have developed a series of such surfaces [19] constructed as follows. First, the OH distance, R_{OH}, is fixed and the lowest energy geometry of the environmental nuclei and the proton is determined with the constrained R_{OH}. This relaxation is done for various OH distances, i. e. for various positions of the transferring proton

along the reaction coordinate R_{rc},

$$R_{rc} = \frac{R_{CH} - R_{OH}}{R_{CO}}. \tag{2.29}$$

We use label C to refer to the donor carbon nucleus, O for the acceptor oxygen nucleus, and H for the transferring proton. Then, the environmental nuclei of each structure, deemed favorable to tunneling, were frozen and a three-dimensional PES dependent on the position of H was constructed. It has been assumed that the environmental motion is negligibly slow compared to the proton motion, and as such the environmental nuclei can be frozen throughout the proton transfer. The electronic structure calculations were performed at the B3LYP/lanl2dz level of theory, and results were compared to a similar set of MP2 calculations which showed good agreement [19]. For practical reasons, the electronic structure data were fit with quadratic functions, V_D and V_A, centered near the minima of the donor (D) and acceptor (A) wells, respectively. The global potential energy surface was defined by switching between the two functions at any position of H,

$$V = \begin{cases} V_D & \text{if} \quad V_D < V_A \\ V_A & \text{if} \quad V_A < V_D \end{cases}. \tag{2.30}$$

In the dynamics study we focus on PESs deemed to be the most important in the tunneling regime as they have nearly equal distribution of the nuclear ground state wavefunctions between the donor and acceptor wells. This regime is referred to as the "quantum transition state" (QTS) because it occurs at a different location along the reaction coordinate than classical transition state theory predicts [19]. The 24 QTS surfaces – several are sketched in Figure 2.2 – cover the range of R_{OH} from 1.49 to 1.70 Å at increments of 0.01 Å. The proton dynamics were performed on all 24 PESs.

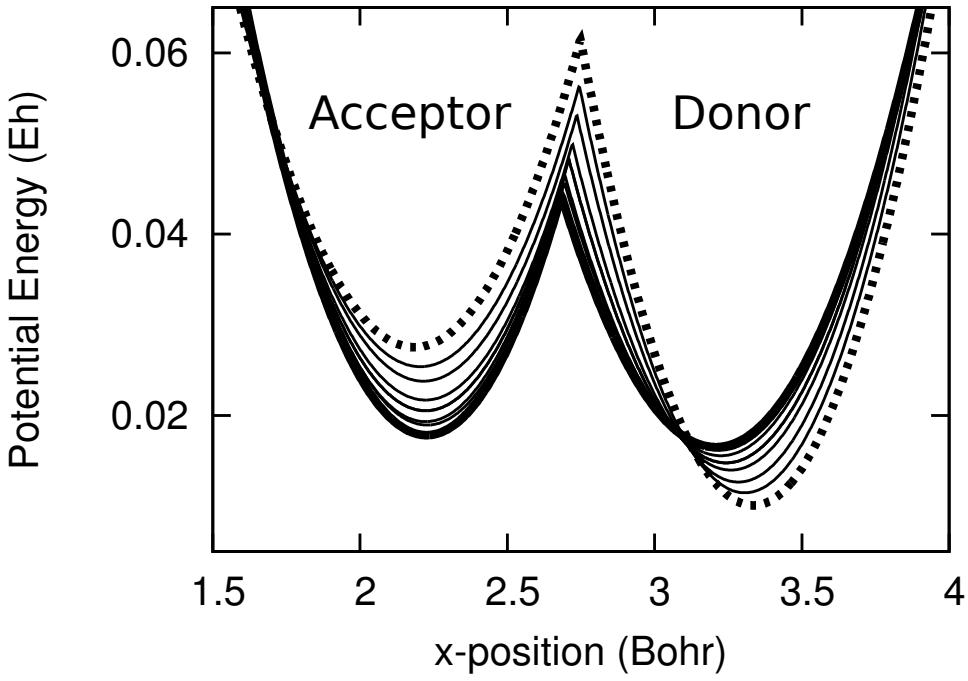

Figure 2.2 Several 1-D slices of potential energy surfaces in the x-direction for $R_{OH} = 1.60$ Å (dashed, close to the pure donor state), $R_{OH} = 1.498$ Å (thick, very close to QTS), and several surfaces in between.

One can expect (using a surface with nearly equal well depths as an example) the donor well to have a fundamental frequency of approximately 3320 K in the direction of the proton transfer, and 1800 K in the two other modes. The ground-state energy of the transferring proton is approximately 3450 K with the first excited state having an energy of 5200 K. When compared to a barrier height of 9500 K it is quite obvious that in terms of the transferring proton, it is not unreasonable to treat the system as if $T = 0$ K even though the experimental measurements were performed at room temperature [17].

2.5.2 Details of implementation

Within the setup of Section 2.5.1 the transferring proton is the only quantum nucleus, described in three-dimensional Cartesian space (x, y, z). The x-coordinate nearly coincides with the reaction coordinate, meaning that a transition between donor and

acceptor states occurs as the wavefunction propagates in the negative x-direction, as shown in Figure 2.2. The approximate QT calculations are compared to exact QM dynamics performed with the conventional Fast Fourier Transform split-operator method [35, 36] implemented on a fixed grid. The parameters are given in Table 2.1.

Table 2.1 Parameters for the FFT quantum dynamics

Grid range [Bohr]	Number of grid points
$x = [-1.0, 5.0]$	128
$y = [-3.0, 3.0]$	64
$z = [-1.5, 4.5]$	64
Final time $t_f = 40000$ a.u.	Time step 2 a.u.

The initial wavepacket was taken as the ground state of V_D, a real Gaussian wavefunction, and propagated up to about 1 ps.

2.5.3 Determination of α-values

To treat the system as simply as possible as well as to prevent interference and node-formation in the total QM simulation, the $A(\boldsymbol{x}, t_0)$ was initialized as the ground state Gaussian function that is an eigenstate of whatever well it originated in. To that end, the parameters for the ground state of each well, namely the α-values for each well had to be identified. This would be very easy if the normal modes of each well were in the direction of our $\boldsymbol{x}$ basis vectors, but each potential energy well is rotated to optimally describe the chemical system, and this direction is arbitrary relative to our set of coordinates. (It should be noted, however, that an effort was made to orient the reaction coordinate in the x-direction of our simulations.) In order to determine which α-values to use in our simulation, where the solution to each well is represented as

$$\phi(\boldsymbol{x},t) = \exp\Big(-\alpha_x(x-x_0)^2 - \alpha_y(y-y_0)^2 - \alpha_z(z-z_0)^2 \qquad (2.31)$$
$$- \alpha_{xy}(x-x_0)(y-y_0) - \alpha_{xz}(x-x_0)(z-z_0) - \alpha_{yz}(y-y_0)(z-z_0)\Big)$$
$$\times \exp\Big(\imath p_x(x-x_0) + \imath p_y(y-y_0) + \imath p_z(z-z_0)\Big),$$

the α-values that correspond to the normal modes had to be calculated, and then rotated into our common system of coordinates. We start with the ground state and zero initial momentum.

For each surface, all of the six alpha values in Eq. 2.31 had to be calculated. The required information from each surface was:

1) The donor and acceptor force constant matrices **FCD** and **FCA** respectively.

2) The rotation matrix **eAD**.

3) The location of well minima ***vminD*** and ***vminA***. The matrix of alpha values is computed by solving the equation

$$\mathbf{eAD} \times \mathbf{Evec} \times \mathbf{Eval}_\alpha \times \mathbf{Evec}^{-1} \times \mathbf{eAD}^{-1} = \mathbf{M}_\alpha \qquad (2.32)$$

where **eAD** is the coordinate rotation matrix, **Evec** is the matrix of eigenvectors for a particular force constant matrix, $\mathbf{Eval}_\alpha$ is the diagonal matrix of square root eigenvalues for a force constant matrix multiplied by $\frac{\sqrt{m}}{2}$ since in normal coordinates for a harmonic oscillator, $\alpha = \frac{1}{2}\sqrt{km}$. In this way, it is possible to initialize the trajectory simulations in a way that is representative of the exact ground state or a linear combination of the donor or acceptor well ground states.

2.5.4 Calculation of reaction probability

The wavepacket reaction probability is defined by the integral of the wavefunction density on the acceptor side for $x < x^0$,

$$P(t)^{\mathrm{QM}} = \langle\psi|h(x^0 - x)|\psi\rangle. \tag{2.33}$$

The value of $x^0 = 2.5$ Bohr defines a plane separating the donor and acceptor wells, close to the barrier top; $h(x^0 - x)$ in Eq. 2.33 is the Heaviside function.

The approximate QT formalism is implemented in the two-component formulation described in Section 2.3.2. The wavefunction is constructed as a superposition of the two functions evolving in the donor (D) and acceptor (A) wells,

$$\psi(x, y, z, t) = c_D(t)\phi_D(x, y, z, t) + c_A(t)\phi_A(x, y, z, t). \tag{2.34}$$

Both states, ϕ_D and ϕ_A, are represented with ensembles of $20,000$ QTs whose initial positions, shown on Figure 2.3, are generated by a quasi-random Sobol sequence [26] with the wavefunction density $|\phi_n(\boldsymbol{x}_i, 0)|^2 > 10^{-15}$. Initially, at $t_0 = 0$, functions ϕ_D and ϕ_A of Eq. 2.34 are real Gaussians describing the eigenstates of the acceptor and donor wells and the population coefficients are $c_D(0) = 1$ and $c_A(0) = 0$. The cost of approximate QT propagation scales essentially linearly with respect to the number of dimensions and trajectories. The exact QM method involved a grid of $524,288$ points. For the present application, the approximate dynamics calculations were 5-20 times faster than the grid-based QM simulations.

In the trajectory calculation we define the probability of the proton being on the product side as

$$P(t)^{\mathrm{QT}} = |c_A(t)|^2, \tag{2.35}$$

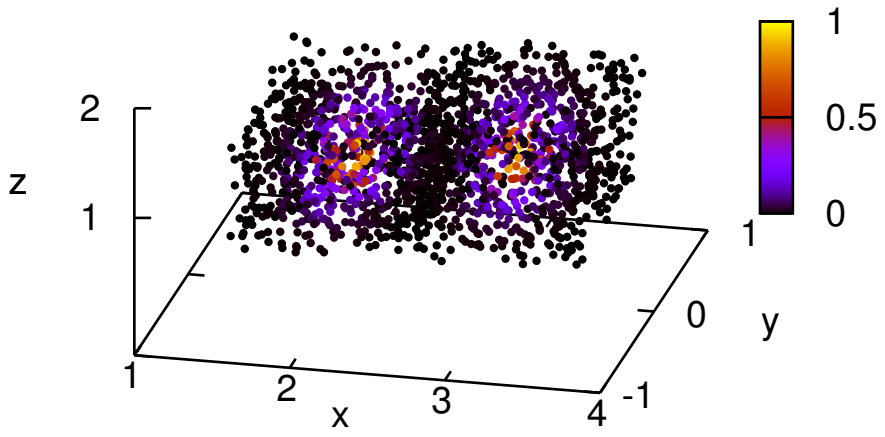

Figure 2.3 Initial sampling of wells with quantum trajectories.

since the typical value of the wavefunction overlap is $\langle\phi_D|\phi_A\rangle \leq 0.005$. For all PESs (that we label by the value of R_{rc} at which the environmental nuclei were frozen) the transmission probability $T(R_{rc})$ is defined as the time-average of the corresponding $P(t)$,

$$T(R_{rc}) = t_f^{-1} \int_0^{t_f} P(t)dt \tag{2.36}$$

In the spirit of Ref. [19] the kinetic isotope effect is defined as the ratio of $T(R_{rc})$ summed over the PESs, computed for the proton to that of the deuteron,

$$\text{KIE} = \frac{\langle T_H \rangle}{\langle T_D \rangle} = \frac{\int T_H(R_{rc})dR_{rc}}{\int T_D(R_{rc})dR_{rc}}. \tag{2.37}$$

The integration is performed over R_{rc} of Eq. 2.29, which measures progress of proton transfer along the reaction path since the surfaces are characterized by different R_{CO} values. It should be noted that definition of the KIE above differs from that in Ref.

[19], where the ratio $T_H(R_{rc})/T_D(R_{rc})$ has been integrated over the surfaces. Eq. 2.37 is more appropriate for computation of absolute transmission probabilities for proton and deuteron, and reduces the error of taking the ratio of two small transmission probabilities.

2.6 Results

The exact and approximate time-evolution was performed for the proton and deuteron on 24 PESs, which contributed to the KIE with varying degree. A comparison of the exact QM and approximate QT probabilities for two representative surfaces is given on Fig. 2.4. The upper panel (a) shows the wavepacket probability and its running average as functions of time for a surface defined by the environment optimized for the reaction coordinate close to the donor side ($R_{OH} = 1.52$ Å). The lower panel (b) shows the same but for a surface with roughly symmetrical donor and acceptor wells ($R_{OH} = 1.505$ Å). The maximal occupations of the acceptor well are 0.07% and 1.5%, respectively. We see that approximate treatment reproduces the exact QM behavior quantitatively in the second case and qualitatively in the first case. The running averages, which give the transmission probability $T(R_{rc})$ at $t = 40,000$ a.u. agree very well in both cases. Overall, we obtain good agreement when the probabilities are greater than 10^{-3}. The relative agreement deteriorates for the surfaces with transmission probabilities of 10^{-4} or below, but these PESs give small contributions to KIE defined by Eq. 2.37.

We attribute this trend to the two-component representation of the approximate wavefunction given by Eq. 2.34, which is more of a limitation for the low probability processes. The evolution of ϕ_D and ϕ_A, however, in this treatment is essentially exact, since V_D and V_A are parabolic wells, in which case the approximate quantum potential of Section 2.3.1 is exact. Some discrepancy can also come from the difference in Eqs 2.33 and 2.35. The latter expression has been introduced, because it does not involve

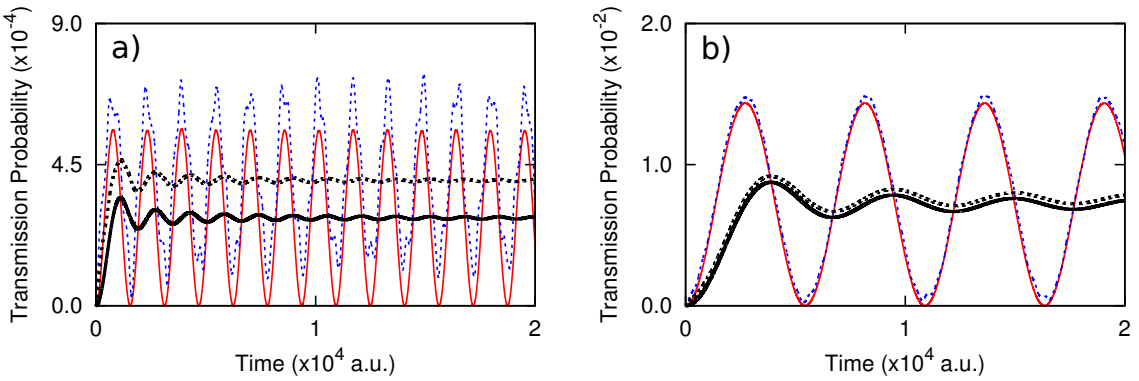

Figure 2.4 Hydrogen transmission probability $P(t)$ for two potential energy surfaces: (a) PES describing proton closer to the donor state for $R_{OH} = 1.52$ Å and (b) nearly symmetric PES for $R_{OH} = 1.505$ Å. On both panels, the results of the approximate QT and exact QM simulations are shown with red solid lines and blue dashes, respectively. The thick solid and dashed (black) lines represent running averages for the same two methods.

estimates of $\langle\phi_D|...|\phi_A\rangle$ terms, with the future anharmonic applications in mind. The values of the total KIE obtained from the exact and approximate simulations are 49 and 51 respectively, which are in good agreement with the experimental results and prior theoretical treatments of SLO-1 system summarized in in Table 2.2. It is clear from the table that our KIE estimates obtained exactly using QM wavepacket dynamics and approximately within the QT formulation, agree with each other and are within the range of experimental results and prior theoretical treatments. Note that our results cannot be directly compared to prior theoretical results since the approaches to this proton transfer problem are very different from one other and from our approach, except the one of Ref. [19]. Even so the KIE of 88 obtained in Ref. [19] for the same PESs was derived from QM dynamics of a wavepacket with the contributions of excited states of the donor well, and was averaged over PESs differently. In the future we will estimate contributions from the excited states of the donor wells and include the zero-point energy and entropy of the active site: our preliminary analysis of these two effects indicates that they may influence the reactivity.

Table 2.2 Comparison of KIE obtained with various methods arranged chronologically. Current work is in bold.

Method	KIE
Experiment [17]	81
CVT with SC tunneling corrections [37]	55
QCP dynamics with EVB surface [18]	100
Hybrid QM(PM3/d-SRP)/MM [38]	90
1-D QCP dynamics with EVB surface [39]	67
ISM/scTST rate calculations [16]	44
Quantum wavepacket, DAF [19]	88
QM dynamics of the donor ground state	**49**
Two-state approximate QT dynamics	**51**

2.7 Conclusions

We have accounted for the QM tunneling effects during the proton transfer step in the hydroperoxidation of linoleic acid by the active site of soybean lipoxygenase-1 within the approximate quantum trajectory framework. The transferring proton was treated as a three-dimensional quantum nucleus whose dynamics initiated in the ground donor state unfolded on a set of 24 electronic potential energy surfaces for different configurations of the active site comprised of 43 atoms. In order to describe the tunneling between the donor and acceptor sites we have employed a two-component formulation of the approximate QT formalism.

An overall KIE of 51 obtained with the approximate QT approach agrees well with our benchmark exact QM calculation (KIE= 49) and with various other simulations ($44 \leq \text{KIE} \leq 100$) as well as with the experimental data (KIE= 81). The agreement between our two-state approximate QT and exact QM dynamics confirms the earlier conclusion of Ref. [19] that the proton transfer in SLO-1 is largely controlled by the overlap of the donor and acceptor eigenstates. Determining the importance of the excited states is in progress. In general, the trajectory-based approach is well suited for allowing of the active site motion and inclusion of water molecules into

future studies. This application contributes to our effort to perform a more complete theoretical study of the proton transfer in SLO-1 and similar systems.

Chapter 3

Excited State Proton Transfer in the SLO-1 Double-Well Potential

3.1 Abstract

The contribution of the first three donor well excited states to the transmission probability and kinetic isotope effect (KIE) in soybean lipoxygenase-1 (SLO-1) has been analyzed. A thermal weighting of each transmission probability is employed which depends on the energy gap between excited states, and therefore included all excited states into the total KIE for the system. The overall KIE only changes by about 1-3, and therefore the excited states do not play a significant role in the rate of proton transfer in SLO-1.

3.2 Introduction

In an effort to include temperature into our simulations of the proton transfer in the active site of SLO-1, the first three (roughly, x, y, and z) excited states were evolved in a similar way as the ground state study in Chapter 2. Average transmission probability was computed for each simulation, and these data were weighted according to a Boltzmann distribution at T = 300 K.

3.3 Inclusion of excited state matrix elements

The inclusion of the excited states in this system is essentially a multi-dimensional, multi-state expansion of the work done by Garashchuk [24]. The matrices increase in

size as a result of an expanded basis. Our wavefunction was previously described as a linear combination of the ground states which comprise each well,

$$\psi(\boldsymbol{x}, t) = c_1\phi_1(\boldsymbol{x}, t) + c_2\phi_2(\boldsymbol{x}, t). \tag{3.1}$$

If we wish to include excited states of the system, the total wavefunction $\psi(\boldsymbol{x}, t)$ needs to include more than just a combination of ground state wavefunctions. To that end, we can redefine our total wavefunction as a linear combination of the ground states of each well plus the first three excited states.

$$\begin{aligned}\psi(\boldsymbol{x}, t) = c_1\phi_1(\boldsymbol{x}, t) + c_2(x - x_0)\phi_1(\boldsymbol{x}, t) + c_3(y - y_0)\phi_1(\boldsymbol{x}, t) \\ + c_4(z - z_0)\phi_1(\boldsymbol{x}, t) + \ c_5\phi_2(\boldsymbol{x}, t) + c_6(x - x_0)\phi_2(\boldsymbol{x}, t) \\ + c_7(y - y_0)\phi_2(\boldsymbol{x}, t) + c_8(z - z_0)\phi_2(\boldsymbol{x}, t)\end{aligned} \tag{3.2}$$

We expect this form of the wavefunction to be capable of fully representing harmonic oscillator excited states, even if the normal modes are in different directions that the unit vectors that define our coordinate system. This is because the first three excited states can be expressed as linear combinations of (x, y, z) modes in our representation of the total wavefunction. This will become apparent in the discussion of the initial conditions of out wavepacket in future sections.

In each state, represented by ϕ_n, the trajectories experience the exact same dynamics as they did in the previous ground state simulations. We only changed the way in which our total wavefunction is expressed as well as how our time-dependent coefficients are propagated. This ensures that the trajectory dynamics will remain as stable as previous simulations. In an attempt to keep the calculations straightforward, we once again attempted to solve the equation

$$\imath\hbar\mathbf{S}\dot{\boldsymbol{c}} = \mathbf{H}\boldsymbol{c} \tag{3.3}$$

where we have redefined the right side of the equation to include kinetic energy terms in the matrix $\mathbf{H}$. We define a basis vector of ϕ interchangeably with how he had previously defined $\boldsymbol{f} = (1, x - x_0, y - y_0, z - z_0)$. Our total basis of ψ can be defined in terms of these smaller basis as

$$\boldsymbol{b} = (f_1\phi_1, ..., f_4\phi_1, f_1\phi_2, ..., f_4\phi_2) \tag{3.4}$$

and we can now define

$$\mathbf{S} = \langle \boldsymbol{b} \otimes \boldsymbol{b} \rangle \tag{3.5}$$

as our total overlap matrix. With this basis defined, the Hamiltonian matrix is defined as

$$\mathbf{H} = \langle \boldsymbol{b} \otimes (\boldsymbol{K} + \boldsymbol{P}) \rangle \tag{3.6}$$

where

$$K_i = -\frac{-\hbar^2}{2m}\left((\nabla^2 f_i)\phi_n + 2\nabla f_i \cdot \nabla \phi_n\right) \tag{3.7}$$

and

$$P_i = (V - V_n)b_i. \tag{3.8}$$

In this representation, we can conveniently express

$$\nabla\phi_n|_{\boldsymbol{x}=\boldsymbol{x}_i} = (\boldsymbol{r}_i + \imath\boldsymbol{p}_i)\phi_n \tag{3.9}$$

and we can therefore calculate all matrix elements in terms of quantum trajectory expectation values.

3.4 Propagation of coefficients to increase stability

For some normalized ψ, the propagation of the coefficients as outlined above and expressed as Eq. 3.3 should be as straightforward as in the ground state calculations.

Previously, we had simply used a single matrix inversion operation to solve Eq. 3.3 and evaluate $\dot{\boldsymbol{c}}$ as

$$\dot{\boldsymbol{c}} = \frac{-\imath}{\hbar}\mathbf{H}\boldsymbol{c}\mathbf{S}^{-1} \tag{3.10}$$

and then the coefficient vector would be propagated as

$$\boldsymbol{c}(t + \Delta t) = \boldsymbol{c}(t) + \dot{\boldsymbol{c}}(t)\Delta t. \tag{3.11}$$

The matrices $\mathbf{S}$ and $\mathbf{H}$ are now 8×8 matrices, whereas they were previously 2×2. This change led to stability issues that caused an accumulation of error that was obvious in the norm conservation. Experimenting with smaller time steps reduced the rate of error accumulation, but the time step size quickly became prohibitively small, and a more exact method of coefficient propagation was required.

In an effort to increase stability, it is helpful to evaluate the matrix equation 3.10 analytically. We can rewrite it as

$$\mathbf{S}\dot{\boldsymbol{c}} = -\imath\mathbf{H}\boldsymbol{c} \tag{3.12}$$

$$\mathbf{S}^{-1}\mathbf{S}\dot{\boldsymbol{c}} = -\imath\mathbf{S}^{-1}\mathbf{H}\boldsymbol{c} \tag{3.13}$$

$$\dot{\boldsymbol{c}} = -\imath\mathbf{S}^{-1}\mathbf{H}\boldsymbol{c} \tag{3.14}$$

$$\boldsymbol{c}(t) = \boldsymbol{c}(0)\exp\left(-\imath\int_0^{t_f}\mathbf{S}^{-1}(t)\mathbf{H}(t)dt\right) \tag{3.15}$$

we now have an analytical integral which we discretize into a sum over time steps.

$$\boldsymbol{c}(t + \Delta t) = \boldsymbol{c}(t)\exp\left(-\imath\mathbf{S}^{-1}(t)\mathbf{H}(t)\Delta t\right) \tag{3.16}$$

Since it is possible to evaluate the exponential of a matrix of eigenvalues as

$$e^{\mathbf{A}} = \mathbf{M} \begin{pmatrix} e^{\lambda_1} & 0 & 0 & 0 \\ 0 & e^{\lambda_2} & 0 & 0 \\ . & . & . & . \\ 0 & 0 & 0 & e^{\lambda_n} \end{pmatrix} \mathbf{M}^{-1} \tag{3.17}$$

where $\mathbf{M}$ is a matrix of eigenvectors that solve the equation

$$\mathbf{AM} = \mathbf{M}\lambda, \tag{3.18}$$

it is possible to write an incremental solution to the coefficient vector,

$$\boldsymbol{c}(t + \Delta t) = \mathbf{M} \exp\left(-\imath\lambda\Delta t\right)\mathbf{M}^{-1}\boldsymbol{c}(t). \tag{3.19}$$

This operation requires several additional steps if we want to propagate coefficients in this way. We need to get eigenvectors and invert complex matrices in several instances. The procedure, with LAPACK routines in parentheses, looks like this:

1) Solve for $\mathbf{S}^{-1}\mathbf{H}$. (ZGESV)

2) Calculate the eigenvalues and eigenvectors $\lambda\mathbf{S}^{-1}\mathbf{H} = \mathbf{M}\lambda$. (ZGEEV)

3) LU decomposition of $\mathbf{M}$ where $\mathbf{LU} = \mathbf{M}$. (ZGETRV)

4) Invert the LU decomposition of $\mathbf{M}$ and performs $(\mathbf{LU})^{-1} = \mathbf{M}^{-1}$. (ZGETRI)

These additional matrix calculation steps approximately double the computational time when compared to the ground-state calculations with 20,000 trajectories. We can expect to see less of an impact when more trajectories are used since the size of the matrices depends on the number of excited states included, and not on the number of trajectories in the simulation.

3.5 Determination of excited states in rotated coordinates

In order to most easily interpret the data generated by these simulations, the system was initialized as eigenstates of the donor well. Since each well was a harmonic oscillator, one can describe excited states as simply a polynomial multiplied by the ground state Gaussian, as shown by each term of Eq. 3.2. Some difficulty arises if the normal modes of the parabolic well are not oriented in the direction of our basis vectors. Ideally, for the x-excited state of the donor well, we would like to simply set all coefficients to 0 except for c_2 which we would set equal to 1 (assuming the inclusion of a normalization constant). Since our normal modes are indeed between our basis vectors, we must express the excited states of this well as a linear combination of our basis-oriented excited states.

Each surface had a unique set of normal modes representing the donor and acceptor state. This means that for each different surface, these modes had to be identified individually in terms of their unique normal mode orientation. In order to accomplish this goal, we need the set of eigenvectors that point in the direction of the normal modes of the well. Conveniently, determination of the α values that were used to describe the ground-state Gaussian wavefunction involved a similar calculation. Where previously we had rotated a diagonal α-matrix into our coordinates, now we must rotate the eigenvectors $\mathbf{F}_v$ that solve the equation

$$\mathbf{F}_\lambda \mathbf{F}_c = \mathbf{F}_v \mathbf{F}_\lambda \tag{3.20}$$

where $\mathbf{F}_c$ is the original force constant matrix, $\mathbf{F}_\lambda$ is the diagonal matrix of eigenvalues, and $\mathbf{F}_v$ is a matrix of normalized eigenvectors

$$\mathbf{F}_v = (\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3)\,. \tag{3.21}$$

By simply rotating each of those three vectors back into our coordinate system, each of the three vectors points in the direction of one of the normal modes. The components of the vectors correspond directly to un-normalized coefficient magnitudes for each excited state. From the force constant matrices and common rotation matrices, we are able to compute the magnitude of all coefficients for each excited state in both the donor and acceptor well.

To include temperature in the overall hydrogen abstraction that occurs in SLO-1, a simulation must be run for each excited state that we are interested in. The dynamics of each excited state are weighted after the simulations according to a Boltzmann distribution.

$$Q = e^{-E_g/kT} + e^{-E_1/kT} + e^{-E_2/kT} + e^{-E_3/kT} \tag{3.22}$$

and we calculate the overall contribution of a particular state n as

$$q_n = \frac{e^{-E_n/kT}}{Q}. \tag{3.23}$$

At first, it would seem that we could simply add these results into our old KIE calculation as corrections to the ground state calculations. This is not the case, as the previous ground state calculations were not capable of accounting for coupling between the excited states that are now available in our extended basis. Therefore, the ground state calculations must be repeated, but now there are three additional excited states available for any nonadiabatic vibrational coupling that may occur. We determine the contributions of each excited state through the assumption that the excited states have the energies

$$E_n = \frac{1}{2}(\omega_1 + \omega_2 + \omega_3) + n_1\omega_1 + n_2\omega_2 + n_3\omega_3 \tag{3.24}$$

which are simply the energy levels of a three-dimensional harmonic oscillator.

3.6 Thermal correction to transmission probabilities

Each excited state was allowed to evolve completely, and average transmission probability was calculated for each simulation. The Boltzmann weighting was applied to the transmission probabilities T_H and T_D for each potential energy surface. Thus, the thermally weighted transmission probability for hydrogen on a surface at R_{rc} is calculated as

$$T_H(R_{rc}) = q_0 T_{H,0}(R_{rc}) + q_1 T_{H,1}(R_{rc}) + q_2 T_{H,2}(R_{rc}) + q_3 T_{H,3}(R_{rc}). \tag{3.25}$$

With these new temperature-dependent transmission probabilities, the total KIE of the system can be calculated in the same way as in Eq. 2.37.

3.7 Results and Conclusions

Transmission probability was recorded for each excited state in the same way as it was for the ground state. The calculated KIE for each excited state is shown in Fig. 3.3. With these thermal corrections, the total calculated KIE of the system changed marginally since for the proton, the system was frequently over 99% ground state, and for deuterium, it was over 97% ground state. The results from these calculations can be seen in Table 3.1.

Table 3.1 KIE results from including first three proton and deuteron excited states

Excited state	Quantum Trajectories	Split-Operator	Relative Error
G(0)	49.78	46.56	0.07
Y(1)	106.05	93.81	0.13
Z(2)	13.11	12.75	0.03
X(3)	10.74	10.16	0.06
Fully Boltzmann Weighted			
Total	47.82	44.60	0.07

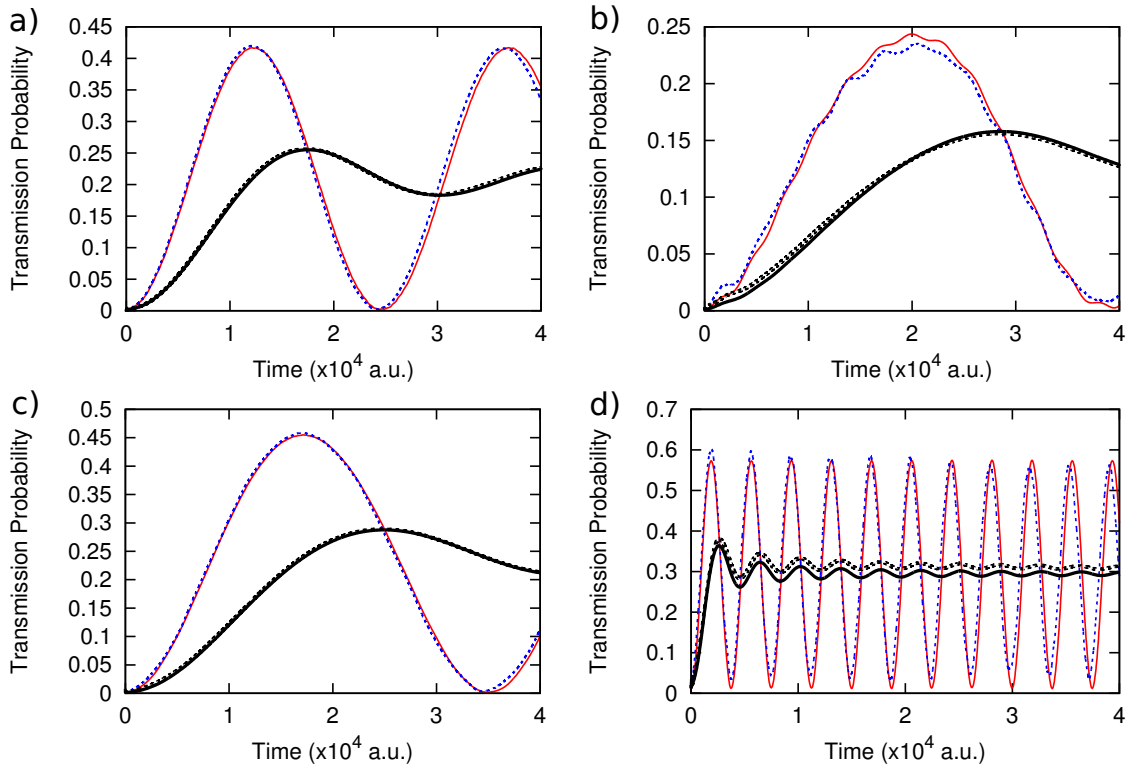

Figure 3.1 Graph showing transmission probability as a function of time for the $R_{OH} = 1.498$ in a) ground-state, b) 1st excited state, c) 2nd excited state, and d) 3rd excited state. Red line is trajectory simulation, blue dashed line is exact split-operator method. Thick black line is running average for trajectory code, thick dashed line is running average for split-operator code.

Although this excited state study did not have a large affect on the overall KIE, we can still take valuable information from the simulations themselves. For example, we can see the degree of coupling between excited states as a function of the absolute value of the excited state coefficients. These results can be seen in Fig 3.4 and 3.5. From these figures it is evident that nuclear motion which couples the various excited states is much more prevalent in the proton case than in the deuteron case. It is also evident that population exchange between these coefficients occurs most rapidly in the ground state an the third excited state in both cases. The large amount of coupling we observe can tell us, in general, that all degrees of freedom influence the rate of transfer, not just the x-direction. This is also shown to be true in Fig. 3.3, where the y-excited state greatly increases the KIE in both of our simulations.

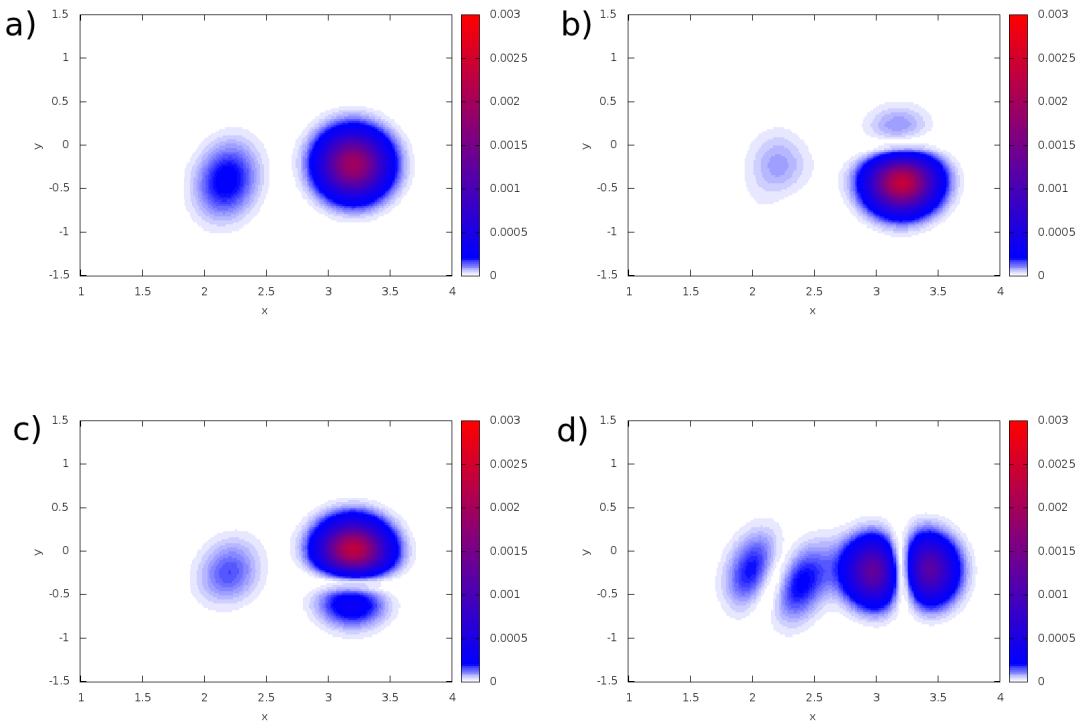

Figure 3.2 Probability density near an area of maximum transient transmission probability, as represented by peaks in Fig 3.1. Snapshots from the split-operator code to reference as physical interpretation of trajectory code. Single frame snapshot from a) ground-state, b) 1st excited state, c) 2nd excited state, and d) 3rd excited state.

We see excellent agreement between trajectory and split-operator methods when transmission probability as a function of time is plotted, as in figure 3.1 and there is never more than a 13% discrepancy between the overall calculation of KIE for any of the excited states, as is shown in Table 3.1. For this particular system, a two-state approximation, as outlined above, appears to be a physically realistic description of the dynamics that we observe experimentally and through exact quantum evolution of the system. This approach is based on stable trajectory dynamics, and it is capable of describing quantum tunneling through a barrier in a bound system.

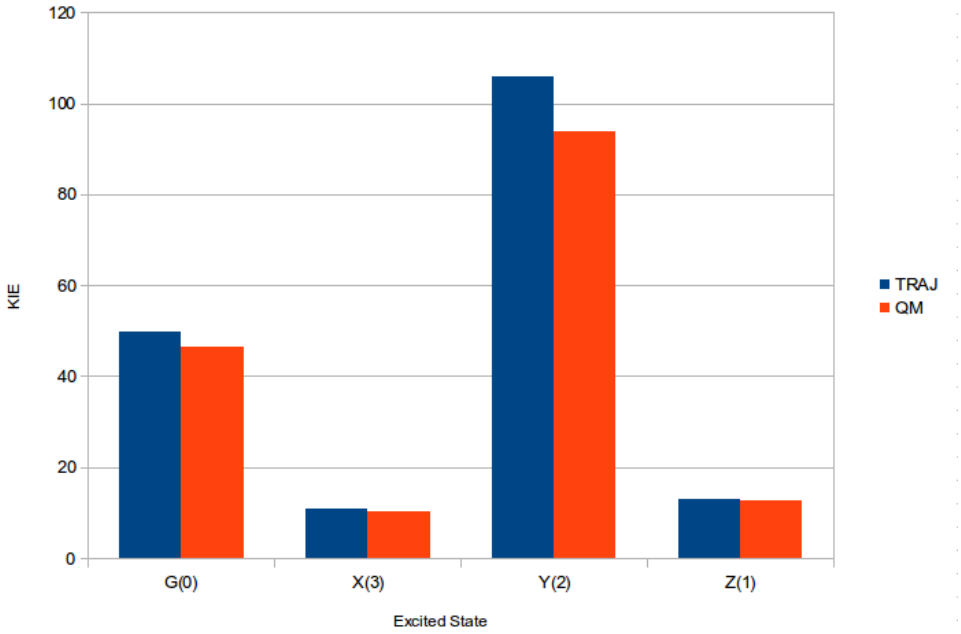

Figure 3.3 Bar graph showing the KIE calculations for each individual excited state between the quantum trajectory method (blue) and the split-operator benchmark (orange).

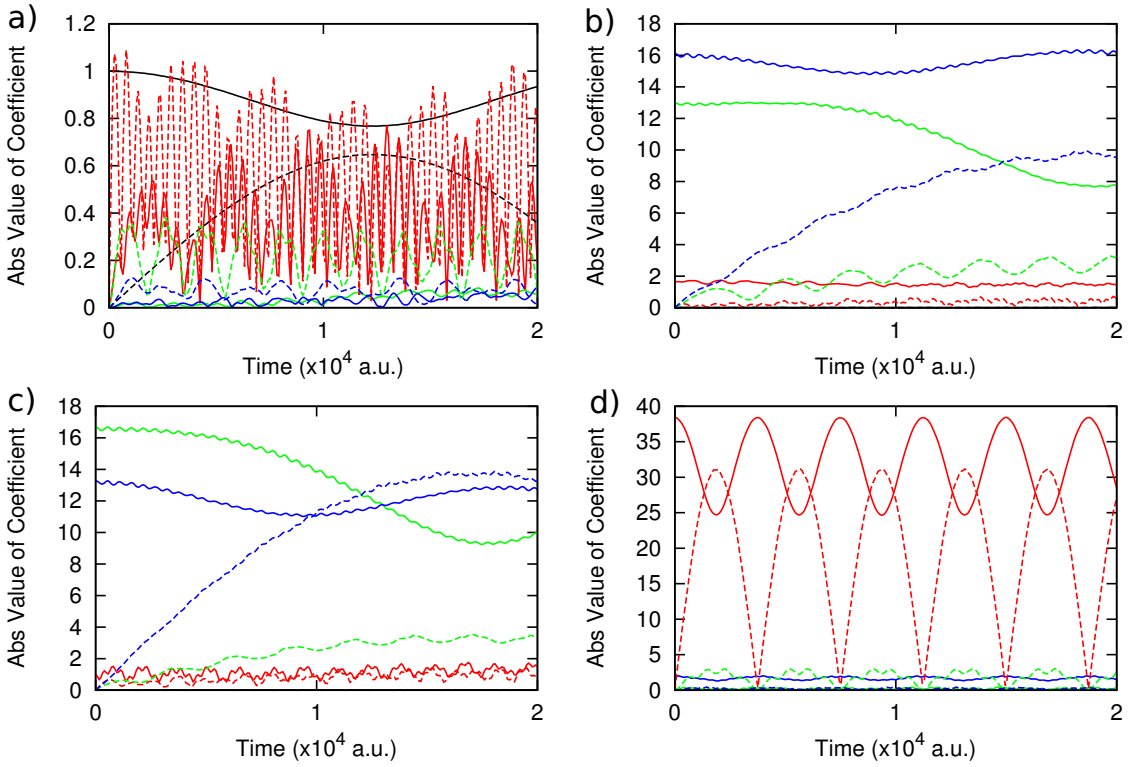

Figure 3.4 Plot of time-dependence of absolute value of excited-state coefficients for proton at $R_{OH} = 1.498$ Å for a) ground-state, b) 1st excited state, c) 2nd excited state, and d) 3rd excited state. solid = (c1(1)-black, c1(2)-red, c1(3)-green, c1(4)-blue), dashed = (c2(1)-black, c2(2)-red, c2(3)-green, c2(4)-blue).

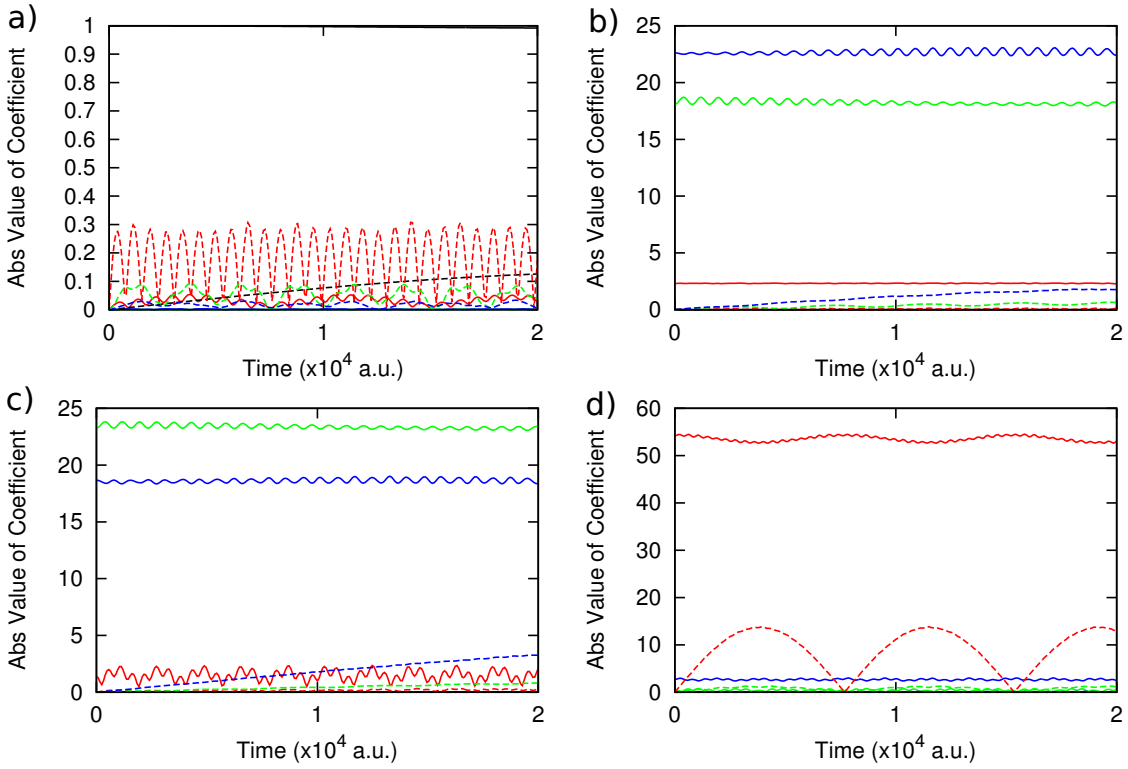

Figure 3.5 Plot of time-dependence of absolute value of excited-state coefficients for deuteron at $R_{OH} = 1.498$ Å for a) ground-state, b) 1st excited state, c) 2nd excited state, and d) 3rd excited state. solid = (c1(1)-black, c1(2)-red, c1(3)-green, c1(4)-blue), dashed = (c2(1)-black, c2(2)-red, c2(3)-green, c2(4)-blue).

Chapter 4

QTES-DFTB Dynamics Study of SLO-1

4.1 Abstract

The proton transfer reaction in the enzymatic active site of soybean lipoxygenase-1 (SLO-1) is dominated by quantum tunneling. This study examines how local substrate vibrations affect the rate constants and H/D kinetic isotope effect (KIE) of this quantum proton transfer. The reaction dynamics are modeled within a quantum trajectory (QT) framework with on-the-fly electronic structure (ES) calculations. The active site of SLO-1 is represented as a truncated 44-atom system, and the electronic structure is calculated using a density-functional tight-binding (DFTB) method. Temperature is included within the quantum thermal flux operator approach. The simulations give a KIE of 17 at 300 K, and as temperature is increased from 250 to 350 K, the KIE gradually decreases by approximately 25% in agreement with experimental results. Substrate vibrations enhance the proton transfer reaction rate constant by 15%, and the KIE is enhanced by 10%. Temperature trends in the rate constants and KIE are observed regardless of substrate vibrations. The reaction rate constants, as well as the KIE, are dominated by quantum motion of the transferring proton in fixed active site configurations, and the substrate vibrations only slightly modify the reaction dynamics.

4.2 Introduction

Enzymes catalyze nearly every important biological chemical reaction, and among these, proton (or Hydrogen/Hydride) transfer is one of the most fundamental. For some of these reactions the quantum tunneling of the transferring proton is essential, and an unusually large primary kinetic isotope effect (KIE) is often an indicator of such a system[40, 20, 41, 42, 43, 44]. One enzyme that has been extensively studied because of its abnormally large kinetic isotope effect ($KIE^{exp} = 81$)[45] is soybean lipoxygenase-1 (SLO-1). Proton abstraction from linoleic acid is the rate-limiting step in the hydroperoxidation reaction which occurs in the active site of SLO-1. It has been shown that quantum tunneling of the transferring proton plays such a large role in this reaction that the active site seems tailored to promote it[12]. The system has been modeled by various theoretical approaches[17, 37, 18, 38, 39, 16, 19]; our study is focused on quantum dynamics of the transferring proton with a possibility of incorporation of temperature and substrate motion into simulations. Following Refs. [46, 47, 19], the dynamics calculations are performed on a truncated active site, represented with 44 atoms as shown in Fig. 4.1. The proton, highlighted in blue, is transferred from the carbon of linoleic acid to the oxygen on the Fe–O–H active site ligand.

In our previous work[15], a KIE of 51 was estimated at zero temperature from the dynamics of the quantum proton/deuteron which was initially localized in the donor well of the potential energy surfaces (PES) associated with fixed active site configurations. While different active site geometries correspond to different PESs of the proton, the dynamic interplay between local active-site vibrations and proton tunneling remains controversial[48, 49, 50]. In this work we examine the dynamics of quantum proton transfer in SLO-1 using a quantum trajectory (QT) method with on-the-fly electronic structure (ES)[51], allowing for simultaneous evolution of the

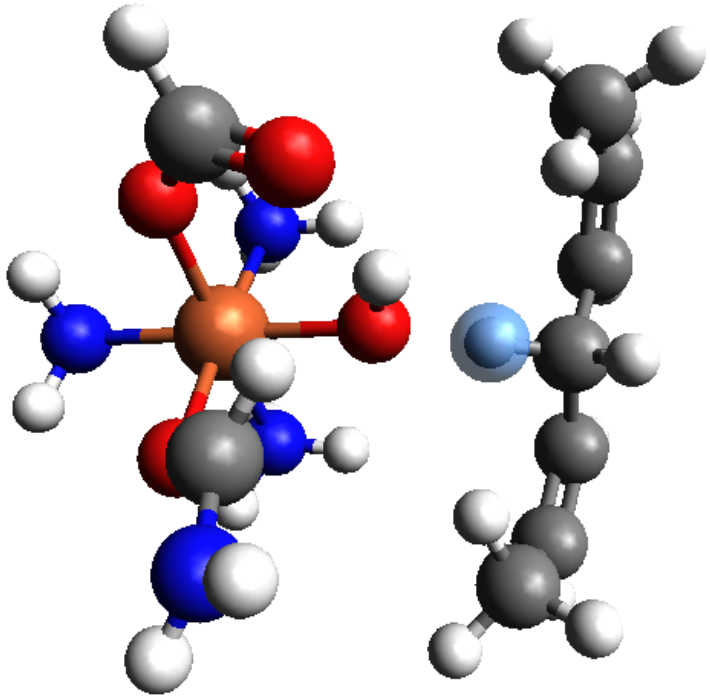

Figure 4.1 A reduced model of the active site of SLO-1. The geometry was optimized using DFTB, keeping the iron center fixed. The transferring hydrogen is highlighted in blue. The acceptor is the Fe–OH ligand. KEY: Red = oxygen, grey = carbon, white = hydrogen, blue = nitrogen, brown = iron.

quantum proton and other atoms in the active site. The effect of linoleic acid (or "substrate") vibrations on the proton transfer can be assessed by comparing time-evolutions with the fixed in space (or "frozen") and moving substrate atoms during the reaction. The electronic energy and the energy gradient, which contributes to the force acting on the atoms in the dynamics, are calculated with a density-functional tight-binding (DFTB) method[10, 11]. The theoretical framework, i. e. the QTES-DFTB dynamics and implementation of the quantum flux operator formalism within the QTES, is described in Section 4.3. Details of simulations and results are discussed in Section 4.4. Section 4.5 concludes.

4.3 Methodology

4.3.1 The electronic structure

The dynamics of nuclei in the active site depends on the electronic potential energy, $V(\boldsymbol{x})$, obtained from electronic structure calculations as a function of $\boldsymbol{x}$, which is a vector of Cartesian coordinates specifying the position of all nuclei. Potential energy is calculated for each trajectory employing a spin-unpolarized self-consistent charge density-functional tight-binding (SCC-DFTB, or just DFTB) method. DFTB is a semi-empirical electronic structure method which scales as n^3 with the number of electrons n. The electronic energy E_{el} is given by the following expression:

$$E_{el} = \sum_i 2 f_i \langle \phi_i | H_0 | \phi_i \rangle + \frac{1}{2} \sum_{\substack{A,B \\ A \neq B}} \gamma^{AB} \Delta q^A \Delta q^B + \sum_{A>B} E_{rep}^{AB}. \tag{4.1}$$

The first term on the right-hand-side is a sum over molecular orbitals labeled i, and f_i is an orbital occupation number that ranges between 0 and 1. The second term describes electron-electron interaction between the Mulliken charges Δq at different atomic centers; this term is treated self-consistently[52]. E_{rep} consists of the pair-wise repulsive interactions between atomic centers A and B, dependent only on the distance between A and B. The E_{rep} term is represented with a cubic spline, and it is parametrized by fitting two-atom energy curves to results from higher levels of theory. In this work we make a system-specific adjustment of E_{rep} to correct the potential energy of the transferring hydrogen interacting with donor C and acceptor O atoms, where the standard pairwise interactions were found inadequate.

All geometry optimization and electronic energy calculations related to validation and correction of the DFTB potential energy curves were performed with the DFTB+ 1.2.1 software package[53]. The *mio* and *trans3d* parameter sets for organic and iron atom-atom interactions, respectively, were used for these calculations[52, 54]. The

benchmark *ab initio* electronic structure calculations were performed using density functional theory, specifically, at the B3LYP/LANL2DZ level using Q-Chem[55]. Although the DFTB geometry of linoleic acid agreed quite well the B3LYP/LANL2DZ results, the reaction profile for the proton transfer reaction generated with standard DFTB parameter files was qualitatively incorrect. We were able to improve the agreement of the B3LYP/LANL2DZ and DFTB electronic energies by adjusting E_{rep}^{OH} of the proton-acceptor interaction.

Corrections to the repulsive spline have been shown to be effective in single- and multi-step proton transfer reactions, but they are system-specific[56, 57, 58, 59]. We achieved satisfactory agreement with the B3LYP proton-transfer profiles by modifying the O–H *mio* repulsive spline with a switching function of the form

$$E_{rep}(r) = E_{rep,0}(r_{\mathrm{OH}}) + E_{max} \times \frac{1}{e^{\alpha(r_{\mathrm{OH}} - r_0)} + 1}. \tag{4.2}$$

$E_{rep,0}$ is the standard *mio* repulsive energy curve, r_{OH} is the O–H distance, r_0 is the center of the switching function, α determines the steepness of the function, and E_{max} is the overall strength of the switching. The parameters yielding the best agreement are: $E_{max} = 31$ $\mathrm{mE_h}$, $\alpha = 6$ $\mathrm{a_0^{-1}}$, and $r_0 = 0.6614$ $\mathrm{a_0}$ where $\mathrm{a_0}$ is the Bohr radius. In this approach, the only modification of the standard parameter sets is between the transferring proton and acceptor oxygen.

The procedure for modifying the parameter sets is described in Appendix B. The potential along the proton reaction coordinate, after the appropriate E_{rep} adjustments, is shown in Fig. 4.2. Each curve was generated by first optimizing the linoleic acid geometry for a chosen R_{OH} constraint and, then, moving the transferring H between the donor carbon and acceptor oxygen for this optimized substrate geometry. Therefore, each curve corresponds to $V(r_{\mathrm{OH}})$, where r_{OH} is the distance between the proton and acceptor oxygen in a particular substrate environment, established by the

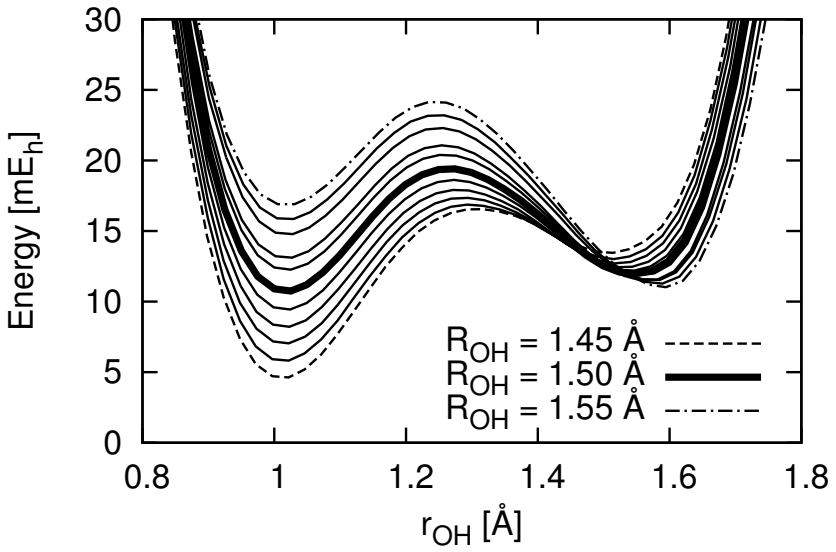

Figure 4.2 DFTB potential energy as the hydrogen moves from carbon to oxygen. The curve shown in bold represents the so-called quantum transition state (QTS). It was used as the starting substrate geometry for trajectory calculations because it is the most conducive to tunneling, and the donor and acceptor well are approximately equal. The curves are obtained under the R_{OH} constraints from 1.45 Å to 1.55 Å, and the bold QTS curve corresponds to $R_{\mathrm{OH}} = 1.50$ Å.

initial R_{OH} constraint. A decrease of R_{OH} describes advancement toward a product state along the minimum energy path. Due to deficiencies in iron-related DFTB parameter sets, the geometry of the iron-coordinated groups was calculated with B3LYP/LANL2DZ, and frozen for all DFTB calculations. For the purpose of this dynamics study, the geometry optimized with $R_{\mathrm{OH}} = 1.50$ Å was chosen for reaction rate calculations as it corresponds to a nearly equal donor and acceptor well depth, and the donor-acceptor barrier produced in this configuration is most conducive to tunneling. This configuration is therefore referred to as a quantum transition state (QTS), represented by the bold curve in Fig. 4.2.

4.3.2 Thermal flux operator formulation for the reaction rate constants

The effect of quantum proton motion and substrate vibrations on the proton transfer in SLO-1 were assessed from rate constants, computed within the flux-flux correlation function formulation. All equations are presented in atomic units, and we set $\hbar = 1$. The mass of a proton is $m_{\mathrm{H}} = 1836$ a.u. The quantum flux operator is defined as

$$\hat{F} = \frac{\imath}{2m}[\hat{p}_x, \delta(x - x_0)] \tag{4.3}$$

where x_0 is the location of the dividing surface along the reaction coordinate x, and $\hat{p}_x$ projects momentum along the reaction coordinate[60]. The truncated active site of SLO-1 consists of 44 atoms that are described by 132 Cartesian coordinates. Three of them, x, y, and z describe the quantum proton. Flux was defined for the x-coordinate of the transferring proton, aligned with the reaction coordinate, and x_0 is at the top of the barrier. Proton coordinates y and z, and the 129 classical degrees of freedom, describing the rest of the active site, influence the rate constant through coupling to the reaction coordinate x.

The flux-flux correlation function,

$$C_{ff}(t) = Tr[e^{-\hat{H}/2k_BT}\hat{F}e^{-\hat{H}/2k_BT}e^{\imath\hat{H}t}\hat{F}e^{-\imath\hat{H}t}], \tag{4.4}$$

gives the rate constant for a temperature T, according to

$$k(T)Q(T) = \int_{-\infty}^{\infty} C_{ff}(t)dt. \tag{4.5}$$

$Q(T)$ is the quantum partition function and k_B is the Boltzmann constant. The trace in Eq. 4.4 is evaluated in the spectral representation of $\hat{F}$, which has two non-zero eigenvalues[60],

$$\hat{F}|\phi^{\pm}\rangle = \pm\lambda|\phi^{\pm}\rangle, \quad |\phi^{+}\rangle^* = |\phi^{-}\rangle. \tag{4.6}$$

Action of the quantum mechanical Boltzmann operator is equivalent to evolution in imaginary time, defined in terms of temperature as $\tau \equiv 1/k_B T$. Since $\hat{F}$ has singular eigenfunctions, it is convenient to define a *thermal* flux operator[61] as

$$\hat{F}_\tau = e^{-\tau \hat{H}/2} \hat{F} e^{-\tau \hat{H}/2}. \tag{4.7}$$

The flux-flux correlation of Eq. 4.7 becomes

$$C_{ff}(t) = Tr[\hat{F}_{\tau/2} e^{\imath \hat{H} t} \hat{F}_{\tau/2} e^{-\imath \hat{H} t}]. \tag{4.8}$$

To compute the thermalized flux eigenfunction for a general barrier, we begin with an eigenfunction of the parabolic barrier at a very high temperature, $\tau_i = (k_B T_i)^{-1}$, and evolve it to the target temperature using the actual barrier. The eigenfunctions for the parabolic barrier, $V(x) = (-\omega^2 m/2)x^2$, are known analytically[62],

$$\phi_\tau^{\pm}(x) = \left(\frac{2\gamma}{\pi}\right)^{1/4} e^{-\gamma x^2} \left(\frac{1}{\sqrt{2}} \pm \imath\sqrt{2\gamma} x\right). \tag{4.9}$$

$$\gamma = \frac{m\omega}{2\,\tan(\tau\omega/2)} \qquad \lambda = \frac{\gamma}{m\sqrt{8\pi}}$$

We initialize them at high temperature (very small τ_i), and evolve to an imaginary time $\tau_f/4$ which is sufficient to calculate kQ at a temperature $T = (k_B \tau_f)^{-1}$[63]:

$$k(T)Q(T) = \frac{\lambda_{\tau_i/2}}{2} \int_{-\infty}^{\infty} \left(2|C^+(t)|^2 - |C^-(t)|^2 - |C^-(-t)|^2\right) dt \tag{4.10}$$

where

$$C^{\pm}(t) \equiv \langle \pm | e^{-\imath \hat{H} t} | + \rangle, \quad |\pm\rangle \equiv e^{-(\tau_f - \tau_i)\hat{H}/4} |\phi^{\pm}_{\tau_i/2}\rangle. \tag{4.11}$$

Thus, there are two wavefunction propagation steps required to calculate kQ for temperature T: (i) evolution of the parabolic barrier flux eigenfunctions $|\phi^{\pm}_{\tau_i/2}(x)\rangle$ to

imaginary time $\tau_f/4$, providing $|\pm\rangle$ for temperature $k_B T = \tau_f^{-1}$, followed by (ii) evolution of $|\pm\rangle$ in real time generating $C_{ff}(t)$, which is the integrand of Eq. 4.10. Both of these evolution steps were accomplished using the quantum trajectory dynamics described in Sec. 4.3.3.

4.3.3 Implementation within the quantum trajectory approach

To implement the thermal flux formalism for a reaction within the molecular environment of the SLO-1 active site, the wavefunction of the transferring proton is described by an ensemble of the approximate quantum trajectories. The nuclei of the substrate (linoleic acid) move according to the Ehrenfest dynamics [64, 65], and their motion contributes to the wavefunction phase[66]. In the approximate quantum trajectory method[22], based on the de Broglie-Bohm representation of the time-dependent Schrödinger equation[8], a globally-defined quantum force acts on the trajectories in addition to the external potential (computed on-the-fly with DFTB). To compute the correlation functions $C^{\pm}$ given by Eq. 4.11 the circumventions of the thermal flux operator should be evolved according to the quantum Boltzmann operator in imaginary time, and propagated according to the Hamiltonian operator in real time. The flux eigenfunction for a parabolic barrier (Eq. 4.9) is a Gaussian function multiplied by a linear function of the reaction coordinate x, and this functional form will be assumed at all times. Defined at a very high temperature, this wavefunction serves as a starting point for calculation of the thermalized flux eigenfunction at a lower temperature on the full potential V, using a hybrid trajectory/basis representation introduced in Ref. [63]. Then, the resulting wavefunction is evolved in real time to obtain $C^{\pm}$ and the reaction rate constants.

The quantum trajectory ensemble is used to evolve a nodeless wavefunction (initially a Gaussian) as outlined in Sections 4.6 and 4.7 for the imaginary and real time dynamics, respectively. Each k^{th} trajectory at time t is characterized by a position

$\boldsymbol{x}_t^{(k)}$, momentum $\boldsymbol{p}_t^{(k)}$, the action function $S_t^{(k)}$ and a weight $w^{(k)}$. The trajectories representing a wavefunction evolve as an ensemble being influenced by the classical potential $V(\boldsymbol{x})$ and the quantum potential $U(\boldsymbol{x})$. The latter is defined by a polynomial function of $\boldsymbol{x}$, which is a fit of the classical momentum $\boldsymbol{p}$ for the imaginary-time dynamics, or of the non-classical momentum $\boldsymbol{r} = A^{-1}\nabla A$ for the real-time dynamics[67, 68, 22]. The fitting is performed once per time step and provides the quantum potential and quantum force for all trajectories in the ensemble. This calculation is a small addition to the cost of trajectory dynamics, and all quantum effects of nuclear motion are due to the quantum potential.

The quantum trajectory ensemble represents the nodeless envelope part of the wavefunction, while the total wavefunction is assumed to be a product of the trajectory part multiplied by a linear function χ of the reaction coordinate x,

$$\psi(\boldsymbol{x}, \tau) = \chi e^{-S(\boldsymbol{x},\tau)}. \tag{4.12}$$

The coefficients of the function χ must be propagated in imaginary time, along with the quantum trajectory ensemble. The linear function χ is a scalar product of two vectors $\boldsymbol{c} = (c_1, c_2)$ and $\boldsymbol{f} = (1, x)$,

$$\chi \equiv \boldsymbol{c} \cdot \boldsymbol{f} = c_1 + c_2 x \tag{4.13}$$

where c_1 and c_2 are complex functions of time. The coefficients evolve as

$$\frac{d\boldsymbol{c}}{d\tau} = -\frac{1}{2m}\mathbf{M}^{-1}\mathbf{D}\boldsymbol{c}, \tag{4.14}$$

where for $i, j = 1, 2$ the matrix elements are

$$M_{ij} = \langle f_i | f_j \rangle, \quad D_{ij} = \langle \nabla f_i | \nabla f_j \rangle. \tag{4.15}$$

After the desired imaginary time is reached, the total wavefunction must be evolved in real time to compute $C^{\pm}$ of Eq. 4.11. The real-time wavefunction has the following form:

$$\psi(\boldsymbol{x},t) = \chi A(\boldsymbol{x},t) e^{\imath S(\boldsymbol{x},t)}. \tag{4.16}$$

The coefficients of χ evolve according to:

$$\frac{d\boldsymbol{c}}{dt} = -\frac{1}{2m}\mathbf{M}^{-1}\left(2\mathbf{\Pi} + \imath\mathbf{D}\right)\boldsymbol{c} \tag{4.17}$$

where

$$\Pi_{ij} = \langle p_x f_i | \nabla f_j \rangle. \tag{4.18}$$

The matrix coefficients D_{ij} and M_{ij} are defined by Eqs. 4.14 and 4.15, and p_x is the momentum in the direction of the reaction coordinate[69]. In imaginary time, the average quantities and action functions S are computed according to Eqs. 4.36 and 4.33, respectively. Real time analogues are calculated according to Eqs. 4.45 and 4.40. When transitioning from imaginary to real time, c_1 and c_2 do not change.

Equation 4.10 gives kQ in terms of correlation functions in a way which simplifies calculations in the trajectory framework. Taking into account the product form of the wavefunctions Eq. 4.9 and the relation $\phi^+ = (\phi^-)^*$, the correlation functions $C^{\pm}$ of Eq. 4.10 can be expressed as

$$\begin{aligned}
C^+(2t) &= \sum_k w^{(k)} \exp(2\imath S_t^{(k)}) \chi^+(x_t^{(k)}) \chi^-(x_t^{(k)}) \\
C^-(2t) &= \sum_k w^{(k)} \exp(2\imath S_t^{(k)}) \left(\chi^+(x_t^{(k)})\right)^2 \\
C^-(-2t) &= \sum_k w^{(k)} \exp(2\imath S_t^{(k)}) \left(\chi^-(x_t^{(k)})\right)^2.
\end{aligned} \tag{4.19}$$

We have defined

$$\chi^{\pm} \equiv c_1 \pm c_2 x. \tag{4.20}$$

The envelope function, represented with trajectories, is real at the start of the real-time propagation $t_i = 0$. Thus evolution to time t_f gives correlation functions at time $2t_f$, reducing the required computation time by a half. Additionally, using symmetries of C_{ff} with respect to $t_i = 0$, evolution of the wavefunction from $t_i = 0$ to t_f gives $C_{ff}(t)$ in the range $-2t_f \leq t \leq 2t_f$.

4.3.4 Dynamics of the quantum proton and substrate atoms

As discussed above, evaluation of rate constants at a given temperature involves (i) dynamics in imaginary time to establish the desired temperature followed by (ii) the real-time propagation of the total wavefunction to obtain the correlation functions $C^{\pm}$. The transferring proton, described in three Cartesian dimensions, $\boldsymbol{x}_Q$, is the only quantum nucleus, but in general its motion may influence motion of the classical substrate, whose position is described by $\boldsymbol{x}_C$. For each trajectory k, the position of all nuclei is written as $\boldsymbol{x}^{(k)} = (\boldsymbol{x}_Q^{(k)}, \boldsymbol{x}_C^{(k)})$. The substrate atoms are treated as point particles, whose initial positions are defined by the equilibrium geometry of the active site shown in Figs. 4.1 and 4.2.

In imaginary time the coordinates of the substrate nuclei remain fixed in space since they are represented by a single classical Ehrenfest trajectory,

$$\boldsymbol{x}_C^{(k)} = \text{constant}, \qquad \boldsymbol{p}_C^{(k)} = 0 \tag{4.21}$$

The subscript C denotes the classical degrees of freedom describing the substrate nuclei. The quantum degrees of freedom, representing the transferring proton evolve according to:

$$\frac{d\boldsymbol{x}_Q^{(k)}}{d\tau} = \frac{\boldsymbol{p}_Q^{(k)}}{m_Q}, \quad \frac{d\boldsymbol{p}_Q^{(k)}}{d\tau} = \nabla_Q \left(V + U\right)\big|_{\boldsymbol{x}=\boldsymbol{x}^{(k)}} \tag{4.22}$$

where the subscript Q represents the quantum degrees of freedom, ∇_Q is the gradient vector in quantum degrees of freedom only, and the momentum-dependent quantum potential U is given in Sec. 4.6.

In the case of real time evolution, the quantum trajectories describing the proton evolve as

$$\frac{d\boldsymbol{x}_Q^{(k)}}{dt} = \frac{\boldsymbol{p}_Q^{(k)}}{m_Q}, \quad \frac{d\boldsymbol{p}_Q^{(k)}}{dt} = -\nabla_Q \left(V+U\right)|_{\boldsymbol{x}=\boldsymbol{x}^{(k)}} \tag{4.23}$$

where the quantum potential U is given in Sec. 4.7. The nuclei of the substrate move according to

$$\frac{d\boldsymbol{x}_C^{(k)}}{dt} = \frac{\boldsymbol{p}_C^{(k)}}{m_C}, \quad \frac{d\boldsymbol{p}_C^{(k)}}{dt} = -\langle \nabla_C V \rangle. \tag{4.24}$$

The substrate nuclei experience the force averaged over the quantum proton,

$$-\langle \nabla_C V \rangle = -\sum_k w^{(k)} \nabla_C V(\boldsymbol{x}^{(k)}). \tag{4.25}$$

The evolution of classical degrees of freedom can be viewed as a single trajectory since $\boldsymbol{x}_C^{(k_1)} = \boldsymbol{x}_C^{(k_2)}$ and $\boldsymbol{p}_C^{(k_1)} = \boldsymbol{p}_C^{(k_2)}$ for all trajectories k_1 and k_2 as a result of the averaging in Eq. 4.24. The quantum nucleus is represented as a swarm of trajectories with different initial positions and different quantum and classical forces throughout the propagation.

4.4 Dynamics and results

The proton wavefunction ψ, a function of coordinates $\boldsymbol{x}_Q = (x, y, z)$, was initialized as a product of the parabolic barrier eigenfunction $\phi_{\tau_i}(x)$ along the reaction coordinate evaluated at temperature $(k_B\tau_i)^{-1} = 10,000$ K, given by Eqs 4.9 and 4.10 and Gaussian functions, approximating the ground state, in the other two degrees of

freedom y and z,

$$\psi(\boldsymbol{x}_Q, \tau_i) = \phi_{\tau_i}(x) \left(\frac{2\alpha_y}{\pi}\right)^{1/4} e^{-\alpha_y y^2} \left(\frac{2\alpha_z}{\pi}\right)^{1/4} e^{-\alpha_z z^2}. \tag{4.26}$$

In these equations, x is the reaction coordinate, pointing toward the oxygen acceptor. Parabolic barrier parameters γ_{τ_i} and λ_{τ_i}, as well as $\alpha_y = m\omega_y/2$ and $\alpha_z = m\omega_z/2$ are given by the Hessian matrix evaluated at the barrier top, and τ_i was chosen sufficiently small not to influence the final results. The quantum proton wavefunction was evolved in imaginary time to $\tau_f/4$, yielding the results at temperature $T = (k_B\tau_f)^{-1}$, which defined the initial real-time wavefunction $\psi_T(\boldsymbol{x}_Q, t = 0) = \psi(\boldsymbol{x}_Q, \tau_f/4)$.

The vibrational partition function $Q(T)$ was calculated for a reactant-state configuration in which the quantum proton/deuteron was localized in the donor well[60]. At our temperature range of interest, the quantum nucleus is almost entirely in its ground vibrational state, and $Q(T)$ was approximated through a diagonalization of the Hessian matrix for a proton at the minimum of the donor well. Therefore, the partition function is $Q(T) = e^{-E_{0,\mathrm{Hess}}/k_BT}$ for each temperature, where $E_{0,\mathrm{Hess}}$ is the ground-state vibrational energy of the quantum particle in the harmonically approximated donor well.

The weights and positions of all trajectories remain the same when the switch to real-time dynamics occurs, and C_{ff} was calculated starting from this configuration at $t = 0$. After the switch to real time, the proton wavefunction becomes dependent on the position of the classical Ehrenfest-type trajectory describing the substrate because the external potential V is evaluated along this trajectory and because the action function associated with this classical trajectory contributes to the wavefunction phase. The total wavefunction phase S along each quantum and a single classical

trajectory is given by

$$S_t = \int_0^t \left(\frac{\boldsymbol{p}_{C,t'}^2}{2m_C} + \frac{\boldsymbol{p}_{Q,t'}^2}{2m_Q} - V(\boldsymbol{x}_{t'}) - U(\boldsymbol{x}_{Q,t'}) \right) dt'. \tag{4.27}$$

We do not compute the amplitude of the wavefunction and use conservation of the trajectory weights to find all average quantities including $C^{\pm}$. Throughout the real-time evolution, the weight of each trajectory is constant. The effect of substrate vibrations is evaluated by freezing linoleic acid atoms and the secondary H atom on the acceptor oxygen. The carbon atoms on either end of the linoleic acid substrate remain fixed in space to maintain the overall confined geometry of the enzymatic active site.

The envelope part of the wavefunction is represented using 9600 trajectories, necessary for convergence of the correlation functions. The trajectories are evolved for 1000 time steps with $d\tau = (\tau_f/4)/1000$ a.u. in imaginary time, followed by 500 time steps with $dt = 1$ a.u. in real time. This timescale was chosen to be comparable to the double-well vibrational frequency in the reduced active site model[15]. The Gaussian sampling of initial trajectory positions gives significantly better convergence with respect to the number of trajectories, than the uniform sampling.

During the imaginary time evolution (unfolding on the inverted external potential) the fringe trajectories tend to "fall off" into the region of high-energy configurations. Given that in imaginary time the wavefunction amplitude associated with each trajectory exponentially depends on the action function (Eq. 4.35), the weight of these trajectories becomes negligible. At longer imaginary times a considerable number of trajectories explore areas of high potential energy corresponding to geometries where the DFTB electronic wavefunction often fails to converge. However, since the weights of such trajectories are vanishingly small, these trajectories can be simply discarded if $w^{(k)} < 10^{-30}$ and they contribute negligibly. Only trajectories with weights higher

than the cutoff value at the end of imaginary-time propagation were used in real time propagation. Typically for final temperatures of 250-350 K, 50-90% of the trajectories survive.

The primary KIE as well as the effect of substrate vibrations were assessed from four sets of calculations: the transferring proton was substituted with deuterium, and the substrate atoms were either constrained in coordinate space, or the linoleic acid nuclei were allowed to move according to Eq. 4.24. The four calculations were (*i*) proton with frozen substrate, (*ii*) deuteron with frozen substrate, (*iii*) proton with substrate vibrations, and (*iv*) deuteron with substrate vibrations. Experimental results are available near $T = 300$ K, so rate constants were generated for $T = [250, 275, 300, 325, 350]$ K.

In each calculation, DFTB potentials were required for 9600 geometries at every time step, which amounted to $9600 \times 1500 = 14.4$ million DFTB calculations. This is equivalent to approximately 50 days of computational time on a typical single core. Since DFTB calculations are independent for each geometry, the QTES-DFTB method lends itself to a nearly ideal scaling with respect to number of compute cores and number of trajectories. The parallel implementation of QTES-DFTB is described in Ref. [51]. The calculations in this study were performed using 2400 cores on the NICS Kraken Cray XT5 supercomputer, resulting in a wall-clock time of 30 minutes.

A typical correlation function $C_{ff}(t)$, obtained from the QTES-DFTB dynamics at $T = 300$ K, is shown on Fig. 4.3 and its time integral, yielding kQ at long times, is shown on Fig. 4.4 as a function of the time-integration limit. Formally, exactly computed C_{ff} should approach 0 and kQ should approach a constant value at long times for scattering systems or for systems fully coupled to the environment. In realistic systems C_{ff} is expected to be quenched due to wavefunction decoherence, as is the case for C_{ff} of Fig. 4.3 calculated for a frozen substrate. For the moving substrate C_{ff} does not go to zero on the same time interval as for the frozen substrate.

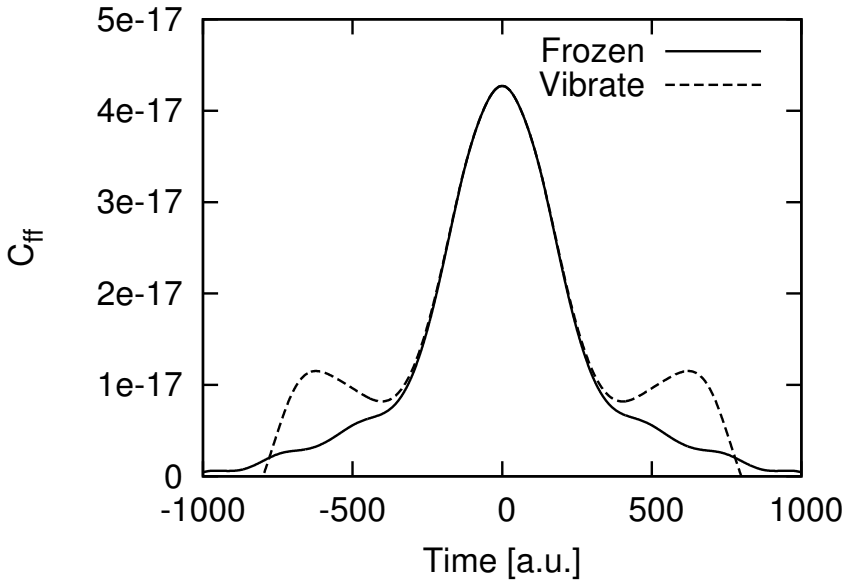

Figure 4.3 The correlation function, C_{ff}, of the proton for the frozen and moving substrates at $T = 300$ K. Additional peaks are present when the substrate vibrations are included, leading to an increase in kQ.

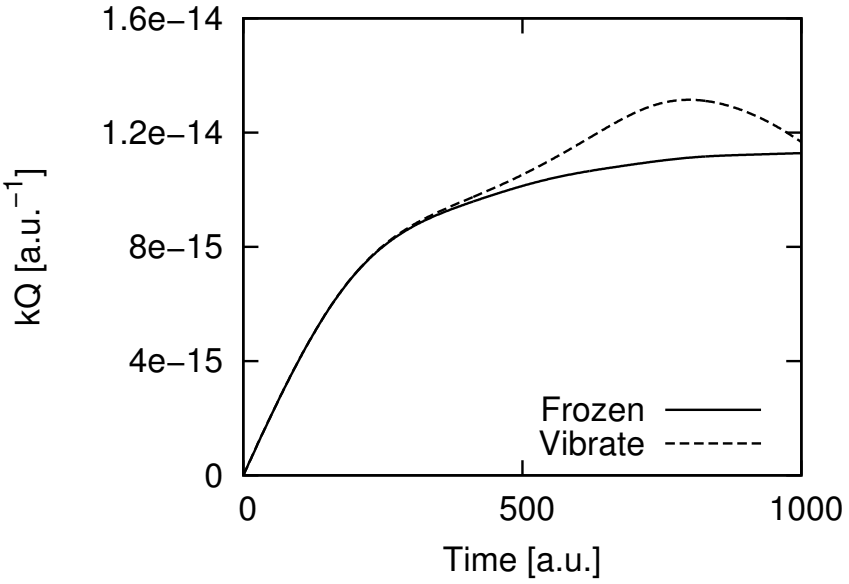

Figure 4.4 Dependence of kQ on the time-integration range obtained from the dynamics of the proton at $T = 300$ K. These results were generated from the integration of C_{ff} in Fig. 4.3. Substrate motion (dashed line) leads to higher rate constants than in cased of the proton transfer in the frozen environment (solid line).

Since the accuracy of the quantum force approximation, which is exact in parabolic potentials, deteriorates with time, we estimate kQ for all calculations as the maximum of the time integral of C_{ff}. In our calculations this time is close to final time of dynamics with fixed classical atoms and is equal to 1000 a.u. at $T = 300$ K. Motion of the classical atoms, when incorporated into dynamics, couples to the reaction coordinate, modifies the potential energy surface, and influences C_{ff} through the action function S of Eq. 4.27.

Comparing the flux-flux correlation functions and the resulting kQ for a proton at $T = 300$ K shown in Figs. 4.3 and 4.4, it is clear that substrate motion changes the system dynamics. The flux-flux correlation functions begin to deviate at $t = 400$ a.u., yielding a modest increae in kQ. This trend is present for every temperature within the examined range, which indicates that for this system the substrate vibrations increase the reaction rate constant. The results for all kQ calculations, as well as partition function values, are given in 4.1. In general, kQ results for the proton are

Table 4.1 Values of kQ from the QTES dynamics for the frozen and moving substrate. kQ always increases when substrate vibrations are included.

Temperature [K]	kQ frozen		kQ moving		$Q(T)$	
	Proton	Deuteron	Proton	Deuteron	Proton	Deuteron
250	3.8217(-17)	9.5261(-17)	4.1840(-17)	1.0979(-16)	1.6506(-6)	8.1519(-5)
275	8.2519(-16)	1.6558(-15)	9.6316(-16)	1.8481(-15)	5.5376(-6)	1.9185(-4)
300	1.1281(-14)	1.8261(-14)	1.3149(-14)	1.9930(-14)	1.5184(-5)	3.9149(-4)
325	1.0524(-13)	1.4202(-13)	1.2230(-13)	1.5315(-13)	3.5648(-5)	7.1584(-4)
350	7.0652(-13)	8.4308(-13)	8.2096(-13)	9.0347(-13)	7.4088(-5)	1.2008(-3)

very close to those calculated for the deuteron, but the rate constant $k(T)$ for the proton is consistently higher than that of the deuteron due to the partition function $Q(T)$. Rate constants are very sensitive to temperature, as shown in Fig. 4.5. This dependence of $k(T)$ on temperature, while over-exaggerated in this study, trends in the direction of experimental results[17, 45]. For every calculation performed, rate constants increased as a result of substrate vibrations by approximately 10-20% as shown in Fig. 4.6.

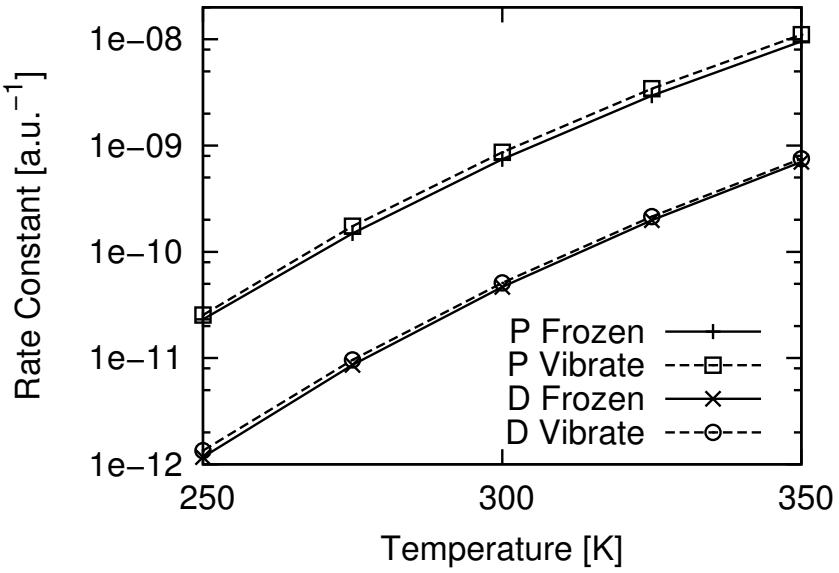

Figure 4.5 The rate constants for a set of temperatures. There is a very large sensitivity to temperature as expressed by the thermalized flux operator. A small increase in the rate constant is seen for every temperature when the substrate motion is included.

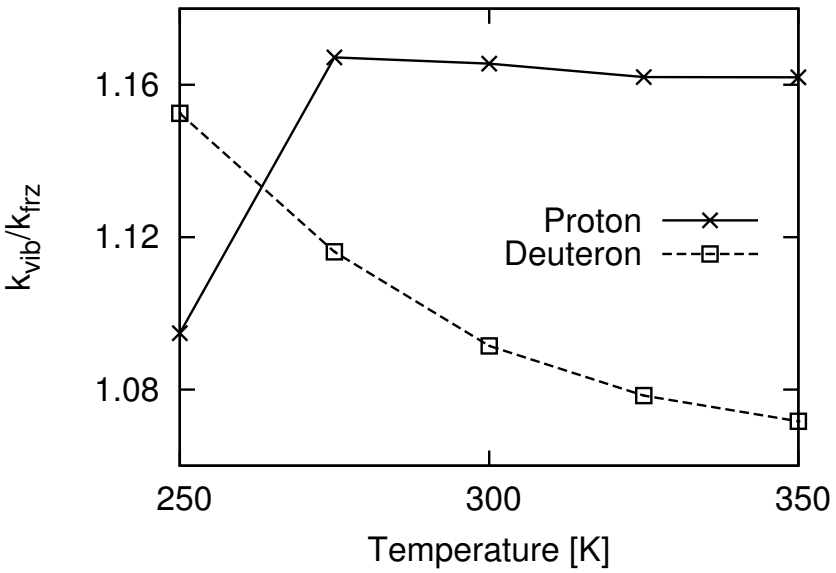

Figure 4.6 Ratio of the rate constants when the substrate vibrates (k_{vib}) or remains frozen (k_{frz}). The motion of the linoleic acid substrate tends to increase the rate constant by 10-20%.

The KIE, i. e. the ratio of rate constant for the proton to that of that the deuteron, for all temperatures can be seen in Fig. 4.7. At $T = 300$ K the calculated

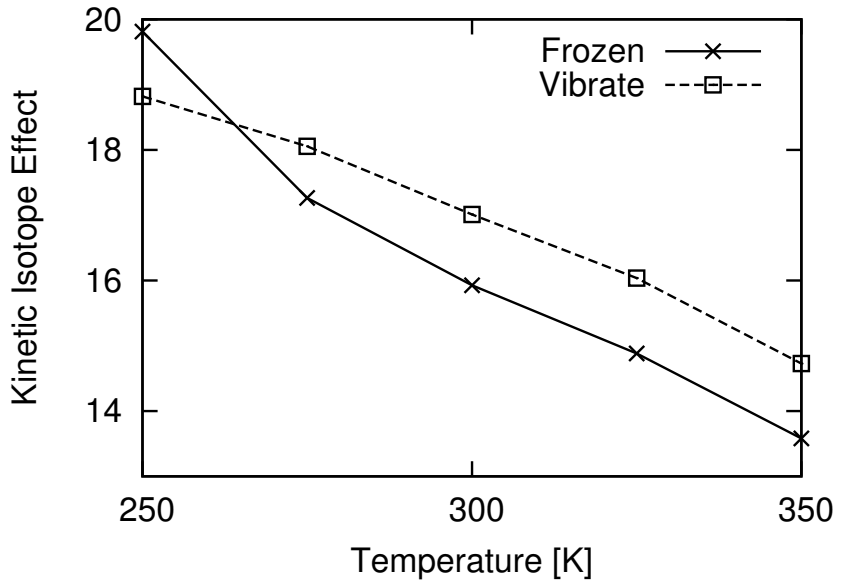

Figure 4.7 Kinetic isotope effect (KIE) as a function of temperature. Substrate vibrations increase the KIE be approximately 5-10% at each temperature. A gradual decrease in the KIE as temperature increases is in good agreement with experimental data. The crossing at low temperature indicates that the vibrations had a larger effect on deuterium transfer at $T = 250$ K.

kinetic isotope effect is KIE = 17, which is within a factor of 5 of the experimentally measured value of 81. A weak temperature dependence over the range of interest, with a decrease in the KIE by 25% as the temperature is increased by 100 K, is in very good agreement with experimental results[17, 50, 45]. The reaction is dominated by the dynamics of the quantum proton, occurring on a short time-scale, as evident in KIE and in the temperature trends of $k(T)$. The substrate vibrations are secondary to the quantum effects of the transferring proton, and the KIE only increases by about 10% as a result.

Motion of the donor carbon as well as the secondary O–H and C–H hydrogens in real time is shown in Fig. 4.8. The donor carbon moves approximately 0.05 Å , while the secondary hydrogens move between 0.1 and 0.2 Å. The only difference between

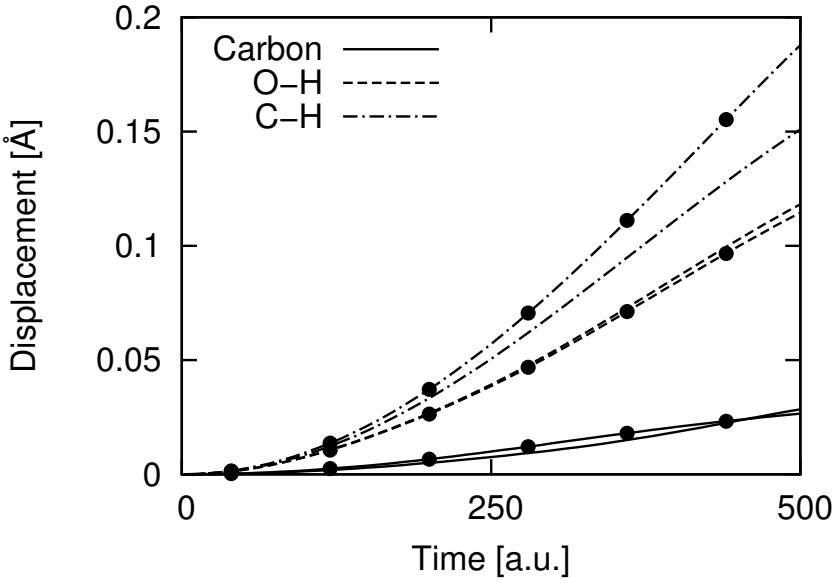

Figure 4.8 Displacement of the donor carbon, secondary O–H, and secondary C–H hydrogens as a function of time for $T = 300$ K. Plain lines mark results for the proton transfer, and lines with circles show results for the deuteron transfer. Displacement is nearly identical between the two simulations, but the secondary hydrogen on the donor carbon is more displaced during the deuteron transfer.

proton and deuteron transfer is the displacement of the secondary donor hydrogen, which moves approximately 0.03 Å further when a deuteron is present. Overall, the substrate vibrations have approximately the same effect on the proton and deuteron rate constants. On the timescale of the proton transfer, the substrate atoms do not move very far, and the vibrations have a relatively small impact on the reaction.

4.5 Conclusions and Future Work

The proton transfer reaction in the active site of SLO-1 was studied with a mixed quantum/classical trajectory dynamics, allowing inclusion of temperature and substrate motion. A density-functional tight-binding method with modified O–H repulsive spline was used to calculate electronic structure on-the-fly. This work is the first application of the QTES-DFTB method to a condensed-phase biological system with a confined proton, and it is the first multi-dimensional application of the quantum me-

chanical flux operator with quantum trajectories. The rate constants calculated with the quantum mechanical flux operator can be conveniently obtained from short-time dynamics.

This study reproduced the experimental kinetic isotope effect, as well as its dependence on temperature. It was shown that the proton transfer reaction in the active site of SLO-1 is dominated by quantum tunneling of the transferring proton, and the linoleic acid substrate vibrations, which enhance the reaction rate by 15%, have a relatively small impact on the reaction. Additionally, the rate constants for both the proton and deuteron increase by two orders of magnitude over the temperature range 250-350 K, and the same trend is observed regardless of substrate vibrations. The KIE shows a weak temperature dependence, increasing by only 10% as a result of substrate vibrations. The substrate vibrations, while they do slightly enhance the reaction rate, do not change the overall picture of a quantum proton evolving in the fixed active site geometry.

The natural improvement of this model is the quantum treatment of the secondary acceptor O–H and donor C–H hydrogen atoms. The option to include motion of heavier acceptor atoms would be beneficial, but deficiencies in the standard DFTB parameter set produce an incorrect iron-coordinated active site geometry. As a result, we were limited to classical motion only in the linoleic acid substrate and the secondary hydrogen atom on the acceptor oxygen. To include motion in the protein atoms (iron, oxygen, nitrogen, etc.), the DFTB parameters must be adjusted to provide a reasonable equilibrium geometry, a major effort. A direct application of temperature to the classical nuclei would also provide insight. In this study, temperature was only included in the quantum flux, and there is currently no clear way to reconcile imaginary time propagation with the Ehrenfest treatment of the classical nuclei other than leaving them at their energy minima. Simply adding kinetic energy to classical nuclei introduces an inconsistent temperature treatment. We are

currently investigating an ensemble-of-ensembles approach for sampling the classical degrees of freedom, relaxing the Ehrenfest approximation[66].

4.6 The Botlzmann evolution with approximate quantum trajectories in imaginary time

Evolution of a wavefunction under the quantum Boltzmann operator, defined in atomic units by

$$\hat{H}\psi(\boldsymbol{x},\tau) = -\frac{\partial\psi(\boldsymbol{x},\tau)}{\partial\tau}, \tag{4.28}$$

is equivalent to the Hamiltonian dynamics on the inverted potential, if the time-variable t is changed to imaginary time,

$$\tau \equiv \imath t. \tag{4.29}$$

The variable τ is interpreted as inverse temperature, $\tau = (k_B T)^{-1}$. Expressing a real wavefunction in the exponential form,

$$\psi(\boldsymbol{x},\tau) = e^{-S(\boldsymbol{x},\tau)}, \tag{4.30}$$

and defining a momentum of a trajectory with position $\boldsymbol{x}_\tau$ as

$$\boldsymbol{p}_\tau = \nabla S|_{\boldsymbol{x}=\boldsymbol{x}_\tau}\,, \tag{4.31}$$

one obtains the following equations of trajectory motion from Eq. 4.28 [70],

$$\frac{d\boldsymbol{x}_\tau}{d\tau} = \frac{\boldsymbol{p}_\tau}{m}, \quad \frac{d\boldsymbol{p}_\tau}{d\tau} = \nabla(V+U)|_{\boldsymbol{x}=\boldsymbol{x}_\tau}\,, \tag{4.32}$$

$$\frac{dS_\tau}{d\tau} = \frac{\boldsymbol{p}_\tau^2}{2m} + (V+U)|_{\boldsymbol{x}=\boldsymbol{x}_\tau}\,, \tag{4.33}$$

where U is the momentum-dependent quantum potential. It is defined as

$$U(\boldsymbol{x}_\tau) = \frac{1}{2m}\nabla^2 S(\boldsymbol{x}_\tau) = \frac{1}{2m}\,\nabla\cdot\boldsymbol{p}|_{\boldsymbol{x}=\boldsymbol{x}_\tau}\,. \tag{4.34}$$

In practical applications, U is found from a global least-squares fitting of $\boldsymbol{p}$ in terms of polynomials of $\boldsymbol{x}$. The calculation of U is a small addition to trajectory dynamics. A trajectory weight, used in calculation of average quantities, are:

$$w_\tau = \psi^*(\boldsymbol{x}_\tau)\psi(\boldsymbol{x}_\tau)d\boldsymbol{x}_\tau = \exp\left(-2\int_0^\tau\left(\frac{\boldsymbol{p}_{\tau'}^2}{2m} + V(\boldsymbol{x}_{\tau'})\right)d\tau'\right)d\boldsymbol{x}_0, \tag{4.35}$$

where $d\boldsymbol{x}_0$ is the initial volume element associated with the given trajectory. Expectation values are calculated as

$$\langle\hat{\Omega}\rangle = \int\psi(\boldsymbol{x},\tau)^*\hat{\Omega}\psi(\boldsymbol{x},\tau)d\boldsymbol{x} = \sum_{k=1}^{Ntraj} w^{(k)}\Omega(\boldsymbol{x}_\tau^{(k)}). \tag{4.36}$$

4.7 The Hamiltonian evolution with approximate quantum trajectories in real time

In real time the time-dependent Schrödinger equation written in atomic units,

$$\hat{H}\psi(\boldsymbol{x},t) = \imath\frac{\partial\psi(\boldsymbol{x},t)}{\partial t} \tag{4.37}$$

with the polar ansatz for the wavefunction,

$$\psi(\boldsymbol{x},t) = A(\boldsymbol{x},t)e^{\imath S(\boldsymbol{x},t)}, \tag{4.38}$$

leads to the quantum Hamilton-Jacobi equation for the wavefunction phase and the Newton's equations of the trajectory motion,

$$\frac{d\boldsymbol{x}_t}{dt} = \frac{\boldsymbol{p}_t}{m}, \quad \frac{d\boldsymbol{p}_t}{dt} = -\nabla\,(V+U)|_{\boldsymbol{x}=\boldsymbol{x}_t} \tag{4.39}$$

$$\frac{dS_t}{dt} = \frac{\boldsymbol{p}_t^2}{2m} - (V+U)|_{\boldsymbol{x}=\boldsymbol{x}_t}\,. \tag{4.40}$$

The momentum of a trajectory whose position is $\boldsymbol{x}_t$ is

$$\boldsymbol{p}_t = \nabla S|_{\boldsymbol{x}=\boldsymbol{x}_t}\,. \tag{4.41}$$

The trajectory ensemble, representing a wavefunction, describes the quantum effects through the quantum potential U,

$$U(\boldsymbol{x}_t) = \frac{-1}{2m}\frac{\nabla^2 A(\boldsymbol{x},t)}{A(\boldsymbol{x},t)} = \frac{-1}{2m}(\boldsymbol{r}\cdot\boldsymbol{r} + \nabla\cdot\boldsymbol{r})\Big|_{\boldsymbol{x}=\boldsymbol{x}_t}. \tag{4.42}$$

In Eq. 4.42,

$$\boldsymbol{r} \equiv \frac{\nabla A}{A}, \tag{4.43}$$

and in practical applications it is approximated by the least-squares fit in terms of linear functions of $\boldsymbol{x}$ (which is exact for Gaussian functions) [22]. Each trajectory carries a weight,

$$w = \psi^*(\boldsymbol{x}_t)\psi(\boldsymbol{x}_t)d\boldsymbol{x}_t = A^2(\boldsymbol{x}_0)d\boldsymbol{x}_0 \tag{4.44}$$

which remains constant through real time propagation as follows from Eq. 4.37. The trajectory weights allow simple evaluation an expectation value of position-dependent (and some other) operators,

$$\langle\hat{\Omega}\rangle = \int \psi(\boldsymbol{x},t)^*\hat{\Omega}\psi(\boldsymbol{x},t)d\boldsymbol{x} = \sum_{k=1}^{Ntraj} w^{(k)}\Omega(\boldsymbol{x}_t^{(k)}). \tag{4.45}$$

Chapter 5

Efficient Quantum Trajectory Representation of Wavefunctions Evolving in Imaginary Time

5.1 Abstract

The Boltzmann evolution describing "cooling" of a wavefunction can be recast as imaginary-time dynamics of the quantum trajectory ensemble. The quantum effects arise from the Momentum-Dependent Quantum Potential – computed approximately to be practical in high-dimensional systems – influencing the trajectories in addition to the external classical potential [JCP 132 (2010) 014112]. For a nodeless wavefunction represented as $\psi(x,t) = \exp(-S(x,t)/\hbar)$ with the trajectory momenta defined by $\nabla S(x,t)$, analysis of the Lagrangian and Eulerian evolution shows that for bound potentials the former is more accurate while the latter is more practical, because the Lagrangian quantum trajectories diverge with time. Introduction of stationary and time-dependent components into the wavefunction representation generates new Lagrangian-type dynamics where the trajectory spreading is controlled improving efficiency of the trajectory description. As illustration different types of dynamics are used to compute zero-point energy of a strongly anharmonic well and low-lying eigenstates of a high-dimensional coupled harmonic system. Some of this work appears in Ref. [70], and it is reproduced here with permission from the publisher.

5.2 Introduction

Some recent research, focused on the quantum or the Madelung–de Broglie–Bohm trajectory formulation of the time-dependent Schrödinger equation [25, 71, 8],

$$\hat{H}\psi(x,t) = \imath\hbar\frac{\partial}{\partial t}\psi(x,t), \tag{5.1}$$

as an alternative to conventional "exact" quantum-mechanical (**QM**) basis methods [1, 72, 3], involves less traditional approaches such as complex-space trajectory dynamics in real time [27, 73] and real-space trajectory dynamics in imaginary time [74, 75, 68]. Here we will consider how the wavefunction representation affects the character of the imaginary-time quantum trajectory dynamics and the accuracy of its numerical implementation.

The connection between the QM Hamiltonian evolution and the Boltzmann evolution via the transformation between the real and imaginary time (the Wick rotation [76]) was used by Feynman in the path integral formulation of statistical mechanics [77], and in a semiclassical context by Miller [78] who related the Boltzmann evolution to the real-time dynamics on the inverted classical potential. The Boltzmann evolution of a wavefunction according to the diffusion equation with the QM Hamiltonian $\hat{H}$,

$$\hat{H}\psi(x,\tau) = -\hbar\frac{\partial}{\partial\tau}\psi(x,\tau). \tag{5.2}$$

is equivalent to Eq. (5.1) where the real time-variable t is replaced by the imaginary time τ, $t \to -\imath\tau$.

As $\tau \to \infty$ any initial wavefunction will evolve into the lowest energy eigenfunction (of the same symmetry if the system has a defined symmetry), since the lowest energy component is the slowest to decay – a feature behind the largest exact QM calculations of the zero-point energies (**ZPE**s) using the Diffusion Monte Carlo [79, 80, 81, 82, 83].

Thus, for any initial ψ the wavefunction energy E converges to the ZPE value, E_0, in the course of evolution,

$$E(\tau) = \frac{\langle\psi|\hat{H}|\psi\rangle_\tau}{\langle\psi|\psi\rangle_\tau}, \quad \lim_{\tau\to\infty} E(\tau) = E_0 \tag{5.3}$$

The zero of energy scale has to be chosen so that $E_0 > 0$ to avoid the exponential growth of the wavefunction norm, $N(\tau) = \langle\psi|\psi\rangle_\tau$. The imaginary time evolution is also encountered in direct calculations of the thermal reaction rate constants [63] involving evaluation of the Boltzmann operator [60], $\exp(-\beta\hat{H})$, where $\tau \equiv \beta = (k_B T)^{-1}$, T being the system temperature.

Recasting of the Schrödinger equation in terms of trajectories whether in real time, Eq. (5.1), or in imaginary time, Eq. (5.2), starts with representing all or part of a wavefunction using an exponent of the "phase" or "action" function S to define the quantum trajectory momenta,

$$p = \nabla S. \tag{5.4}$$

The formalism below is given for a particle of mass m in one Cartesian dimension x; multidimensional generalizations can be found in Refs [84, 85]. For clarity, ∇ is used to denote spatial derivatives, including the one-dimensional case $\nabla = \partial/\partial x$.

The real-time quantum or Bohmian trajectory formalism [25, 71, 8], for which the theory and implementations are described in detail in Ref. [86], is derived from the polar representation of a wavefunction,

$$\psi^{Bohm} = A(x,t)\exp\left(\frac{\imath}{\hbar}S(x,t)\right), \tag{5.5}$$

with $A(x,t)$ and $S(x,t)$ assumed to be real functions. This particular choice of representation is justified by the physical meaning of the evolution equations following

from Eqs (5.1), (5.4) and (5.5) in the Lagrangian frame-of-reference,

$$\frac{d}{dt} = \frac{\partial}{\partial t} + \frac{p}{m}\frac{\partial}{\partial x}. \tag{5.6}$$

These evolution equations are the continuity equation for the probability density, which contains all the terms proportional to $\imath$ and linear in $\hbar$,

$$\frac{dA^2}{dt} = -\frac{\nabla p}{m}A^2, \tag{5.7}$$

and the classical-like equations of trajectory motion,

$$\frac{dS}{dt} = \frac{p^2}{2m} - (V+U) \tag{5.8}$$

$$\frac{dx}{dt} = \frac{p}{m}, \quad \frac{dp}{dt} = -\nabla(V+U). \tag{5.9}$$

All QM effects stem from the non-local quantum potential U,

$$U = -\frac{\hbar^2}{2m}\frac{\nabla^2 A}{A}, \tag{5.10}$$

which formally vanishes in the classical limit $\hbar \to 0$ or $m \to \infty$ inviting approximate and semiclassical implementations of the formalism [9, 87, 88, 89, 22]. For numerical implementation a system of equations (5.7-5.10) is solved, exactly or approximately, as reviewed in Ref. [86]. Despite the conceptual appeal, however, for general classical potentials and wavefunctions, U of Eq. (5.10) becomes singular near wavefunction nodes ($A(x,t) = 0$), or whenever QM interference is present, giving rise to unstable trajectory dynamics. To alleviate the ensuing difficulties of numerical implementation the Eulerian and arbitrary Lagrangian/Eulerian grids have been used [28, 90, 91, 30] to prevent the trajectories, i. e. the time-dependent grid points, from moving into the node region. The strategy helps with numerical stability, but does not change

the singular character of underlying quantum trajectory dynamics, which is traceable back to the polar representation given in Eq. (5.5). The bipolar wavefunction representation has the capability of reproducing QM interference via superposition of multiple sub-wavepackets [34, 92], but is more complicated in concept and in practice (especially in high dimensions).

The method of Bohmian trajectories with Complex Action (**BOMCA**) [75] is based on the representation of a wavefunction simpler than Eq. (5.5),

$$\psi^{BOMCA} = \exp\left(\frac{\imath}{\hbar} S(x,t)\right), \tag{5.11}$$

substituted into Eq. (5.1). The BOMCA trajectory dynamics defined in the Lagrangian frame given by Eq. (5.6) unfolds in complex (x,p)-space. It captures QM interference via multiple complex-space paths contributing to $\psi(x,t)$ evaluated at a single real x, thus, avoiding the node problem. However, from the numerical point-of-view BOMCA is very challenging due to analytical continuation of all quantities into the complex plane and due to the initial/final value search of the complex-space trajectories contributing to the wavefunction on real x [75, 93, 73].

5.3 Quantum trajectory dynamics in imaginary time

5.3.1 Formalism

In contrast to the real-time/complex-valued quantum trajectory dynamics of BOMCA, the diffusion equation (5.2), which is real since it does not mix real and imaginary parts of ψ, allows construction and investigation of different types of real-space trajectory dynamics. As shown by Liu and Makri [74], substitution $t \to -\imath\tau$ in the original real-time Bohmian Eqs (5.5) and (5.9) results in a singular trajectory representation of a Gaussian wavepacket, a feature attributed to the non-uniqueness of representing $\psi(x,\tau)$ with the amplitude, $A(x,\tau)$, and "phase" factors, $\exp(-S(x,\tau)/\hbar)$. The

repartitioning of $\psi(x,\tau)$ between these two factors was successfully introduced to generate non-singular and, in fact, nearly stationary trajectories enabling first multidimensional ZPE calculations from the imaginary-time Bohmian trajectories. The approach was implemented using the "independent" trajectory approach [94] based on the truncated hierarchy of equations for S, A and their high-order derivatives.

Perhaps, the simplest route to the imaginary-time quantum trajectory dynamics begins with the positive wavefunction expressed as a single exponential function,

$$\psi(x,\tau) = \exp\left(-S(x,\tau)/\hbar\right). \tag{5.12}$$

Note that since Eq. (5.2) is real, the imaginary and real parts of an initially complex wavefunction do not mix and can be propagated independently. Thus, without loss of generality we take $\psi(x,\tau)$ to be real. In addition, for $S(x,0)$ to be smooth we take $\psi(x,0)$ to be nodeless. Substitution of Eq. (5.12) into Eq. (5.2) (followed by division by $\psi(x,\tau)$) gives the following Hamilton-Jacobi equation,

$$\frac{\partial S}{\partial \tau} = -\frac{(\nabla S)^2}{2m} + V + \frac{\hbar}{2m}\nabla^2 S. \tag{5.13}$$

Defining the momentum according to Eq. (5.4) the last term in Eq. (5.13) is interpreted as the Momentum-Dependent Quantum Potential (**MDQP**) [68, 63],

$$U(x,\tau) = \frac{\hbar \nabla p}{2m}. \tag{5.14}$$

In the Lagrangian frame Eq. (5.13) gives quantum trajectory dynamics on the inverted classical potential with MDQP of Eq. (5.14) added to it,

$$\frac{dS}{d\tau} = \frac{p^2}{2m} + V + U \tag{5.15}$$

$$\frac{dx}{d\tau} = \frac{p}{m}, \quad \frac{dp}{d\tau} = \nabla(V+U). \tag{5.16}$$

For practicality of multi-dimensional implementation we compute MDQP of Eq. (5.14), which formally vanishes in the classical limit (as was the case for the quantum potential in real time) approximately from the global Least Squares Fit [95] to p in the Taylor basis $\vec{f} = (1, x, x^2 \dots)$. The optimal expansion coefficients $\vec{c}$, minimizing $\langle (p - \vec{f} \cdot \vec{c})^2 \rangle$, are found from the system of linear equations,

$$\mathbf{M}\vec{c} = \vec{b}, \quad \mathbf{M} = \langle \vec{f} \otimes \vec{f} \rangle, \quad \vec{b} = \langle p\vec{f} \rangle. \tag{5.17}$$

The linear basis is exact for Gaussian wavefunctions (for the harmonic potentials) and produces zero quantum force. The quadratic basis is the smallest one affecting trajectory dynamics. The expectation values are evaluated according to Eq. (5.21). Eq. (5.17) will be used to approximate MDQP (and other terms containing spatial derivatives of p) needed in the Eulerian and in the modified Lagrangian dynamics described below.

Besides the classical-like appearance of the MDQP Eqs (5.15) and (5.16), in the Lagrangian frame-of-reference the contribution of the quantum potential of Eq. (5.14) to expectation values is cancelled by the time-dependence of the volume element, δx_τ, associated with each trajectory (x_t, p_t) [68],

$$\delta x_\tau = \delta x_0 \exp\left(\int_0^\tau \frac{\nabla p_t}{m} dt \right). \tag{5.18}$$

For an operator $\hat{\Omega}$ and the wavefunction (5.12) the change of integration variable from coordinate x to the time-dependent trajectory positions x_τ gives

$$\langle \Omega \rangle = \int \Omega(x, \tau) e^{-2S(x,\tau)} dx = \int \Omega(x_\tau) e^{-2S(x_\tau)} dx_\tau. \tag{5.19}$$

The action function computed along the trajectory (x_t, p_t) from Eq. (5.15) is

$$S(x_\tau) = S(x_0) + \int_0^\tau \left(\frac{p_t^2}{2m} + V(x_t) + \frac{\nabla p_t}{2m} \right) dt. \tag{5.20}$$

The last term cancels the time-dependence of δx_τ in the "trajectory-specific" probabilities, $P(x_\tau) = \exp(-2S(x_\tau))\delta x_\tau$, contributing to the expectation values. Discretizing the wavefunction via a trajectory ensemble, the expectation value becomes

$$\begin{aligned} \langle \Omega \rangle &= \int \Omega(x_\tau) \exp\left(-2 \int_0^\tau \left(\frac{p_t^2}{2m} + V(x_t) \right) dt \right) \delta x_0 \\ &= \sum_j \Omega(x_\tau^{(j)}) \exp(-2\tilde{S}_\tau^{(j)}) w^{(j)}. \end{aligned} \tag{5.21}$$

In Eq. (5.21), index j labels trajectories; the trajectory weight $w^{(j)}$ denotes the initial contribution of the j^{th} trajectory to the average,

$$w^{(j)} = \exp\left(-2S(x_0^{(j)}) \right) \delta x_0^{(j)}. \tag{5.22}$$

$\tilde{S}_\tau$ denotes the "classical" part of the action function computed along the quantum (influenced by the classical and quantum potentials) trajectory,

$$\tilde{S}_\tau = \int_0^\tau \left(\frac{p_t^2}{2m} + V(x_t) \right) dt. \tag{5.23}$$

The MDQP formulation given by Eqs (5.12), (5.15) and (5.16) has been shown [68, 63] to give accurate ZPE estimates for anharmonic systems, including the double well, and for the triatomic molecules with a reasonably small (quadratic) fitting basis determining U and to converge to the QM result for larger bases (3-6 functions). For multidimensional bound systems, however, the sampling, or the trajectory representation problem was identified: in bound classical potentials ($-V$ is a barrier) the Lagrangian trajectories fall off the barrier top and leave the region of high wave-

function density. Consequently, to represent a wavefunction at long times we needed several thousand randomly distributed at $\tau = 0$ trajectories already in three dimensions, even when introducing importance sampling that emphasized low energy region of V.

To avoid the sampling (or representation) problem in the ZPE calculations, evolution in the Eulerian frame appears as a logical alternative to the Lagrangian evolution. Since the ZPE is determined by the energy eigenstate, which is a stationary object, the fixed-in-time points placed in the low-energy regions of the potential will remain in the ares of high ground state density, and thus will be adequate for the description of the ground state wavefunction (and of a wavefunction decaying to it) at all times. To determine the evolving wavefunction at stationary points, instead of the Lagrangian Eqs (5.15) and (5.16), functions S and p are evolved in the Eulerian frame according to Eq. (5.13) and its gradient,

$$\frac{\partial p}{\partial \tau} = -\frac{p\nabla p}{m} + \nabla(V + U). \tag{5.24}$$

Instead of the quantum trajectories one initializes random grid-points fixed in space; for each point function p is defined by Eq. (5.4) as before. In Ref. [27] the independent "zero-velocity" trajectories were invoked to implement the imaginary-time Hamiltonian-Jacobi equation and high-order gradients of S in one dimension. For reasons of scalability, we implement the imaginary-time Eulerian evolution of S and p once again approximately, computing ∇p and $\nabla^2 p$ on the right-hand-side of Eqs (5.13), (5.14) and (5.24) from the global Least Squares Fit to p given by Eq. (5.17).

5.3.2 Comparison of the approximate MDQP dynamics in the Eulerian and Lagrangian frames of reference

Let us start with an illustration of the MDQP dynamics performed in the Eulerian frame for a system of dimensionality N_{dim} comprised of linearly coupled harmonic oscillators. The potential,

$$V = \frac{1}{2}\,\vec{x}\cdot\mathbf{A}\cdot\vec{x}, \tag{5.25}$$

is defined by the symmetric positive definite matrix $\mathbf{A}$. The diagonal elements of the matrix are taken as $A_{ii} = 1$; the off-diagonal elements, set to a constant value, $A_{ii\pm1} = \gamma$, change the mode frequencies of the system; the remaining matrix elements are set to zeros. All particles have unit mass: $m_i = 1$, where $i = 1 \ldots N_{dim}$. For this potential, regardless of the off-diagonal elements, evolution of a wavefunction defined at $\tau = 0$ as a multidimensional Gaussian function is exact for the linear basis $\vec{f}$ (of the size $N_b = N_{dim}+1$) used to fit the components of $\vec{p}(\vec{x})$. The initial wavefunction is a direct product of Gaussian wavefunctions,

$$\psi(x,0) = \left(\frac{2a}{\pi}\right)^{1/4} \exp\left(-a(x-x_0)^2\right), \tag{5.26}$$

for each dimension. The parameter values listed in Table 5.1 describe the ground state with zero coupling, $\gamma = 0$, for $N_{dim} = 40$. We have considered the coupling constants $\gamma = 0.2$ and $\gamma = 0.4$. These values define the Hamiltonians for which the ratio of the highest to the lowest mode frequencies are 1.53 and 2.98 respectively, so that our $\psi(\vec{x},0)$ is noticeably different from the ground state of the coupled system. The ZPEs are obtained with five-digit accuracy from dynamics of an ensemble of $50-100$ points as shown in Table 5.1. Convergence to the ZPE value despite very sparse sampling is explained by the fact that energies of individual trajectories,

$$\varepsilon(\vec{x}) = \psi^{-1}(\hat{H}\psi) = -\frac{\vec{p}\cdot\vec{p}}{2m} + V + \frac{\hbar\nabla\cdot\vec{p}}{2m}, \tag{5.27}$$

become essentially the same with time as shown on Fig. 5.1. Consequently, the wavefunction norm cancels in the normalized energy expression of Eq. (5.3). Representation of a wavefunction itself or estimates of quantities involving higher moments over the trajectory distributions, such as energies of the excited states, require about ten thousand trajectories as seen from Table 5.2.

Table 5.1 The zero-point energy from the Eulerian quantum trajectory dynamics for a system of 40 coupled harmonic oscillators. The coupling constant γ (the first column) defines the ratio of the eigen-energies (the second column). E_0^{QM} is the analytical ZPE; E_0^{MDQP} is given by Eq. (5.3); $\bar{\varepsilon} = \sum_k \varepsilon(\vec{x}_k)/N_{traj}$ is the average energy value for the trajectory ensemble at the final time. The bottom row contains initial wavepacket parameters, final propagation time and time step in a.u.

γ	w_{40}/w_1	N_{traj}	E_0^{QM}	E_0^{MDQP}	$\bar{\varepsilon}$
0.2	1.53	50	19.7973	19.7972	19.7975
0.4	2.98	100	19.0537	19.0535	19.0584
$x_0 = 0$ bohr		$a = 0.5$ bohr^{-2}		$\tau = 5.0$	$d\tau = 0.005$

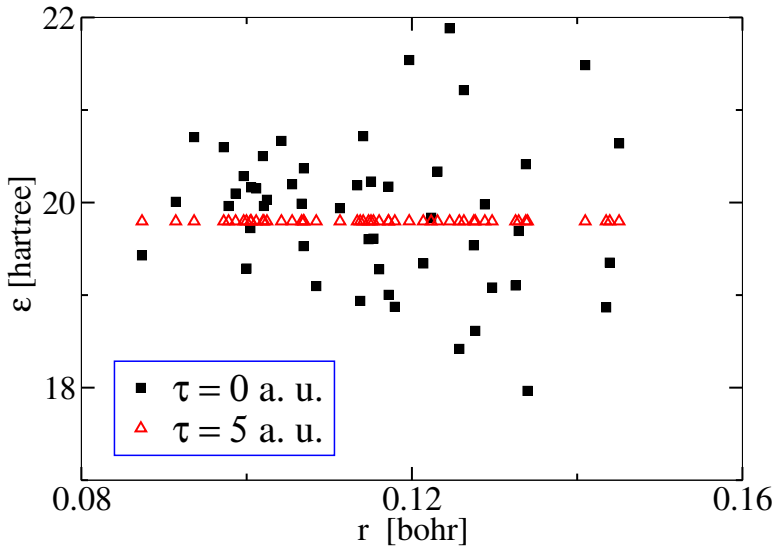

Figure 5.1 The Eulerian imaginary-times evolution for 40 coupled harmonic oscillators. The initial (squares) and final (triangles) energy of the points, ε of Eq. (5.27), are shown for $\gamma = 0.2$ as a function of the average distance from the center of the well per dimension, $r = |\vec{x}|/N_{dim}$.

The low-lying excited energy levels were obtained as the *generalized eigenvalues* of the Hamiltonian evaluated in a small basis, such as the Taylor fitting basis $\vec{f}$

multiplied by $\psi(x,\tau)$,

$$\mathbf{H}\psi_n = E_n\mathbf{M}\psi_n. \tag{5.28}$$

(The approximate MDQP evolution with *projections*, a more complicated approach, was used to evaluate the excited states in a double well [68].) As shown in Sec. 5.6 the Hamiltonian matrix elements in Eq. (5.28) are:

$$H_{ij} = \langle f_i|\varepsilon|f_j\rangle + \frac{\hbar}{2m}\langle\nabla f_i|\nabla f_j\rangle. \tag{5.29}$$

The overlap matrix $\mathbf{M}$ is defined in Eq. (5.17).

The low-lying excited state energies for a system of 15 linearly coupled oscillators, defined by Eq. (5.25) for $\gamma = 0.2$ are shown in Table 5.2. The limited (linear) basis $\vec{f}$ gives more accurate estimates of eigenstates if $\psi(\vec{x},\tau)$ is close to the ground state. Generally, Eq. (5.28) can be defined in a basis different from the fitting basis of the MDQP approximation and solved for several values of τ during the propagation to monitor the convergence of eigenvalues with time. If the bases in Eqs (5.28) and (5.17) are the same, then estimation of the excited states requires little effort in addition to the approximate MDQP evolution.

Our application of MDQP dynamics in the Eulerian frame to the high-dimensional system of coupled harmonic oscillators shows that the approach is feasible and scales linearly with the number of trajectories. However, since a general multidimensional Gaussian evolving in a quadratic potential remains a Gaussian, the linear fitting basis gives correct functional form of $\vec{p}$. The fitting procedure is exact within the Monte Carlo integration error responsible for the discrepancies between the analytical and numerical results in Table 5.2. Thus, this application does not illustrate the approximate MDQP regime analyzed in the remainder of this section. Before proceeding, we note that the earlier ZPE calculation for two coupled Morse oscillators mimicking H_2 bond in the Eulerian frame was quite efficient [63]; the quadratic fitting gave the

ZPE accurate within 0.5% in units of the harmonic ZPE value and overestimated the ZPE shift due to anharmonicity by a third of the exact QM result.

Table 5.2 Selected energy levels ($0 \leq \boldsymbol{n} \leq 15$) for a system of 15 linearly coupled harmonic oscillators. Data in rows 1-3 are obtained after imaginary-time evolution to $\tau = 4.0$ a.u. for which the CPU time is listed. The last row contains eigenvalues obtained in the linear basis at $\tau = 0.0$

$N_{\mathrm{traj}}/10^3$	CPU [s]	$\boldsymbol{n} = 0$	1	2	7	8	14	15
2.5	4.96	7.4273	8.1660	8.1969	8.3920	8.4262	8.6149	8.6406
25	52.1	7.4273	8.1932	8.2202	8.3893	8.4460	8.5945	8.6079
125	259.5	7.4273	8.2077	8.2244	8.3879	8.4212	8.5951	8.5990
QM	analyt	7.4273	8.2069	8.2213	8.3875	8.4273	8.5976	8.6073
2.5	$\tau = 0$	7.5088	8.2677	8.3055	8.4748	8.5091	8.7081	8.7390

To start analyzing the accuracy of the approximate MDQP approach we notice that there are two differences between the Eulerian and Lagrangian dynamics. (i) The Eulerian "trajectories" give efficient ground-state representation only if the ground state is localized and the trajectories are placed near the global minimum of V, whereas the Lagrangian trajectories explore the potential as they move. (ii) To implement the Eulerian evolution, the gradient of the momentum in the term $p\nabla p/m$ of Eq. (5.24) is approximated in addition to the MDQP present in both, Lagrangian and Eulerian formulations. This term is *not* an $\hbar$-quantity and therefore generally does not vanish in the classical limit.

To quantify these differences let us examine a strongly anharmonic one-dimensional potential,

$$V = \frac{x^2}{2} + x^4, \tag{5.30}$$

for a particle mass $m = 1$, studied for example in Refs [74, 68]. The initial wavefunction defined by Eq. (5.26) and parameters $a = 0.5$ bohr^{-2} and $x_0 = 0$, describes the ground state of the harmonic part of V. The initial positions of the quantum trajectories are taken on a uniform grid, their initial momenta are $p = 2a(x - x_0)$. The momentum was fitted using the Taylor bases through 5th order, $N_b = 2, 4$ and 6. Evolution was

performed up to $\tau = 2.0$ a.u. At this time the wavefunction energy for calculations converged with respect to N_b, reached plateau values. 1000 trajectories initially spaced by 0.006 bohr and 100 points space by 0.06 bohr were used for the Lagrangian and Eulerian evolution, respectively. The wavefunction energies and their convergence to the ZPE values are given in Table 5.3.

While both, the Lagrangian and the Eulerian formulations give ZPE values of the same accuracy for $N_b = 6$ (when p is fitted with high accuracy), the low-order polynomial fittings show significant differences. The linear basis fit generating zero quantum force, yields unphysical energies at the end of the Eulerian evolution; at the same time the significant part of the ZPE change due to anharmonicity (79%) is captured in the Lagrangian implementation. (The exact ground state energy for the system is $E_0 = 0.804$ hartree. The harmonic ZPE value is 0.5 hartree.) The wavefunction energy of the correct order of magnitude, though not reaching a plateau value with time, is achieved in the Eulerian formulation only for the cubic fitting basis, $N_b = 4$, whereas the Lagrangian formulation already gives a well-converged result for the same basis. To verify that the source of inferior performance of the Eulerian formulation is, indeed, the $p\nabla p/m$ term of Eq. (5.13), which was incorporated exactly in the Lagrangian frame, we combined accurate ∇p evaluation by finite difference with the MDQP evaluated from the global polynomial fitting of p. The obtained ZPE values were similar in accuracy to the Lagrangian ZPE for the same fitting basis size. For a potential with quartic anharmonicity $V^{(4)} = kx^4$ a simple error analysis shows, that after incrementing all quantities by a single time-step t from their initial values and fitting the resulting $p(t)$ with a polynomial in x_t for the Lagrangian frame and in x for the Eulerian frame, the lowest (cubic) fitting coefficient due to anharmonicity is smaller for the Lagrangian frame,

$$c_3^{Eul} = 4kt, \quad c_3^{Lagr} = \frac{4kt}{(1 + ta/m)^3}. \tag{5.31}$$

We also tried evolution with the momentum gradients described in Sec. 5.7, expecting that evolution of ∇p for each quantum trajectory in the ensemble would improve the accuracy of the Eulerian formulation implemented in a small fitting basis. We found that for the lowest order fitting, giving zero approximate quantum force, this was not the case: there was no improvement on the unphysical values of the ZPE estimate in the Eulerian formulation. For the next basis size considered the accuracy was improved by a factor of 2, but was still worse than for the Lagrangian dynamics with the same fitting basis. This also points to the importance of treating the $p\nabla p/m$ term accurately.

Table 5.3 The zero-point energy of the quartic oscillator from the approximate MDQP trajectory evolution up to $\tau = 2.0$ a.u. Asterisks mark the values that are not converged with respect to τ. The exact and harmonic ZPE values for the system are 0.804 and 0.5 hartree respectively. Results for the Eulerian evolution with exact ∇p as described in text are given in the last column.

dynamics	Lagrangian		Eulerian		Eulerian$^{\nabla p}$	
N_{traj}	1001		101		101	
N_b	E	$dE/d\tau$	E	$dE/d\tau$	E	$dE/d\tau$
2	0.742	-3×10^{-3}	0.009*	-0.51	0.794	-7×10^{-5}
4	0.804	2×10^{-5}	0.578*	-0.18	0.804	-1×10^{-6}
6	0.804	3×10^{-6}	0.803	-7×10^{-4}	0.804	-1×10^{-7}

5.4 Modifications of the Lagrangian quantum trajectory formulation

Our experience with the Eulerian frame can be summarized as follows. While the stationary trajectories have important advantages for the ground state calculations – (i) trajectories started in the low energy region continue to contribute to the ground state at later times allowing sparse sampling in high-dimensional space and (ii) the classical potential V has to be evaluated only once, which gives big computational savings for on-the-fly calculations – the small-basis fitting is accurate only for mildly anharmonic systems because a non-$\hbar$ term is approximated. The Lagrangian dynam-

ics, which approaches classical dynamics as $\hbar \to 0$, gives more stable and accurate ZPE estimates. Therefore, we will reformulate the Lagrangian dynamics to reduce divergence of trajectories by changing the potential acting on the trajectories, rather than by simply postulating $dx/d\tau = 0$ which is effectively done in the Eulerian formulation.

The single function in the exponent in Eq. (5.12) is the simplest representation of a real nodeless wavefunction, but it is not a unique one. Ideally, we want a formulation where the Lagrangian trajectories describing ground states do not move, at least for the quadratic potentials. This can be achieved by introducing the time-independent function $S_0(x)$ into ψ,

$$\psi(x,\tau) = \exp\left(-S_0(x) - S(x,\tau)/\hbar\right). \tag{5.32}$$

The trajectory momentum is $p = \nabla S$ as given by Eq. (5.4) and used in all types of dynamics discussed here. Substitution of Eq. (5.32) into Eq. (5.2) and division by $\psi(x,\tau)$ in the Lagrangian frame of Eq. (5.6) gives

$$\frac{dS(x,\tau)}{d\tau} = \frac{p^2}{2m} + V + U + V_s + V_c. \tag{5.33}$$

U is the MDQP given by Eq. (5.14). The stationary potential V_s is defined by the time-independent analytical function S_0,

$$V_s = \frac{\hbar^2}{2m}\left(\nabla^2 S_0 - (\nabla S_0)^2\right). \tag{5.34}$$

Term V_c couples the time-dependent and time-independent components,

$$V_c = -\frac{\hbar p \nabla S_0}{m}. \tag{5.35}$$

The term ∇p needed for evaluation of ∇V_c (an $\hbar/m$ quantity), will be determined from the global fitting of p, along with the MDQP, approximately. The gradient of Eq. (5.33) leads to the following equations of motion:

$$\frac{dp}{d\tau} = \nabla(V + U + V_s + V_c), \quad \frac{dx}{d\tau} = \frac{p}{m}. \tag{5.36}$$

The most efficient representation of the known ground state, $\psi_0(x)$, evolving in time is based, of course, on the separation of time and space variables,

$$S_0(x) = -\ln \psi_0(x) \tag{5.37}$$

$$S(x,\tau) = S(\tau) = S(0) + E_0\tau. \tag{5.38}$$

In this case V_s of Eq. (5.34) cancels the classical potential V up to a constant E_0 and the initial trajectory momenta are zeros. Therefore, the coupling potential V_c and the total force acting on the trajectories are equal to zeros and remain so at later times: the trajectories are stationary and the only time-dependence is in the function S of Eq. (5.38).

In practice, $\psi_0(x)$ is not known and we choose $S(x,0)$ and $S_0(x)$ as quadratic functions: $S_0(x)$ defines a time-independent Gaussian localized where ψ_0 is presumed to be significant; $S(x,\tau)$ captures the rest of space- and time-dependence of $\psi(x,\tau)$ as it decays into the ground state. The trajectory spreading is reduced in two ways. (i) The quadratic function, $S_0 = a_0x^2$, generates a parabolic barrier V_s which counteracts the divergence due to a classical well V. (ii) For the same Gaussian initial wavefunction of width a (Eq. (5.26)), the representation given by Eq. (5.32) results in smaller initial trajectory momenta, $p = 2(a - a_0)x$, compared to the momenta of the original single-function representation given by Eq. (5.12). The choice of a and a_0 should be guided by the normal mode frequencies, so that the total wavefunction $\psi(x,0)$ is more localized than the eigenstate and a_0, $a_0 < a$, defines a Gaussian

more delocalized than the eigenstate to avoid convergent trajectory dynamics that is unphysical and will lead to numerical problems.

The illustration below is given for the quartic potential of Eq. (5.30): 100 trajectories spaced by 0.04 bohr uniformly sample the initial wavefunction given by Eq. (5.26) for $a = 0.5$ bohr^{-2}. The propagation was performed up to $\tau = 2.0$ a.u. for $a_0 = 0.25$ and 0.5 bohr^{-2} and compared to the original MDQP setup corresponding to $a_0 = 0$. The Least Squares Fit of p with the cubic polynomial is analogous to Eq. (5.17) and determines all the necessary spatial derivatives of p in the equations of motion. The divergence of trajectories with time (Fig. 5.2(a)) is reduced as we go from $a_0 = 0$ to $a_0 = 0.5$ bohr^{-2}. The effect of the dynamics modification is clearly seen in the description of the wavefunction (Fig. 5.2(b)): plotting only trajectories with the wavefunction density above 10^{-16} one has 5, 11 and 23 points at the end of propagation for $a_0 = 0$, 0.25 and 0.5 bohr^{-2}, respectively. Convergence to the ZPE value is shown on Fig. 5.2(c): an ensemble of 100 quantum trajectories is enough to obtain converged E_0 for $a_0 = 0.5$ bohr^{-2} which is not the case for $a_0 = 0$. After $\tau > 1.5$ a.u. we observe oscillations of the wavefunction energy as the original Lagrangian trajectory representation of $\psi(x, \tau)$ becomes inadequate.

Similar to Ref. [74], the wavefunction representation via S_0 and S given by Eq. (5.32) can also be used to repartition $\psi(x, \tau)$ between the stationary and dynamic components without approximations. The function Δ corresponding to the fitted function $\tilde{p}$,

$$\Delta = \int \tilde{p} dx + \langle S \rangle - \langle \int \tilde{p} dx \rangle, \tag{5.39}$$

can be subtracted from the time-dependent "dynamic" component S and added to the time-independent "stationary" function S_0, with the appropriate change in the

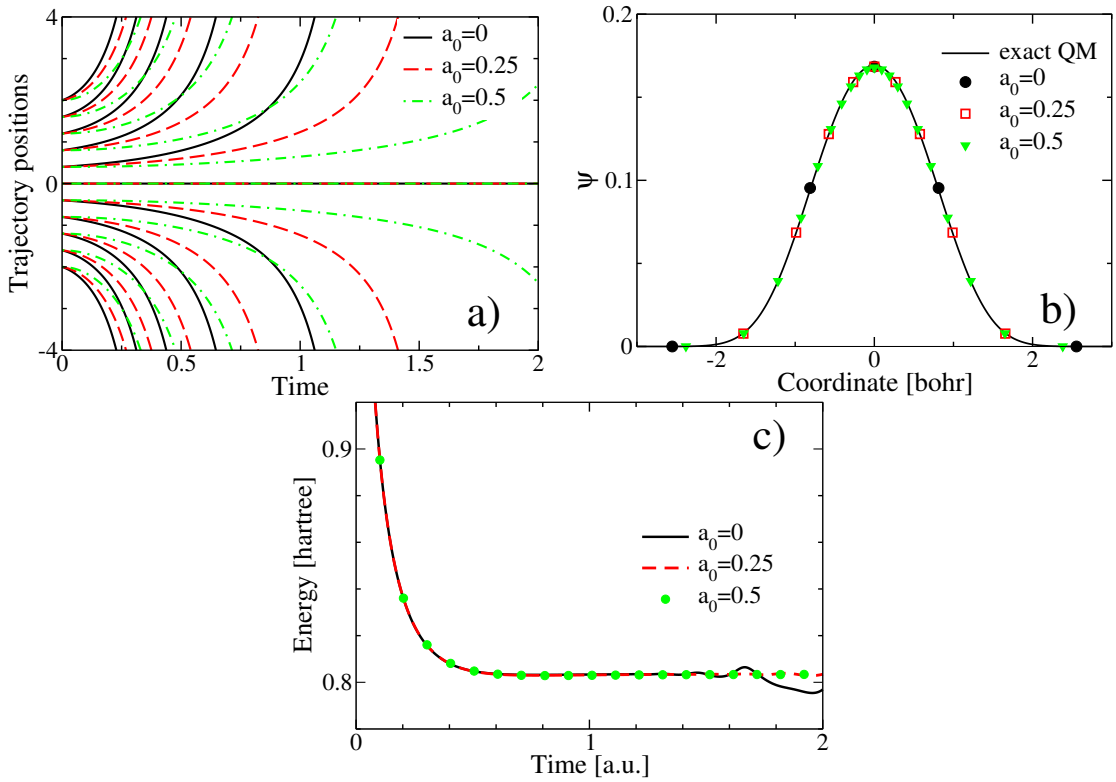

Figure 5.2 Modification of the Lagrangian dynamics in the quartic well. The considered stationary width values are $a_0 = 0,\ 0.25$ and 0.5 bohr^{-2}; $\psi(x,0)$ is given by Eq. (5.26) for $a = 0.5$ bohr^{-2}. a) Selected trajectories; b) $\psi(x,\tau)$ for $\tau = 2$ a.u.; c) The wavefunction energy.

trajectory momenta:

$$\begin{aligned} S_0^{new} &= S_0 + \Delta \\ S^{new} &= S - \Delta \\ p^{new} &= p - \tilde{p}. \end{aligned} \tag{5.40}$$

For the model system described in this Section, the effect of repartitioning (Eq. (5.40)) on dynamics is found to be qualitatively similar to calculations with various a_0 described above. We defined Δ by the linear part of the cubic fit to p so that V_s remained parabolic in x at all times. The results shown on Fig. 5.3 are obtained for initial $a_0 = 0$ with the wavefunction repartitioning at intervals of $T = 0.5$ a.u. and

$T = 0.125$ a.u. and compared to the calculation with no repartitioning. The remaining parameter values are the same as already described. As seen from Fig. 5.3(a) the repartitioning procedure resets the trajectory momenta to smaller values and introduces analytical V_s reducing the trajectory divergence and improving wavefunction sampling. As shown on Fig. 5.3(b) at the end of propagation 21 and 33 trajectories (for $T = 0.5$ and $T = 0.125$ a.u. respectively) contribute to the wavefunction compared to 5 trajectories for dynamics with no repartitioning. Better wavefunction representation improved the ZPE convergence shown on Fig. 5.3(c). Of course,

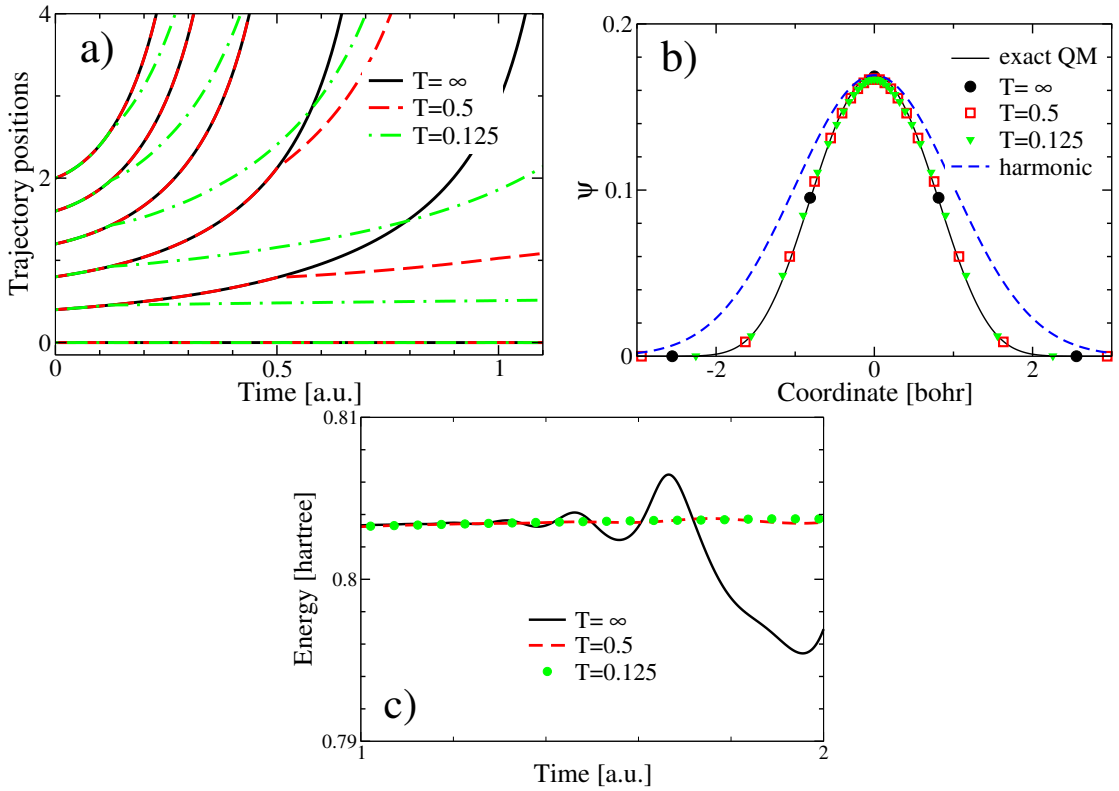

Figure 5.3 Lagrangian dynamics with wavefunction repartitioning for $\psi(x, 0)$ given by Eq. (5.26). Solid lines mark the results without repartitioning on all panels. a) Selected trajectories. b) $\psi(x, \tau)$ for $\tau = 2$ a.u. The ground state wavefunction for the harmonic oscillator, scaled to match the maximum of exact QM wavefunction, is shown with the dash. c) The wavefunction energy. Exact ZPE is $E_0 = 0.804$ hartree.

nonzero a_0 and wavefunction repartitioning can be used in the same calculation: the

goal is to balance efficient trajectory representation of the evolving wavefunction with the exploration of the classical potential with the quantum trajectories.

5.5 Summary

We have analyzed the accuracy of the imaginary-time quantum trajectory evolution with the approximate Momentum-Dependent Quantum Potential (MDQP) in the Eulerian and Lagrangian frames of reference in the context of the ground state calculations. The problem of the Lagrangian quantum trajectory evolution is that the trajectories leave the region of localization of the ground state. We have concluded that while the Eulerian evolution has very appealing features – classical potential and force are evaluated only once and the coordinate space sampling points do not move – there is also a significant drawback: for strongly anharmonic systems the Eulerian approach is not as accurate as the Lagrangian quantum trajectory dynamics, because it does not have classical dynamics as its $\hbar \to 0$ limit. The error is traced to the approximation of ∇p in the non-$\hbar$ term in the Eulerian formulation (Eqs (5.13) and (5.24)), which is incorporated exactly in the Lagrangian formulation. To counteract the divergent dynamics of the Lagrangian trajectories at long τ, we have introduced (i) a stationary component into the wavefunction form, Eq. (5.32). This modification generates analytical time-independent potential which reduces the total force acting on the trajectories, and results in smaller initial momenta defined by the dynamic component. Both consequences improve the wavefunction sampling with trajectories. If the stationary function described the ground state, the new Lagrangian trajectories would become stationary. The stationary/dynamic wavefunction representation also allows (ii) repartitioning of $\psi(x, \tau)$ between the two components without approximations. This repartitioning resets the trajectory momenta to smaller values improving the wavefunction representation even further. Work-in-progress includes development

of the criterion for a balanced application of the two modifications of the Lagrangian quantum trajectory dynamics and multidimensional chemical applications.

5.6 The Hamiltonian matrix elements

In the context of the imaginary-time quantum-trajectory dynamics the low-lying excited eigenstates can be determined as the generalized eigenvalues of the Hamiltonian matrix $\mathbf{H}$ as given by Eq. (5.28). The matrix elements are evaluated in a basis of functions $\vec{f}(\vec{x})$, multiplied by the time-dependent wavefunction, $\psi(\vec{x},\tau)$. For sufficiently long τ, the initially nodeless wavefunction $\psi(\vec{x},\tau) = \exp(-S(\vec{x},\tau))$ approaches the ground state. The simplest basis describing – for anharmonic potentials approximately – one excitation per degree of freedom is linear, $\vec{f} = (x_1, x_2 \ldots, x_{N_{dim}}, 1)$. Larger bases can be used to estimate higher eigenstates and/or for better accuracy. Using one spatial dimension and $\hbar = 1$ for clarity, the Hamiltonian matrix elements are:

$$H_{ij} = \int e^{-S} f_i \left(V - \frac{\nabla^2}{2m} \right) e^{-S} f_j \; dx \tag{5.41}$$

$$= \int e^{-S} f_i \left(V f_j \; e^{-S} - \frac{f_j}{2m} \nabla^2 e^{-S} \right) dx \tag{5.42}$$

$$- \frac{1}{2m} \int e^{-S} f_i \; 2 \; (\nabla e^{-S})(\nabla f_j) \; dx - \frac{1}{2m} \int e^{-S} f_i \; e^{-S} \; (\nabla^2 f_j) \; dx$$

For infinite integration range, integration of the last integral in Eq. (5.42) by parts gives

$$\int e^{-2S} f_i \; (\nabla^2 f_j) dx = \tag{5.43}$$

$$e^{-2S} f_i \; (\nabla f_j) \Big|_{-\infty}^{\infty} - \int e^{-2S} (\nabla f_i)(\nabla f_j) dx - \int f_i (\nabla f_j)(\nabla e^{-2S}) dx.$$

With that, assuming $\psi(x,\tau)$ vanishes at $\pm\infty$ and using definition $\nabla S = p$ in Eq. (5.41), the expression for H_{ij} becomes

$$\begin{aligned} H_{ij} &= \int_{\infty}^{\infty} e^{-2S} f_i f_j \left(V - \frac{p^2}{2m} + \frac{\nabla p}{2m} \right) dx + \frac{1}{2m} \int_{\infty}^{\infty} e^{-2S} \nabla f_i \nabla f_j dx \\ &= \langle f_i | \varepsilon | f_j \rangle + \frac{1}{2m} \langle \nabla f_i | \nabla f_j \rangle. \end{aligned} \tag{5.44}$$

Definitions (5.27) and (5.21) of the trajectory energies ε and of the trajectory ensemble averages were used to obtain Eq. (5.44).

5.7 The imaginary-time evolution of momentum gradients

The real and imaginary time propagation of the derivatives (above the first order) of the wavefunction phase and amplitude was considered by several research groups attracted by the idea of *independent* quantum trajectories, as a way to deal with the node problem in Bohmian dynamics, to reduce the number of trajectories to one or very few and to give their methods semiclassical flavor. Thus, the Derivative Propagation Method [32, 96], BOMCA and semiclassical approximation with zero velocity trajectories [27, 75] and Bohmian dynamics from trajectory stability properties [94, 74] have emerged. All of them are based on the hierarchy of equations obtained by successive differentiation of the Hamilton-Jacobi Eqs (5.8) and (5.13) truncated at some finite order, such as 4th or 6th, with higher order derivatives set to zero. These independent quantum trajectory methods were successfully applied to several one-dimensional systems and to ZPE calculation of H_2O and SO_2 [74]. However, the high-order truncation strategies might have certain shortcomings. The truncated set of equations is not guaranteed to converge to QM result, and it is expensive in multiple dimensions due to the large number of high-order derivatives of S and A and due to high-order derivatives of the classical potential V for realistic chemical systems.

Our goal is to capture the dominant effects of non-locality of quantum mechanics using the quantum trajectory ensemble and to do this in a practical manner. Thus, we considered evolution of equations of motion for the imaginary-time quantum trajectories only through the second order (momentum gradient) with higher order derivatives found approximately from the *global* Least Squares Fit to ∇p. The second order scheme requires the gradients and the Hessians of V, which is typical for semiclassical methods, such as for example the Herman–Kluk propagator and Frozen Gaussians [97, 98]. We have implemented the momentum gradient approach in the Lagrangian and Eulerian formulations, with the expectation of achieving higher-accuracy ZPE estimates within a small fitting basis, because now we approximate ∇p rather than p, as was done in the approximate MDQP. The Eulerian formulation is obtained by taking the gradient of Eq. (5.24) which, denoting $g \equiv \nabla p$, gives the following equations of motion:

$$\begin{aligned} \frac{\partial S}{\partial \tau} &= -\frac{p^2}{2m} + \frac{\hbar g}{2m} + V \\ \frac{\partial p}{\partial \tau} &= -\frac{pg}{m} + \frac{\hbar \nabla g}{2m} + \nabla V \\ \frac{\partial g}{\partial \tau} &= -\frac{p \nabla g}{m} - \frac{g^2}{m} + \frac{\hbar \nabla^2 g}{2m} + \nabla^2 V \end{aligned} \tag{5.45}$$

The Lagrangian formulation is obtained by transforming Eqs (5.45) into the Lagrangian frame given by Eq. (5.6).

To implement the quantum trajectory dynamics with Hessians given by Eq. (5.45) approximately we performed the Least Squares Fit of the momentum gradient, g, in terms of monomials, as described by Eq. (5.17) with p replaced by g. Numerical studies of one-dimensional model systems of Ref. [68] (the Morse potential and the double well) in the Lagrangian frame showed that while there were instances when the momentum gradient fitting using N_b functions yielded more accurate ZPEs (relative errors were twice smaller) than those obtained with the momentum fitting in a basis

$N_b + 1$, we also had counter-examples. All-in-all, there was no obvious advantage to the more expensive dynamics with Hessians. Dynamics with Hessians for the quartic well of Sec. 5.3.2 did not improve the accuracy of the Eulerian frame implementation, which supports the conclusion at the end of Sec. 5.3.2.

Chapter 6

Rate Constant Calculations for the HO–H–CH_3 Double Well Potential

6.1 Abstract

A new expression for the calculation of reaction rate constants in a double-well potential is analyzed in the context of the proton-transfer reaction HO–H–CH_3. A QM wavepacket propagation method as well as the WKB method was used to calculate the transmission probability, and these methods agreed quite well with one another. The QM wavepacket method regularly produced a larger rate constant at low energy, and the KIE calculations between these two methods had a similar level of agreement, but no noticeable trend. An exponential relationship between KIE and barrier height at low temperature points to a tunneling mechanism near the ground state, and therefore quantum effects are very important when describing the reaction in this system. Some of this work appears in Ref. [99], and it is reproduced here with permission from the publisher.

6.2 Introduction

The double-well potential is very common in chemical reactions such as proton-transfer reactions [100, 19, 24, 101, 48, 102, 103, 43]. If the transferring nucleus is light, a classical approximation may prove inaccurate when calculating observables such as reaction rate. Quantum effects must be taken into account for accurate results. In the case of a double well which is uncoupled to environmental degrees of freedom,

a quantum mechanical object will oscillate indefinitely between reactant and product wells [15, 24]. This property causes difficulty in defining the rate of reaction as well as deriving an expression do describe it. We have developed an expression which takes into account these features of the uncoupled double well. In this chapter, properties of the new rate constant formula are evaluated for the model proton-transfer system HO–H–CH_3 and kinetic isotope effect information is extracted from the rates.

6.3 Rate constant formula

The expression for a quantum mechanical rate constant is

$$k(T) = \frac{\int_0^\infty T(E) e^{-E/(k_B T)} dE}{\int_0^\infty e^{-E/(k_B T)} dE} \tag{6.1}$$

for a scattering system. In this equation, $T(E)$ is the transmission probability from reactant to product for a wavepacket of energy E that originates in the asymptotic region of the reactant state. This expression is extended to a bound system by taking eigenstate projections into account,

$$k(T) = \frac{\sum_n T(E_n) \rho_n e^{-E_n/(k_B T)}}{\sum_n e^{-E_n/(k_B T)}}, \tag{6.2}$$

where E_n is the energy of bound eigenstate n and ρ_n is the projection of eigenstate n onto the reactant region of the potential.

The calculation of $T(E_n)$ for a system which oscillates indefinitely is circumvented by treating the system as a scattering system for dynamics. In order to calculate rate, the potential energy surface must be transformed into an asymptotic surface with a definite reactant and product region. This transformation prevents re-crossing in the dynamics. In this way, the calculation of $T(E)$ would be identical to a scattering state in which the barrier is one from a bound double-well potential. All values of ρ_n

and E_n are calculated for the bound system (the true potential), and these factors are responsible for the main difference between Eq. 6.1 and 6.2.

6.4 Potential energy surfaces for HO–H–CH_3

The proton transfer reaction that occurs in HO–H–CH_3 is treated as a collinear reaction for simplicity. Surfaces are generated for a fixed R_{CO}, and energy is calculated as a function of R_{OH}. All electronic structure calculations were performed using Q-Chem ver. 4.01 at the DFT (B3LYP/6-31G(d,p)) level of theory, and 40 points were generated for each surface. These points were then fit to a function of the form

$$V(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + a_5x^5 + a_6x^6 \tag{6.3}$$

and the parameters for each surface are in table 6.1.

Table 6.1 Potential energy surface parameters

	$R_{CO} = 2.7$	$R_{CO} = 2.8$	$R_{CO} = 2.9$
a_0	-110.128	-110.511	-110.834
a_1	-22.1385	-20.5493	-19.2373
a_2	31.2018	28.7141	26.7153
a_3	-21.738	-19.9869	-18.6161
a_4	7.84274	7.35175	6.97034
a_5	-1.41449	-1.40449	-1.38909
a_6	0.112944	0.121945	0.126298

Potential energy surfaces for three $R_{CO} = 2.7,\ 2.8,\ 2.9$ Å can be seen in Fig. 6.1. These are the three surfaces that will be used in this study. Comparative rates between the surfaces, as well as different methods for calculating $T(E)$ and its effect on the rate were evaluated. The variety in barrier shape between these surfaces makes for an optimal evaluation on the effect of rate calculation on tunneling probability.

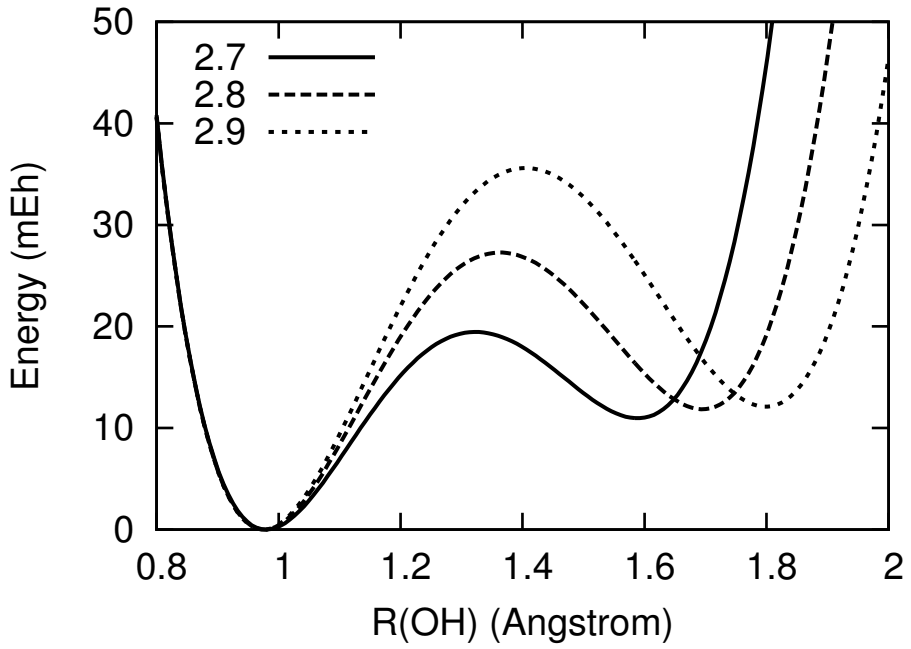

Figure 6.1 Three *bound* potential energy surfaces for the collinear proton transfer in HO–H–CH_3. One curve is generated for each fixed R_{CO} distance. The curves have been aligned by setting the product (O–H) minimum to 0 mEh. It is clear that as R_{CO} is increased, the barrier height also increases.

6.5 Calculation of $T(E)$

The transmission probability was calculated as a function of energy for each potential energy surface. To avoid the problem of oscillations between reactant and product state, the surfaces were treated as 1-D scattering potentials. The reactant and product minima become asymptotes as shown in Fig. 6.2. A side-by-side comparison of the bound and asymptotic surfaces can be seen in Fig. 6.3.

The calculation of the quantum mechanical transmission probability is accomplished through the time-evolution of quantum wavepacket which originates in the reactant region and collides with the barrier, and this method is described in detail in Appendix C. The form of the reactant wavepacket is

$$\psi_R^+(x) = e^{-\alpha(x-x_R)^2+\imath p_0(x-x_R)} \tag{6.4}$$

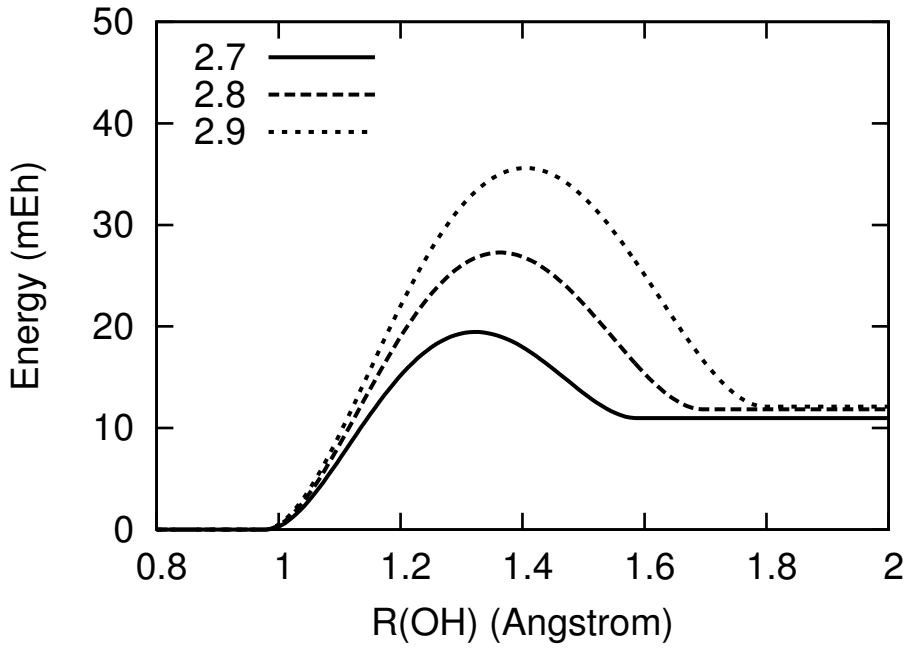

Figure 6.2 Three *scattering* potential energy surfaces for the collinear proton transfer in HO–H–CH_3. One curve is generated for each fixed R_{CO} distance. The curves have been aligned by setting the product (O–H) asymptote to 0 mEh.

and this wavepacket evolves through time. A product wavefunction is also defined as

$$\psi_P^-(x) = e^{-\alpha(x-x_P)^2+\imath p_0(x-x_P)} \tag{6.5}$$

and this wavefunction remains stationary throughout the calculation. For all calculations, α and p_0 are the same for reactant and product. A schematic of this configuration is shown in Fig. 6.4.

The reaction rate is calculated using a form of S-matrix theory [104]. The transmission probability is defined as

$$T(E) = |S_{RP}|^2 \tag{6.6}$$

where S_{RP} is an S-matrix element composed from reactant (R) and product (P) scattering eigenstates. This matrix element is defined as shown in Eq. 6.7

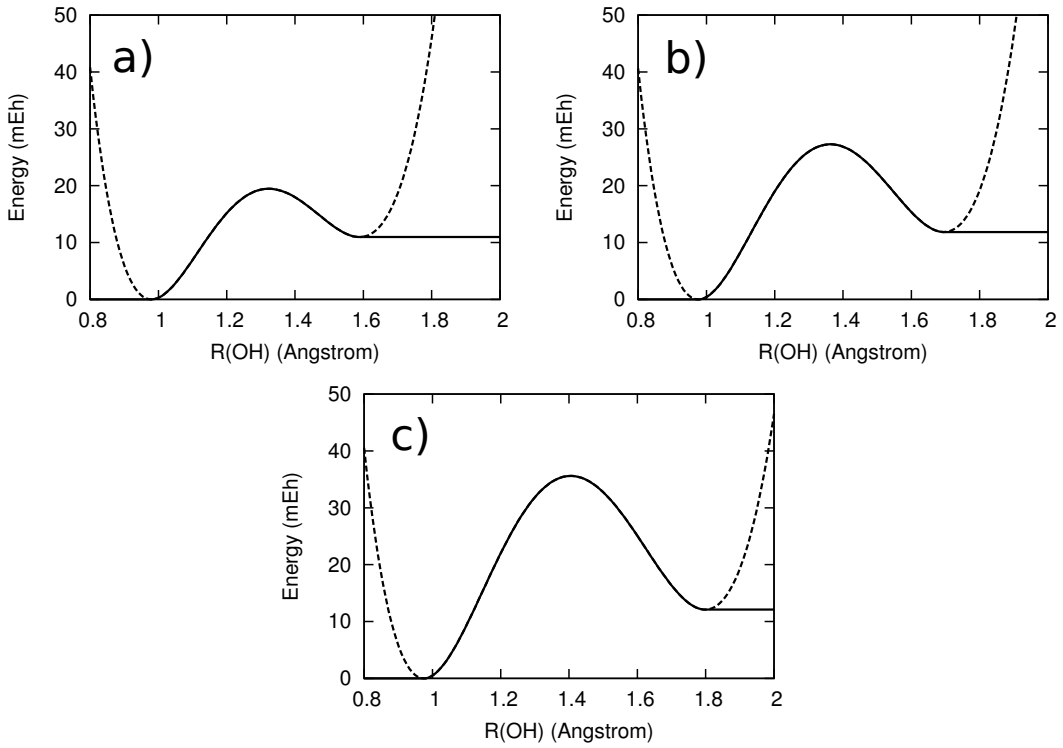

Figure 6.3 A comparison of the asymptotic transformation on the three potential energy surfaces. R_{CO} = a) 2.7, b) 2.8, and c) 2.9 Å.

$$S_{RP}(E) = \frac{(2\pi\hbar)^{-1}}{\eta_P^*(E)\eta_R^*(E)} \int_{-\infty}^{+\infty} \langle \psi_P^- | e^{-i\hat{H}t/\hbar} | \psi_R^+ \rangle e^{iEt/\hbar} dt \tag{6.7}$$

and the details of this equation are described in Appendix C.4. It is important for the wavepacket dynamics that: (*i*) the correlation function $\langle \psi_P^- | e^{-i\hat{H}t/\hbar} | \psi_R^+ \rangle$ is accurate, and (*ii*) the energy eigenstates which contribute to the time dependent wavepacket are suitably large in the energy range of interest. Parameters of the incident wavepacket ψ_R^+ (and therefore ψ_P^-) had to be chosen accordingly and are shown in Table 6.2.

Table 6.2 Wavefunction parameters

R_{CO} (Å)	α_P	$p_{0,P}$ (a.u.)	α_D	$p_{0,D}$ (a.u.)
2.7	9	-8	13	-10
2.8	9	-12	13	-10
2.9	9	-12	13	-16

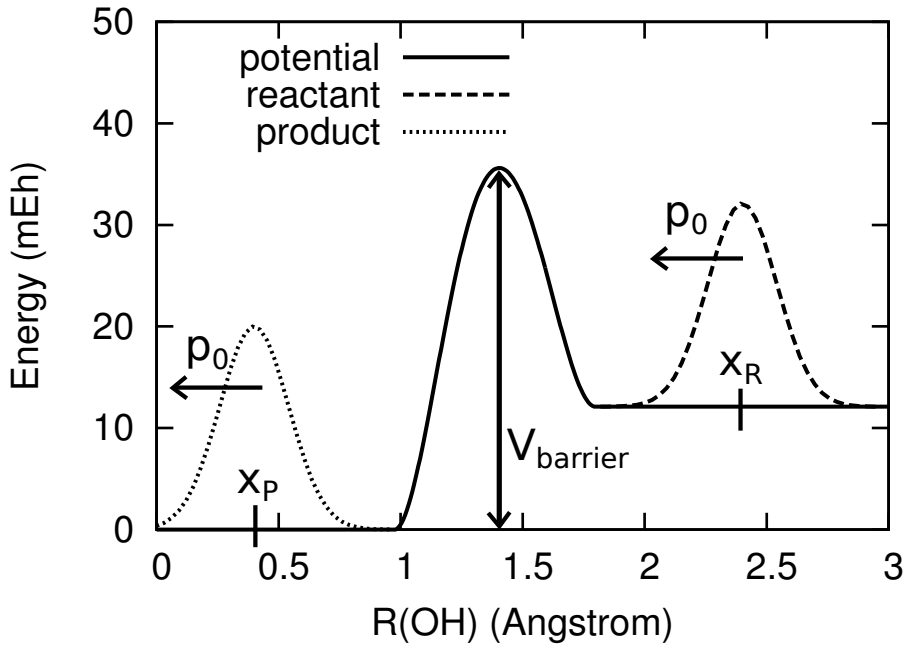

Figure 6.4 Initial conditions for quantum wavepacket scattering simulations. ψ_R^+ is on the right of the barrier, and ψ_P^- is on the left.

Transmission probability was also calculated, for comparison sake, with the WKB approach [105, 106, 101] as shown in Eq. 6.8

$$T(E) = \begin{cases} \exp\left[-2\int_{x_1}^{x_2}\sqrt{2m(V(x)-E)}dx\right] & \text{if } E < V_{barrier} \\ 1 & \text{if } E \geq V_{barrier} \end{cases} \tag{6.8}$$

where x_1 and x_2 are the classical turning points on either side of the barrier for a particle with total energy E. There is additionally a "classical" description of transmission probability, and it is simply a step function

$$T(E) = \begin{cases} 0 & \text{if } E < V_{barrier} \\ 1 & \text{if } E \geq V_{barrier} \end{cases} \tag{6.9}$$

which states that the only time a particle will make it over the barrier is if its total energy is greater than or equal to the barrier maximum. In that case, it will overcome

the barrier every time it encounters it. These various definitions of $T(E)$ are then applied directly to Eq. 6.2, and a rate constant can be calculated.

6.6 Calculation of eigenstate projections $\boldsymbol{\rho_n}$

The calculation of a rate constant also requires the projection of relevant (based on the energy range of interest) eigenstates onto the reactant state of the reaction. The calculations of bound-surface eigenstates were performed through a discrete variable representation (DVR) calculation [107, 108]. One can expect better agreement for lower energies since the highest eigenstates will contribute less, and therefore the finite number of eigenstates that are chosen will appear more complete. For these calculations, there were, on average, 15-20 eigenstates (for proton or deuteron calculations) which were below our maximum energy threshold of 100 $\mathrm{mE_h}$. At $T = 20,000$ K, which was the highest temperature considered, the highest energy eigenstate contributed about 1-5%. Lower temperatures were of much more interest since that is where tunneling occurs, and at those temperatures, high energy eigenstates did not contribute at all.

Once eigenstates are calculated, the projection is calculated as

$$\rho_n = \int_{x_{cut}}^{\infty} \phi_n^* \phi_n dx \tag{6.10}$$

where x_{cut} is the location of the barrier maximum on a particular surface, and ϕ_n is the eigenstate for energy level n. All space to the right of this point is considered a "reactant" region, and therefore this quantity represents the probability of finding the system in a reactant configuration at energy E_n. The lowest-energy eigenstate included in these calculations is the first one which has an energy higher than the reactant minimum. These projection terms are responsible for any bound-potential properties that were neglected in the dynamics. These terms, along with transmission

probability terms, can now be directly inserted into Eq. 6.2 and rate constants can be calculated.

6.7 Rate constant calculation results

Rate constants for each potential energy surface were calculated as a function of temperature. In addition to proton transfer, deuteron transfer was also taken into account so that a KIE could be calculated. Rate constants are plotted on a logarithmic scale to clearly see differences at low energies. Results for proton and deuteron transfer can be seen in Fig. 6.5 and 6.6 respectively. From these figures, it is clear

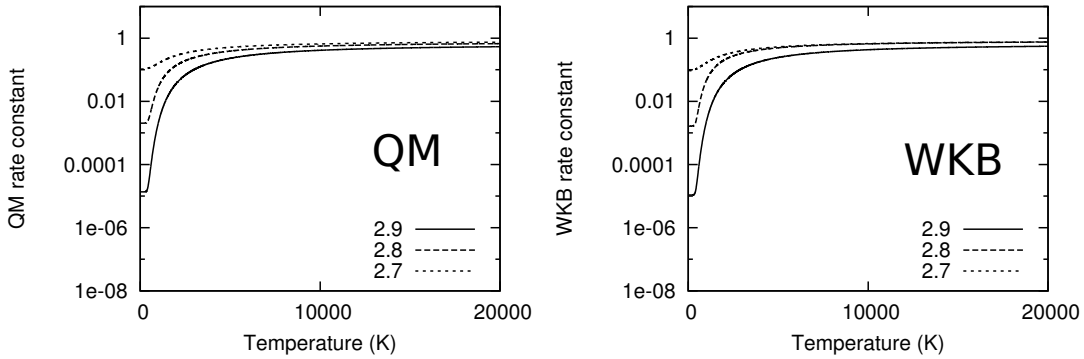

Figure 6.5 Proton rate constant calculation results for each potential energy surface. Each line represents a different R_{CO} constraint.

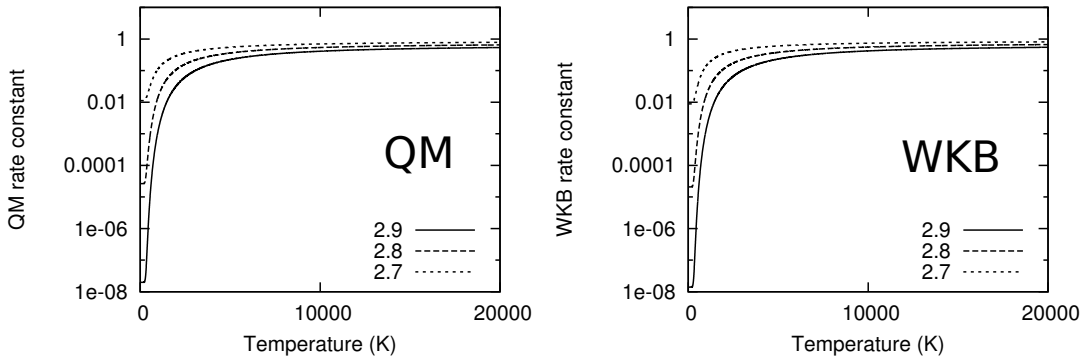

Figure 6.6 Deuteron rate constant calculation results for each potential energy surface. Each line represents a different R_{CO} constraint.

that agreement is always within an order of magnitude between the two methods.

Smaller differences are more easily visualized in the form of a ratio of QM/WKB rate constants as shown in Fig. 6.7.

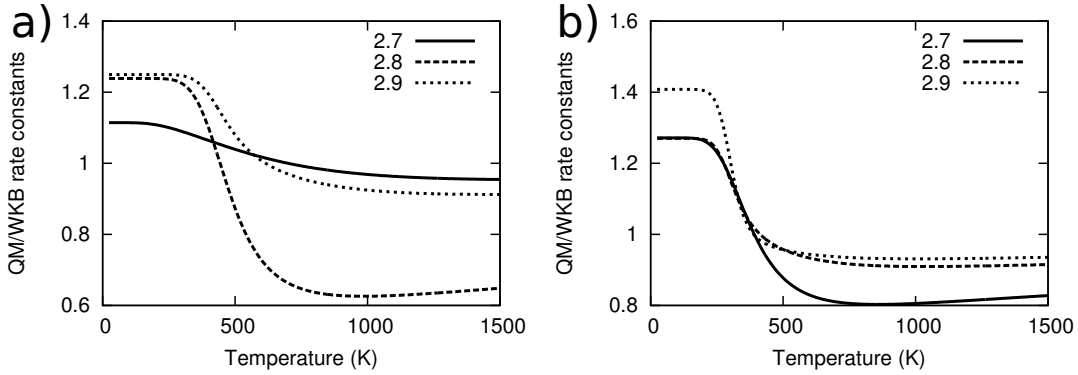

Figure 6.7 Ratio of QM/WKB rate constants for a) proton and b) deuteron transfer reactions. At low energies, the quantum rate constant is always higher than the WKB rate constant.

At low energies, the QM wavepacket calculations predict a higher reaction rate than the WKB equivalents. This is likely due to the inclusion of certain quantum effects that are neglected in the WKB calculations. The otherwise close similarities between the QM and WKB rates comes partially from the fact that $\rho_n(E)$ for each method is identical, and it comes from the DVR calculations of eigenstates. Only several points from the $T(E)$ curves are actually included (one per eigenstate). It is also instructive to examine the kinetic isotope effect KIE $= k_P/k_D$ for various potentials as a function of temperature. These results can be seen in Fig. 6.8. The ratio of these results can be seen in Fig. 6.9.

Each surface that was studied had a different barrier height and width. A relationship between barrier height and low-temperature KIE was noticed, and this relationship can be seen in Fig. 6.10. The KIE increases exponentially as the barrier is raised. This trend suggests that tunneling plays an increasingly important role in the reaction as the barrier becomes larger.

Additionally, a comparison of k_{QM} and k_{QC} at their respective ground states is given in Table 6.3. The QM ground state is the lowest-energy double-well eigenstate

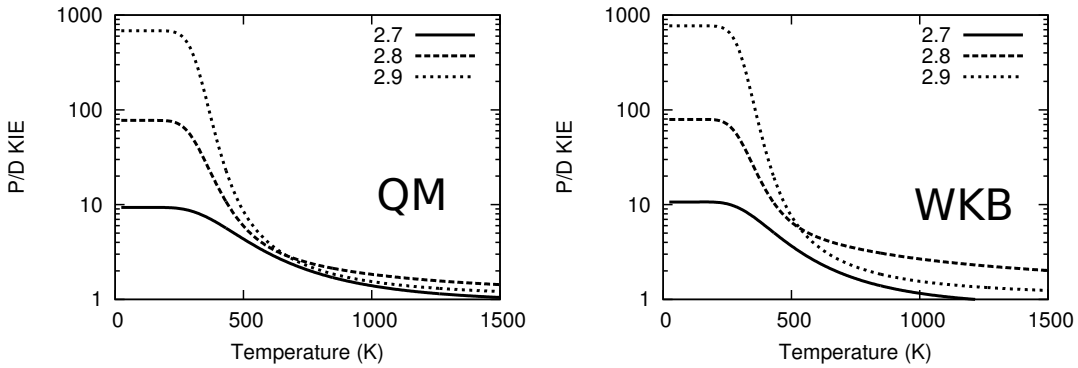

Figure 6.8 P/D kinetic isotope effect calculated as a function of temperature for QM wavepacket and WKB method.

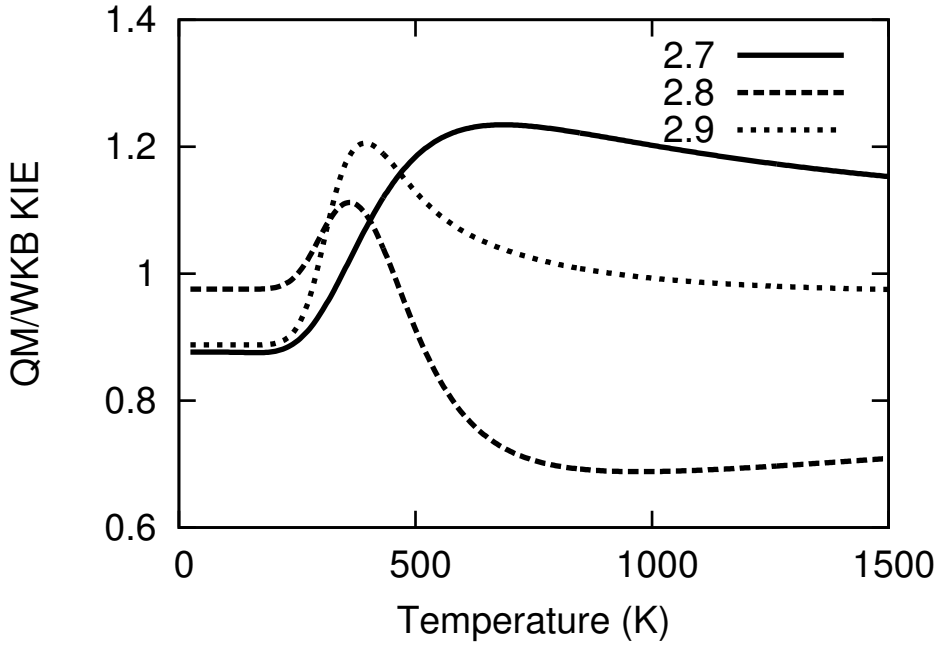

Figure 6.9 Ratio of QM/WKB KIE as a function of temperature. Better agreement is seen at low temperatures. Since the KIE is a ratio of rates within each method, a cancellation of errors is possible.

with a non-zero ρ_n as was described previously. The QC (quazi-classical) ground state is now the ground state of the isolated parabolic reactant well. In our example the QC ground state energy is always lower than the QM energy, and this produces lower rate constants. The KIE predicted by the QC approach is always higher than the QM results, but is nevertheless a good estimate.

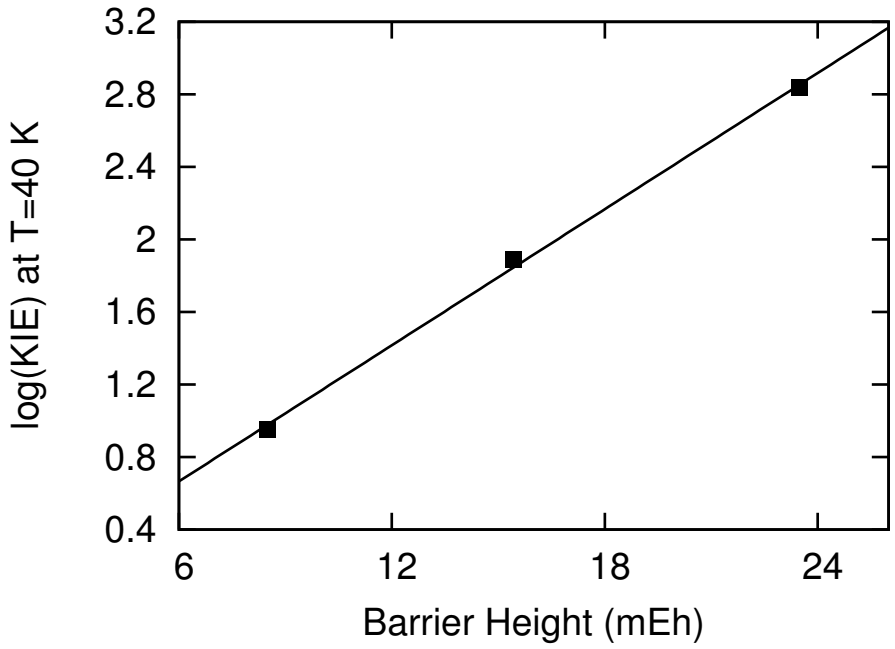

Figure 6.10 Comparison of low-energy P/D KIE to barrier height. When plotted on a logarithmic scale, a nearly linear relationship is evident.

R_{CO}		k^{QM}	k^{QC}	k^{QC}/k^{QM}	E_0^{QM} [mE$_h$]	E_0^{QC} [mE$_h$]
2.7 Å	H	0.9909(-1)	0.4463(-1)	0.4505	16.0194	15.0527
	D	0.1125(-1)	0.3232(-2)	0.2874	14.7793	13.8555
	KIE	8.8114	13.8086	1.5671*	–	–
2.8 Å	H	0.2037(-2)	0.5481(-3)	0.2691	17.7501	16.2825
	D	0.2383(-4)	0.5669(-5)	0.2379	16.1081	14.9856
	KIE	85.4646	96.6882	1.1313*	–	–
2.9 Å	H	0.1262(-4)	0.2804(-5)	0.2222	18.4041	16.7282
	D	0.1894(-7)	0.3310(-8)	0.1748	16.6112	15.3731
	KIE	666.0938	847.0630	1.2716 *	–	–

Table 6.3 Contribution of the ground state to the rate constants for the CH_3–H–OH proton transfer model obtained using fully quantum and fully quasiclassical approaches. H and D label quantities relevant to reactions with hydrogen (proton) and deuterium respectively; KIE= k_H/k_D. The ground state energies of the QM and QC descriptions are listed in the last two columns. Asterisk marks KIE^{QC}/KIE^{QM}.

6.8 Conclusions

It has been shown that the reaction rate formulation, as outlined by equation 6.2, produces reasonable results for the proton transfer reaction HO–H–CH_3. In general,

WKB results were close to QM wavepacket results, always within an order of magnitude. Therefore, WKB is a good approximation when calculating $T(E)$ in this system. When the transferring hydrogen is replaced with a deuterium atom, the KIE shows an exponential dependence on barrier height, and this suggests a tunneling mechanism in the regime of low temperatures and high barriers.

The current formulation of $k(T)$ is generalizable to any bound 1-D system in which the reactant and product state can be represented as asymptotes (for dynamics calculations), and the full potential contains eigenstates at discrete energy levels. As long as transmission probability $T(E)$ and eigenstate reactant projections ρ_n can be calculated, a rate constant can be calculated as well. Time-dependent QM wavepacket calculations should be capable of producing an accurate $T(E)$ in an arbitrary 1-D system, regardless of barrier shape, and DVR should be capable of producing accurate low-energy eigenstates likewise.

Chapter 7

Rotational Isomers of N-methylacetamide

7.1 Abstract

A laboratory experiment for an undergraduate physical chemistry course is outlined. The experimental portion may be done in the semester in which thermodynamics is covered, and the computational chemistry portion is well-suited for the semester in which quantum mechanics is covered. The enthalpy change between the two rotational isomers of N-methylacetamide was calculated using NMR spectroscopy at various temperatures. Computational chemistry software was able to reproduce the experimentally measured enthalpy change, and the high and low-energy conformers were identified.

7.2 Introduction

Nuclear magnetic resonance (NMR) spectroscopy is a tool used in nearly every field of chemical study. The most common application of NMR in chemistry is the structure determination of molecules dissolved in solvent, and this laboratory technique is typically taught in undergraduate organic chemistry courses. A study of this type typically involves the collection of a single spectrum at room temperature. The peak structure/integral intensity data, when examined under the correct pretext, can be used to construct a possible molecular structure. Typically, other spectral techniques, such as Fourier-transform infrared spectroscopy (FT-IR), are used to resolve ambiguities in the NMR results and settle on one particular molecular structure.

When the sample's temperature can be controlled, thermodynamic and kinetic properties of molecules can be examined with NMR. One such property, which we investigate in this experiment, is the enthalpy difference between molecular conformers of N-methylacetamide in thermodynamic equilibrium. The difference in enthalpy between the two conformers, ΔH, is determined from ^{1}H NMR data, and electronic structure calculations are performed to support these results. This experiment is ideal for undergraduate physical chemistry laboratories, and it includes content from thermodynamics and quantum chemistry.

7.3 Chemical system

N-methylacetamide has two rotational conformers, pictured in Fig. 7.1. This molecule

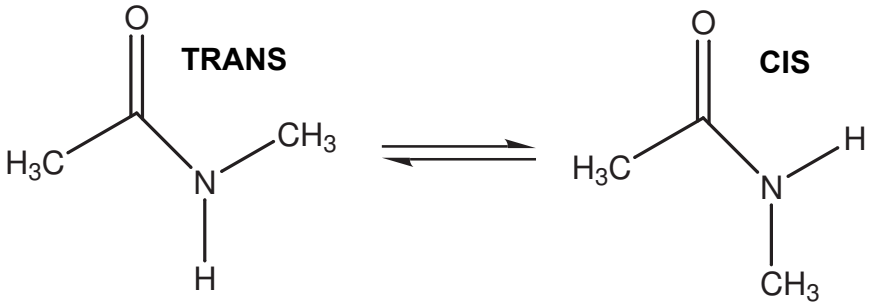

Figure 7.1 There are two rotational conformers of N-methylacetamide which pivot around the central N–C bond atom.

is the ideal candidate for investigation with ^{1}H NMR because (1) the two conformers are close enough in energy that they can both be observed near room temperature, and (2) the exchange rate between these conformers is slow enough that we can distinguish the conformer peaks with ^{1}H NMR. Conformer populations at a variety of temperatures can be used to estimate the enthalpy difference between the two forms.

7.4 Theory

This molecule contains two different rotational conformers with respect to the central N–C bond. Because this bond contains partial $\pi-$bond characteristics, there is a very high barrier between these conformers, and they are both observable in NMR near room temperature. As seen in Fig. 7.1, one conformer has the N–methyl in close proximity to the carbonyl oxygen (labeled *trans*), and the other conformer has the N–H hydrogen in close proximity to the carbonyl oxygen (labeled *cis*). Both of these conformers are known to be planar. Writing mole fractions of each conformer x_{cis} and x_{trans}, respectively, the equilibrium constant can be expressed as

$$K(T) = \frac{x_{cis}}{x_{trans}} = e^{\Delta S/R} e^{-\Delta H/RT} \tag{7.1}$$

where the entropy and enthalpy differences represent a transition from the *trans* to *cis* conformer[109, 110]. In the case of small molecules, it is usually a very good approximation to assume a negligible entropy change [109, 111], thus reducing Eq. 7.1 to

$$\frac{x_{cis}}{x_{trans}} = e^{-\Delta H/RT} \tag{7.2}$$

and we recognize that both x_{cis} and x_{trans} are dependent on temperature.

Although these equations can be solved as a function of coupling constants J, it is much simpler to just recognize that the peak integrals are proportional to the total population of ^{1}H nuclei in a particular configuration [112, 113, 114]. If we therefore assume x_i is proportional to A_i, where A_i is the area under the ^{1}H NMR peak representing conformer i, Eq. 7.2 can be written as

$$\frac{A_{cis}}{A_{trans}} = e^{-\Delta H/RT}. \tag{7.3}$$

If we have measurements over a range of temperatures, it is easy to solve for ΔH

$$\ln\left(\frac{A_{cis}}{A_{trans}}\right) = \left(\frac{-\Delta H}{R}\right)\left(\frac{1}{T}\right) \tag{7.4}$$

based on several measurements of A_{cis}/A_{trans} and a least-squares fit over a range of temperatures.

7.5 NMR Experiment

An NMR tube containing a 10 mg/mL solution of N-methylacetamide in DMSO was prepared and the tube was sealed. NMR spectra were collected at six temperatures ranging from 25°C to 50°C. The two conformers can be distinguished by focusing on the large peak at 1.76 ppm. This is a singlet representing the methyl group attached to the carbonyl (the leftmost methyl group in Fig. 7.1) in its low-energy conformer. The other conformer of this molecule is seen at 1.83 ppm, and its signal is much smaller, indicating it is a higher-energy conformer. These peaks can be seen in Fig. 7.2. As temperature increases, the high-energy conformer peak becomes more pronounced and its area increases when compared to the low-energy conformer peak. Since the peaks begin to coalesce as the temperature increases, it is necessary to perform a peak deconvolution to determine the contribution of each conformer.

As computational chemistry results will show, the *cis* conformer has a higher enthalpy than the *trans* conformer. We now adopt this notation for discussing peak calculations. Peak integrals were calculated through a deconvolution routine which includes Lorentzian and Gaussian contributions. It is very important that the deconvolution is a good fit for both peaks, as one is much smaller than the other and we will be taking a ratio, as shown in Eq. 7.4. The integrals and conformer ratios for all temperatures can be seen in Table 7.1. The far right column in Table 7.1 is inserted

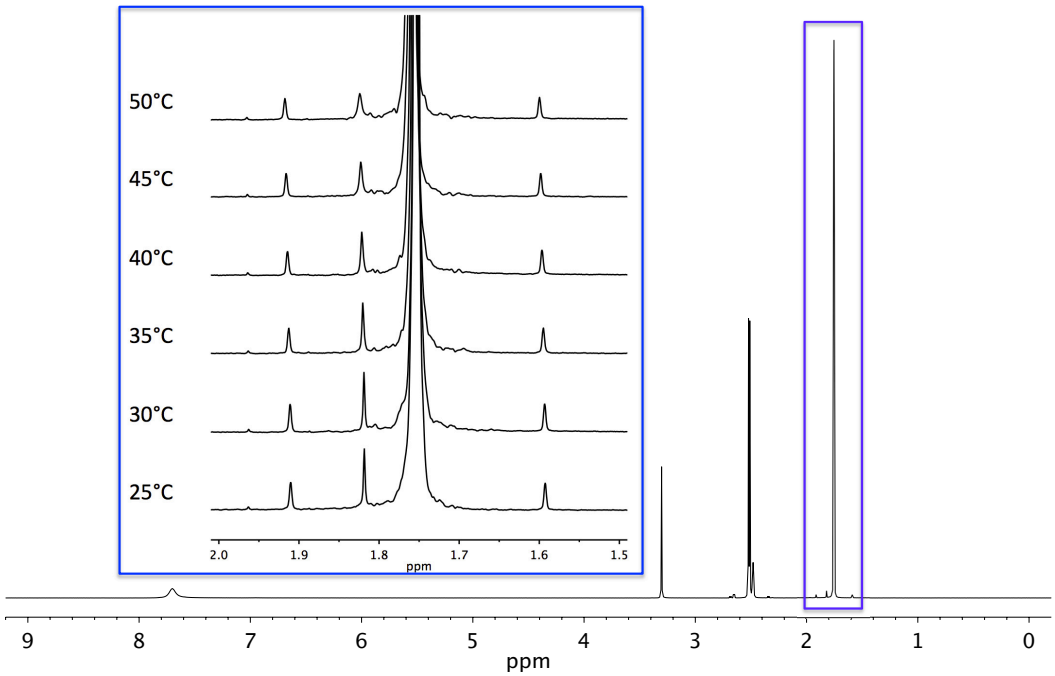

Figure 7.2 Plots of NMR spectra for all temperatures. The inset shows how the two peaks of interest change as a function of temperature. As temperature increases, the peak for the *cis* conformer becomes more pronounced, but it begins to drift closer to the large *trans* peak. The peaks at 1.59 and 1.91 ppm are ^{13}C satellites of the main resonance and can be ignored for this experiment.

Table 7.1 Areas of conformer peaks from NMR measurements. Peak areas are normalized to 100.

Temp. (°C)	A_{trans}	A_{cis}	A_{cis}/A_{trans}	$\ln(A_{cis}/A_{trans})$
25	99.1664	0.8336	0.008406	-4.7788
30	99.1001	0.8999	0.009081	-4.7016
35	98.9772	1.0228	0.010333	-4.5724
40	98.9511	1.0489	0.010600	-4.5469
45	98.9461	1.0539	0.010651	-4.5421
50	98.8173	1.1827	0.011969	-4.4254

directly into Eq. 7.4, and the slope of the line is calculated through a least-squares fit, resulting in ΔH between the two conformers.

Eq. 7.4, as approximated by our data, is shown in Fig. 7.3. The equation of the line

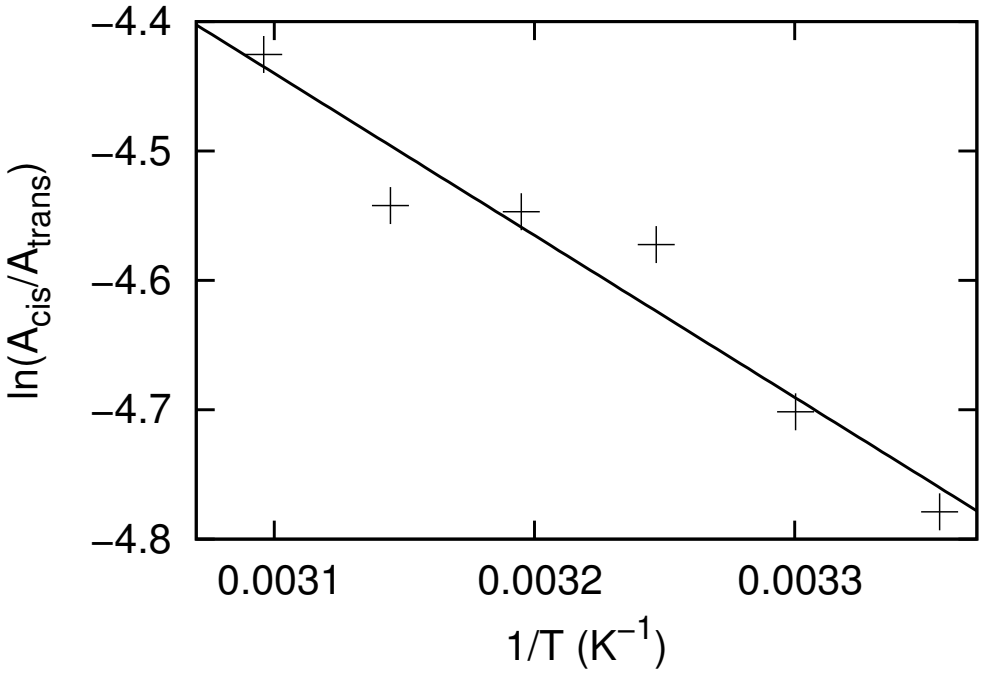

Figure 7.3 The data from Table 7.1 were plotted and fit to a line. The slope of the fit is used to calculate ΔH.

is

$$\frac{A_{cis}}{A_{trans}} = -1252.265 \text{ K} \left(\frac{1}{T}\right) - 0.558 \tag{7.5}$$

with a correlation coefficient of $R^2 = 0.931$. For the purpose of this study, we assume the non-zero y-intercept is an artifact of the fitting, and it is ignored. It is, however, a good approximation to assume that within our temperature range, a linear relationship is accurate. The slope was used to solve for ΔH,

$$-1252.265 \text{ K} = \frac{-\Delta H}{R} \tag{7.6}$$

which yielded $\Delta H = 2.489$ kcal/mol.

7.7 Computational methods and results

The results from the NMR measurements can provide a value for ΔH, but the identity of each conformer cannot be identified with this technique. Electronic structure calculations were performed to determine the geometry of the conformers and calculate the energy of each one. SPARTAN '10 software on a desktop workstation was used for these calculations[115]. A conformer search was performed at the B3LYP/6-31G** level of theory, and only two conformers were identified. The lower-energy conformer corresponded to *trans*, and the higher-energy conformer corresponded to *cis*, as shown in Fig. 7.1. The electronic structure calculations produced $\Delta H \approx \Delta E_0 = 2.612$ kcal/mol, where ΔE_0 is the change in total electronic energy at 0 K.

7.8 Conclusions

NMR spectroscopy supplemented with computational chemistry techniques were used to determine the molecular structure and enthalpy difference between the two conformers of N-methylacetamide. Electronic structure calculations show that the *cis* conformer is 2.612 kcal/mol higher in energy than *trans*, as seen in Fig. 7.1. Results from ^{1}H NMR spectra from 25°C to 50°C show two conformers with $\Delta H = 2.489$ kcal/mol, in excellent agreement with computational results, differing by only 0.123 kcal/mol. These results also agree very well with previous NMR and computational work in which the authors have found $\Delta H = 2.1 - 2.5$ kcal/mol[116, 117, 118, 119].

This experiment is excellent for an undergraduate physical chemistry course, and it can be separated into two sections. The NMR portion of the experiment should be performed during the semester in which thermodynamics is covered. When quantum chemistry is being taught, the computational chemistry section can be performed. In any case, students should eventually complete both parts, and then compare the results between them. The predictive power of computational chemistry is highlighted,

and insight into the molecular structure can create a picture of the conformer conversion mechanism. The calculation of ΔH using NMR will build on the students' experience from previous organic chemistry classes, and it will provide an opportunity to apply their skills to calculate a thermodynamic property in a physical chemistry setting.

References

[1] Light, J. C.; Jr., T. C. *Adv. Chem. Phys.* **2000**, *114*, 263.

[2] Webb, S. P.; Hammes-Schiffer, S. *J. Chem. Phys.* **2000**, *113*, 5214–5227.

[3] Meyer, H.; Worth, G. A. *Theor. Chem. Acc.* **2003**, *109*, 251–267.

[4] Nyman, G.; Yu, H. *Int. Rev. Phys. Chem.* **2013**, *32*, 39–95.

[5] Fang, J.; Guo, H. *J. Molec. Str.* **1995**, *341*, 201–215.

[6] Miller, W. H. *J. Phys. Chem.* **2001**, *105*, 2942–2955.

[7] Craig, I. R.; Manolopoulos, D. E. *J. Chem. Phys.* **2005**, *122*, 084106.

[8] Bohm, D. *Phys. Rev.* **1952**, *85*, 166–179.

[9] Garashchuk, S.; Rassolov, V. A. *Chem. Phys. Lett.* **2002**, *364*, 562–567.

[10] Zheng, G.; Lundberg, M.; Jakowski, J.; Vreven, T.; Frisch, M. J.; Morokuma, K. *Int. J. Quantum Chem.* **2009**, *109*, 1841–1854.

[11] Elstner, M. *Theor. Chem. Acc.* **2006**, *116*, 316–325.

[12] Klinman, J. P. *Chem. Phys. Lett.* **2009**, *471*, 179–193.

[13] Garashchuk, S.; Dixit, V.; Gu, B.; Mazzuca, J. *J. Chem. Phys.* **2013**, *138*, 054107.

[14] Geer, M. F.; Mazzuca, J.; Smith, M. D.; Shimizu, L. S. *Cryst. Eng. Comm.* **2013**, *15*, 9923–9929.

[15] Mazzuca, J.; Garashchuk, S.; Jakowski, J. *Chem. Phys. Lett.* **2012**, *542*, 153–158.

[16] Barroso, M.; Arnaut, L. G.; Formosinho, S. J. *J. Phys. Org. Chem.* **2008**, *21*, 659–665.

[17] Knapp, M. J.; Rickert, K.; Klinman, J. P. *J. Am. Chem. Soc.* **2002**, *124*, 3865–3874.

[18] Olsson, M. H. M.; Siegbahn, P. E. M.; Warshel, A. *J. Am. Chem. Soc.* **2004**, *126*, 2820–2828.

[19] Iyengar, S. S.; Sumner, I.; Jakowski, J. *J. Chem. Phys. B.* **2008**, *112*, 7601–7613.

[20] Meyer, P. M.; Klinman, J. P. *Chem. Phys.* **2005**, *319*, 283–296.

[21] Ranaghan, K. E.; Mulholland, A. J. *Interdiscip. Sci. Comp. Life Sci.* **2010**, *2*, 78–97.

[22] Garashchuk, S.; Rassolov, V. A. *J. Chem. Phys.* **2004**, *120*, 1181–1190.

[23] Venkataraman, C.; Miller, W. H. *J. Chem. Phys.* **2007**, *126*, 094104.

[24] Garashchuk, S. *J. Phys. Chem. A.* **2009**, *113*, 4451–4456.

[25] Madelung, E. *Z. Phys.* **1926**, *40*, 322–326.

[26] Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling, W. T. *Numerical Recipes: The Art of Scientific Computing*, 3rd ed.; Cambridge University Press: Cambridge, 2007.

[27] Goldfarb, Y.; Degani, I.; Tannor, D. J. *J. Chem. Phys.* **2006**, *125*, 231103.

[28] Wyatt, R. E.; Bittner, W. R. *J. Chem. Phys.* **2000**, *113*, 8898–8907.

[29] Gindensperger, E.; Meier, C.; Beswick, J. A. *J. Chem. Phys.* **2000**, *113*, 9369–9372.

[30] Kendrick, B. K. *J. Chem. Phys.* **2003**, *119*, 5805–5817.

[31] Lopreore, C. L.; Wyatt, R. E. *Phys. Rev. Lett.* **1999**, *82*, 5190–5193.

[32] Trahan, C. J.; Hughes, K.; Wyatt, R. E. *J. Chem. Phys.* **2003**, *118*, 9911–9914.

[33] Garashchuk, S.; Rassolov, V.; Prezhdo, O. In *Reviews in Computational Chemistry*; Wiley, 2011; Vol. 27, Chapter Semiclassical Bohmian dynamics, pp 111–210.

[34] Park, K.; Poirier, B.; Parlant, G. *J. Chem. Phys.* **2008**, *129*, 194112.

[35] Leforestier, C.; Bisseling, R. H.; Cerjan, C. *J. Comp. Phys.* **1991**, *94*, 59–80.

[36] Feit, M. D.; J. A. Fleck, J.; Steiger, A. *J. Comp. Phys.* **1982**, *47*, 412–433.

[37] Tresadern, G.; McNamara, J. P.; Mohr, M.; Wang, H.; Burton, N. A.; Hillier, I. H. *Chem. Phys. Lett.* **2002**, *358*, 489–494.

[38] Tejero, I.; Garcia-Viloca, M.; Gonzalez-Lafont, A.; Lluch, J. M.; York, D. M. *J. Phys. Chem. B.* **2006**, *110*, 24708–24719.

[39] Mavri, J.; Liu, H.; Olsson, M. H. M.; Warshel, A. *J. Phys. Chem.* **2008**, *112*, 5950–5954.

[40] Klinman, J. P. *Proced. Chem.* **2011**, *3*, 291–305.

[41] Sutcliffe, M. J.; Scrutton, N. S. *Eur. J. Biochem.* **2002**, *269*, 3096–3102.

[42] Basran, J.; Sutcliffe, M. J.; Scrutton, N. S. *Biochem.* **1999**, *38*, 3218–3222.

[43] Hammes-Schiffer, S.; Hatcher, E.; Ishikita, H.; Skone, J. H.; Soudackov, A. V. *Coord. Chem. Rev.* **2008**, *252*, 384–394.

[44] Liang, Z.; Klinman, J. P. *Curr. Op. Str. Bio.* **2004**, *14*, 648–655.

[45] Rickert, K. W.; Klinman, J. P. *Biochemistry* **1999**, *38*, 12218–12228.

[46] Borowski, T.; Krol, M.; Chruszcz, M.; Broclawik, E. *J. Phys. Chem. B* **2001**, *105*, 12212–12220.

[47] Tejero, I.; Eriksson, L. A.; Gonzalez-Lafont, A.; Marquet, J.; Lluch, J. M. *J. Phys. Chem. B.* **2004**, *108*, 13831–13838.

[48] Klinman, J. P. *Pure Appl. Chem.* **2003**, *75*, 601–608.

[49] Boekelheide, N.; Salomòn-Ferrer, R.; III, T. F. M. *PNAS* **2011**, *108*, 16159–16163.

[50] Hatcher, E.; Soudackov, A. V.; Hammes-Schiffer, S. *J. Am. Chem. Soc.* **2007**, *129*, 187–196.

[51] Garashchuk, S.; Jakowski, J.; Wang, L.; Sumpter, B. G. *J. Chem. Theory Comput.* **2013**, *9*, 5221–5235.

[52] Elstner, M.; Porezag, D.; Jungnickel, G.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. *Phys. Rev. B* **1998**, *11*, 7260–7268.

[53] Aradi, B.; Hourahine, B.; Frauenheim, T. *J. Phys. Chem. A* **2007**, *111*, 5678–5684.

[54] Zheng, G.; Witek, H. A.; Bobadova-Parvanova, P.; Irle, S.; Musaev, D. G.; Prabhakar, R.; Morokuma, K. *J. Chem. Theory Comput.* **2007**, *3*, 1349–1367.

[55] Shao, Y. et al. *Phys. Chem. Chem. Phys.*

[56] Walewski, L.; Krachtus, D.; Fischer, S.; Smith, J. C.; Bala, P.; Lesyng, B. *Int. J. Quant. Chem.* **2006**, *106*, 636–640.

[57] Maupin, C. M.; Aradi, B.; Voth, G. A. *J. Phys. Chem. B* **2010**, *114*, 6922–6931.

[58] Yang, Y.; Yu, H.; York, D.; Cui, Q.; Elstner, M. *J. Phys, Chem. A* **2007**, *111*, 10861–10873.

[59] Goyal, P.; Elstner, M.; Cui, Q. *J. Phys. Chem. B* **2011**, *115*, 6790–6805.

[60] Miller, W. H.; Schwartz, S. D.; Tromp, J. W. *J. Chem. Phys.* **1983**, *79*, 4889–4898.

[61] Park, T. J.; Light, J. C. *J. Chem. Phys.* **1988**, *88*, 4897–4912.

[62] Pollak, E. *J. Chem. Phys.* **1997**, *107*, 64–69.

[63] Garashchuk, S. *Chem. Phys. Lett.* **2010**, *491*, 96–101.

[64] Ehrenfest, P. *Z. Phys.* **1927**, *45*, 455.

[65] Tully, J. C. *Faraday Discuss.* **1998**, *110*, 407–419.

[66] Garashchuk, S.; Volkov, M. *Mol. Phys.* **2012**, *110*, 985–993.

[67] Garashchuk, S.; Rassolov, V.; Prezhdo, O. *Reviews in computational chemistry* **2011**, *27*,.

[68] Garashchuk, S. *J. Chem. Phys.* **2010**, *132*, 014112.

[69] Chattaraj, P. K. *Quantum Trajectories*; CRC Press, 2010; pp 361–379.

[70] Garashchuk, S.; Mazzuca, J.; Vazhappilly, T. *J. Chem. Phys.* **2011**, *135*, 034104.

[71] de Broglie, L. *An introduction to the study of wave mechanics*; E. P. Dutton and Company, Inc.: New York, 1930.

[72] Meyer, H. D.; Manthe, U.; Cederbaum, L. S. *Chem. Phys. Lett.* **1990**, *165*, 73–78.

[73] Wyatt, R. E.; Chou, C. C. *J. Chem. Phys.* **2008**, *128*, 234106.

[74] Liu, J.; Makri, N. *Molec. Phys.* **2005**, *103*, 1083–1090.

[75] Goldfarb, Y.; Degani, I.; Tannor, D. J. *Chem. Phys.* **2007**, *338*, 106–112.

[76] Raymond, P. *Field theory: a modern primer*; Addison-Wesley, 1990.

[77] Feynman, R. P.; Hibbs, A. R. *Quantum Mechanics and Path Integrals*; McGraw-Hill, 1965.

[78] Miller, W. H. *J. Chem. Phys.* **1971**, *55*, 3146.

[79] Blume, D.; Lewerenz, M.; Niyaz, P.; Whaley, K. B. *Phys. Rev. E* **1997**, *55*, 3664–3675.

[80] Ceperley, D. M.; Mitas, L. *Monte Carlo methods in quantum chemistry*; Wiley, 1996; Vol. 93.

[81] Jr., W. A. L.; Mitas, L.; Hammond, B. *Chem. Phys. Lett.* **2009**, *478*, 1–10.

[82] Viel, A.; Coutinho-Neto, M. D.; Manthe, U. *J. Chem. Phys.* **2007**, *126*, 024308.

[83] Hinkle, C. E.; McCoy, A. B. *J. Phys. Chem. A* **2008**, *112*, 2058–2064.

[84] Rassolov, V. A.; Garashchuk, S.; Schatz, G. C. *J. Phys. Chem. A* **2006**, *110*, 5530–5536.

[85] Garashchuk, S.; Vazhappilly, T. *J. Phys. Chem. C* **2010**, *114*, 20595–20602.

[86] Wyatt, R. E. *Quantum Dynamics with Trajectories: Introduction to Quantum Hydrodynamics*; Springer-Verlag, 2005.

[87] Maddox, J. B.; Bittner, E. R. *J. Phys. Chem. B* **2002**, *106*, 7981–7990.

[88] Maddox, J. B.; Bittner, E. R. *J. Chem. Phys.* **2003**, *119*, 6465–6474.

[89] Donoso, A.; Zheng, Y.; Martens, C. *J. Chem. Phys.* **2003**, *119*, 5010–5020.

[90] Wyatt, R. E. *J. Chem. Phys.* **2002**, *117*, 9569–9573.

[91] Trahan, C. J.; Wyatt, R. E. *J. Chem. Phys.* **2003**, *118*, 4784–4790.

[92] Ppoirier, B. *J. Theor and Comp. Chem.* **2007**, *6*, 99–125.

[93] David, J. K.; Wyatt, R. E. *J. Chem. Phys.* **2008**, *128*, 094102.

[94] Liu, J.; Makri, N. *J. Phys. Chem. A* **2004**, *108*, 5408–5416.

[95] Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling, W. T. *Numerical Recipes: The Art of Scientific Computing*, 3rd ed.; Cambridge University Press: Cambridge, 1992.

[96] Bittner, E. R. *J. Chem. Phys.* **2003**, *119*, 1358–1364.

[97] Herman, M. F.; Kluk, E. *Chem. Phys.* **1984**, *91*, 27–34.

[98] Heller, E. J. *J. Chem. Phys.* **1981**, *75*, 2923–2931.

[99] Garashchuk, S.; Gu, B.; Mazzuca, J. W. *J. Theor. Chem.* **2014**.

[100] Weiner, J. H. *J. Chem. Phys.* **1978**, *68*, 2492–2506.

[101] Hay, S.; Johannissen, L. O.; Sutcliffe, M. J.; Scrutton, N. S. *Biophys. J.* **2010**, *98*, 121–128.

[102] Weiner, J. H. *J. Chem. Phys.* **1978**, *69*, 4743–4749.

[103] Kohen, A.; Klinman, J. P. *Chemis. and Biol.* **1999**, *6*, R191–R198.

[104] Tannor, D.; Weeks, D. *J. Chem. Phys.* **1993**, *98*, 3884–3893.

[105] Vigneron, J. P.; Lambin, P. *J. Phys. A: Math.* **1980**, *13*, 1135–1144.

[106] Miller, W. H. *Science* **1986**, *233*, 171–177.

[107] Light, J. C.; Hamilton, I. P.; Lill, J. V. *J. Chem. Phys.* **1985**, *82*, 1400–1409.

[108] Colbert, D. T.; Miller, W. H. *J. Chem. Phys.* **1992**, *96*, 1982–1991.

[109] Govil, G.; Bernstein, H. J. *J. Chem. Phys.* **1967**, *47*, 2818–2823.

[110] Abraham, R. J.; Bernstein, H. J. *Can. J. Chem.* **1961**, *39*, 39–41.

[111] Fessenden, R. W.; Waugh, J. S. *J. Chem. Phys.* **1962**, *37*, 1466–1467.

[112] Kraszni, M.; Szakacs, Z.; Noszal, B. *Anal. Bioanal. Chem.* **2004**, *378*, 1449–1463.

[113] Hutowsky, H. S.; Belford, G. G.; McMahon, P. E. *J. Chem. Phys.* **1962**, *36*, 3353–3368.

[114] Pople, J. A. *Molec. Phys.* **1957**, *91*, 385–390.

[115] Inc., W. **2010**.

[116] Song, S.; Asher, S. A.; Krimm, S.; Shaw, K. D. *J. Am. Chem. Soc.* **1991**, *113*, 1155–1163.

[117] Jorgensen, W. L.; Gao, J. *J. Am. Chem. Soc.* **1988**, *110*, 4212–4216.

[118] Radzicka, A.; Pedersen, L.; Wolfenden, R. *Biochemistry* **1988**, *27*, 4538–4541.

[119] Kang, Y. K. *J. Mol. Struct.* **2001**, *546*, 183–193.

[120] Eckart, C. *Phys. Rev.* **1930**, *35*, 1303–1309.

[121] Johnston, H. S.; Heicklen, J. *Advances in Chem. Phys.* **1961**, *3*, 532–533.

Appendix A

Adjusting the DFTB Repulsive Spline for HO–H–CH_3 Model System

A.1 Abstract

In an effort to improve agreement between DFTB and B3LYP energy curves for the collinear HO–H–CH_3 reaction, a procedure to change Slater-Koster (SK) parameter files was implemented. Repulsive spline terms were fit using a combination of B3LYP results for C–H and O–H interaction as well as analytical functions. Agreement for the HO–H–CH_3 system was greatly improved, and the curves are nearly indistinguishable after the fitting. Portability of the parameter files was demonstrated by replacing the an external hydrogen with a methyl group, and then testing the new parameters on H_3CO–H–CH_3. The quality of the curves improved greatly with the new parameter files present in this new system, suggesting that some form of parameter file portability between systems may be possible.

A.2 Introduction

Density functional tight-binding (DFTB) is a semi-empirical electronic structure method. Empirical parameters are included in the form of Slater-Koster (SK) files (file extension `.skf`), and one file is required for each type of atom-atom interaction in the system. It was found that the energy profile for proton transfer in the active site of soybean lipoxygenase-1 (SLO-1) did not agree with B3LYP results. We investigate a model system to fix this problem, and this process is detailed in this section.

Agreement between DFTB and B3LYP energy profile curves for the hydrogen transfer in OH–H–CH_3 is achieved through the parameterization of the repulsive spline term in the SK file.

A.3 Standard energy profiles

The collinear hydrogen transfer from carbon to oxygen in HO–H–CH_3 is examined, and this reaction is shown in Fig. A.1. This system is a prototype for hydrogen

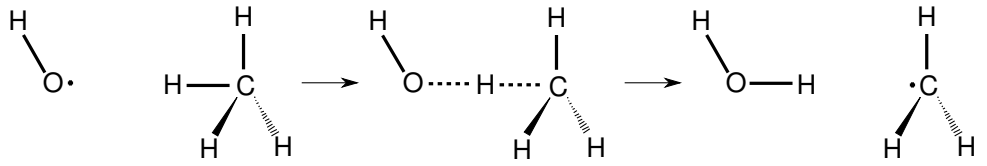

Figure A.1 Schematic of collinear reaction that is being used as a benchmark.

transfer with a donor carbon and acceptor oxygen (much like the active site of soybean lipoxygenase-1). Configurations with C–O fixed at 2.4, 2.45, 2.5, 2.55, 2.6, 2.65, 2.7, 2.8, 2.9, and 3.0 Å were examined. Optimized geometries for every O–H–C frozen configuration were obtained using DFT with the B3LYP exchange-correlation functional and LANL2DZ basis using Q-Chem ver. 4.01. These geometries were then imported into DFTB+ ver. 1.2. and the total energy was calculated. These energy curves were then compared to the B3LYP/LANL2DZ results, which we take as the target potential energy curves. Standard DFTB calculations are compared to B3LYP calculations in Fig. A.2. The goal is to adjust DFTB parameter files to minimize the trend discrepancies between the energy profiles shown.

A.4 Procedure for adjustment of SK files

The total DFTB energy can be written as

$$E_{el} = \sum_i 2 f_i \langle \phi_i | H_0 | \phi_i \rangle + \frac{1}{2} \sum_{\substack{A,B \\ A \neq B}} \gamma^{AB} \Delta q^A \Delta q^B + \sum_{A>B} E_{rep}^{AB}. \tag{A.1}$$

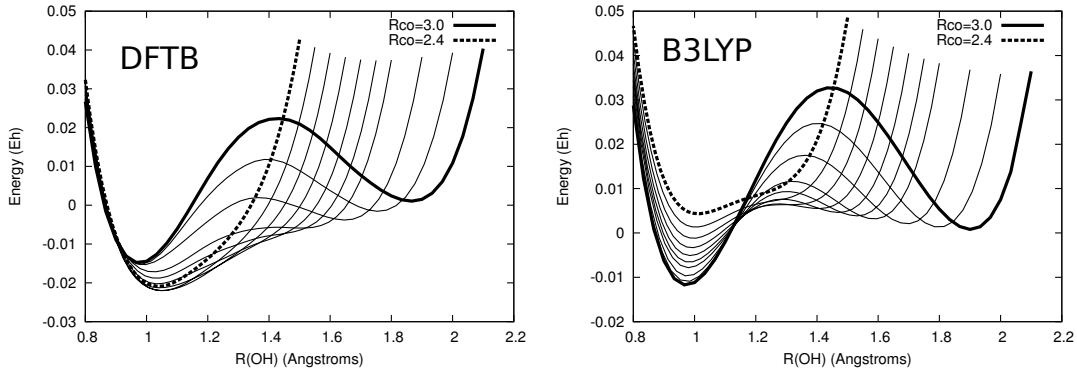

Figure A.2 Comparison of DFTB curves to those generated with B3LYP. The most obvious problems are the energy ordering of the O–H minima is reversed, and the same goes for the C–H minima.

The first term on the right-hand-side is a sum over molecular orbitals labeled i, and f_i is an orbital occupation number that ranges between 0 and 1. The second term describes electron-electron interaction between the Mulliken charges Δq at different atomic centers; this term is treated self-consistently[52]. E_{rep} consists of the pair-wise repulsive interactions between atomic centers A and B, dependent only on the distance between A and B. We define

$$E_{band}^{AB} = \sum_i 2 f_i \langle \phi_i | H_0 | \phi_i \rangle + \frac{1}{2} \sum_{\substack{A,B \\ A \neq B}} \gamma^{AB} \Delta q^A \Delta q^B \tag{A.2}$$

for simplicity. This term contains all of the angular properties of inter-atomic interactions. The E_{rep} term is a cubic spline function of atom-atom distance. All SK files are specific ONLY for two-atom interactions, and for a given system, an SK file is required for every possible two-atom combination. For example, for ethane, we require `cc.skf`, `ch.skf`, `hc.skf`, and `hh.skf`. The files `ch.skf` and `hc.skf` may contain different E_{band} parameters, but identical E_{rep} (spline) parameters.

It is much simpler to adjust E_{rep} because it is a one-dimensional cubic spline, a function of the distance between two atomic centers, whereas the generation of E_{band} Hamiltonian terms is much more involved. In this study, it is assumed that the E_{band}

terms are sufficiently good, and any inaccuracies can be accounted for in the E_{rep} term. For now, the transferring hydrogen is considered. The most important terms are a result of C–H interactions and O–H interactions. We would like to, for the sake of simplicity, only change the way in which oxygen and carbon interact with the transferring hydrogen. The parameters from the *mio* set [52] have been more thoroughly optimized to work in a variety of organic systems, and changing these parameters for the other hydrogen atoms may be detrimental to the description of the chemical system as a whole. Therefore, the transferring hydrogen will be treated as a special hydrogen which we denote as **X** where appropriate. The reaction then looks like HO–**X**–CH_3.

The corrected E_{rep} for a particular atom pair can be calculated as

$$E_{rep} = E_{B3LYP} - E_{DFTB}^{no\ rep} \tag{A.3}$$

where E_{B3LYP} is the total B3LYP energy for C–H or O–H at varying distances, and $E_{DFTB}^{no\ rep}$ is the energy calculated with DFTB, but with the repulsive spline set to 0. In order to correct the splines for the transferring hydrogen, energy profiles must be generated for a C–H and O–H bond using B3LYP. The version of DFTB which we are using ignores spin-polarization and dispersion effects. A C–H energy profile was calculated using H_3C–H, and the O–H energy profile was calculated using HO–H. Figure A.3 shows $E_{B3LYP}^{rep} = E_{B3LYP} - E_{DFTB}^{no\ rep}$ and $E_{DFTB}^{rep} = E_{DFTB} - E_{DFTB}^{no\ rep}$ and it is clear that at long distances, some features are being ignored by DFTB.

It is expected that when E_{rep} is calculated using DFTB, that the results should be identical to the solid lines in Fig. A.3. Using E_{B3LYP} and $E_{DFTB}^{no\ rep}$ over the inter-atomic distances as input, a new repulsive spline was produced for O–H interactions, and it can be seen in Fig. A.4. This spline is directly implemented into the SK file, so the file must simply be referenced in the DFTB calculation for O–**X** interactions. No post-calculation fitting is performed to use these repulsive splines once they are

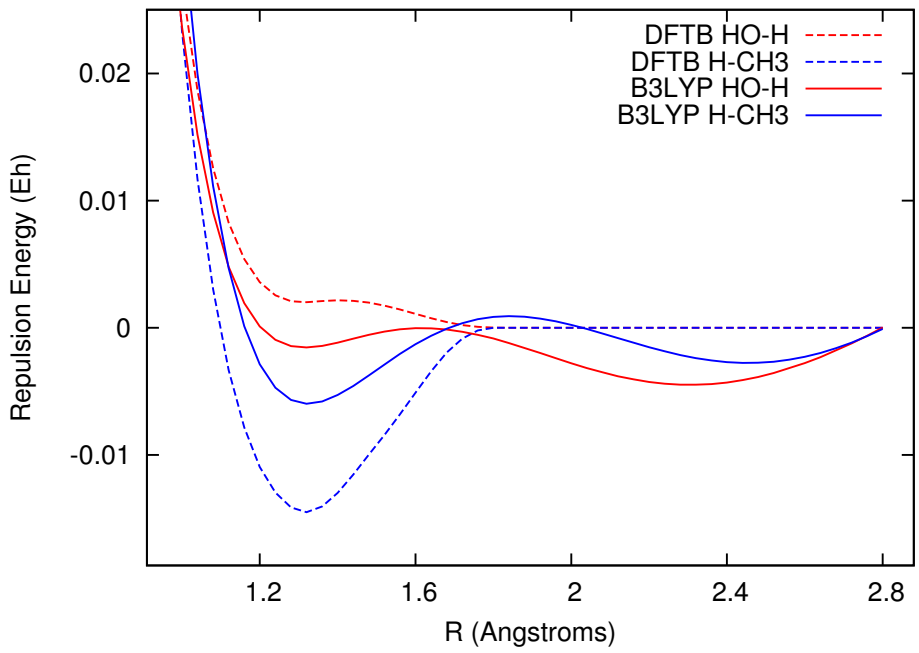

Figure A.3 The repulsion energy E_{rep} is shown for B3LYP calculations (solid lines) and DFTB calculations (dashed lines) for O–H and C–H bonds.

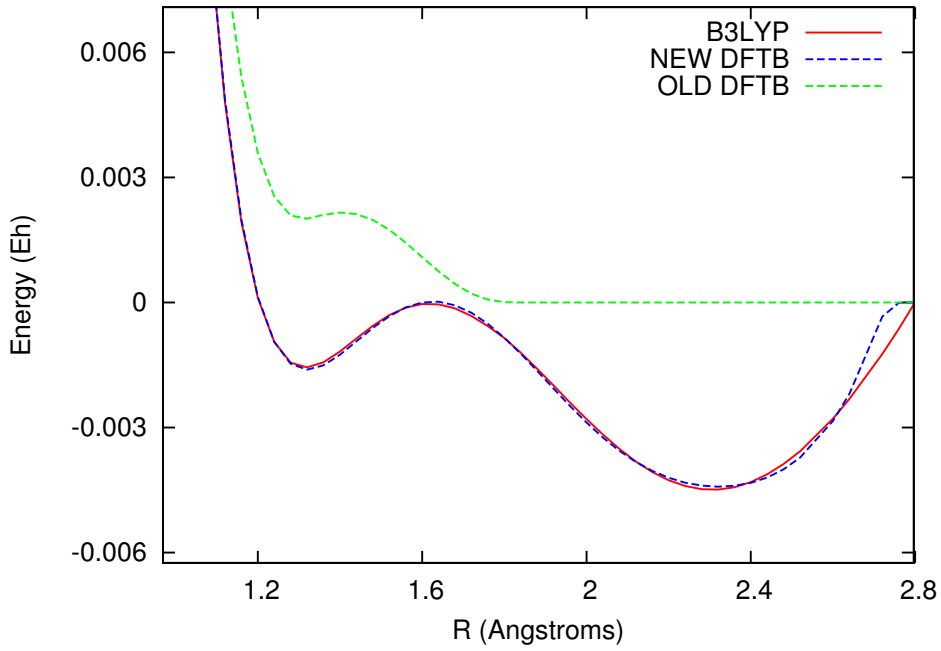

Figure A.4 Plot of E_{rep} for HO–H before and after fitting to B3LYP results.

implemented. Since the primary goal is agreement in hydrogen transfer reactions, specifically HO–H–CH_3, the effectiveness of these two-atom parameter files must be evaluated for this, "three-atom" configuration.

A.5 Splines generated using only B3LYP curves

The new parameter files are designed *only* for C–X and O–X interactions in the HO–X–CH_3 reaction. All exterior hydrogens are described by the standard *mio* parameter set. Utilizing the parameter files designed in the previous section, the curves shown in Fig. A.5 are produced. The trend previously seen at the O–H minima is no

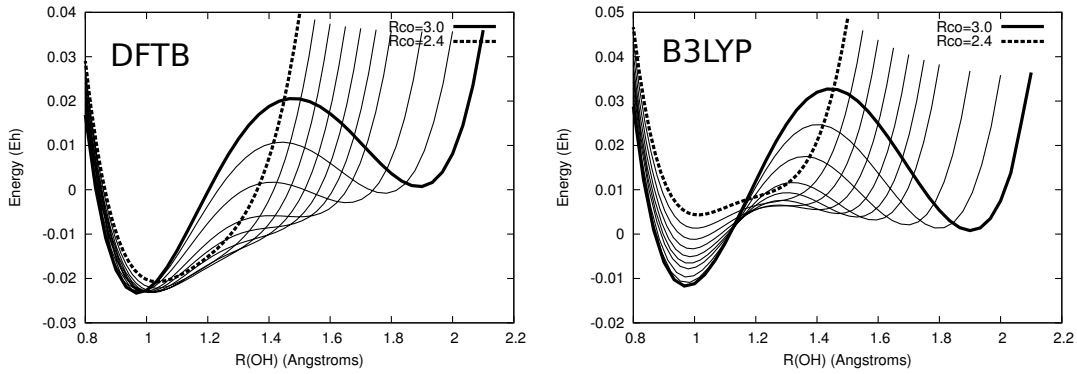

Figure A.5 Comparison of DFTB curves with new B3LYP E_{rep} terms to those generated with B3LYP. While the O–H minima trend is mitigated a bit, overall the curves are not much better than before.

longer there, but there is still substantial disagreement with the B3LYP results. It seems that these two-atom terms, at least for this type of reaction, are not directly portable to this three-atom ..O–H–C.. configuration. Since all parameter files and splines are, by definition, two-atom objects, we must work within this framework to achieve portability. It is possible, however, to modify these two-atom parameters in systems which contain our other atoms of interest, thereby implicitly accounting for three-atom interactions.

A.6 Achieving agreement in the HO–H–CH_3 system

For portability and practical reasons, it is desirable to continue treating only C–X and O–X interactions in a special way, and avoid parameterizing O–C interactions directly. Let's use the trends of the B3LYP curves in Fig. A.5 and construct an additive energy function which will improve the corresponding DFTB curves. Once these functions are constructed, they are included in the spline-fitting procedure and included directly into the C–X and O–X SK files. Let us examine the trend on the right of the curves, specifically the ordering of the C–H minima. In the DFTB results, this minimum decreases in energy as C–O (and thus, O–X) decreases. It makes sense to construct a function of R_{OX} to correct this problem. A function of the form

$$E_{rep} = E_{rep,0} + f(R_{OH}) \tag{A.4}$$

where

$$f(R_{OH}) = E_{max} \times \frac{1}{e^{\alpha(x-x_0)} + 1} \tag{A.5}$$

for these two-atom interactions consists of a switching function with an intermediate region whose slope depends on α. The function that was used to reconstruct the O–H spline, and thus correct the trend of C–H minima is

$$f(R_{OH}) = \frac{25\text{mEh}}{e^{4(x-(1.164+1.905)/2)} + 1} \tag{A.6}$$

and the parameters of x_0 were chosen to correspond to the range of the plots which we wish to adjust. 1.905 Bohr corresponds to the furthest C–H minimum, and 1.164 Bohr corresponds to the smallest x we want to adjust. The parameter of 25 mEh was chosen to optimize the trend correction. This function can be seen over the curves it is meant to correct in Fig. A.6.

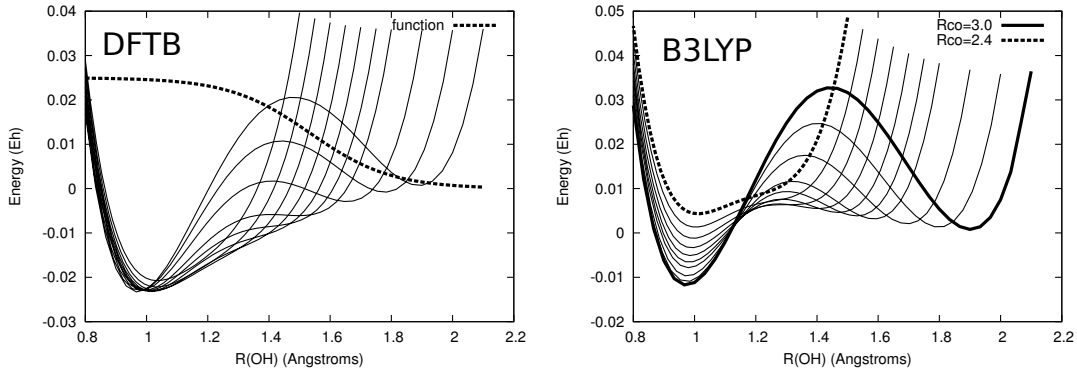

Figure A.6 A plot of Eq. A.6 plotted over the curves it is meant to fix, and the B3LYP (goal) curves for comparison.

A function was also generated for C–H interactions to adjust the ordering of O–H minima. Each O–H minimum becomes progressively higher in energy as C–O (and thus, C–H) decreases. The function for this term, as well as the function for the one in Eq. A.6 are shown in Eq. A.7 and A.8 in units of Bohr rather than Å since the SK files are in atomic units.

$$f(R_{OH}) = \frac{25\text{mEh}}{e^{4(x-(2.2+3.6)/2)}+1} \tag{A.7}$$

$$f(R_{CH}) = \frac{18\text{mEh}}{e^{5(x-(3.97+2.83)/2)}+1} \tag{A.8}$$

where

$$3.97\text{ Bohr} = \frac{(3.0-0.9)\text{Å}}{0.52918\text{ Å Bohr}^{-1}}, \qquad 2.83\text{ Bohr} = \frac{(2.4-0.9)\text{Å}}{0.52918\text{ Å Bohr}^{-1}} \tag{A.9}$$

and (3.0-0.9) Å corresponds to the R_{CH} distance at the lowest energy O–H minimum, and (2.4-0.9) Å corresponds to the R_{CH} at the highest energy O–H minimum. Now the O–H minima are moved up somewhat linearly depending on the C–H distance when $R_{OH} = 0.9$ Å. When functions A.7 and A.8 are included into the generation of SK files, the curves in Fig. A.7 are generated for the HO–**X**–CH_3 system.

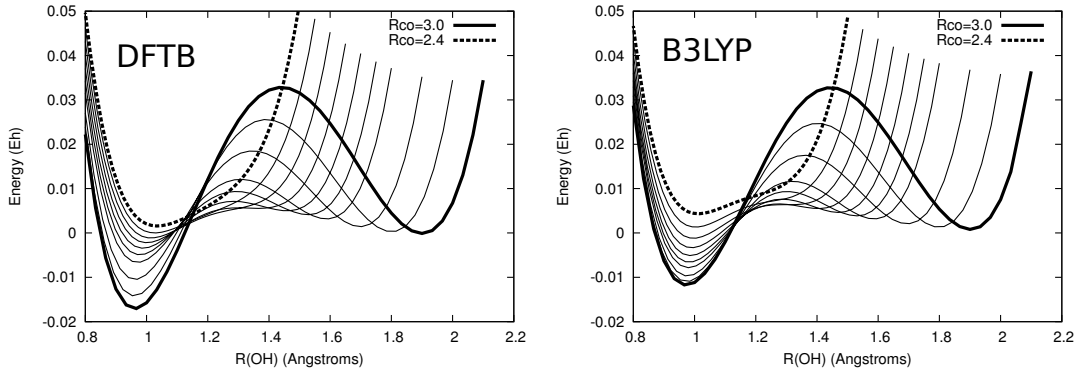

Figure A.7 DFTB curves show excellent agreement with B3LYP curves once the functions shown in Eq. A.7 and A.8 are used to generate repulsive splines.

The agreement is much better than what we see in Figs. A.2 and A.5. Both the barrier height and local minima are in quantitative agreement with the B3LYP results. The most promising aspect of this result is that a simple one-dimensional function, applied *only* to atom-atom interactions which involve the special hydrogen (C–X and O–X) have been changed. There have been no changes to the C–O SK files in this system, and the O–H–C interactions are implicit in the C–X and O–X SK files. Since no changes were made to E_{band}, it is unclear how these new parameter files will perform if the non-transferring hydrogens are replaced with something else.

A.7 Application of new parameter files to H_3CO–H–CH_3

Portability of the new parameter files was examined by calculation of energy profiles for the H_3CO–H–CH_3 hydrogen transfer reaction. Recognizing that we are only adjusting the transferring hydrogen, we can write the reaction as H_3CO–**X**–CH_3. The un-corrected DFTB curves can be seen in Fig. A.8. The new parameter files were used to generate the curves in Fig. A.9. The agreement shown in Fig. A.9 is much better than when the standard *mio* set is used, but not quite as good as the agreement in the system without the methyl group for which the parameters were designed. It is, however, good enough to suggest that adjusted SK parameter files in

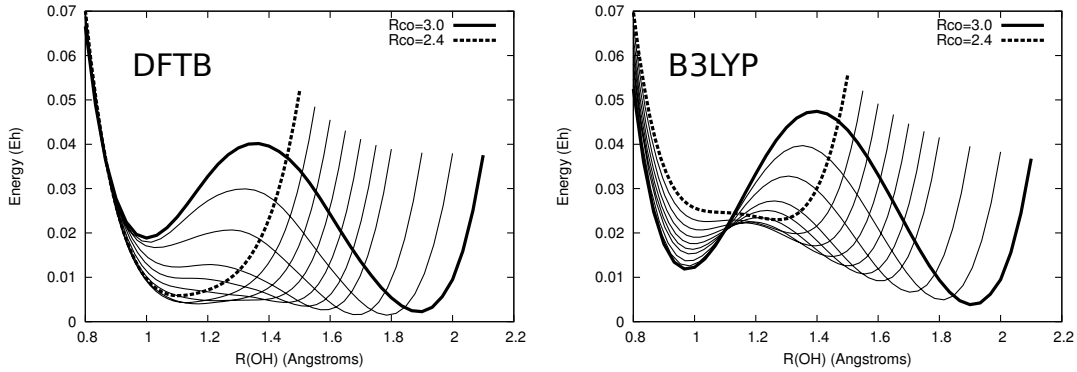

Figure A.8 Energy profiles for the H_3CO–H–CH_3 system. The standard *mio* parameters were used, and agreement is bad in similar ways to the HO–H–CH_3 system. It should be noted that the O–H minima are higher in the B3LYP calculations than they were in the HO–H–CH_3 system, but overall the curves have similar features.

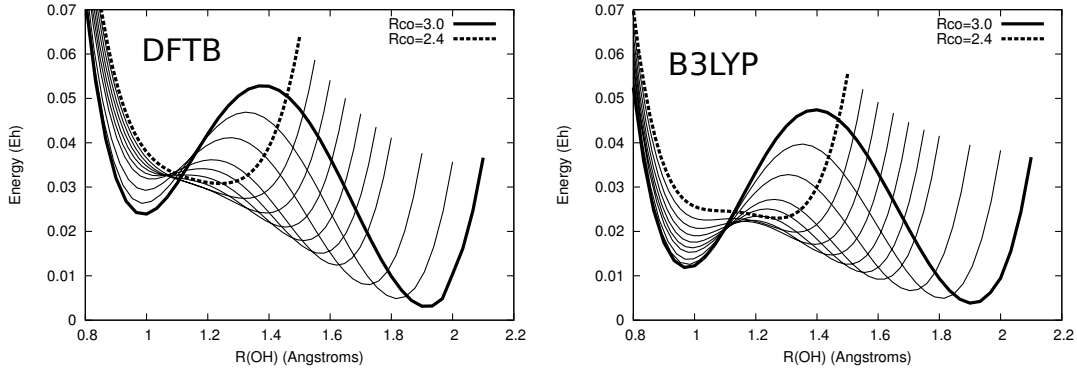

Figure A.9 Energy profiles for the H_3CO–H–CH_3 system with the NEW parameter files. Agreement is much better than when *mio* parameters were used, but it is not as good as the HO–H–CH_3 results for which they were designed.

may be used for the donor and acceptor atoms remain the same, but the environment changes.

A.8 Summary

New DFTB parameter files were generated by adjusting the repulsive spline term in O–H and C–H parameter files to fit the B3LYP results for the hydrogen transfer in the HO–H–CH_3 system. A combination of B3LYP results for C–H and O–H interaction as well as analytical functions were used to produce the best agreement.

Portability was demonstrated by replacing the an external hydrogen with a methyl group, and then testing the new parameters on H_3CO–H–CH_3. The quality of the curves improved greatly with the new parameter files present. It is possible that new SK files can be generated for prototype donor-acceptor combinations and then used in larger/different systems with the same donor and acceptor atoms as long as the donor–acceptor axis remains relatively clear of other atoms.

Appendix B

DFTB Benchmark and Spline Adjustment for the Active Site of SLO-1

B.1 Abstract

A correction to the MIO repulsive spline was generated for the active site of soybean lipoxygenase-1 (SLO-1). The corrective function was applied at T_{elec} = 10,000 K in the same way as outlined in Appendix A. In this case, the corrective spline was applied only to the O–H interaction between the transferring proton and acceptor oxygen atom. These new parameter files were used for the QTES-DFTB study in Chapter 4.

B.2 Introduction

The reaction profile for hydrogen transfer in the active site of SLO-1 was reproduced with DFTB. In order to generate the correct energy profile with DFTB, the *mio* parameter file for O–H interactions had to be adjusted. Specifically, the repulsive spline was fit to B3LYP results in the active site of SLO-1. In this study, the reduced active site of SLO-1 consisted of 44 atoms in a +1 charge, 6 multiplicity state. Collinear curves for the hydrogen transfer were generated for 30 different substrate geometries.

Geometry optimizations were performed at the B3LYP/LANL2DZ level of theory. The distance between the transferring proton and acceptor oxygen, R_{OH}, was constrained and all other nuclei were optimized. Various properties of these geometry optimizations are shown in Table B.1. These geometries were used to calculate

hydrogen-transfer reaction profiles, and all nuclei remained fixed in space throughout the hydrogen transfer reaction.

Table B.1 Geometry optimization results

R_{OH} Å	R_{CO} Å	R_{CH} Å	∠ OHC	Energy (E_h)
1.40	2.669	1.271	175.9	-1001.78268215068
1.41	2.665	1.256	176.4	-1001.78247380917
1.42	2.662	1.243	176.9	-1001.78234358146
1.43	2.660	1.231	177.3	-1001.78227716931
1.44	2.660	1.221	177.4	-1001.78227162178
1.45	2.660	1.211	177.4	-1001.78232041267
1.46	2.661	1.202	177.2	-1001.78242411211
1.47	2.663	1.194	177.0	-1001.78257987101
1.48	2.666	1.187	176.6	-1001.78277716344
1.49	2.670	1.181	176.3	-1001.78301334589
1.50	2.673	1.175	175.8	-1001.78328029579
1.51	2.677	1.169	175.1	-1001.78358333909
1.52	2.683	1.165	175.5	-1001.78388633664
1.53	2.687	1.160	174.3	-1001.78425160683
1.54	2.693	1.156	173.9	-1001.78461361228
1.55	2.699	1.153	173.6	-1001.78498566009
1.56	2.705	1.150	173.3	-1001.78536522377
1.57	2.712	1.147	173.0	-1001.78575045263
1.58	2.719	1.144	172.7	-1001.78613700483
1.59	2.726	1.142	172.5	-1001.78652379397
1.60	2.733	1.139	172.3	-1001.78692269276
1.61	2.740	1.137	172.0	-1001.78731273916
1.62	2.748	1.135	171.8	-1001.78769876313
1.63	2.756	1.133	171.6	-1001.78807985456
1.64	2.764	1.132	171.5	-1001.78845788729
1.65	2.772	1.130	171.3	-1001.78882399869
1.66	2.781	1.129	171.1	-1001.78918218677
1.67	2.789	1.127	170.9	-1001.78953204944
1.68	2.797	1.126	170.7	-1001.78987344501
1.69	2.805	1.125	170.3	-1001.79020149988
1.70	2.814	1.124	170.3	-1001.79053021740

B.3 B3LYP reaction profiles

Reaction profiles were generated for collinear hydrogen transfer at the geometries in table B.1. Geometries are indexed based on the frozen R_{OH} constraint that was used

for the optimization. All environmental nuclei were frozen for these energy profile calculations as the transferring hydrogen was moved between donor C and acceptor O. Figure B.1 shows reaction profiles from fixed R_{OH} ranging from $1.40 - 1.70$ Å. These potential energy profiles are in good agreement with other work [19, 47], and they will be used as the standard to which DFTB results will be compared.

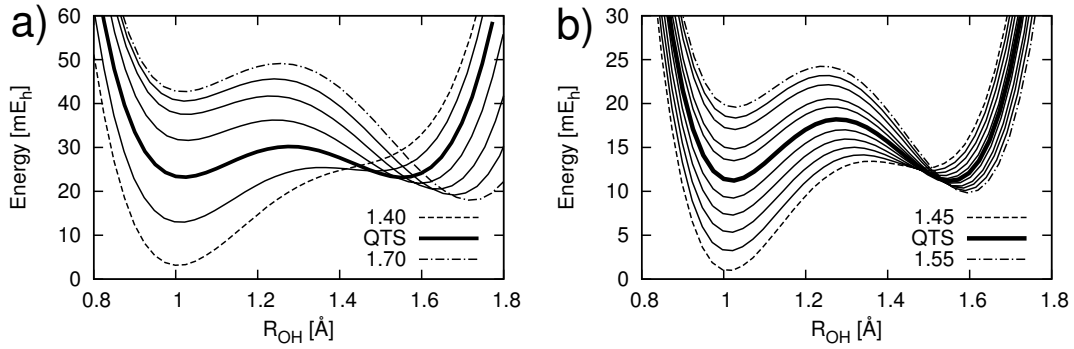

Figure B.1 (a) B3LYP potential energy profiles for geometries with fixed $R_{OH} = 1.40 - 1.70$Å. Donor state (larger fixed $R_{OH} = 1.70$ Å) has deeper C–H well, and acceptor state (smaller fixed $R_{OH} = 1.40$ Å) has a deeper O–H well. The quantum transition state (QTS) has equal depths for both wells. (b) Geometries closer to the quantum transition state. It is clear that the QTS corresponds to switching from a donor to acceptor state.

B.4 DFTB agreement using standard parameter sets

Potential energy profiles were generated using DFTB in the same way as the B3LYP calculations using the standard *mio* set as well as the parameters previously generated for the model system HO–H–CH_3. Electronic temperature (T_{elec}) and existing parameter sets were varied to produce optimal agreement. To improve agreement with *ab initio* quantum chemistry data, T_{elec} is used as an adjustable parameter, not a physical temperature. We observe in Figures B.2-B.4 that the best agreement is produced, given the parameter sets already available to us, when the electronic temperature is set to $T_{elec} = 10,000$ K. Furthermore, the O–H acceptor well is most evident when the corrective HO–**X**–CH_3 function is present, but the barrier is much

too high from the donor's perspective. A trend in all cases is that the acceptor state is much higher in energy than it should be.

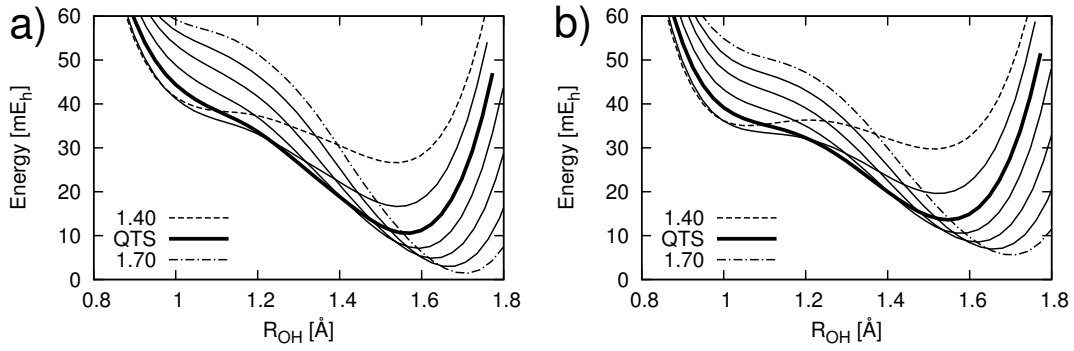

Figure B.2 (a) Standard *mio/trans3d* set with $T_{elec} = 300$ K. (b) B3LYP parameters from HO–H–CH_3 model system with $T_{elec} = 300$ K. In both cases, an acceptor well is lacking.

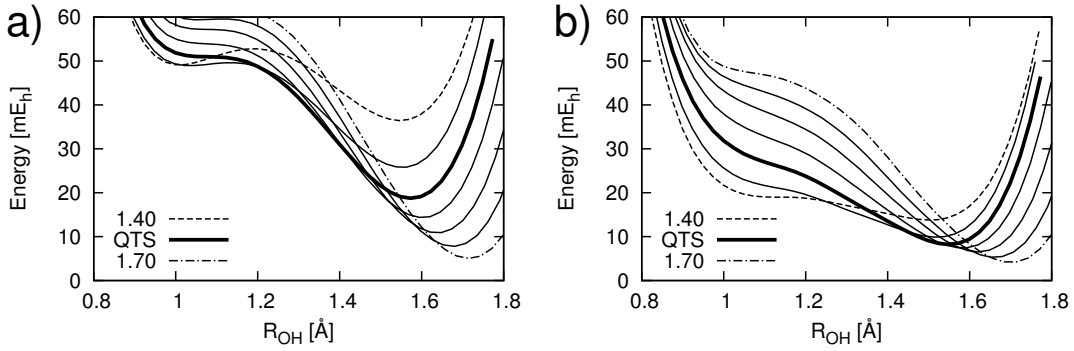

Figure B.3 (a) Switching function from HO–H–CH_3 model system with $T_{elec} = 300$ K. (b) Standard *mio/trans3d* set with $T_{elec} = 10,000$ K. An improvement is seen at a higher electronic temperature.

B.5 Donor and acceptor dissociation curves

To identify the problem specifically, the system was separated into donor and acceptor molecules and dissociation curves were compared. Curves were generated for the dissociation of a hydrogen atom from 1) linoleic acid donor molecule, and 2) iron-containing active-site acceptor molecule using the same level of theory as the full

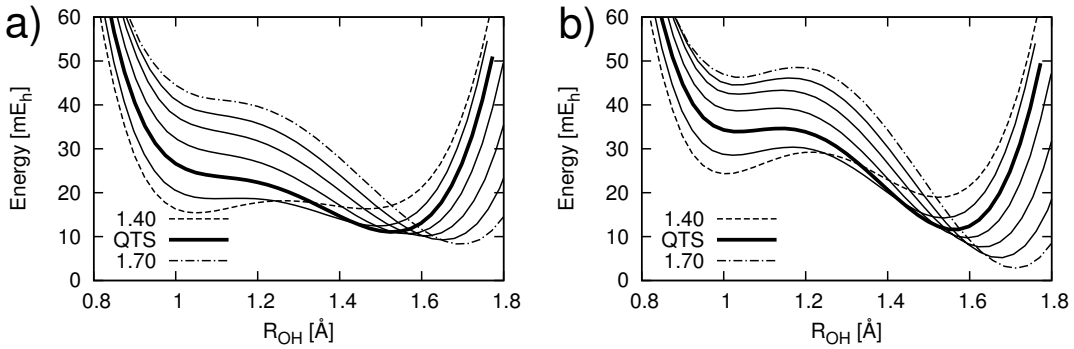

Figure B.4 (a) B3LYP parameters from HO–H–CH_3 model system with $T_{elec} = 10,000$ K. (b) Switching function from HO–H–CH_3 with $T_{elec} = 10,000$ K. These are the best results achievable with the standard SK files.

system. To provide the option of fitting E_{rep} to these results, only the QTS geometry ($R_{OH} = 1.50$ Å) was used here. The dissociation curves for donor and acceptor molecule are shown in Fig. B.5.

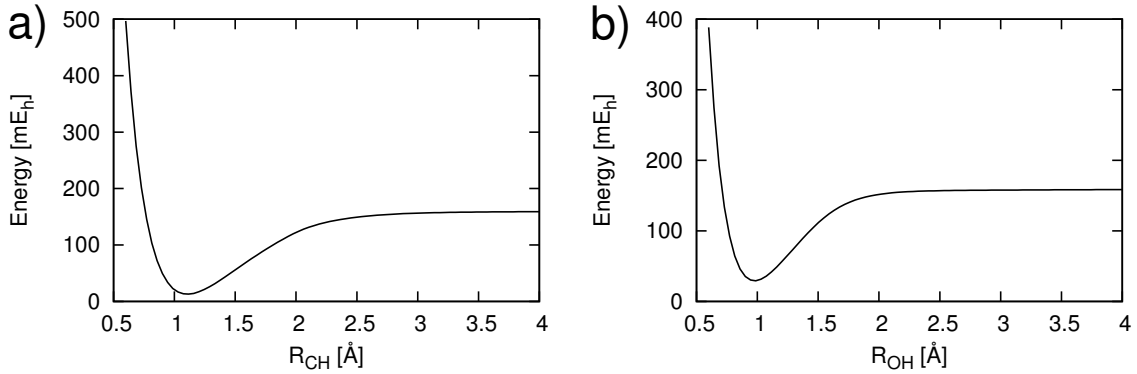

Figure B.5 B3LYP curve for hydrogen atom dissociation from (a) linoleic acid and (b) iron-oxygen active site.

We now examine how DFTB results, varying T_{elec} and parameter sets as before, compare to these curves. The resulting energy curves are shown in Fig. B.6. For the linoleic acid system, setting $T_{elec} = 5,000$ K produces the best agreement. Setting $T_{elec} = 10,000$ K as was done in the full system greatly over-corrects. In the iron-oxygen acceptor system, setting $T_{elec} > 300$ K worsens the agreement, with $T_{elec} = 10,000$ K being particularly bad and introducing an additional barrier at $R_{OH} \approx 1.4$

Å. This evidence shows that the improvement in the large system that was obtained by finding $T_{elec} = 10,000$ K is NOT portable to smaller pieces of the system. While the curves themselves looked better, it is very unclear what effect changing this variable had on the electronics of the rest of the system.

However, changing 2-body parameter files to work only in a donor–acceptor type of system has the same sort of effect when the system is split up. It therefore makes sense to look at the system as a whole, accepting that these parameter adjustments (whether they be SK files or T_{elec}) will generally not be portable to individual pieces of the donor-acceptor system. We now focus on the full 44-atom system, and after much experimentation with C–H and O–H switching functions and changing electronic temperature, a combination was found that produces the best agreement and can be implemented as a new SK parameter file.

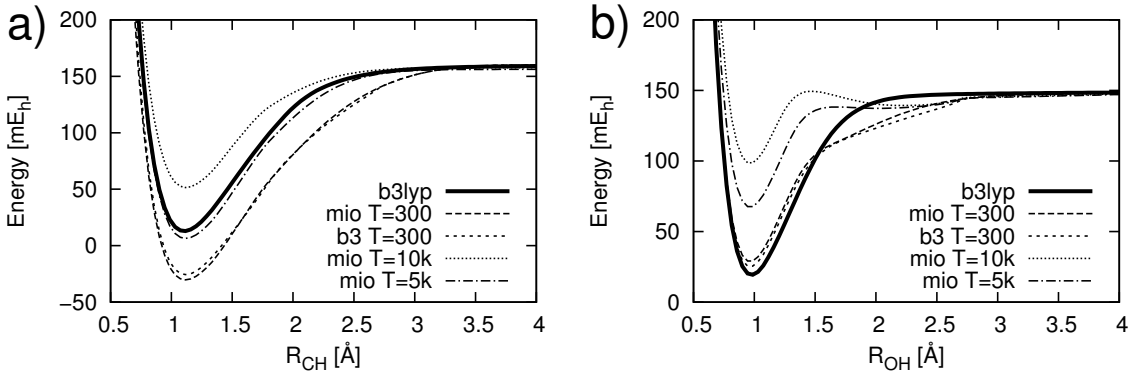

Figure B.6 Comparison of DFTB results to B3LYP results for the dissociation of hydrogen from (a) linoleic acid and (b) the iron-oxygen active site.

B.6 New parameter files for SLO-1

It was found that, given our current set of available parameter files, the *mio/trans3d* set at $T_{elec} = 10,000$ K produced the best agreement with the full system. It spread out the O–H minima in a realistic way, and the trend of increasing energy in the C–H minima with respect to decreasing C–O distance is also observed. These curves are

shown in Fig. B.7. In this case, the main problem is the acceptor oxygen's ability to form a bond to the transferring hydrogen. It appears that an O–H minimum only forms when the C–O distance is very small, and when compared to Fig. B.1, the energy in the O–H bond area ($\approx$ 1 Å) is approximately 20 mE_h too high. Since the C–H area already agrees well, the approach will be to increase the O–H bonding affinity by using a switching function that depends *only* on O–H. For reference, all plots are shown in Fig. B.7.

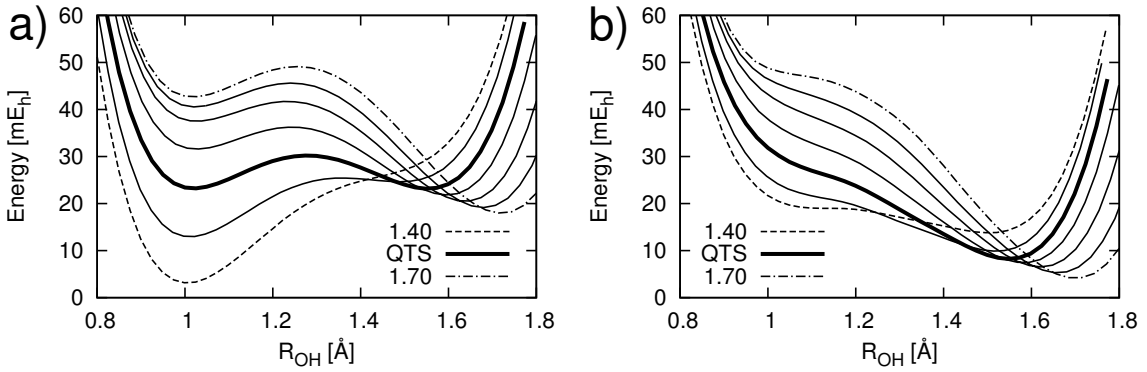

Figure B.7 SLO-1 reaction profiles calculated with (a) B3LYP and (b) standard *mio/trans3d* parameter set at $T_{elec} = 10,000$ K.

The following switching function produced good agreement with the B3LYP results.

$$f(R_{OH}) = \frac{31 \text{ mE}_\text{h}}{e^{6(x-(1.7952+2.4566)/2)} + 1} \tag{B.1}$$

Eq. B.1 was applied *only* to the acceptor oxygen atom using *mio* curves. The origin of this O–H bonding problem is very likely the presence of the iron atom nearby and the high multiplicity of the system. In another set of calculations, it was found that the DFTB geometry of the iron-containing active-site molecule is very different than the one calculated from B3LYP, and therefore some deficiency in DFTB (parameter set or otherwise) is treating the iron in a strange way. This switching function corrects for the iron-oxygen interaction problems and correctly increases the oxygen's affinity

for the transferring hydrogen. The effect of the new O–H parameter file can be seen in Figs. B.8 and B.9.

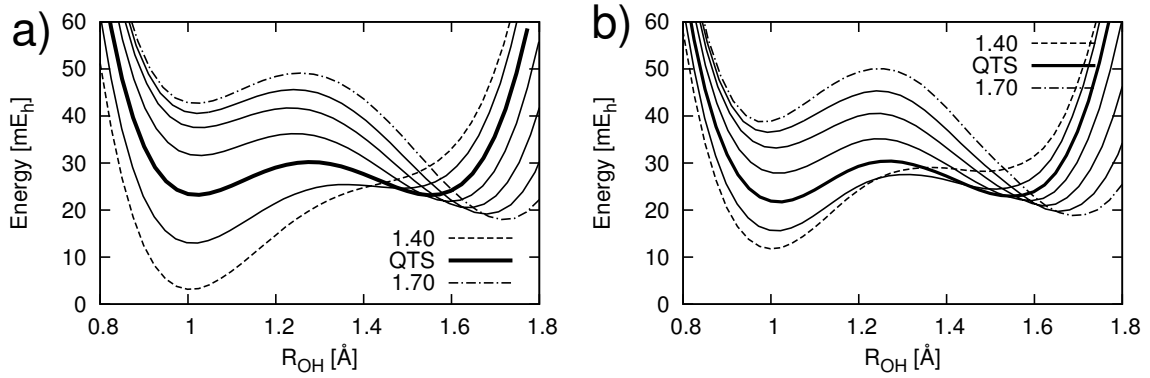

Figure B.8 SLO-1 reaction profiles calculated with (a) B3LYP/LANL2DZ and (b) new switching function with *mio/trans3d* parameter set at $T_{elec} = 10,000$ K. Agreement is much better than with just the standard *mio/trans3d* set, and the only adjustment was the oxygen bonding affinity for hydrogen

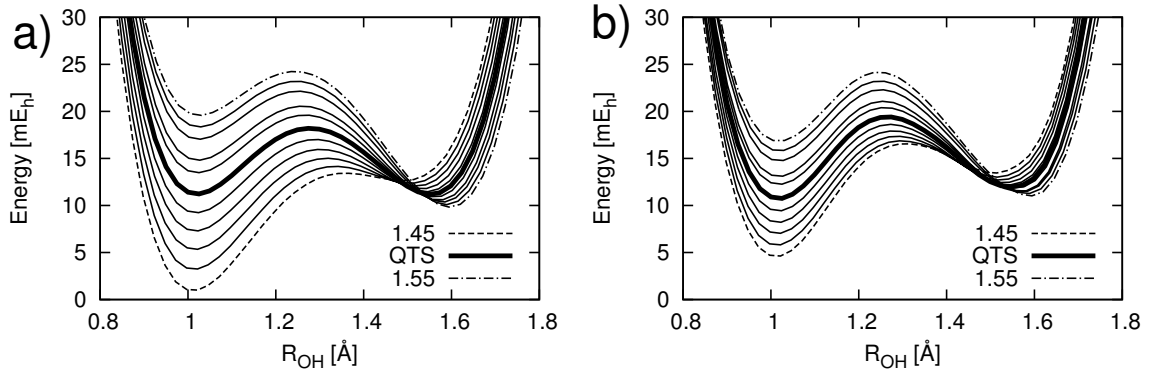

Figure B.9 For fixed R_{OH} in the range of 1.45-1.55 Å, which closely surround the QTS, agreement between (a) B3LYP and (b) new parameter set are very good. The parameters were fit to a larger range of fixed R_{OH}, but reproduce the correct trend on more finely separated curves as well.

Very good agreement is seen in all cases, and the trends are reproduced correctly. The main area that could be further improved is the energy spacing of the O–H minima, especially when the system is in a product state. This spacing should be increased. Increasing T_{elec} to 10,000 K did greatly improve this trend, but a higher

T_{elec} made the trends worse. At T_{elec} =10,000K, the QTS is in the correct place, and barrier heights are what they should be. The $R_{OH} = 1.65, 1.70$ Å curves still correspond to a product state, and are therefore reasonably good.

Each curve that was shown represents a different, frozen geometry which was generated by fixing *only* O–H. This means that the other atoms have, as far as the switching function is concerned, an arbitrary configuration. Yet, it produced the correct trends, minima locations, and barrier heights for all of them. Thus, this adjustment is reasonable, and will remain so should the hydrogen explore other degrees of freedom, such as in dynamics. While further improvement is possible, it is better to use the simple corrections shown here and avoid side-effects from over-parameterization (problems in other degrees of freedom, etc.) and to keep the physical significance of increased hydrogen affinity dependent only on the O–H distance.

B.7 Summary

A new SK parameter file was created for O–H interactions in the active site of SLO-1. It was designed to improve the proton-transfer reaction profile in the active site, and it has been shown to be accurate over a wide range of active site geometries. This new parameter file was used in the QTES-DFTB calculations, and it is implemented in the exact same way as the standard parameter files. This parameter file only changes the bonding affinity between the acceptor oxygen atom and the transferring hydrogen, and it should be implemented accordingly. All other atom-atom interactions in this system are treated with a combination of *mio* and *trans3d* standard parameter files.

Appendix C

Transmission Probability Calculations on an Eckart Barrier

C.1 Abstract

Transmission probability $T(E)$ was successfully calculated at a range of energies for a 1-D asymmetric Eckart barrier. QM calculations were performed using a split operator method, and $T(E)$ was obtained through an S-matrix element calculation. WKB transmission was also calculated, and both methods were compared to the analytical solution. Agreement is excellent between all three methods at very low energies, but as the energy rises, WKB becomes systematically worse. The QM method should be valid for any energy situation as long as the wavepacket contains appreciable contributors from eigenstates at the energy of interest. This method was used for calculations of $T(E)$ in the HO–H–CH_3 system in Chapter 6.

C.2 Introduction

There are many cases in which a double-well potential describes a chemical reaction such as hydrogen transfer. In the case where the reaction coordinate couples to the other nuclear degrees of freedom, calculation of reaction rate is quite straightforward. In the case where there is no coupling, and the double-well system is isolated from the rest of the environment, the quantum system will oscillate from a reactant to a product state indefinitely, as can be seen in previous studies involving the active site of SLO-1 [15]. There is currently no formal way to describe the rate for such a system due

to these oscillations, and there is a need for one such expression due to the prevalence of these systems. The form of the expression which we are developing, as seen in section C.3 depends on dynamics which are performed on an open system which contains the barrier from the potential energy surface of interest. The asymmetric Eckart barrier is used to benchmark this portion of the rate formula.

C.3 Rate constant formula

Consider a one-dimensional scattering system for a potential with flat asymptotic regions which describe the reactant and product. The quantum thermal reaction rate is defined as

$$k(T) = \frac{\int_0^\infty T(E) e^{-E/(kT)} dE}{\int_0^\infty e^{-E/(kT)} dE} \tag{C.1}$$

where $T(E)$ is the transmission probability from reactant to product for a wavepacket of energy E that originates in the asymptotic region of the reactant state. This expression is quite straightforward because there will be no re-crossing of the barrier, and after the incident collision, all necessary transmission probability information will be available.

This equation can be extended to a bound system as

$$k(T) = \frac{\sum_n T(E_n) \rho_n e^{-E_n/(kT)}}{\sum_n e^{-E_n/(kT)}}, \tag{C.2}$$

where E_n is the energy of bound eigenstate n and ρ_n is the projection of eigenstate n onto the reactant region of the potential. The main problem remains of how to calculate $T(E_n)$ for a system which oscillates indefinitely. One way to circumvent this problem is to temporarily treat the system as if it were a scattering system. In the example of a one-dimensional double-well, one could locate the two minima and make them asymptotic. This means that the calculation of $T(E)$ would be identical to a scattering state in which the barrier is one from a bound double-well potential.

Before testing Eq. C.2 on a bound potential, we must first be sure that $T(E)$ can be accurately calculated for arbitrary scattering states.

C.4 Expressions for calculating $T(E)$

In order to calculate a rate constant for a double-well surface of arbitrary form, it must be possible to calculate the transmission probability numerically. The formulation that we are using to calculate transmission probability is influenced very heavily by work done by Tannor and Weeks [104]. Using this formalism, the transmission probability is defined as

$$T(E) = |S_{RP}|^2 \tag{C.3}$$

where S_{RP} is an S-matrix element composed from reactant (R) and product (P) scattering eigenstates. This matrix element is defined as shown in Eq. C.4.

$$S_{RP}(E) = \frac{(2\pi\hbar)^{-1}}{\eta_P^*(E)\eta_R^*(E)} \int_{-\infty}^{+\infty} \langle \phi_P^- | e^{-i\hat{H}t/\hbar} | \phi_R^+ \rangle e^{iEt/\hbar} dt \tag{C.4}$$

This equation represents a decomposition of a correlation function $\langle \phi_P^- | e^{-i\hat{H}t/\hbar} | \phi_R^+ \rangle$ into contributors of incoming and outgoing eigenstates. We define ϕ_P^- as a product wavepacket which is moving away from the barrier (denoted by "$-$"), and ϕ_R^+ as a reactant wavepacket which is moving toward the barrier (denote by "+"). The correlation function is therefore a time-dependent projection of the donor-originated wavepacket onto the product state. For a clean energy spectrum, the dynamics must begin and end with this correlation function being zero. In the coefficient, the denominator terms for the reactant are defined as

$$\eta_R(E) = \sqrt{\frac{m}{2\pi\hbar k}} \int_{-\infty}^{+\infty} e^{-ikx} \phi_R^+(x) dx \tag{C.5}$$

and η_P is calculated analogously. These η coefficients are proportional to the amount of contributing eigenstates to the reactant or product state.

C.5 Assymetric Eckart barrier setup

Since we expect our scattering systems to be, in general, asymmetric as well as smooth to the first derivative, the asymmetric Eckart barrier was chosen as a suitable test system [120, 121]. This potential has the form

$$V = -\frac{Ay}{1-y} - \frac{By}{(1-y)^2} \tag{C.6}$$

where we define

$$y = -\exp\left(\frac{2\pi x}{L}\right) \tag{C.7}$$

where x is the variable dimension and L is a characteristic length. A and B are related to the asymptotic limits on either side of the barrier where $A = \Delta V_1 - \Delta V_2$ and $B = \left(\sqrt{\Delta V_2} + \sqrt{\Delta V_1}\right)^2$. ΔV_1 is the difference in energy from the asymptotic region on the left to the top of the barrier, and ΔV_2 is the same thing for the right side of the barrier. If we have $\Delta V_1 = \Delta V_2$ then the system is symmetric. The potential used to benchmark can be seen in Fig C.1.

The primary motivator to use a system with this functional form is that there exists an analytical form for $T(E)$ which we can compare our numerical results against. This expression is shown in Eq. C.8.

$$T(E) = 1 - \frac{\cosh 2\pi(a-b) + \cosh 2\pi d}{\cosh 2\pi(a+b) + \cosh 2\pi d} \tag{C.8}$$

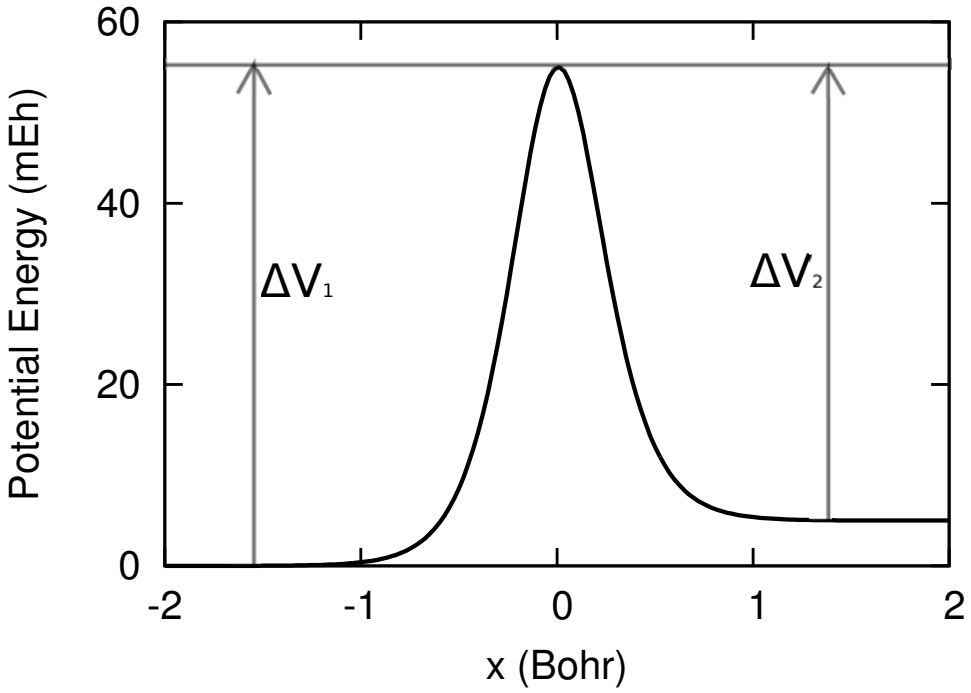

Figure C.1 A potential with $\Delta V_1 = 55$ mEh, $\Delta V_2 = 50$ mEh, and $L = 1$ Bohr.

where

$$2\pi a = 2(\alpha_a \xi)^{1/2}(\alpha_1^{-1/2} + \alpha_2^{-1/2})^{-1}$$

$$2\pi b = 2((1+\xi)\alpha_1 - \alpha_2)^{1/2}(\alpha_1^{-1/2} + \alpha_2^{-1/2})^{-1}$$

$$2\pi d = 2(\alpha_1\alpha_2 - 4\pi^2/16)^{1/2}$$

$$\xi = E/\Delta V_1$$

and

$$\nu = (1/2\pi)(-F^*/m)^{1/2}$$

$$\alpha_1 = 2\pi\Delta V_1/h\nu$$

$$\alpha_2 = 2\pi\Delta V_2/h\nu$$

where F^* denotes the second derivative of the potential at the maximum,

$$F^* = -\pi^2(A^2 - B^2)^2/2L^2B^3. \tag{C.9}$$

We now compare the analytical transmission probability to results from a numerical wavepacket propagation as described in section C.6 and the WKB approximation[105, 106].

C.6 Calculation of $T(E)$ using numerical wavepacket propagation

The asymmetric Eckart potential described in Fig. C.1 was examined using the split-operator method to treat a time-dependent quantum wavepacket. The reactant wavepacket ϕ_R^+ originates in the asymptotic region to the right of the barrier, and it has initial momentum toward the barrier (left). The mass of hydrogen (1836 atomic units) was used, and is Gaussian in shape with a width parameter of 9.58. The product state ϕ_P^- which is used to calculate the correlation function is identical to the reactant wavepacket, only it is on the other side of the barrier in the asymptotic region to the left. In an effort to maximize agreement for low energy cases, p_0 was set to -4.0 a.u. This corresponds to a kinetic energy of approximately 7 $\mathrm{mE_h}$.

The primary piece of information extracted from the dynamics is the cross-correlation function $\langle\phi_P^-|e^{-i\hat{H}t/\hbar}|\phi_R^+\rangle$. The results of this calculation are shown in Fig. C.2. The Fourier transform of this function and its absolute value squared (which is equivalent to S_{RP}) is shown in Fig. C.3. Before the transmission probability can be calculated, the reactant and product wavepacket must be decomposed into its contributing eigenstates so that η can be calculated as shown in Eq. C.5. The transmission probability calculated over approximately 100 mEh can be seen in Fig. C.4.

It is clear that very high energies, the QM results tend to be worse. This is simply a result of numerical instability because the eigenstates in the high energy

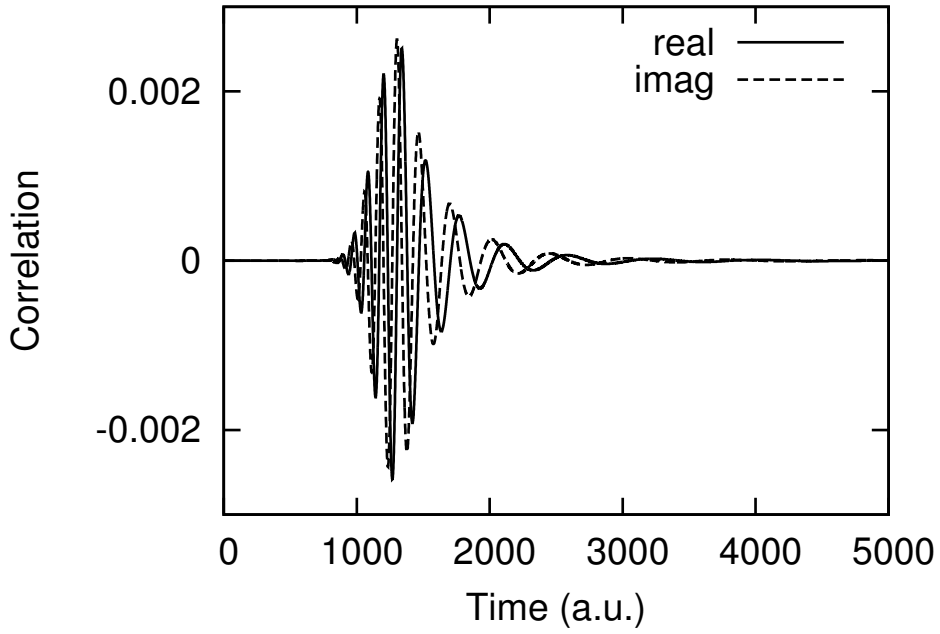

Figure C.2 Correlation function that is used to calculate S_{RP} as shown in Eq. C.4 It is necessary for this function to go to 0 at both beginning and end time of the simulation.

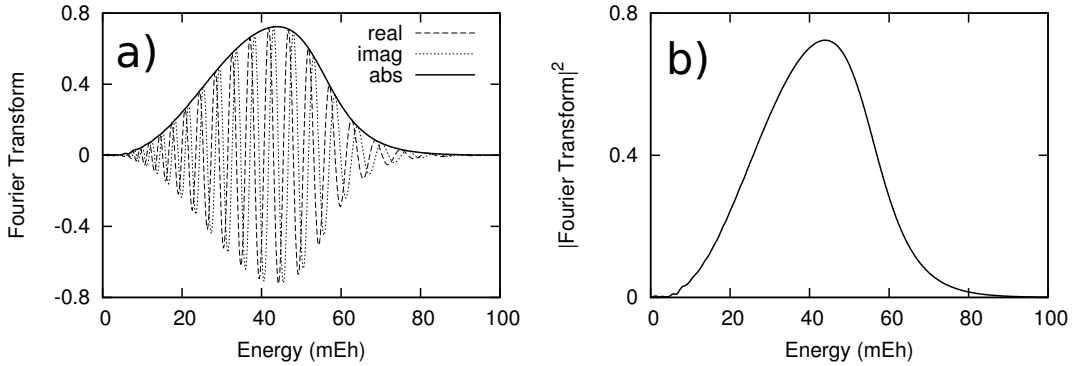

Figure C.3 Fourier transform results of the correlation function in Fig. C.2. The regular Fourier transform and its absolute value are in a), and the absolute value squared, which will be directly used in the calculation of that transmission probability, is in b).

regime contribute very little to the dynamics. WKB produces poor results even at moderately low energy, and only worsens as the energy gets higher. Once the WKB energy is higher than the barrier, 100% transmission is assumed. The agreement at low energies (Fig. C.4), with $T(E) < 10^{-6}$ is very good for both methods, and exact

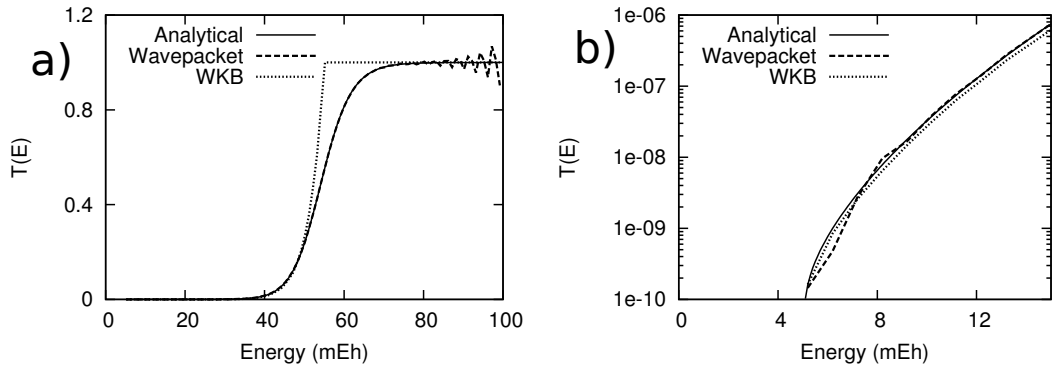

Figure C.4 A plot of T(E) as expressed in Eq. C.3 by solving S_{RP} as shown in Eq. C.4. The a) overall $T(E)$ calculated by both a QM method and WKB method, and a view of $T(E)$ at b) low energy.

agreement is seen more frequently with the QM approach although some oscillations are present. These problems can be mitigated by improving the quality of the QM propagation.

C.7 Conclusions

It has been shown that propagation of a time-dependent QM wavepacket using the split-operator method coupled with S-matrix element calculations can accurately calculate transmission probability for an asymmetric Eckart potential. The analytical $T(E)$ is nearly identical to the QM results, even at very low energies, and WKB also shows good agreement for energy much lower than the barrier top. This S-matrix QM calculation of transmission probability will be used to calculate rate constants for bound, asymmetric double-well potentials, and it is expected that the dynamics of the unbound barrier system will be quite similar to the results shown here. Based on the results for the Eckart potential, it is expected that the results for other asymmetric scattering potentials will produce accurate rate constants for the bound systems.

Appendix D

Permission to Reprint

D.1 Chapter 2: Ground State Proton Transfer in the SLO-1 Double-Well Potential

Permission is granted to all authors to include published work in a dissertation or thesis. As seen on: http://www.elsevier.com/journal-authors/author-rights-and-responsibilities accessed on 04/08/14.

Authors can use either their accepted author manuscript or final published article for:	
✓	Use at a conference, meeting or for teaching purposes
✓	Internal training by their company
✓	Sharing individual articles with colleagues for their research use* (also known as 'scholarly sharing')
✓	Use in a subsequent compilation of the author's works
✓	Inclusion in a thesis or dissertation
✓	Reuse of portions or extracts from the article in other works
✓	Preparation of derivative works (other than for commercial purposes)

D.2 Chapter 5: Efficient Quantum Trajectory Representation of Wavefunctions Evolving in Imaginary Time

Permission is granted to all authors to include published work in a dissertation or thesis. As seen on: http://publishing.aip.org/authors/copyright-reuse accessed on 04/08/14.

Q: May I include my AIP article in my thesis or dissertation?

AIP permits authors to include their published articles in a thesis or dissertation. It is understood that the thesis or dissertation may be published in print and/or electronic form and offered for sale, as well as included in a university's repository. Formal permission from AIP is not needed. If the university requires written permission, however, we are happy to supply it.

D.3 Chapter 6: Rate Constant Calculations for the HO–H–CH_3 Double Well Potential

Permission is granted to all authors to include published work in a dissertation or thesis. As seen on: http://www.hindawi.com/journals/jtc/guidelines/ accessed on 04/08/14.