The Pennsylvania State University

The Graduate School

Department of Chemistry

HYBRID QUANTUM/CLASSICAL MOLECULAR DYNAMICS SIMULATIONS OF HYDROGEN TRANSFER REACTIONS IN ENZYMES

A Dissertation in

Chemistry

by

Dhruva Kumar Chakravorty

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

May 2010

The dissertation of Dhruva Kumar Chakravorty was reviewed and approved* by the following:

Sharon Hammes- Schiffer
Eberly Professor of Biotechnology
Professor of Chemistry
Dissertation Advisor
Chair of Committee

Philip C. Bevilacqua
Professor of Chemistry

James B. Anderson
Evan Pugh Professor
Professor of Chemistry

James D. Kubicki
Associate Professor of GeoChemistry

Barbara G. Garrison
Professor of Chemistry
Head of the Department of Chemistry

*Signatures are on file in the Graduate School

ABSTRACT

This thesis describes the study of hydrogen transfer reactions in enzymes. For this purpose the thesis is divided into two parts. The first part is work undertaken in the development of new methods for the study of such reactions, while the second part of the thesis focuses on the study of the isomerization reaction in ketosteroid isomerase (KSI).

In part one of this thesis, the implementation of the umbrella integration method for calculating the potential of mean force (PMF) for a chemical reaction within the empirical valence bond (EVB) framework is presented. The umbrella integration method is based on the derivative of the PMF with respect to the reaction coordinate. An analytical expression for this derivative applicable to certain types of EVB potentials is presented. The umbrella integration method reduces the statistical errors, converges efficiently, and does not require significantly overlapping windows compared to the Weighted Histogram Analysis method. A modified version of the weighted histogram analysis method that shares these advantages is also proposed and implemented.

The second part of this thesis presents the study of proton transfer reactions in ketosteroid isomerase (KSI) using hybrid quantum/classical methodologies. The reaction requires two proton transfer reactions. Tyr14 and Asp99 form catalytically important hydrogen bonds that help stabilize the dienolate intermediate in these reactions. Chapter 3 describes the study of these proton transfer reactions in wild-type KSI. Chapter 4 extends this analysis to the Tyr14Phe, Asp99Leu and Tyr14Phe/Asp99Leu mutant varieties of KSI. Our simulations suggest a pre-organized active site in which relatively small changes occur to strengthen the hydrogen bonds that stabilize the intermediate, thereby facilitating the proton transfer reactions.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

It has been a long journey and I would like to thank every person who helped me over the last five years. I would like to thank my research advisor Professor Sharon Hammes-Schiffer whose guidance, patience and constant support made this thesis possible. I thank Dr. Alexander Soudackov and Dr. Ari Chakraborty for finding solutions to my myriad problems. I thank Dr. Yolanda Small, Dr. Malika Kumarasiri , Michele Ludlow and Phil Hanioan for maintaining the computer systems on which I ran my simulations. I thank Dr. James Watney for introducing me to the science of computational chemistry. I thank Dr. Jonathan Skone, Dr. Chet Swalina, Dr. Rajarshi Guha and Dr. Elizabeth Hatcher for being a constant source of encouragement and advice over the course of my graduate work. I thank Sarah Edwards and Narayan Veeraraghavan for all their help. I thank Dr. Anirban Hazra, Dr. Mike V. Pak and Dr. Ben Auer for answering my inane questions. I thank Laura Fernandez, Brian Solis, Andrew SirjooSingh and Abir Ganguli for providing the group with an infectious air of energy. In addition to them I would like to thank all past and present members of the Hammes-Schiffer group for their support.

I would like to thank my committee members, Prof. J.B. Anderson, Prof. P. C. Bevilacqua and Prof. J.D. Kubikci for their guidance and support over the course of my graduate work. I would also like to thank Prof. S.J. Booker for teaching me my first biochemistry class and for being a member of my committee for five years.

I thank Prof. T.P. Radhakrishnan for his encouragement and support over the years. I also thank him for allowing me to discover chemistry at my own pace. I would like to thank the science's faculty at St. Stephen's College, Delhi. I am especially grateful to Dr. S.K. Vig,

the late Dr. D.P. Goel, Dr. V. Jain and Dr. Tara Chand for their mentorship during undergraduate years. They never stopped believing in what I could achieve. I would also like to thank my high school teacher Mr. S. Majumdar who laid the foundation for my interests in the natural world.

I thank my parents Major General P.K. Chakravorty and Ms. Debrani Chakravorty who helped me become who I am. They provided me with everything I needed and a lot more. Their never ending support and love have kept me going through the toughest of times. I also thank my extended family Lee and Grace Gordon who welcomed me into their lives.

I would like to thank Conan O' Brien who kept me entertained through the nights I spent working. I also thank Kelsey Kramer, Steven Colbert, Jeremy Clarkson, "The Stig", Richard Hammond and James May for making me smile.

Lastly, I would like to thank my wife Sarah Leigh Gordon whose remarkably optimism, generosity and encouragement have helped me bring this work to fruition.

Chapter 1

Introduction

1-1. Introduction

Enzymes are generally globular proteins that catalyze chemical reactions.[1] They are some of the most efficient catalysts known today and have been shown to effectively catalyze reactions by preferentially stabilizing the transition state or by providing an alternate pathway to a reaction. Recent studies have found that internal motions within enzymes play an important role in their catalytic abilities. Enzymes are specific to their substrates and the kind of reactions they catalyze. These reactions are generally known to follow Michaelis-Menten kinetics.[2] The ability of enzymes to catalyze reactions faster than their non-enzymatic counterparts has opened up the field of bio-mimetic catalysis. Enzymes help maintain homeostasis in living organisms, and the malfunction of a single enzyme can cause a disruption in this balance leading to severe illness. This makes enzymes very interesting for medicinal purposes.

Hydrogen transfer reactions commonly occur in enzymatic reaction pathways. In these reactions, a hydrogen atom moves from a donor atom to an acceptor atom within the enzyme complex. Hydrogen transfer reactions have been shown to exhibit quantum mechanical behavior and are of significant fundamental interest. These reactions are important from a medical point of view as they occur in enzymes such as dihydroorotate dehydrogenase that are targeted to find a potential cure for cancer.[3] An understanding of hydrogen transfer reactions is also crucial to successfully produce hydrogen for fuel. It is possible that hydrogen transfer reactions could well be the cornerstone of our future economy.

The field of computational chemistry has played a significant role in understanding hydrogen transfer reactions in enzymes. Computational methods are often used to calculate transition states of reactions, rates of reactions, NMR shifts, vibrational spectra and other similar properties. Simulations have also been instrumental in elucidating the mechanistic pathways of enzymatic reactions such as the isomerization reaction catalyzed by Δ^5-3-Ketosteroid isomerase (KSI). A variety of computational methodologies have been developed to answer questions related to enzymatic catalysis. The scope of these approaches range from very precise ab-initio calculations that can only treat a few atoms at a time, to coarse-grained simulations that can calculate properties of systems the size of hundreds of thousands of atoms within certain limitations. Classical all-atom molecular dynamics (MD) simulations allow large systems to be studied without losing chemical insight. These MD simulations provide dynamic and equilibrium information but are severely hampered by their inability to describe the breaking and formation of covalent bonds.

This has led to the emergence of hybrid Quantum Mechanical/Molecular Mechanical (QM/MM) methods that allow for the simulation of chemical reactions. In a typical QM/MM simulation of an enzymatic reaction, the active site of the enzyme is treated with quantum mechanical methods while the remaining parts of the enzyme and surrounding water molecules are treated with molecular mechanical methods. This hybrid approach allows QM/MM calculations to benefit from the accuracy of quantum calculations and the speed of MD simulations. It is often convenient to study the reaction along a collective reaction coordinate that incorporates the motions of the protein and solvent molecules. Ideally, the treatment of hydrogen transfer reactions should include nuclear quantum effects. Nuclear quantum effects such as zero point energy and tunneling are exhibited by light atoms such as hydrogen and have been found to contribute to the rate of proton transfer reactions.

A QM/MM method with these characteristics was the principal investigative tool used to study the hydrogen transfer reactions presented in this thesis.[4-6] In this approach, enzymatic reactions are simulated on a potential energy surface described by the Empirical Valence Bond (EVB) formalism. A two state EVB model is used to describe the hydrogen transfer reaction. In the first valence bond state, the transferring hydrogen atom is bonded to the donor, while in the second valence bond state the hydrogen atom is bonded to the acceptor atom. An energy gap reaction coordinate defined as the energy difference between the two valence bond states is used as the collective reaction coordinate. A series of mapping potentials are used to drive the reaction from reactant to product state along this collective reaction coordinate. These mapping potentials efficiently cover the entire relevant range of the collective reaction coordinate. Treating the transferring hydrogen atom with path integral calculations incorporates nuclear quantum effects into these simulations.

This thesis describes the study of proton transfer reactions in enzymes using hybrid quantum/classical methodologies. For this purpose the thesis is divided into two parts. The first part is work undertaken in the development of new methods for the study of such reactions. Chapter 2 describes the implementation of Umbrella Integration within the framework of the EVB approach.[7] This method allows for the calculation of more accurate free-energy profiles for enzymatic reactions with less sampling compared to traditional methods. This chapter also presents the application of this method for the calculation of the free-energy profile for the hydride transfer reaction catalyzed by dihydrofolate reductase.

The second part of the thesis describes the study of the proton transfer reactions catalyzed by KSI. [8-27] KSI has been found in numerous mammalian and bacterial species. In mammals, KSI is membrane bound, and its activity is associated with transferases. [12, 21-24] Previous investigations of the KSI reaction have studied bacterial strains of KSI from *Commamonas testosteroni* (TI) and *Psuedomonas putida*. KSI from these bacterial sources are homologous and

share 34% sequence identity.[10] Exhaustive structural studies have shown the placement of the catalytically important groups in the active sites of KSI from the two strains to be similar. This has allowed transferability in results from investigations across the species, though differences exist in their kinetics. This thesis focuses on the study of the KSI reaction from TI and residues are numbered accordingly. TI KSI has 125 residues in each monomer with the active site buried in a deep hydrophobic pocket. A biologically active unit is a homo-dimer as depicted in Figure 1-1, in which the isomerization reaction takes place in each monomer unit.

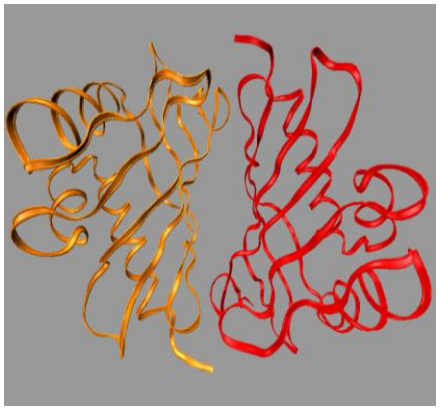


Figure**1-1**: Snapshot of the 1BUQ NMR structure[14] of KSI generated using VMD [28] program.

KSI (E.C. 5.3.3.1) is a very proficient enzyme and catalyzes the allylic isomerization reaction of 3-oxo-Δ^5-steroids to their conjugate Δ^4-isomers depicted in Figure 1-2. This reaction requires two proton transfer steps and involves the formation of a dienolate intermediate. In the first step of the reaction, Asp38 abstracts a proton from the C4β position of the steroid substrate to form a dienolate intermediate. In the following step of the reaction, the abstracted proton is transferred to the C6β position of the substrate from Asp38. This returns the substrate to its ketonic state and results in a product that is more stable than the reactant. The ability of a strong acid group such as Asp38 to act as a general base, and abstract a proton from a considerably weaker acid group like the substrate makes this reaction a perennial mechanistic question. The commonly occurring carbon-hydrogen bond cleavage reaction with 5-androstene-3, 17-dione as

substrate is catalyzed by KSI at a rate approaching the diffusion controlled limit. [11,12] Additionally, considerable rate enhancement is achieved in the KSI catalyzed reaction as compared to the corresponding non-enzymatic reactions catalyzed by acetate ion, acid and hydroxide.

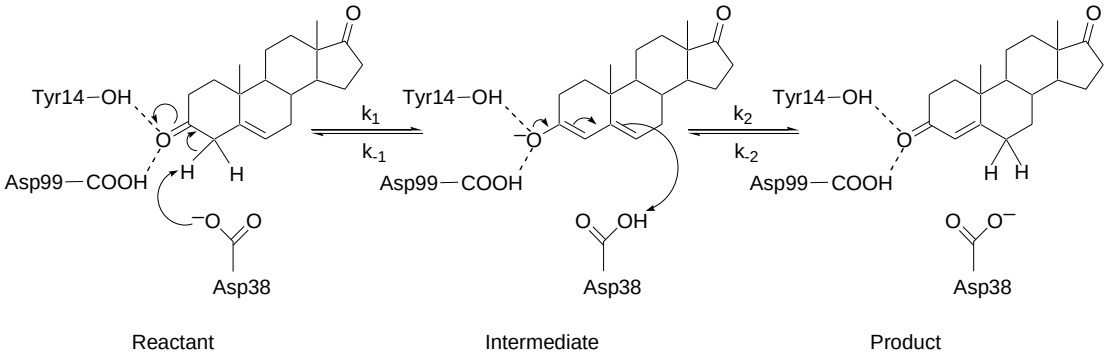


Figure **1-2**: Schematic depiction of the proton transfer reactions catalyzed by KSI. In the first step, the proton transfers from the C4 atom of the substrate to the Asp38 residue. In the second step, the proton transfers from the Asp38 residue to the C6 atom of the substrate. The reactant, intermediate, and product states of the overall reaction are labeled.

Fundamental queries about the reaction mechanism and the enzyme's catalytic efficiency have made the KSI catalyzed isomerization reaction the subject of numerous kinetic studies. Free energy profiles for the reaction catalyzed by acetate ion, and wild-type (WT) and mutant forms of *TI* KSI were experimentally determined by Pollack et. al. [9] The enzymatic free-energy profiles for the isomerization reaction have four kinetically significant barriers corresponding to the binding of the substrate, proton transfer from substrate to Asp38, proton transfer reaction from Asp38 to the substrate, and product release, respectively. The WT free-energy profile shows the intermediate state (bound intermediate) to be stabilized in the KSI reaction, making it almost iso-energetic with the reactant state (bound substrate). In contrast, the intermediate state is stabilized as much in the corresponding non-enzymatic reaction catalyzed by acetate ion. The preferential stabilization of the intermediate state by WT KSI has been hypothesized to be one of the major reasons for its catalytic efficiency.

The stabilization of the intermediate state and the catalytic efficiency of the KSI reaction have been attributed to the formation of catalytically important hydrogen bonds by Tyr14 and Asp99 residues to the substrate. The hydrogen bonding scheme involving the substrate, Tyr14 and Asp99 has been of considerable interest over the past decade. It was debated whether Asp99 formed a hydrogen bond directly to the substrate [17] or to Tyr14, which then formed a hydrogen bond to the substrate.[14] Evidence from structural, kinetic, and computational studies have identified the scheme in Figure 1-2 as the correct hydrogen bonding scheme in the KSI reaction. [16-24] In this scheme, both Tyr14 and Asp99 form hydrogen bonds directly with the substrate. Structural studies have found the active site of KSI to be lined with hydrophobic residues. Hydrogen bonds formed by Tyr14 and Asp99 to the substrate in such a hydrophobic environment are postulated to be stronger than similar hydrogen bonds formed to the substrate in an aqueous environment. Recent studies have attempted to understand the pre-organized nature of the active site and its role in the formation of these strong stabilizing hydrogen bonds. [25,]

Chapter 3 describes the study of these proton transfer reactions in wild-type KSI. [27] It presents calculations of nuclear quantum effects, transmission coefficients, thermally averaged distances and angles, thermally averaged structures, root mean square fluctuations of the backbone residues and an analysis of the non-bonding contributions. Chapter 4 extends this analysis to the Tyr14Phe, Asp99Leu and Tyr14Phe/Asp99Leu mutant varieties of KSI. These mutant forms of KSI were chosen to disrupt the catalytically important hydrogen bonds formed by Tyr14 and Asp99. Specifically, mutation of Asp99 to Leu decreases k_{cat} by $10^{2.1}$-fold, mutation of Tyr14 to Phe decreases k_{cat} by $10^{3.4}$-fold, and the corresponding double mutation decreases k_{cat} by $10^{4.6}$-fold in *Pseudomonas putida* KSI. Over the course of this analysis the rate constants for the mutant forms of KSI are calculated. We also identify new hydrogen bonding patterns in these mutants. Our simulations of the two proton transfer reactions in KSI provide a greater insight into the origins of catalysis in the enzyme.

1-2. References

(1) Garett, R.H.; Grisham, C.M.;, *Biochemistry.* Second ed.; Harcourt Brace College Publishers: Orlando, 1999; Protein Dynamics, p425-460.

(2) Michaelis, L.; Menten, M. *Biochem. Z.* **1913,** *49,* 333.

(3) Chen, S.F.; Perella, F.W.; Behrens, D.L.; Papp, L.M. *Cancer Research* **1992,** *52,* 3521.

(4) Bileter, S.R.; Webb, S.P.; Iordanov; T., Agarwal; P.K., Hammes-Schiffer, S. *J. Chem. Phys.* **2001,** *114,* 6925.

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

(6) Agarwal, P.K**;** Bileter, S.R; Hammes-Schiffer, S. *J. Phys. Chem. B* **2002,** *106,* 328

(7) Chakravorty, D.K.; Kumarasiri, M.D.; Soudackov, A.; Hammes-Schiffer, S. *J. Chem. Theory Comput.* **2008,** *4,* 1974.

(8) Oh, K. S.; Cha, S.-S.; Kim, D.-H.; Cho, H.-S.; Ha, N.-C.; Choi, G.; Lee, J. Y.; Tarakeshwar, P.; Son, H. S.; Choi, K. Y.; Oh, B.-H.; Kim, K. S. *Biochemistry.* **2000,** *39,* 13891.

(9) Pollack, R. M. *Bioorg. Chem.* **2004,** *32*, 341.

(10) Kim, S.M.; Choi, K.Y. J. Bacteriol. **1995,** *177,* 2602.

(11) Hawkinson, D. C.; Eames, T. C. M.; and Pollack, R. M. *Biochemistry* **1991,** *30,* 10849.

(12) Radzicka, A.; Wolfenden R.; *Science* **1995,** *267,* 90

(13) Zeng, B.; Pollack, R. M. *J. Am. Chem. Soc.* **1991,** *113*, 3838.

(14) Zhao, Q.; Abeygunawardana, A.G.; Gittis, A.S.; Mildvan, A.S. *Biochemistry* **1997,** *36,* 14616

(15) Massiah, M.A..; Abeygunawardana, C.; Gittis, A.G.; Mildvan, A.S. *Biochemistry* **1998,** *37,* 14701.

(16) Cho, H.S.; Ha, N.C.; Choi, G.; Kim, H.J.; Lee, D.; Oh, K.S.; Kim, K.S.; Lee, W.; Choi, K.Y.; Oh, B.H. *J. Biol. Chem.* **1999,** *274*, 32863.

(17) Wu, Z.R.; Ebrahimian, S.; Zawrotny, M.E.; Thornburg, L.D.; Perez-Alvarado, G.C.; Brothers, P.; Pollack, R.M.; Summers, M.F. *Science* **1997,** *276,* 415

(18) Park, H.; Merz, K. M., Jr. *J. Am. Chem. Soc.* **2003**, *125*, 901-911.

(19) Mazumder, D.; Kahn, K.; Bruice, T. C. *J. Am. Chem. Soc.* **2003**, *125*, 7553-7561.

(20) Kim, K. S.; Oh, K. S.; Lee, J. Y. *Proc. Nat. Acad. Sc. USA* **2000**, *97*, 6373-6378.

(21) Pan, Y.; McAllister, M. A. *Journal of Molecular Structure (Theochem)* **2000**, *504*, 29-33.

(22) Feierberg, I.; Aqvist, J. *Biochemistry* **2002**, *41*, 15728-15735.

(23) Feierberg, I.; Aqvist, J. *Theoretical Chemical Accounts* **2002**, *108*, 71-84.

(24) Warshel, A.; Sharma, P. K.; Chu, Z. T.; Aqvist, J. *Biochemistry* **2007**, *46*, 1466-1476.

(25) Kamerlin, S. C. L.; Sharma, P. K.; Chu, Z. T.; Warshel, A. *Proceedings of the National Academy of Sciences (USA)* **2010,** *107*, 4075-4080.

(26) Kraut, D. A.; Sigala, P. A.; Pybus, B.; Liu, C. W.; Ringe, D.; Petsko, G. A.; Herschlag, D. *PLoS Biology* **2006,** *4*, 501-519.

(27) Chakravorty, D.K.; Soudackov, A.; Hammes-Schiffer, S. *Biochemistry* **2009,** *48*, 10608.

(28) Humphrey, W.; Dalke, A.; Schulten, K. *Journal of Molecular Graphics* **1996**, *14*, 33-38.

(29) Benson, A.M.; Talalay P.; Keen, J.H.; Jakoby, W.B. *PNAS, USA* **1977,** *74*, 158

(30) Benson, A. M.; Talalay, P. *Biochem. Biophys. Res. Commun.* **1976,** *69*, 1073.

(31) Sulimovic, S.; Lunefeld, B. *J. Steroid Biochem.* **1971,** *2*, 203.

(32) Oleinick, N.L.; Koritz, S.L. *Biochim. et Biophys. Acta – Enzy. Bi.l Oxid.* **1966,** 1*22*, 333.

Chapter 2

Implementation of Umbrella Integration Within the Framework of the Empirical Valence Bond Approach

Reproduced in part with permission from Dhruva K. Chakravorty, Malika Kumarasiri, Alexander V. Soudackov, and Sharon Hammes-Schiffer, *Journal of Chemical Theory and Computation*, **2008,** *4,* 1974

2-1. Introduction

The calculation of free energy barriers for chemical reactions is critical for predicting reaction rates. The free energy barrier is typically obtained by generating the potential of mean force (PMF) along a specified reaction coordinate. In umbrella sampling,[1] the PMF is generated by performing molecular dynamics or Monte Carlo simulations with a series of biasing potentials that enable sampling of the entire relevant range of the reaction coordinate. The probability distribution along the reaction coordinate for each biasing potential is obtained using standard binning techniques. Various methods have been developed for combining the probability distributions for the different biasing potentials to obtain the complete PMF for the unbiased system. The weighted histogram analysis method (WHAM) has been used extensively for this purpose.[2-7] Recently, Kästner and Thiel presented the alternative umbrella integration (UI) method.[8, 9] The advantages of the UI method are that it avoids the iterative procedure inherent to WHAM, reduces the statistical errors, and converges more efficiently.[8, 9] The previous implementation of UI considered only biasing potentials in the form of harmonic restraints along the reaction coordinate.[8, 9]

In this chapter, we implement UI within the framework of the empirical valence bond (EVB) approach, in conjunction with an energy gap reaction coordinate and non-harmonic

biasing potentials defined in terms of mapping potentials. The empirical valence bond (EVB) approach has been used successfully to describe a wide range of chemical reactions in solution and proteins.[10-14] In this approach, the chemical reaction is described in terms of a small number of valence bond states, and the EVB electronic ground state is obtained by diagonalizing the Hamiltonian matrix formed in the basis of these VB states. Single proton, hydride, and electron transfer reactions are often described in terms of two valence bond states, and the energy gap reaction coordinate is defined to be the difference between the energies of these two valence bond states. When umbrella sampling is used to generate the PMF along the energy gap reaction coordinate, the biasing potential may be chosen to be the energy difference between a mapping potential, which is a linear combination of the energies of the two valence bond states, and the EVB electronic ground state energy. Previously, we used thermodynamic integration and WHAM to generate the PMF within the framework of this EVB approach for charge transfer reactions in enzymes.[15-18] We also proposed and utilized an approach for calculating the rate constant from this PMF.[19] The implementation of UI within this framework provides an alternative method with the advantages enumerated above.

An outline of the chapter is as follows. In Section 2-2, we summarize the WHAM and UI approaches and present the equations required for the implementation of UI within the framework of the EVB approach. In Section 2-3, we use both WHAM and UI to generate the PMF for the hydride transfer reaction catalyzed by the enzyme dihydrofolate reductase (DHFR). Our analysis of these calculations illustrates the advantages of UI over WHAM for this type of system. We also propose and implement a modification to WHAM that leads to similar advantages. The conclusions are presented in Section 2-4.

2-2. Methods

In umbrella sampling,[1] simulations are performed with a series of biasing potentials $w_i(\xi)$, where ξ is the reaction coordinate. The distribution $P_i^{\mathrm{b}}(\xi)$ of the biased system along the reaction coordinate is typically obtained by standard binning procedures to generate a histogram. Specifically, the relevant range of the reaction coordinate is divided into bins, and $P_i^{\mathrm{b}}(\xi_{\mathrm{bin}})$ is the fraction of sampled configurations in the bin centered at the reaction coordinate ξ_{bin} for the window corresponding to the biasing potential $w_i(\xi)$. The PMF for the biased system along the reaction coordinate is given by:

$$A_i^{\mathrm{b}}(\xi) = -\frac{1}{\beta}\ln P_i^{\mathrm{b}}(\xi), \tag{1}$$

where $\beta = 1/k_B T$. The PMF for the unbiased system in each window is

$$A_i^{\mathrm{u}}(\xi) = -\frac{1}{\beta}\ln P_i^{\mathrm{b}}(\xi) - w_i(\xi) + F_i, \tag{2}$$

where F_i are constants that differ for each biasing potential or window.

In WHAM,[2-6] the constants F_i are calculated iteratively to combine the unbiased potentials of mean force for different windows. The following two equations are solved iteratively:

$$P(\xi) = \sum_i^{windows} N_i P_i^{\mathrm{b}}(\xi) \Big/ \sum_j^{windows} N_j e^{\left[F_j - w_j(\xi)\right]\beta} \tag{3}$$

$$e^{-F_i\beta} = \int d\xi e^{-w_i(\xi)\beta} P(\xi) \tag{4}$$

where N_i is the total number of configurations sampled for window *i* used to construct $P_i^{\mathrm{b}}(\xi)$. After these equations are solved to self consistency, the PMF $A(\xi)$ is obtained directly from $P(\xi)$ using the relation $A(\xi) = -\ln P(\xi)/\beta$.

In UI,[8, 9] the derivative of the unbiased PMF with respect to the reaction coordinate is calculated for each window:

$$\frac{\partial A_i^{\mathrm{u}}(\xi)}{\partial \xi} = -\frac{1}{\beta}\frac{\partial \ln P_i^{\mathrm{b}}(\xi)}{\partial \xi} - \frac{dw_i(\xi)}{d\xi} \quad . \tag{5}$$

The data from different windows are combined according to a weighted average:

$$\frac{\partial A(\xi)}{\partial \xi} = \sum_i^{\text{windows}} p_i(\xi)\left(\frac{\partial A_i^{\mathrm{u}}(\xi)}{\partial \xi}\right) \tag{6}$$

where

$$p_i(\xi) = N_i P_i^{\mathrm{b}}(\xi) \Big/ \sum_i N_i P_i^{\mathrm{b}}(\xi) \quad . \tag{7}$$

Subsequently, $A(\xi)$ is obtained by numerical integration over ξ. In previous applications of UI, the biasing potential is assumed to be of the form $w_i(\xi) = K(\xi - \xi_i)^2/2$. Moreover, the biased PMF is expanded in a power series and truncated after the quadratic term, which is equivalent to assuming a normal distribution for $P_i^{\mathrm{b}}(\xi)$:

$$P_i^{\mathrm{b}}(\xi) = \frac{1}{\sigma_i^{\mathrm{b}}\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\xi - \overline{\xi_i^{\mathrm{b}}}}{\sigma_i^{\mathrm{b}}}\right)^2\right], \tag{8}$$

where the mean $\overline{\xi_i^{\mathrm{b}}}$ and the variance σ_i^{b} for each window are determined from the simulation data. These approximations lead to an analytical expression for the derivative of the unbiased PMF given in Eq. 5.

The UI method differs from WHAM in two important aspects. First, the UI method is based on the derivative of the PMF, rather than the PMF itself, so it does not involve offsets and therefore avoids the iterative procedure inherent to WHAM. Second, UI does not require a binning procedure because the mean and variance of the normal distribution for each window are determined directly from the raw simulation data, so a binning procedure is not required to obtain

the derivative of the PMF given in Eq. 6. Specifically, the values of the reaction coordinate for all configurations sampled are collected during the simulation, and the mean and variance of the reaction coordinates collected for each window are determined directly from these data without generating a histogram. Moreover, in our implementation, the numerical integration of this derivative to generate the PMF is performed using an adaptive integration method that is converged to a specified precision without requiring the specification of a bin width. These numerical integrals are evaluated using the global adaptive strategy[20] in conjunction with the Gauss-Kronrod quadrature rule[21] as implemented in the Mathematica software package.[22]

To facilitate a meaningful comparison of the WHAM and UI methods, we propose a modified version of the WHAM method, denoted WHAM(n), that also avoids the binning procedure. In WHAM(n), the biased distribution $P_i^{\mathrm{b}}(\xi)$ for each window is represented by the normal distribution given in Eq. 8, where the mean and variance of ξ for each window are determined directly from the simulation data. The WHAM equations given in Eqs. 3 and 4 are still solved iteratively, but $P_i^{\mathrm{b}}(\xi)$ in Eq. 3 is represented by the analytical normal distribution rather than the histogram obtained from a binning procedure. The integration in Eq. 4 is performed numerically using the adaptive integration method discussed above, thereby eliminating the necessity of specifying a bin width. Statistical methods[23] may be used to determine the error bars for the mean and variance of ξ used in Eq. 8 for both UI and WHAM(n). In addition to these statistical errors, a truncation error is introduced for both of these methods due to the approximation of the biased distribution by a normal distribution. A detailed analysis of the different sampling errors associated with UI has been performed for an analytical model potential.[9]

The main objective of this chapter is to implement the UI method within the framework of a two-state EVB potential using an energy gap reaction coordinate and a mapping potential. For a two-state EVB model, the ground state EVB energy is

$$V_{\mathrm{EVB}} = \frac{1}{2}\left(V_{11} + V_{22}\right) - \frac{1}{2}\sqrt{\left(V_{11} - V_{22}\right)^2 + 4V_{12}^2} \tag{9}$$

where V_{11} and V_{22} are the energies of VB states 1 and 2, respectively, and V_{12} is the coupling between these two states. In general, all of these quantities depend on the nuclear coordinates of the system. The energy gap reaction coordinate is defined as $\xi = V_{11} - V_{22}$. The simulations are performed with mapping potentials

$$V_{\mathrm{map}}^{i} = (1-\lambda_i)V_{11} + \lambda_i V_{22}, \tag{10}$$

where the mapping parameter λ_i is varied from zero to unity. The biasing potential is then of the form

$$w_i(\xi) = V_{\mathrm{map}}^{i} - V_{\mathrm{EVB}} = \left(\frac{1}{2} - \lambda_i\right)\xi + \frac{1}{2}\sqrt{\xi^2 + 4V_{12}^2}\,. \tag{11}$$

Note that this biasing potential is a function of only ξ if V_{12} is a function of only ξ. In this chapter, we assume that V_{12} is a constant, although the extension to the case in which V_{12} is a function of ξ is straightforward. Using this form for the biasing potential, the derivative of the unbiased PMF given in Eq. 5 is expressed as

$$\frac{\partial A_i^{\mathrm{u}}}{\partial \xi} = -\frac{1}{\beta}\frac{\partial \ln P_i^{\mathrm{b}}(\xi)}{\partial \xi} - \left(\frac{1}{2} - \lambda_i + \frac{\xi}{2\sqrt{\xi^2 + 4V_{12}^2}}\right). \tag{12}$$

Approximating $P_i^{\mathrm{b}}(\xi)$ by a normal distribution, we have obtained an analytical form for the derivative of the unbiased PMF for each window. The data for the different windows can be combined using Eq. 6, followed by numerical integration of the derivative of the PMF over ξ to obtain the PMF $A(\xi)$.

We also explore the use of different forms for the biased distribution $P_i^{\mathrm{b}}(\xi)$ because the mapping potential could lead to deviations from a normal distribution. We present results for the Gram-Charlier and the asymptotic Edgeworth expansions, which are expansions in terms of Chebyshev-Hermite polynomials. The Gram-Charlier expansion is of the form[24]

$$P_i^{\mathrm{b}}(\xi) = \frac{1}{\sigma_i^{\mathrm{b}}\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\xi - \overline{\xi_i^{\mathrm{b}}}}{\sigma_i^{b}} \right)^2 \right] \times \left[1 + \frac{\kappa_3}{3!\ {\sigma_i^{\mathrm{b}}}^5} He_3\left(\frac{\xi - \overline{\xi_i^{\mathrm{b}}}}{\sigma_i^{\mathrm{b}}} \right) + \frac{\kappa_4}{4!\ {\sigma_i^{\mathrm{b}}}^6} He_4\left(\frac{\xi - \overline{\xi_i^{\mathrm{b}}}}{\sigma_i^{\mathrm{b}}} \right) + \frac{10}{6!}\left(\frac{\kappa_3}{{\sigma_i^{\mathrm{b}}}^5} \right)^2 He_6\left(\frac{\xi - \overline{\xi_i^{\mathrm{b}}}}{\sigma_i^{\mathrm{b}}} \right) + \ldots \right] \quad (13)$$

where

$$He_n\ x\ = \ -1\ ^n e^{x^2/2} \frac{d^n}{dx^n} e^{-x^2/2}$$

are Chebyshev-Hermite polynomials, σ_i^{b} is the variance, and κ_n are the cumulants of the distribution $P_i^{\mathrm{b}}(\xi)$. The asymptotic Edgeworth expansion can be presented in the following compact form[25]

$$P_i^{\mathrm{b}}(\xi) = \frac{1}{\sigma_i^{\mathrm{b}}\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\xi - \overline{\xi_i^{\mathrm{b}}}}{\sigma_i^{\mathrm{b}}} \right)^2 \right] \left\{ 1 + \sum_{s=1}^{\infty} {\sigma_i^{\mathrm{b}}}^s \times \sum_{k_m} He_{s+2r}\ x\ \prod_{m=1}^{s} \frac{1}{k_m!} \left(\frac{S_{m+2}}{m+2\ !} \right)^{k_m} \right\}, \quad (14)$$

where $S_n \equiv \kappa_n \Big/ {\sigma_i^{\mathrm{b}}}^{2n-2}$, k_m are the solutions of the Diophantine equation $k_1 + 2k_2 + \ldots + sk_s = s$, and $r = k_1 + k_2 + \ldots + k_s$. These asymptotic expansions are useful when the biased distributions for some windows differ from the normal distributions. The derivatives of the asymptotic expansions can still be evaluated analytically, and the moments and cumulants of the biased distributions can be calculated directly from the raw sampling data.

2-3. Application

We use both WHAM and UI to calculate the PMF for hydride transfer in the enzyme DHFR. In this reaction, the hydride is transferred from the NC4 position of the NADPH (nicotinamide adenine dinucleotide phosphate) cofactor to the C6 position of the protonated dihydrofolate substrate. This reaction is depicted in Figure 2-1. We studied this reaction previously with a hybrid quantum-classical molecular dynamics approach, which includes the nuclear quantum effects of the transferring hydrogen with grid-based or path integral methods.[16-18] Here we use the same simulation system and EVB potential but do not include the nuclear quantum effects for simplicity. Since the simulations details are given elsewhere,[16, 17] we provide only a brief summary in the present chapter.

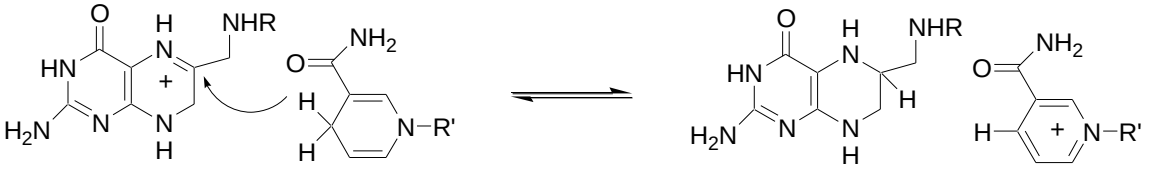


Figure **2-1**: Hydride transfer reaction from the NADPH cofactor to the protonated dihydrofolate substrate H_3F^+ to form the products tetrahydrofolate H_4F and $NADP^+$.

The simulation system includes the entire protein, the substrate, and the cofactor solvated by 4122 explicit water molecules in a truncated octahedral periodic box. The initial coordinates were obtained from a crystal structure of *Escherichia coli* DHFR complexed with $NADP^+$ and folate (PDB code 1rx2).[26] The potential energy surface is represented by a two-state EVB potential,[10] where state 1 corresponds to the transferring hydrogen atom bonded to the donor, and state 2 corresponds to the transferring hydrogen atom bonded to the acceptor. The diagonal elements of the EVB Hamiltonian are based on the GROMOS force field[27] with the EVB parameters given in Ref. [17]. The two EVB parameters corresponding to the relative energy of the

two valence bond states and the coupling between these states were fit to the experimental free energies of reaction and activation.[28]

In previous simulations, we used a set of 20 mapping parameters and performed 4.5 ns of molecular dynamics for each window with an additional 2 ns for the four windows near the transition state.[17] For the analysis in the present chapter, we generated new data, starting with a snapshot from a reactant window in the previous simulation. We used a set of 19 mapping parameters from λ_i = 0.05 to 0.95 with a spacing of 0.05. The starting configuration for each window was obtained from the previous window after 20 ps of equilibration. Each window was equilibrated for a total of 350 ps, followed by 300 ps of data collection. We also generated two other independent sets of data with 50 ps of equilibration followed by 300 ps of data collection. The free energy barriers determined from these three data sets, as well as the previous simulations,[17, 18] differ by less than 0.5 kcal/mol.

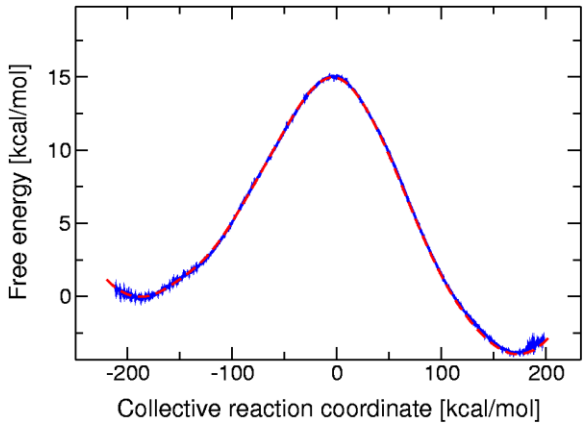


Figure **2-2**: PMF for the hydride transfer reaction generated with UI (red dashed) and WHAM (blue solid) with a bin size of 1.0 kcal/mol.

Figure 2-2 illustrates that the PMF curves generated with UI and WHAM are very similar. The free energy barriers of 15.0 kcal/mol and 15.3 kcal/mol determined with UI and WHAM, respectively, are consistent with the classical barriers determined from previous simulations using both thermodynamic integration and WHAM. However, the WHAM curve

exhibits more numerical noise, particularly in the reactant and product wells. The WHAM curve in Figure 2-2 was generated with a bin size of 1 kcal/mol. The impact of bin size on the systematic and statistical errors in WHAM has been discussed in the literature.[29]

(a) (b)

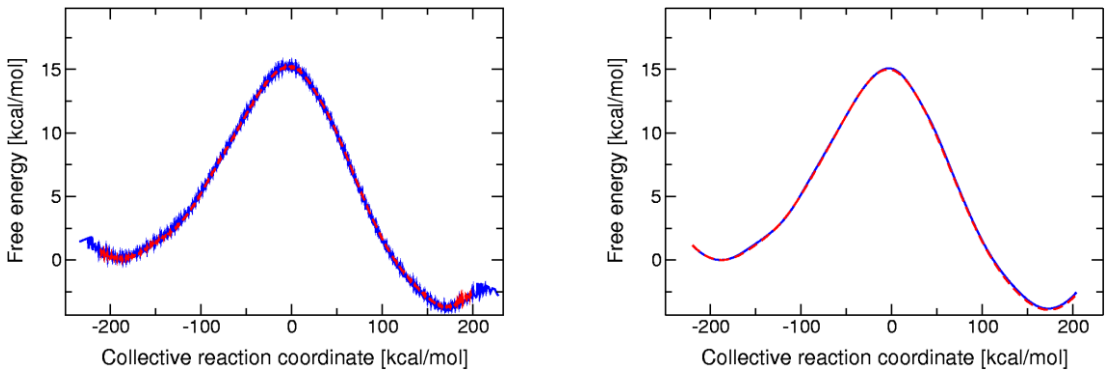


Figure **2-3**: PMF for the hydride transfer reaction generated with (a) 1.0 kcal/mol (red dashed) and 0.2 kcal/mol (blue solid) and (b) UI (red dashed) and WHAM(n) (blue solid). The UI and WHAM(n) PMF curves are virtually indistinguishable.

As discussed above, an advantage of UI is that it does not require a binning procedure for the simulation data, although it does require numerical integration to generate the PMF from its derivative. In contrast, WHAM relies on a binning procedure to generate the biased distributions used in the iterative procedure to determine the overall unbiased distribution. Moreover, WHAM does not converge as the number of bins increases (i.e., as the bin width decreases) because the statistical error increases as the number of bins increases.[29] In particular, the bin width must be sufficiently large to ensure that a sufficient number of configurations are sampled for each bin. Insufficient sampling per bin leads to large statistical fluctuations that can result in substantial inaccuracies in the probability densities generated with WHAM. These difficulties with statistical error are avoided in UI because the biased distribution is represented by the analytical normal distribution function given in Eq. 8, where the mean and variance of the reaction coordinate for each window are obtained directly from the simulation data. In addition, a low

weight is assigned to the tails of the distribution from each window in UI, as indicated by Eq. 7. As mentioned above, statistical methods[23] may be used to provide well-defined error bars for the mean and variance of the reaction coordinate, which can be propagated to estimate the sampling error for the resulting PMF.

Figure 2-3a illustrates the impact of bin size on the PMF curve generated with WHAM. Decreasing the bin size from 1.0 kcal/mol to 0.2 kcal/mol significantly increases the statistical noise of the PMF generated with WHAM. For comparison, Figure 2-3b depicts the PMF curve generated with the WHAM(n) method. This figure indicates that the statistical errors in WHAM are significantly reduced when the biased distribution for each window is represented by the analytical normal distribution function rather than the histogram obtained from the binning procedure. This figure also illustrates that the PMF curve generated with WHAM(n) is virtually indistinguishable from the PMF curve generated with UI.

Another advantage of UI is that it does not require overlap between the distributions of the windows, although such overlap is desirable to enhance the accuracy. In contrast, WHAM requires sufficient overlap between the distributions of the windows. Figure 2-4 depicts the PMF generated with UI, WHAM, and WHAM(n) using only five windows corresponding to λ_i = 0.05, 0.15, 0.50, 0.85, and 0.95 (i.e., two windows in the reactant and product regions and one window in the barrier region). The PMF curve generated with UI using only five windows is virtually identical to the curve generated with all 19 windows. In contrast, the PMF curve generated with WHAM using only five windows is clearly problematic, as illustrated by Figure 2-4b. The barrier improves as the convergence criterion for the constants F_i determined during the iterative procedure is tightened from a maximum change of 10^{-4} to 10^{-8}, but the number of iterations required for convergence increases to more than 7.6×10^{7} for a convergence criterion of 10^{-8}, which still does not generate a smooth PMF. As shown in Figure 2-4c, the PMF curve generated with WHAM(n) using only five windows is better than that generated with WHAM for the same

a)

b)

c)

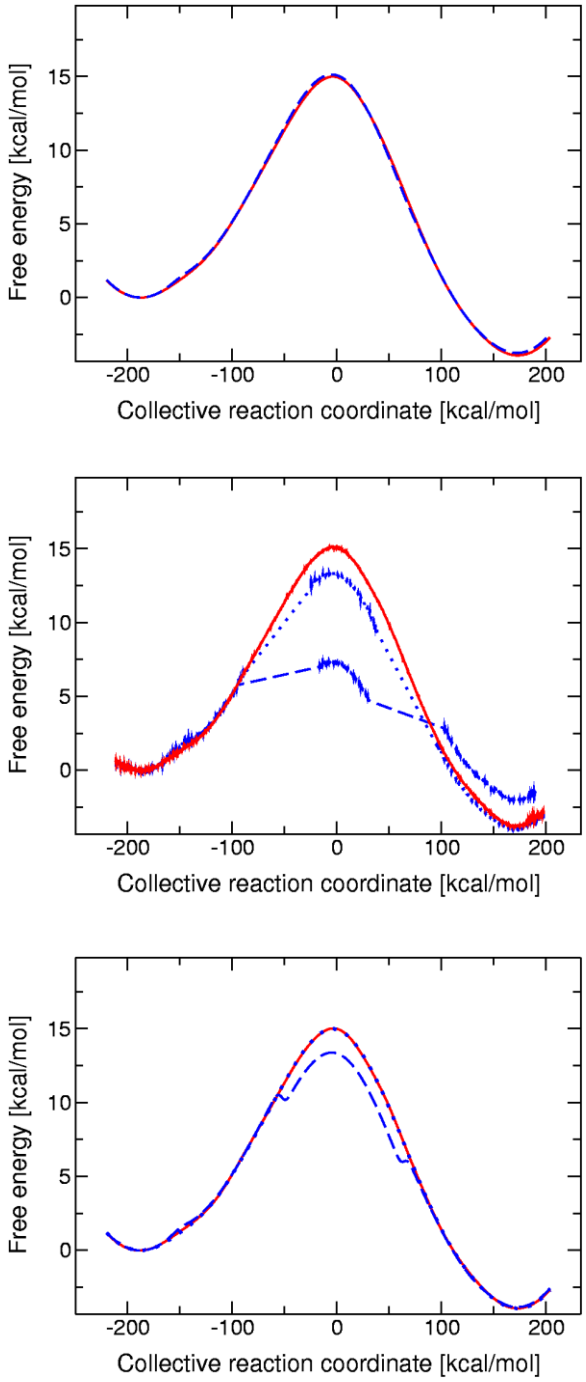


Figure **2-4**: PMF for the hydride transfer reaction generated using 19 windows (red solid) and five windows (blue dashed or dotted) with (a) UI, (b) WHAM, and (c) WHAM(n). For UI, the two PMF curves are virtually indistinguishable. The PMF curves generated with WHAM are shown for a convergence criterion of 10^{-4} (dashed) and 10^{-8} (dotted). The PMF curves generated with WHAM(n) are shown for a convergence criterion of 10^{-4} (dashed) and 10^{-6} (dotted). The 19 windows correspond to equally spaced values of λ_i in the range λ_i = 0.05 to 0.95, and the five windows correspond to λ_i = 0.05, 0.15, 0.50, 0.85, 0.95.

convergence criterion, but WHAM(n) still requires more than 8.6×10^{4} iterations for a convergence criterion of 10^{-6}, which generates a PMF that is indistinguishable from the PMF generated with WHAM(n) using all 19 windows.

In principle, given sufficient sampling within each window, WHAM and UI should converge to the same results if the distributions are Gaussian. However, the convergence of the iterative procedure in WHAM becomes slow for small overlap between the distributions of the windows, and insufficient sampling of the tail regions of the distributions combined with very small overlap could preclude convergence. An advantage of UI is that it utilizes an analytical expression for the distributions, thereby decreasing the statistical noise. Moreover, UI does not require an iterative procedure, so convergence is not an issue. These advantages become particularly pronounced for small overlaps between the distributions of the windows, although additional windows will enhance the accuracy of both methods.

Lastly, we test the approximation of the biased distribution function $P_i^{\mathrm{b}}(\xi)$ by a normal distribution, as given in Eq. 8. For this purpose, we explore the use of the Gram-Charlier and Edgeworth expansions.[25] The data and biased distribution functions for a representative window in the reactant region are shown in Figure 2-5a. The Gram-Charlier expansion is virtually indistinguishable from the normal distribution, whereas the Edgeworth expansion slightly improves the fit of the distribution obtained from the simulation data. As shown in Figure 2-5b, however, all three distribution functions lead to indistinguishable PMF curves. These data indicate that the approximation of the biased distribution by a normal distribution function is sufficient for generating quantitatively accurate PMF curves for this system. Note that this approximation may not be valid for certain systems, particularly when weak biasing potentials are used for free energy surfaces with high barriers or extended flat regions. In these cases, the

WHAM method based on histograms obtained from a binning procedure could be more effective than the UI method.

(a) (b)

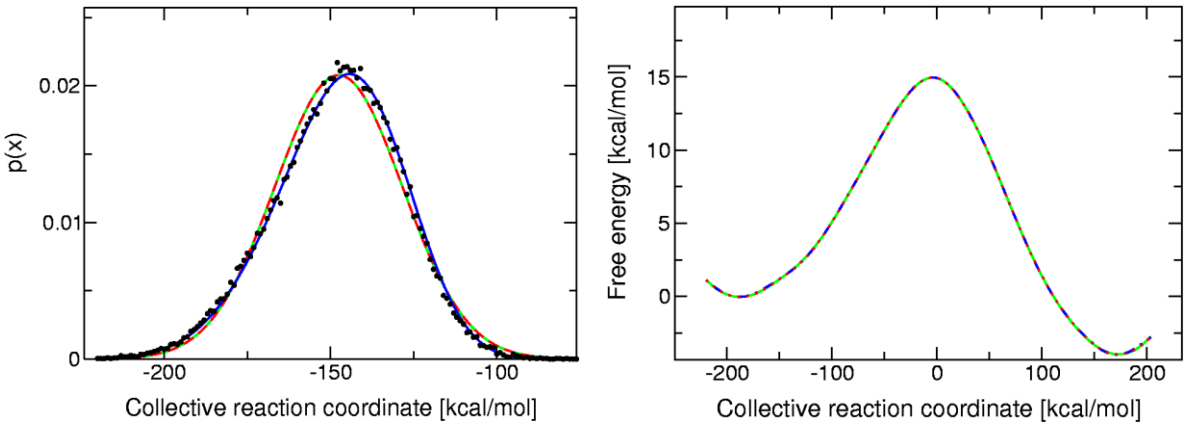


Figure **2-5**: (a) Biased probability distribution function for the window with λ_i = 0.10, where the filled circles represent the normalized histogram constructed from the simulation data and the solid and dashed lines represent fits to a normal distribution (red), a Gram-Charlier expansion to third order (green), and an Edgeworth expansion with three terms (blue). (b) PMF curves generated with the three fits in (a). The three PMF curves are virtually indistinguishable.

2-4. Conclusions

In this chapter, we implemented the UI method for calculating the PMF along an energy gap reaction coordinate within the EVB framework. The UI method is based on the derivative of the PMF with respect to the reaction coordinate rather than the PMF itself. In this implementation, the biasing potential is the difference between the mapping potential, which is defined to be a linear combination of the valence bond state energies, and the EVB ground state energy. This biasing potential can be expressed as an analytical function of the energy gap reaction coordinate for a two-state EVB model in which the coupling between the two states is constant or is a function of the reaction coordinate. In this case, the derivative of the biasing

potential with respect to the reaction coordinate can be expressed analytically, and the implementation of the UI method is straightforward.

We applied the UI and WHAM methods to the hydride transfer reaction catalyzed by DHFR. We showed that the UI and WHAM methods generate very similar PMF curves, although the PMF curve generated with UI exhibited less statistical noise. We also showed that the representation of the biased probability distributions as normal distributions is reasonable by comparison to expansions including non-Gaussian effects. Furthermore, our analysis illustrated two significant advantages of UI over WHAM. The first advantage is that UI does not rely on a binning procedure to generate histograms and therefore reduces the statistical error and converges efficiently. We proposed a modified version of WHAM that shares these advantages by representing the biased probability distribution for each window as an analytical normal distribution function rather than the histogram obtained from a binning procedure. The second advantage is that UI can provide accurate PMF curves efficiently even with a small number of windows that do not overlap significantly. In this case, the modified version of WHAM can also provide accurate PMF curves but is more computationally expensive because it requires a large number of iterations for convergence. Thus, UI is a promising method for generating accurate PMF curves for large systems for which sampling may be limited.

2-5. References

(1) Torrie, G. M.; Valleau, J. P. *Chem. Phys. Lett.* **1974**, *28*, 578-581.

(2) Ferrenberg, A. M.; Swendsen, R. H. *Phys. Rev. Lett.* **1988**, *61*, 2635-2638.

(3) Ferrenberg, A. M.; Swendsen, R. H. *Phys. Rev. Lett.* **1989**, *63*, 1195-1198.

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

(5) Roux, B. *Comput. Phys. Commun.* **1995**, *91*, 275-282.

(6) Souaille, M.; Roux, B. *Comput. Phys. Commun.* **2001**, *135*, 40-57.

(7) Bartels, C.; Karplus, M. *J. Comput. Chem.* **1997**, *18*, 1450-1462.

(8) Kastner, J.; Thiel, W. *J. Chem. Phys.* **2005**, *123*, 144104.

(9) Kastner, J.; Thiel, W. *J. Chem. Phys.* **2006**, *124*, 234106.

(10) Warshel, A. *Computer Modeling of Chemical Reactions in Enzymes and Solutions*; John Wiley & Sons, Inc.: New York, 1991.

(11) Schmitt, U. W.; Voth, G. A. *J. Phys. Chem. B* **1998**, *102*, 5547-5551.

(12) Vuilleumier, R.; Borgis, D. *Chem. Phys. Lett.* **1998**, *284*, 71-77.

(13) Cembran, A.; Gao, J. *Theor. Chem. Acc.* **2007**, *118*, 211-218.

(14) Truhlar, D. G. *J. Comput. Chem.* **2007**, *28*, 73-86.

(15) Billeter, S. R.; Webb, S. P.; Iordanov, T.; Agarwal, P. K.; Hammes-Schiffer, S. *J. Chem. Phys.* **2001**, *114*, 6925-6936.

(16) Agarwal, P. K.; Billeter, S. R.; Hammes-Schiffer, S. *J. Phys. Chem. B* **2002**, *106*, 3283-3293.

(17) Wong, K. F.; Watney, J. B.; Hammes-Schiffer, S. *J. Phys. Chem. B* **2004**, *108*, 12231-12241.

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

(19) Watney, J. B.; Soudackov, A. V.; Wong, K. F.; Hammes-Schiffer, S. *Chem. Phys. Lett.* **2006**, *418*, 268-271.

(20) Malcolm, M. A.; Simpson, R. B. *ACM Transactions on Mathematical Software* **1975**, *1*, 129-146.

(21) Piessens, R.; Branders, M. *Mathematics of Computation* **1974**, *28*, 135-139.

(22) *Mathematica, Version 6.0*; Wolfram Research, Inc.: Champaign, IL, 2007.

(23) Schiferl, S. K.; Wallace, D. C. *J. Chem. Phys.* **1985**, *83*, 5203-5209.

(24) Stuart, A.; Ord, J. K. *Kendall's Advanced Theory of Statistics*; Sixth ed.; John Wiley & Sons, Inc.: New York, 1994; Vol. 1. Distribution Theory.

(25) Blinnikov, S.; Moessner, R. *Astronomy & Astrophysics Supplement Series* **1998**, *130*, 193-205.

(26) Sawaya, M. R.; Kraut, J. *Biochemistry* **1997**, *36*, 586-603.

(27) van Gunsteren, W. F.; Billeter, S. R.; Eising, A. A.; Hunenberger, P. H.; Kruger, P.; Mark, A. E.; Scott, W. R. P.; Tironi, I. G. *Biomolecular simulation: The GROMOS96 manual and user guide*; VdF Hochschulverlag, ETH Zurich: Zurich, 1996.

(28) Fierke, C. A.; Johnson, K. A.; Benkovic, S. J. *Biochemistry* **1987**, *26*, 4085-92.

(29) Kobrak, M. N. *J. Comput. Chem.* **2003**, *24*, 1437-1446.

Chapter 3

Hybrid Quantum Classical Molecular Dynamics Simulations of the Proton Transfer Reactions in Ketosteroid Isomerase: Analysis of Hydrogen Bonding, Conformational Motions and Electrostatics Analysis

Reproduced in part with permission from Dhruva K. Chakravorty, Alexander V. Soudackov, and Sharon Hammes-Schiffer, *Biochemistry*, **2009,** *48,* 10608.

3-1. Introduction

Δ^5-3-Ketosteroid isomerase (KSI) catalyzes the isomerization of 3-oxo-Δ^5-steroids to their Δ^4-conjugated isomers. The reaction occurs by a two-step general acid-base mechanism involving a dienolate intermediate. [1] Kinetic studies of KSI have provided estimates of the free energy barriers for the four steps corresponding to substrate binding, two proton transfer reactions, and dissociation of product. [2] A variety of mutant KSIs have been studied to identify the catalytically important residues. [3-8] The three-dimensional structures of KSIs from two bacterial strains, *Commamonas testosteroni* and *Pseudomonas putida*, have been determined by X-ray crystallography and NMR. [3, 4, 9-13] These two KSIs are homologous with 34% sequence identity, and the key catalytic residues of the active sites are conserved.

The catalytic efficiency of KSI approaches the diffusion-controlled limit. [1] This high catalytic efficiency has been explained in terms of a wide range of factors, including electrostatic stabilization and hydrogen bonding. [3, 4, 14, 15] Electronic structure calculations on model systems have illustrated the importance of hydrogen bonding in KSI. [13, 16, 17] Molecular dynamics (MD) simulations of KSI have provided additional insights into the mechanistic role of hydrogen

bonding. [14, 15, 18-20] Experimental studies of the primary and secondary kinetic isotope effects have suggested that hydrogen tunneling is significant in these reactions. [21] Moreover, systematic studies of a series of phenolates binding to KSI have enabled the evaluation of both geometric and electrostatic contributions to binding and catalysis. [20, 22] Additional recent experimental studies have been directed at clarifying the role of electrostatics, geometrical discrimination, and hydrogen bond coupling in the active site of KSI. [23-26]

In this chapter, we present hybrid quantum/classical MD simulations aimed at elucidating the geometrical, conformational, and electrostatic changes occurring during the isomerization reaction catalyzed by KSI. Our simulations focus on KSI from *Commamonas testosteroni* bacteria with the Δ^5-androstene-3,17-dione (5-AND) substrate. The two-step mechanism is depicted in Figure 3-1. In the first step, Asp38 abstracts the axial β–proton from the C4 carbon atom of the substrate to form a dienolate intermediate. The negative charge on the dienolate intermediate is thought to be stabilized by hydrogen bonds with Tyr14 and Asp99. In the second step, Asp38 transfers a proton to the axial β-position at the allylic C6 carbon atom of the substrate to form the Δ^4-androstene-3,17-dione (4-AND) product.

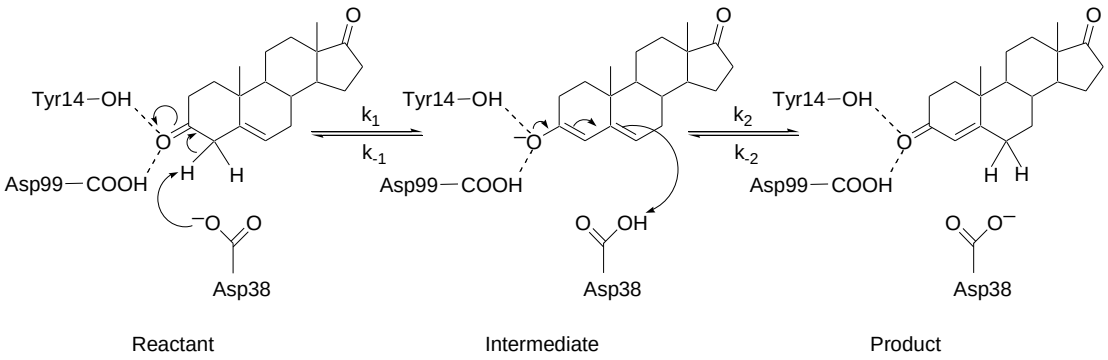


Figure **3-1**: Schematic depiction of the proton transfer reactions catalyzed by KSI. In the first step, the proton transfers from the C4 atom of the substrate to the Asp38 residue. In the second step, the proton transfers from the Asp38 residue to the C6 atom of the substrate. The reactant, intermediate, and product states of the overall reaction are labeled.

To study this isomerization reaction, we performed MD simulations using an empirical valence bond (EVB) potential to generate the free energy profiles for both proton transfer steps along a collective reaction coordinate. We also calculated the contributions to the proton transfer rates from the nuclear quantum effects of the transferring hydrogen and the dynamical barrier recrossings. Nuclear quantum effects are thought to be important for this reaction based on the experimentally measured primary deuterium kinetic isotope effect of ~6. [27] In addition, we analyzed changes in key geometrical properties within the active site along the collective reaction coordinate to clarify the roles of the catalytically important hydrogen bonds. Moreover, we examined thermally averaged structures along the reaction pathway to identify conformational changes in the protein that are associated with the proton transfer reactions. An analysis of the electrostatic potentials for these thermally averaged structures provides additional insights into the changes in the electrostatic environment throughout the protein that are occurring during the proton transfer reactions.

The present work is related to several previous computational studies of KSI. In Refs. [14] and [15], MD simulations were used to characterize hydrogen bonding in the active site of KSI bound to the substrate and intermediate, but these simulations did not probe the proton transfer reactions. In Ref. [18], the EVB method was combined with classical MD free energy perturbation simulations to study the proton transfer reactions catalyzed by KSI. The methodology in the present work differs from that used in Ref. [18] because the present work includes contributions from the nuclear quantum effects of the transferring hydrogen and the dynamical barrier recrossings. The methodologies also utilize different forcefields and EVB parametrizations, and the present MD simulations were propagated for substantially longer times. In addition, the previous simulations focused on KSI from *Pseudomonas putida*, while the present simulations focus on KSI from *Commamonas testosteroni*. The previous work analyzed hydrogen bonding, as well as the electrostatic and van der Waals interaction energies, in the active site for the

reactant and intermediate species. In the present work, we extend this analysis to study hydrogen bonding, conformational motions in the active site and distal loop regions, and electrostatics throughout the enzyme along the entire collective reaction coordinate for both proton transfer reactions. Our results are consistent with the previous study [18] and provide additional insights into these aspects of the KSI enzymatic reaction.

An outline of this chapter is as follows. Section 3-2 describes the computational methods used in these studies. Section 3-3 presents the results of applying these methods to the two proton transfer reactions catalyzed by KSI. In the first part, we present the calculated free energy profiles, nuclear quantum effects, and transmission coefficients. In the second part, we present an analysis of the geometrical changes in the active site during the two proton transfer reactions, as well as the conformational and electrostatic changes in the entire enzyme along the reaction pathway. The conclusions are presented in Section 3-4.

3-2. Methods

A. EVB Model

A two-state empirical valence bond (EVB) potential was used to model the electronic potential energy surface for each proton transfer reaction catalyzed by KSI. [28] The two proton transfer reactions are thought to be sequential, where the product of the first proton transfer reaction is the reactant of the second proton transfer reaction. Moreover, early experiments established conservation of the transferred proton (i.e., the same proton is transferred in both steps). [29] These two proton transfer steps are depicted in Figure 3-1. Since the Intermediate is thermally stable and the electronic coupling between the Reactant and Product states is negligible, we describe these reactions with two two-state EVB models rather than one three-state EVB model. For the first step, the two EVB states correspond to the Reactant and the Intermediate. For the second step, the two EVB states correspond to the Intermediate and the Product. The

partial atomic charges for the substrate in all three states were calculated using the RESP methodology. [30] For each two-state EVB model, the diagonal matrix elements V_{11} and V_{22} of the EVB Hamiltonian were described by the AMBER99 forcefield [31, 32] with modifications described below, and the energy difference Δ and the coupling V_{12} between the two states were obtained by fitting to the experimentally determined rate constants for the forward and backward reactions. [2, 33] The ground state electronic potential energy surface for each proton transfer reaction is obtained from the lowest energy eigenvalue of the corresponding 2×2 EVB Hamiltonian.

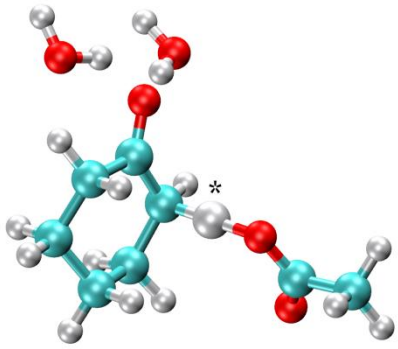


Figure **3-2**: Model system of the KSI active site used to assist in the parameterization of the EVB potential. In this simple model, the substrate is represented by cyclohexone, Asp38 is represented by ethanoic acid, and Tyr14 and Asp99 are represented by water molecules to include the key hydrogen bonding interactions. The transferring hydrogen is identified with an asterisk.

As mentioned above, minor modifications to the AMBER99 forcefield were implemented for the diagonal matrix elements of the EVB Hamiltonians. The charge on the transferring hydrogen was found to be small, so this charge was added to the charge of the donor atom in the reactant and to the charge of the acceptor atom in the product. The resulting zero charge on the transferring hydrogen enhances the computational efficiency for the path integral calculations

because the electrostatic interactions do not need to be recalculated when only the hydrogen coordinates change. In addition, small Lennard-Jones parameters were assigned to the transferring hydrogen atom to provide a more accurate description of the breaking and forming of the bonds. We determined the values of these parameters by studying the model system shown in Figure 3-2, where cyclohexone is used to mimic the substrate, ethanoic acid is used to mimic aspartic acid, and two water molecules are used to mimic the hydrogen-bonding network in the active site. We calculated the transition state for this model system using density functional theory (DFT) at the B3LYP/6-31G** level with Gaussian03. [34] The transition state structure donor-hydrogen distance, acceptor-hydrogen distance, and donor-acceptor distance were calculated to be 1.45 Å, 1.20 Å, and 2.65 Å, respectively. The Lennard-Jones parameters for the hydrogen were chosen so that the donor-acceptor distance for the midpoint of the first proton transfer step generated with the EVB potential was consistent with the distance of 2.65 Å obtained for the transition state of this model system. The resulting Lennard-Jones parameters for the transferring hydrogen were $\varepsilon = 4.0 \times 10^{-6}$ kcal/mol and $\sigma = 1.487$ Å. The resulting EVB hydrogen potential energy surfaces were found to be qualitatively consistent with those obtained from DFT calculations at the B3LYP/6-31G** level.

We calculated the potential of mean force (PMF) for each proton transfer reaction as a function of a collective reaction coordinate that includes motions of the entire solvated enzyme. This energy gap reaction coordinate, Λ, is defined as

$$\Lambda = V_{11} - V_{22} . \tag{1}$$

A series of biasing potentials (i.e., mapping potentials) was used to drive the reaction from the reactant to the product state. These mapping potentials are defined as [28]

$$V_{\text{map}}^{\lambda} = (1-\lambda)V_{11} + \lambda V_{22}\,, \tag{2}$$

where the mapping parameter λ is varied between zero and unity. Classical MD trajectories were propagated according to these mapping potentials, and the PMF for each mapping potential was calculated using standard binning procedures. These individual pieces of the PMF were combined to obtain the complete PMF for the unbiased EVB potential using the weighted histogram analysis method (WHAM) [35-38] or umbrella integration. [39-41]

B. Calculation of rate constants

Within the framework of transition state theory (TST), the rate constant is expressed in the following form:

$$k = \kappa k_{\text{TST}}\,, \tag{3}$$

where k_{TST} is the TST rate constant determined by the equilibrium forward flux through the dividing surface and κ is the transmission coefficient that accounts for dynamical recrossings of the dividing surface. The TST rate constant can be expressed as

$$k_{\text{TST}} = \frac{1}{\beta h} e^{-\beta \Delta G^{\ddagger}} \tag{4}$$

where $\beta = 1/k_{\text{B}}T$, k_{B} is the Boltzmann constant, and $\Delta G^{\ddagger}$ is the free energy barrier for the reaction. An alternative form for the TST rate constant in terms of the PMF along a general reaction coordinate has been derived. [42-44] For simplicity, here we use Eq. (4) with the PMF barrier as the free energy barrier. The results using the alternative expression derived in Ref. [42] are very similar and are provided in Appendix A.

The reactive flux approach can be used to calculate the transmission coefficient κ [45-49]. In this approach, the transmission coefficient is calculated as the flux-weighted average of the

quantity ξ for a canonical ensemble of trajectories initiated at the dividing surface and propagated backward and forward in time. The quantity ξ is defined so that $\xi = 1/\alpha$ for trajectories with α forward crossings and $\alpha-1$ backward crossings and $\xi = 0$ otherwise. In practice, an equilibrium distribution of configurations near the dividing surface is selected from a simulation based on the mapping potential in Eq. (2) with the mapping parameter $\lambda = 0.5$. Initial velocities are sampled from a Maxwell-Boltzmann distribution of velocities chosen so that the initial velocity component normal to the dividing surface is positive. Starting with these initial conditions, MD trajectories are propagated backward and forward in time until they reach either the reactant or product region. The transmission coefficient is calculated from the expression [50, 51]

$$\kappa = \frac{\sum_{i=1}^{N_{\text{traj}}} \dot{\mathbf{R}}_i \cdot \hat{\mathbf{n}}_i \;\; w_i^{\text{can}} \xi}{\sum_{i=1}^{N_{\text{traj}}} \dot{\mathbf{R}}_i \cdot \hat{\mathbf{n}}_i \;\; w_i^{\text{can}}}, \tag{5}$$

where $\dot{\mathbf{R}}_i$ denotes the initial velocities, $\hat{\mathbf{n}}_i$ is the unit vector normal to the dividing surface, and w_i^{can} is the weighting to ensure a canonical distribution for the EVB potential at the dividing surface. In our calculations, the dividing surface is defined as $\Lambda = 0$, where Λ is the energy gap reaction coordinate defined in Eq. (1).

The nuclear quantum effects associated with the transferring hydrogen nucleus are included using the quantized classical path (QCP) method. [52-54] In this method, free particle path integral calculations are performed along a classical MD trajectory, where the transferring hydrogen is represented by a ring of beads with the centroid constrained to the classical position of the hydrogen. This method provides corrections to the classical free energy barrier that account for nuclear quantum effects such as zero point energy and hydrogen tunneling. This method has been shown to lead to similar results as Fourier grid methods. [55]

C. Simulation Details

The 2.3 Å resolution crystal structure (PDB 1QJG)[12] of *Commonas testosteroni* Δ^5-3-ketosteroid isomerase complexed with equilenin (EQU) was selected as the starting structure for the simulations. This structure has six monomeric units (i.e., three dimers) with the two mutations Asp38Asn and Ile83Thr. For computational efficiency, we used only the second dimer, monomeric units C and D, for our simulations. For simplicity, the proton transfer reactions in only one monomeric unit (Chain C) were studied. Allosteric effects between the two identical monomeric units are not thought to be significant. [15] Moreover, we changed the mutated residues 38 and 83 back to Asp and Ile, respectively. Finally, we replaced the EQU, which mimics the intermediate state of the substrate, 3-hydroxy-androsta-3,5-dien-17-one, back to the naturally occurring substrate, 5-AND.

The initial state of the enzyme was prepared as follows. The mutated residues 38 and 83 were restored to those in the wild-type enzyme using the profix utility in the JACKAL protein modeling package. [56, 57] The protonation states for wild-type KSI were determined from the 1BUQ solution structure[11] of KSI complexed with 19-nortestosterone–hemisuccinate. This NMR structure was calculated at a pH of 7.0, which is the pH at which the experimental rates were obtained. The EQU substrate analogue was replaced with 5-AND in both active sites of the dimeric unit. For this purpose, 5-AND was optimized in the gas phase with DFT at the B3LYP/6-31G** level with Gaussian03. [34] The optimized structure of 5-AND was mapped onto EQU in the active sites of the enzyme by superimposing the C4 atom and the carbonyl oxygen of 5-AND onto the C4 atom and hydroxyl oxygen of EQU. The partial atomic charges for the 5-AND substrate, the 4-AND product, and the dienolate intermediate were determined using the RESP methodology, [30] and the other forcefield parameters for these ligands were obtained from the General Amber Force Field (GAFF). [58]

The protein was immersed in a truncated octahedron box of pre-equilibrated TIP3P rigid water molecules. [59, 60] All crystallographic waters within 5.4 Å of units C and D were included to preserve any significant hydrogen-bonding interactions. Six water molecules were replaced by Na^+ ions to maintain charge neutrality. The total solvated enzyme system consisted of 3844 protein atoms, 6837 water molecules, and six Na^+ ions, leading to a total of 24361 atoms in the system. All atoms in the system, including the water molecules, were treated explicitly. The hydrogen atoms bonded to heavy atoms were constrained to their equilibrium distances with the SHAKE algorithm. [61] The Smooth Particle Mesh Ewald method [62] was used to calculate the electrostatic interactions. A modified version of the DLPROTEIN package [63] was used for these simulations.

The system was prepared for the simulations using a well-defined equilibration procedure. In the first step, the water molecules were equilibrated with 50 ps of MD for an isobaric, isothermal ensemble (NPT) at 300K with the protein fixed. Then the residues Asp38 and Thr83, which were mutated in the crystal structure but restored to the wild-type residues for the simulations, were optimized (i.e., the energy was minimized with respect to their coordinates) with the protein and water molecules fixed. Subsequently, the water molecules and ions were optimized with the protein fixed, and then the entire system was optimized. During this last optimization step and for the remainder of the equilibration procedure, three constraints were applied to maintain the catalytically important hydrogen bonds. These constraints were applied to the distances between the Tyr14 hydroxyl oxygen atom and the substrate O3 atom, between the Asp99 hydroxyl oxygen atom and the substrate O3 atom, and between the carboxylic carbon of Asp38 and the substrate C4 atom. These distances were constrained to their crystal structure values of 2.56 Å and 2.57 Å and 4.42 Å, respectively, during the equilibration. Note that the hydrogen bond lengths obtained by NMR chemical shifts and fractionation factors of the dihydroequilenin complex were 2.49 ± 0.02 Å and 2.72 ± 0.02 Å for Tyr14 and Asp99,

respectively. [8, 64, 65] The errors in these distances in the crystal structure impact the starting configuration and initial equilibration of our MD simulations. As discussed below, however, these constraints were released prior to the equilibration for each window (i.e., for each mapping potential) and during data collection. Thus, the hydrogen bond distances are allowed to change according to the EVB potential, providing qualitatively meaningful but not quantitatively accurate values.

The system was equilibrated according to a simulated annealing procedure consisting of 50 ps of NPT MD at each temperature, where the temperature was increased from 50 K to 300 K by increments of 50K. The last step of equilibration consisted of 200 ps of MD for a canonical ensemble (NVT) at 298 K. For all of the MD simulations, the temperature and pressure were maintained with Nosé-Hoover thermostats. [66, 67] The three constraints mentioned above were removed after this last step of the initial equilibration. The length of the side for the final cell was 78.77 Å. The root mean squared deviation (RMSD) between the protein backbone atoms of the final equilibrated structure and the initial crystal structure was calculated to be 1.42 Å. All MD simulations for data collection were propagated at 298 K for a canonical ensemble (NVT) with a time step of 1 fs.

For the first proton transfer step, we propagated MD trajectories using 19 different mapping potentials with values of λ = 0.05, 0.09, 0.125, 0.190, 0.250, 0.310, 0.375, 0.400, 0.440, 0.500, 0.565, 0.625, 0.690, 0.750, 0.80, 0.815, 0.875, 0.910 and 0.950. These mapping potentials span the region of the reaction coordinate from the Reactant to the Intermediate state. The window corresponding to λ = 0.05 was started from the equilibrated system described above. This structure was further equilibrated for 100 ps prior to data collection. The trajectory in the next window (λ = 0.09) was started from this equilibrated structure. Each subsequent window was started from the configuration of the previous window after 10 ps of equilibration. For each window, 100 ps of equilibration was performed prior to data collection. Two independent sets of

trajectories were propagated, starting from the same equilibrated configuration but different initial velocities in the first window. In the second set of trajectories, a procedure involving gradually decreasing restraints on the hydrogen bonds between the substrate and Asp99 and Tyr14 was performed prior to the 100 ps of equilibration in each window. Over 1 ns of data was collected for each mapping potential in both sets of trajectories, leading to a total of over 40 ns of data collection for this first proton transfer step. For each of the two data sets acquired, the PMF curves were generated with WHAM using a bin size of 1.0 kcal/mol.

We followed a similar procedure to study the second proton transfer step. In this case, we propagated MD simulations using 19 different mapping potentials with values of λ = 0.05, 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40, 0.45, 0.50, 0.55, 0.60, 0.65, 0.70, 0.75, 0.80, 0.85, 0.90 and 0.95. These mapping potentials span the region of the reaction coordinate from the Intermediate to the Product state. The initial configurations for the two independent sets of trajectories were chosen to be the protein conformations obtained after 1 ns of MD from the window corresponding to λ = 0.95 for the first proton transfer step. The energies of these conformations were minimized according to the mapping potential corresponding to λ = 0.05 for the second proton transfer step and were subsequently equilibrated with 100 ps of MD prior to data collection. The remainder of the procedure was identical to the one described above for the first proton transfer step. Over 1 ns of data was collected for each mapping potential for the two independent sets of trajectories, leading to a total of over 40 ns of data collection, and the PMF curves were generated for the two sets of trajectories with WHAM.

In the path integral QCP calculations, each quantum hydrogen nucleus was represented by a ring of 16 beads. The quantum distribution for each classical configuration was equilibrated over 100 three-level bisection steps. Data was collected over 1000 three-level bisection steps with quantum energies stored every 10 steps. We confirmed that the number of beads, equilibration steps, and data collection steps were sufficient for this system. We performed QCP

calculations treating only the transferring hydrogen nucleus quantum mechanically and treating the two hydrogen nuclei involved in the hydrogen bonds between the substrate and Tyr14 and Asp99, as well as the transferring hydrogen nucleus, quantum mechanically. For the latter QCP calculations, the SHAKE constraints applied to the O–H bonds in Tyr14 and Asp99 during the generation of classical configurations were removed.

We used the reactive flux approach described above to calculate the transmission coefficients. For these calculations, we selected 100 configurations that were within 2.5 kcal/mol of the dividing surface. For each configuration, 100 sets of initial velocities were assigned according to a Maxwell-Boltzmann distribution. Each trajectory was propagated forward and backward in time with a time step of 0.01 fs until it reached the reactant or product region. The transmission coefficient was calculated for these 10^4 trajectories using Eq. (5) for each of the two data sets for each proton transfer step.

3-3. Results

A. Rate calculations: PMFs, nuclear quantum effects, transmission coefficients

As described above, we generated 40 ns of MD data using a series of mapping potentials for each proton transfer step. We used this data to calculate the classical free energy profiles (i.e., the PMF curves) for each proton transfer step using both WHAM and umbrella integration and found that the barrier is the same to within 0.5 kcal/mol for these two methods. The PMF curves for the two proton transfer steps are depicted in Figure 3-3. For each step, we generated separate PMF curves for the two data sets to test convergence and reproducibility. For the first proton transfer step, the PMF barrier is 10.3 and 9.7 kcal/mol for the first and second data sets, respectively. For the second proton transfer step, the PMF barrier is 10.1 and 9.7 kcal/mol for the first and second data sets, respectively. A comparison between the two data sets, as well as

between the use of WHAM and umbrella integration, suggests that the error in the PMF barrier due to numerical procedures is ~1.0 kcal/mol.

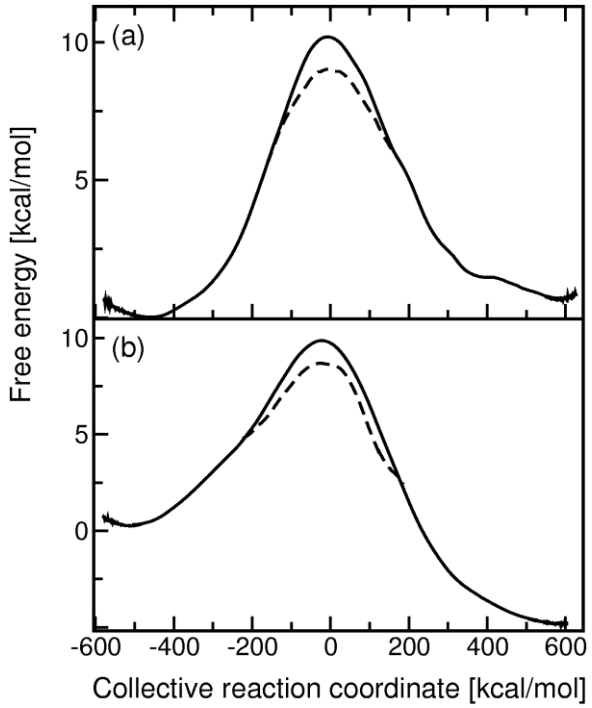


Figure **3-3**: Free energy profiles calculated for (a) the first proton transfer step and (b) the second proton transfer step catalyzed by KSI. These PMF curves were generated along an energy gap collective reaction coordinate using two-state EVB models with a series of mapping potentials. The QCP free energy profiles, which include the nuclear quantum effects of the transferring hydrogen, are indicated by the dashed lines in the barrier region.

We also analyzed the impact of nuclear quantum effects and dynamical barrier recrossings. We found that treatment of the transferring hydrogen nucleus quantum mechanically with the QCP method lowers the classical free energy barrier by 1.2 kcal/mol and 1.3 kcal/mol for the first and second proton transfer steps, respectively. For the first proton transfer step, we also performed the QCP calculations with a quantum treatment of the two hydrogen nuclei involved in the hydrogen bonds between the substrate and Tyr14 and Asp99, as well as the transferring hydrogen nucleus. The quantum treatment of these additional two hydrogen nuclei does not

significantly impact the free energy barrier. We calculated the transmission coefficients κ using the reactive flux approach described above. For the first proton transfer step, κ = 0.26 and 0.28 for the first and second data sets, respectively. For the second proton transfer step, κ = 0.25 and 0.26 for the first and second data sets, respectively. The magnitudes of the transmission coefficients indicate that dynamical barrier recrossings occur but only impact the overall rate constant by a factor of ~3.

Table **3-1**: Rate constants[f] for the two proton transfer reactions catalyzed by KSI.

Rate constants[a]	Experimental[b]	k_{TST}[c]	k_{TST} with QCP[d]	k_{tot}[e]
k_1	1.7×10^5	1.7×10^5	1.3×10^6	3.6×10^5
k_{-1}	5.6×10^5	5.6×10^5	4.3×10^6	1.2×10^6
k_2	2.1×10^5	2.1×10^5	1.8×10^6	4.6×10^5
k_{-2}	40	40	343.2	89.2

a) The EVB parameters used in all calculations for this table were V_{12} = 102.6 kcal/mol and Δ = 3.8 kcal/mol for the first step and V_{12} = 98.2 kcal/mol and D = -14.45 kcal/mol for the second step.
b) The experimental rate constants were obtained from Ref. [2, 33].
c) k_{TST} was calculated with Eq. (4).
d) The QCP method includes the nuclear quantum effects of the transferring H.
e) k_{tot} was calculated with Eq. (3) and includes nuclear quantum effects and barrier recrossings.
f) Rate constants given in s^{-1}.

Table 3-1 provides the calculated rate constants. The experimentally measured values are also provided for comparison. These values indicate that both proton transfer steps are partially rate-limiting in the overall two-step reaction (i.e., the two rate constants are similar). The EVB parameters V_{12} and Δ were fit so that the TST rate constants reproduce the experimentally estimated forward and backward rate constants for each proton transfer step. [2, 33] Since the nuclear quantum effects increase the rate and the dynamical barrier recrossings decrease the rate, the two effects counteract each other, and the inclusion of both effects together does not

significantly impact the rate constant. The results obtained using the alternative rate constant expression given in Ref. [42] are given in Appendix A. Our objective is to generate a qualitatively reasonable potential energy surface that enables us to analyze the geometrical, conformational, and electrostatic changes in the enzyme during the proton transfer reactions. We found that these analyses were virtually indistinguishable using the parameters obtained with the two different rate constant expressions.

B. Geometrical changes in the active site

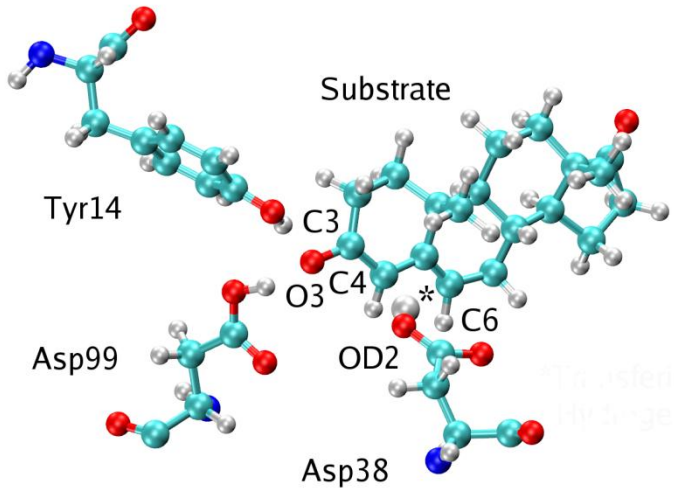


Figure **3-4**: Snapshot of the KSI active site for the Intermediate state shown in Figure 3-1. The snapshot was obtained from the MD simulation of the first proton transfer step with the mapping potential corresponding to λ = 0.95. The proton has transferred from the substrate to Asp38, and residues Tyr14 and Asp99 are hydrogen bonded to the substrate, which is in the dienolate form. The transferring hydrogen is identified with an asterisk.

By construction, the EVB potential energy surface provides a qualitatively reasonable description of the changes in bonding and charge localization during the two proton transfer steps. During the first proton transfer step, the donor carbon atom smoothly transitions from exhibiting characteristics associated with sp^3 to those associated with sp^2 hybridization.

Moreover, the substrate C3-O3 bond increases and the magnitude of the negative charge on the O3 atom increases during the first proton transfer step, which results in the dienolate intermediate form of the substrate. A snapshot of the dienolate intermediate is depicted in Figure 3-4. The reverse process occurs during the second proton transfer step, in which the dienolate intermediate is transformed into the final product. Specifically, the substrate C3-O3 bond decreases and the magnitude of the negative charge on the O3 atom decreases during the second proton transfer step. Concurrently, the acceptor carbon atom smoothly transitions from exhibiting characteristics associated with sp^2 to those associated with sp^3 hybridization.

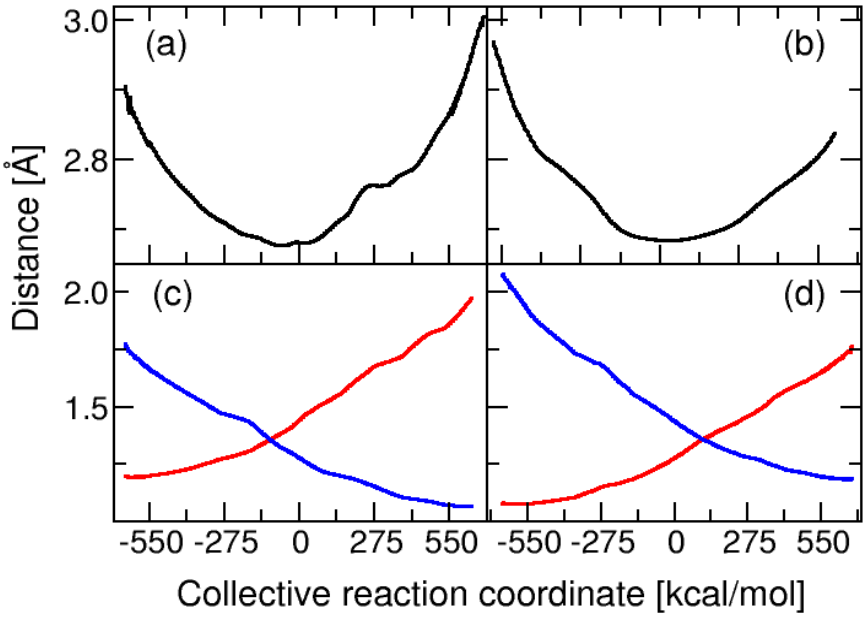


Figure **3-5**: Thermally averaged distances within the proton transfer interface calculated along the collective reaction coordinate for the two proton transfer reactions catalyzed by KSI. For the first proton transfer step, the substrate C4 atom is the donor and the Asp38 OD2 atom is the acceptor. For the second proton transfer step, the Asp38 OD2 atom is the donor and the substrate C6 atom is the acceptor. The donor-acceptor distance is depicted for the first step in (a) and the second step in (b). The donor-hydrogen (red) and acceptor-hydrogen (blue) distances are depicted for the first step in (c) and the second step in (d).

In order to study the conformational changes occurring during these two proton transfer steps, we calculated thermally averaged distances and angles in the active site along the collective reaction coordinate. Distances and angles involving catalytically important residues were sampled with the mapping potentials, and standard binning procedures were used to calculate biased distributions for these quantities. Subsequently, the contributions from the individual mapping potentials were unbiased and combined using the WHAM procedure that was used to generate the PMFs. This procedure resulted in the generation of curves corresponding to the thermally averaged distances and angles over the entire relevant range of the collective reaction coordinate.

Figure 3-5 depicts the thermally averaged distances between the donor, acceptor, and transferring hydrogen atoms for the first and second proton transfer steps. For the first proton transfer step, the donor atom is the substrate C4 atom, and the acceptor atom is the Asp38 carboxylate oxygen atom closest to the substrate C4 atom. For the second proton transfer step, the donor atom is the protonated carboxylate oxygen atom in Asp38, and the acceptor is the substrate C6 atom. The donor-hydrogen distance increases and the acceptor-hydrogen distance decreases for both proton transfer reactions. The donor-acceptor distance exhibits a minimum value at the transition state, which is defined as $\Lambda = 0$. For the first proton transfer step, the donor-acceptor distance starts at 2.87 Å, decreases to 2.66 Å at the transition state, and then increases to 2.97 Å at the end of the first proton transfer step. The transition state donor-acceptor distance of 2.66 Å is in agreement with the value obtained from DFT calculations for the model system described above. For the second proton transfer step, the donor-acceptor distance starts at 2.97 Å, decreases to 2.67 Å at the transition state, and then increases to 2.85 Å. Note that the first and last mapping potentials for each step represent mixtures of the two EVB states dominated by the first and second states, respectively, to prevent the substrate from dissociating.

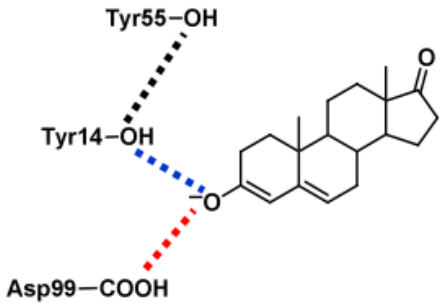


Figure **3-6**: Schematic depiction of the hydrogen-bonding interactions analyzed in Figure 3-7.

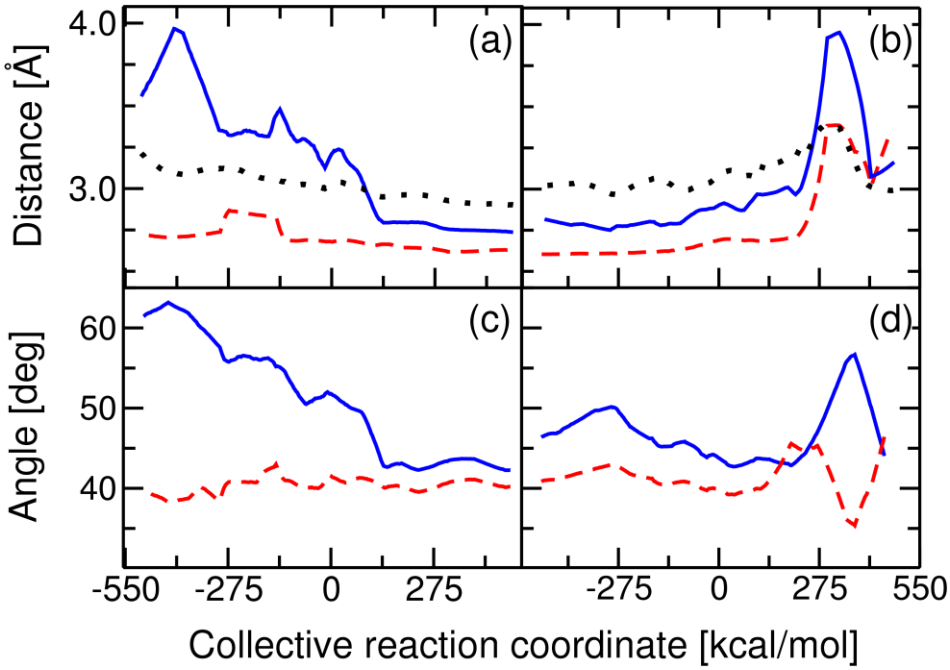


Figure **3-7**: Thermally averaged distances and angles within the active site calculated along the collective reaction coordinate for the two proton transfer reactions catalyzed by KSI. The hydrogen bond donor-acceptor distances between the substrate O3 atom and Tyr14 (solid blue), between the substrate O3 atom and Asp99 (dashed red), and between Tyr14 and Tyr55 (dotted black) are depicted for the first step in (a) and the second step in (b). The angles Tyr14-Asp99-SubstrateO3 (solid blue) and Asp99-Tyr14-SubstrateO3 (dashed red) are depicted for the first step in (c) and the second step in (d). These angles are defined in terms of the heavy atoms involved in the hydrogen bonds between the substrate and both Tyr14 and Asp99.

We also examined the catalytically important hydrogen bonds formed in the active site. In particular, we analyzed the hydrogen bonds between the substrate O3 atom and both Asp99 and Tyr14. These hydrogen bonds stabilize the dienolate intermediate and are thought to enhance the enzymatic rate relative to the non-enzymatic rate. In addition to analyzing these distances, we also analyzed the hydrogen bond between Tyr14 and Tyr55. Analysis of the crystal structure of the Tyr55Phe mutant suggested that Tyr55 may play a minor role in positioning Tyr14 to optimize its hydrogen bond with the substrate. [68] Since the Tyr55Phe mutation decreases k_{cat} by only four-fold, while the Tyr14Phe mutation decreases k_{cat} by $10^{4.7}$-fold, [69] however, Tyr55 is not thought to contribute significantly to catalysis. Figure 3-7 depicts the changes in these hydrogen bonds during the isomerization reaction. Figures 3-5 and 3-7 were generated from the first data set. The analogous figures generated from the second data set are qualitatively similar and are provided in Appendix A.

Figures 3-7a and 3-7b depict the changes in hydrogen bond distances during isomerization. Figure 3-7a depicts the thermally averaged hydrogen bond distances along the collective reaction coordinate for the first proton transfer step. These data indicate the presence of a hydrogen bond between the substrate and Asp99 throughout this proton transfer reaction. In contrast, the hydrogen bond between Tyr14 and the substrate is virtually absent for the first part of the reaction but is formed by the end of the reaction. Figure 3-6b depicts the thermally averaged hydrogen bond distances for the second proton transfer step. In this step, the hydrogen bond distances between the substrate and both Asp99 and Tyr14 remain relatively constant during the first part of the reaction but increase in the later part of the reaction. This increase and the numerical fluctuations at the end of the second step are due to partial dissociation of the product. The Tyr14-Tyr55 hydrogen bond distance remains relatively constant and independent of the hydrogen bond distance between Tyr14 and the substrate for both steps.

Note that the hydrogen bond distances between the substrate and both Asp99 and Tyr14 in the Intermediate state are larger than the values obtained from the crystal structure [12] and NMR chemical shift experiments [8, 64, 65] discussed above. Moreover, the NMR chemical shift experiments suggest that the hydrogen bond with the substrate is shorter for Tyr14 than for Asp99, and this trend is not observed in the simulations. These quantitative discrepancies in average distances may be due to limitations of the forcefield in describing hydrogen bonding. Furthermore, experiments have shown that k_{cat} is diminished more by the Tyr14Phe mutation than by the Asp99Ala mutation.[3, 4, 64, 65] The later development of the hydrogen bond between the substrate and Tyr14 observed in our simulations is not inconsistent with these experimental data. One possibility is that the catalytic rate is more sensitive to the mutation of Tyr14 than Asp99 because the hydrogen bond forms during the proton transfer reaction and thus impacts the free energy barrier more.

To monitor the changes in the orientation of the substrate during the reaction, we analyzed the thermally averaged angles for the heavy atoms involved in the hydrogen bonds between the substrate O3 atom and both Asp99 and Tyr14. Figures 3-7c and 3-7d depict these thermally averaged angles for the first and second proton transfer steps, respectively. For the first proton transfer step, the Tyr14-Asp99-O3(Substrate) angle decreases during the initial part of the reaction, and the Asp99-Tyr14-O3(Substrate) angle only fluctuates slightly during the initial part of the reaction. For the later part of the first step, these angles remain relatively constant. For the second proton transfer step, these angles remain relatively constant for the initial part of the reaction but then fluctuate significantly in the later part of the reaction. These structural rearrangements in the later part of the second proton transfer step are due to partial dissociation of the product.

This analysis of the thermally averaged distances and angles indicates that the enzyme and substrate undergo conformational changes to facilitate the proton transfer reactions. The

results are consistent with the hydrogen bonding of both Asp99 and Tyr14 to the substrate. The hydrogen bond between the substrate and Asp99 appears to be present from the beginning of the first proton transfer step. In contrast, the hydrogen bond between the substrate and Tyr14 appears to be absent at the beginning of the first proton transfer step and to form nearly concurrently with the formation of the transition state for this step. Thus, the active site appears to be pre-organized to bind the substrate, but relatively small conformational changes of both the enzyme and substrate occur during the proton transfer reactions to strengthen the hydrogen bonds that stabilize the intermediate.

C. Conformational and electrostatic changes in enzyme

In addition to analyzing the thermally averaged distances, we also analyzed the thermally averaged structures of the entire enzyme-substrate complex at various stages of the proton transfer reactions. This analysis provides insight into the overall structural changes in the protein occurring during the catalytic reaction. We generated these thermally averaged structures for the reactant state, transition state, and product state of both proton transfer reactions. These structures were generated from MD trajectories obtained with mapping potentials corresponding to λ = 0.05, 0.50, and 0.95 for each proton transfer step. The thermally averaged structure for each mapping potential was obtained by minimizing the RMSD of the protein backbone with respect to a reference structure for the thousands of configurations sampled with more than 1 ns of MD using the g_confrms utility in GROMACS. [70] Subsequently, the configurations were weighted and averaged in a manner that produced the thermally averaged structure for the EVB potential. Since this process may lead to non-physical bond distances and angles in the side chains, we used these average structures to analyze only the global changes in protein structure. The more specific changes were analyzed from the thermally averaged distances and angles described in Section 3-3B.

Based on these thermally averaged structures, we identified conformational changes of the protein backbone that are associated with the proton transfer steps. The loop regions that exhibit significant conformational changes during the proton transfer reactions are illustrated in Figure 3-8. The Tyr88-Lys92 and Phe103-Gly107 loops exhibit the greatest structural rearrangements during the proton transfer reactions. We calculated RMSD values between these structures using the VMD program. [71] The average RMSD between these six thermally averaged

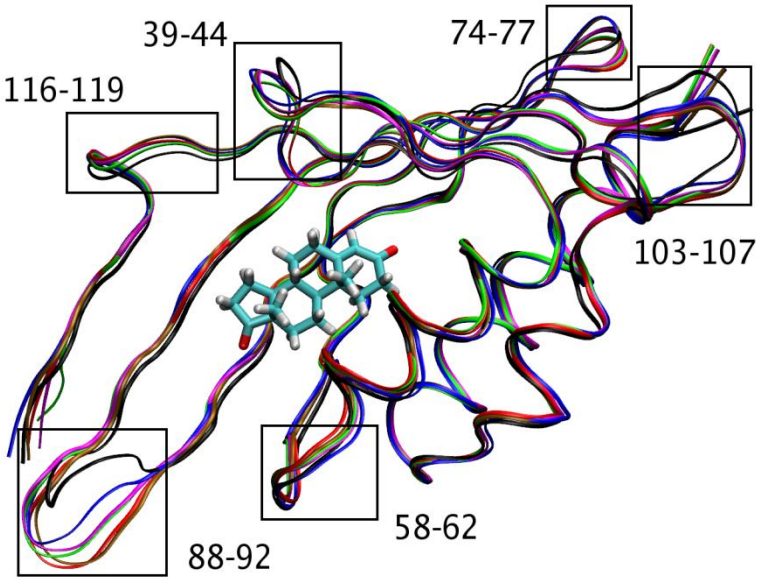


Figure **3-8**: Thermally averaged structures of KSI along the reaction pathway for both proton transfer reactions. The reactant state (black), transition state (red) and product state (green) for the first proton transfer step and the reactant state (brown), transition state (magenta) and product state (blue) for the second proton transfer step are presented. The loop regions exhibiting significant structural changes are labeled.

structures and an average structure for the backbone of the Tyr88-Lys92 and Phe103-Gly107 loops are 1.43 Å and 1.02 Å, respectively. The average RMSD for the backbone of the entire enzyme is 0.53 Å. Thus, the structural rearrangements in these two loop regions are significantly larger than those in the rest of the enzyme during the two proton transfer reactions. In addition, our analysis of the thermally averaged structures suggests an absence of large structural changes

in the active site. The average RMSD between these six thermally averaged structures and an average structure for the backbone residues within 5 Å of the substrate, Tyr14, and Asp99 residues is 0.21 Å. This relatively low average RMSD value is consistent with experimental studies. [24, 26, 72, 73]

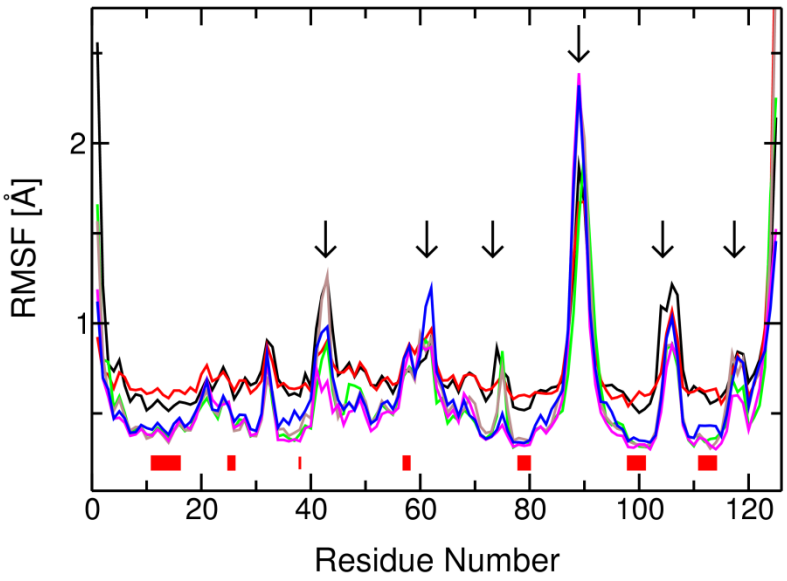


Figure **3-9**: RMSF values for the C_α atoms of the protein backbone corresponding to the reactant state (black), transition state (red), and product state (green) for the first proton transfer step and the reactant state (brown), transition state (magenta), and product state (blue) for the second proton transfer step. The peaks indicated with arrows correspond to the loop regions labeled in Figure 3-8. The red blocks under the data identify the active site residues as defined in the text.

We also calculated the root mean square fluctuations (RMSF) for all C_α atoms of the protein backbone for the reactant state, transition state, and product state of both proton transfer reactions. Analogous to the generation of the thermally averaged structures, these RMSF values were generated from MD trajectories obtained with mapping potentials corresponding to $\lambda = 0.05$, 0.50, and 0.95 for each proton transfer step. The results are depicted in Figure 3-9. The peaks indicated by arrows correspond to the loop regions identified in Figure 3-8 as exhibiting significant conformational changes during the proton transfer reactions based on the analysis of the thermally averaged structures. The more mobile loops are found to exhibit greater conformational changes during the proton transfer reactions. The red blocks underneath the data

indicate the backbone residues within 5 Å of the substrate, Tyr14, and Asp99 residues. Note that the reactant state and transition state for the first proton transfer step have greater RMSF values for these active site residues. These data suggest that the active site is more mobile during the first proton transfer step than for the second proton transfer step. As mentioned above, however, the average RMSD for these active site residues among all six of the thermally averaged structures is only 0.21 Å. Thus, although the fluctuations within the active site are greater during the first proton transfer step, the average structure of the active site does not change significantly over the two proton transfer steps. The greater mobility of the active site residues in the initial stage of the chemical process may facilitate the strengthening of the hydrogen bonds between the substrate and the enzyme, as well as other relatively minor conformational changes that accompany the proton transfer reactions.

Figure 3-10 illustrates the conformational change that occurs in the Asp38 side chain to accommodate proton transfer to the C6 atom of the substrate after accepting a proton from the C4 atom of the substrate. The dihedral angle comprised of the C4 carbon on the substrate, the protonated carboxylate oxygen (OD2) on Asp38, the neighboring carbon atom on Asp38, and the other carboxylate oxygen (OD1) on Asp38 changes by 117°. These structures were obtained from the thermally averaged structure for the product of the first proton transfer step (i.e., the mapping potential with λ = 0.95 for the first step) and the reactant of the second proton transfer step (i.e., the mapping potential with λ = 0.05 for the second step). Although both structures correspond to the Intermediate state defined in Figure 3-1, the product of the first step has 5% contribution from the overall Reactant state, while the reactant of the second step has 5% contribution from the overall Product state. Thus, the differences between the structures shown in Figure 3-10 arise from the mixed nature of the states in the EVB models describing the first and second proton transfer steps. Nevertheless, this figure provides a qualitative illustration of the motion of Asp38 that occurs between the two proton transfer steps.

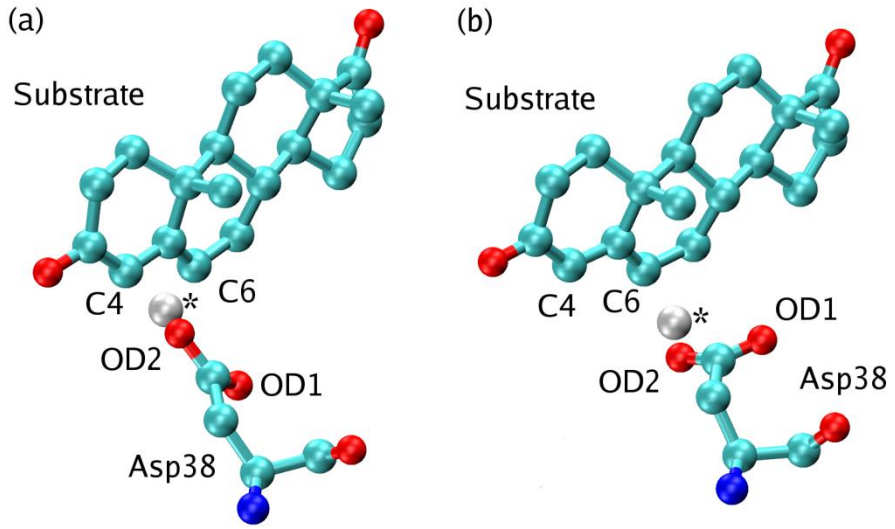


Figure **3-10**: Qualitative illustration of the conformational changes of Asp38 occurring between the first and second proton transfer steps catalyzed by KSI. The thermally averaged structures of the substrate and Asp38 are depicted for (a) the product of the first step and (b) the reactant of the second step. Although both structures correspond to the Intermediate state defined in Figure 3-1, each structure arises from a mixture of states in the two-state EVB models describing the two proton transfer steps. The transferring hydrogen is identified with an asterisk.

Furthermore, we calculated the electrostatic potential on the solvent accessible surface for the thermally averaged Reactant, Intermediate, and Product for the overall reaction pathway shown in Figure 3-1. This Intermediate structure is the thermally averaged product structure for the first proton transfer step. The electrostatic potentials were calculated by solving the non-linear Poisson-Boltzmann equation numerically at 298 K using the Adaptive Poisson-Boltzmann Solver (APBS) utility [74] in the Pymol program. [75] The PDB2PQR utility [76] was used to generate the input files for this calculation, and the charges for the residues were obtained from the AMBER99 forcefield. The dielectric constant was chosen to be 6 inside the cavity and 80 outside the cavity determined by the solvent-accessible surface. We found that the results did not change qualitatively when the dielectric constant inside the cavity was varied between 2 and 18. Note that the substrate was not included in these calculations and that the Intermediate structure has one more proton than the Reactant and Product structures.

The results of these electrostatic potential calculations are presented in Figure 3-11. The active site is comprised of predominantly hydrophobic residues, and the dielectric constant near Tyr14 has been estimated to be ~18 from pK_a measurements. [77] Figure 3-11 depicts the electrostatic potential on the solvent-accessible surface due to all of the charges in the system, and the red and blue colors depict the relative potentials for each particular system. Figures 3-11a, 3-11b, and 3-11c illustrate that the electrostatic potential in the active site is negative for the Reactant and Product but positive for the Intermediate. This trend is due mainly to the protonation of Asp38 in the Intermediate. The negatively charged dienolate form of the substrate will interact favorably with this positive electrostatic potential in the Intermediate structure. Figures 3-11d, 3-11e, and 3-11f depict the other side of the protein and identify regions distal from the active site exhibiting significant changes in the electrostatic potential during the proton transfer reactions. These regions correspond to some of the loop regions exhibiting structural changes in Figure 3-8. This analysis suggests an association between changes in the structure and electrostatic potential of these loop regions with the catalyzed proton transfer reactions, but it does not distinguish between cause and effect. Further insight into the possible catalytic role of these loop regions could be obtained by mutation experiments.

To further analyze the interaction between the substrate and the enzyme during the proton transfer steps, we calculated the average van der Waals and electrostatic interaction energies between the substrate ligand and the enzyme for the Reactant, first transition state, Intermediate, second transition state, and Product. These quantities were calculated by averaging the interaction energies over configurations sampled with more than 1 ns MD for mapping potentials with λ = 0.05, 0.5, and 0.95 for the first proton transfer step and λ = 0.5 and 0.95 for the second proton transfer step. The results are given in Table 3-2. The van der Waals interaction energy remains virtually constant during both proton transfer reactions, suggesting that no major structural rearrangements occur in the active site. The electrostatic interaction energy is similar

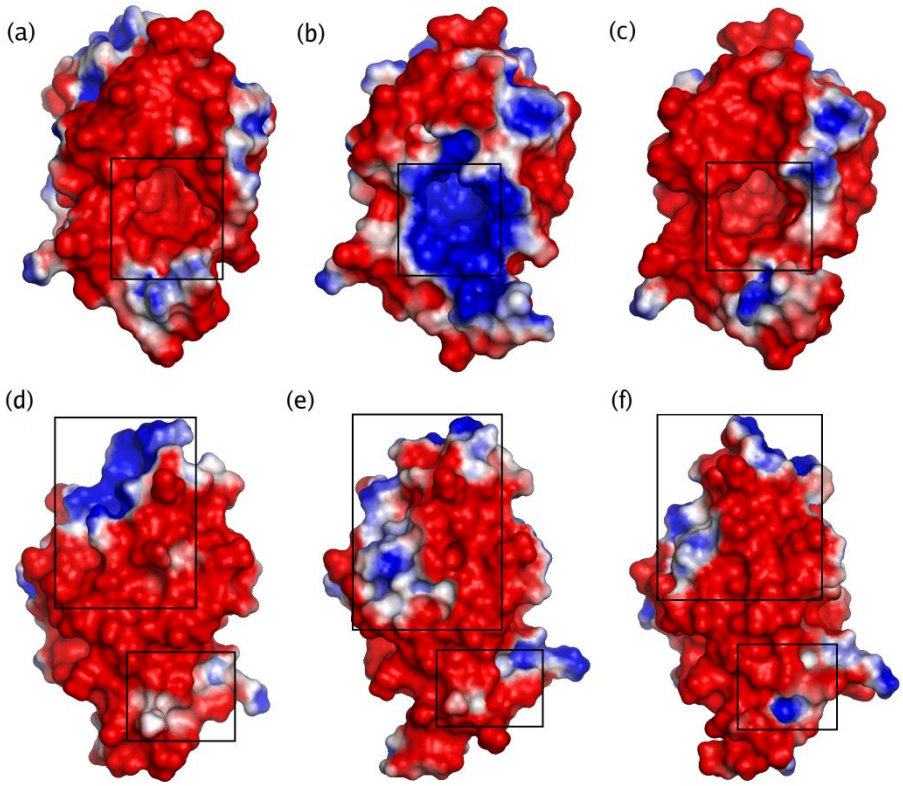


Figure **3-11**: Illustration of the changes in the electrostatic potential on the solvent accessible surface in KSI during the two proton transfer reactions. The electrostatic potential is depicted for the thermally averaged structures of the Reactant in (a) and (d), the Intermediate in (b) and (e), and the Product in (c) and (f) for the overall reaction pathway shown in Figure 3-1. The active site is identified in (a), (b), and (c). The regions exhibiting significant changes in electrostatic potential on the side of the protein opposite the active site are identified in (d), (e), and (f). The red regions correspond to negative potential (-1) and the blue regions correspond to positive potential (+1). Since the substrate is not included in the electrostatic potential calculations, the Intermediate includes an extra positive charge in the active site from the transferred proton.

for the Reactant and Product but is significantly stronger for the Intermediate. This stronger electrostatic interaction energy for the Intermediate is due to the transfer of the positively charged proton to Asp38 and the resulting redistribution of electronic charge in the substrate, as well as strengthening of the hydrogen bonds between the substrate oxygen and Asp99 and Tyr14. Note that Table 3-2 provides the total electrostatic interaction between the ligand and the entire enzyme, and a significant fraction of this total value is comprised of contributions from a large

number of residues far from the active site due to the long-range nature of these electrostatic interactions.

Table **3-2**: Electrostatic and van der Waals interaction energies[b] between the substrate and KSI for different states along the reaction pathway.

State	Electrostatic	Van der Waals	Hydrogen bond electrostatic[a]
Reactant	-29	-35	-6.8
Transition State 1	-58	-31	-11
Intermediate	-99	-32	-16
Transition State 2	-57	-29	-12
Product	-27	-33	-6.1

a) The hydrogen bond interaction energies were calculated between the carbonyl group at the C3 position of the substrate and the OH atoms on Tyr14 and Asp99 involved in hydrogen bonding with the substrate.

d) Interaction energies given in kcal/mol.

The specific electrostatic interaction energies between the carbonyl group at the C3 position of the substrate and the OH groups of Asp99 and Tyr14 are also given in Table 3-2. The electrostatic interaction energy for each hydrogen bond is –8 kcal/mol in the Intermediate state, and the van der Waals interaction energy for each hydrogen bond in the thermally averaged Intermediate state is ~2 kcal/mol, leading to an approximate interaction energy of –6 kcal/mol for each hydrogen bond. These values provide evidence for a reasonably strong hydrogen-bonding interaction in the Intermediate state. [64] Moreover, the strengthening of the electrostatic interaction in the Intermediate relative to the Reactant for each hydrogen bond is 4.5 kcal/mol, which is consistent with the value of 6 kcal/mol obtained in previous EVB calculations. [18] In future studies, a more quantitative analysis of the geometrical changes, electrostatic interactions, and hydrogen bonding will be performed with a quantum mechanical/molecular mechanical

(QM/MM) approach in which the substrate and neighboring residues are treated quantum mechanically with a method such as DFT.

3-4. Conclusions

In this chapter, we performed hybrid quantum/classical MD simulations of the two proton transfer reactions catalyzed by KSI. We generated the free energy profiles for both reactions along a collective reaction coordinate. We also analyzed changes in thermally averaged distances and angles in the active site along the collective reaction coordinate. For both proton transfer reactions, the donor-acceptor distance decreases to ~2.66 Å at the transition state and then increases again.

In addition, we examined the hydrogen bonding interactions within the active site of KSI during the two proton transfer reactions. Our results are consistent with the hydrogen bonding of both Asp99 and Tyr14 to the substrate and provide additional information about the formation of these hydrogen bonds. Specifically, our analysis suggests that a hydrogen bond between Asp99 and the substrate is present from the beginning of the first proton transfer step, whereas the hydrogen bond between Tyr14 and the substrate is virtually absent in the first part of this step but forms nearly concurrently with the formation of the transition state. Both of these hydrogen bonds are present throughout the second proton transfer step until partial dissociation of the product. We also found that the hydrogen bond between Tyr14 and Tyr55 is present throughout both proton transfer steps. This observation is consistent with the proposal that Tyr55 plays a minor role in positioning Tyr14 to optimize its hydrogen bond with the substrate.[68] In addition, we found that the active site residues are more mobile during the first proton transfer step than during the second proton transfer step. Overall, these results suggest that relatively small conformational changes of both the enzyme and substrate strengthen the hydrogen bonds that stabilize the intermediate, thereby facilitating the proton transfer reactions.

Furthermore, we analyzed the conformational and electrostatic changes occurring throughout the protein during the proton transfer reactions. The van der Waals interaction energy between the substrate and the enzyme remains virtually constant along the reaction pathway, but the electrostatic interaction energy is significantly stronger for the Intermediate than for the Reactant and Product. The mobility of the active site residues is lower for the Intermediate than for the Reactant. Moreover, we identified mobile loop regions distal to the active site that exhibit significant structural rearrangements and, in some cases, qualitative changes in the electrostatic potential during the two proton transfer reactions. These observations imply that the conformational and electrostatic changes associated with the catalyzed proton transfer reactions are not limited to the active site but rather extend throughout the entire enzyme. The possible catalytic role of these distal regions could be probed with mutation experiments.

3-5. References

(1) Pollack, R. M. *Bioorg. Chem.* **2004**, *32*, 341-353.

(2) Hawkinson, D. C.; Eames, T. C. M.; Pollack, R. M. *Biochemistry* **1991**, *30*, 10849-10858.

(3) Wu, Z. R.; Ebrahimian, S.; Zawrotny, M. E.; Thornburg, L. D.; Perez-Alvarado, G. C.; Brothers, P.; Pollack, R. M.; Summers, M. F. *Science* **1997**, *276*, 415-418.

(4) Choi, G.; Ha, N.-C.; Kim, S. W.; Kim, D.-H.; Park, S.; Oh, B.-H.; Choi, K. Y. *Biochemistry* **2000**, *39*, 903-909.

(5) Brothers, P. N.; Blotny, G.; Qi, L.; Pollack, R. M. *Biochemistry* **1995**, *34*, 15453-15458.

(6) Qi, L.; Pollack, R. M. *Biochemistry* **1998**, *37*, 6760-6766.

(7) Thornburg, L. D.; Henot, F.; Bash, D. P.; Hawkinson, D. C.; Bartel, S. D.; Pollack, R. M. *Biochemistry* **1998**, *37*, 10499-10506.

(8) Zhao, Q.; Abeygunawardana, C.; Gittis, A. G.; Mildvan, A. S. *Biochemistry* **1997**, *36*, 14616-14626.

(9) Kim, S. W.; Cha, S.-S.; Cho, H.-S.; Kim, J.-S.; Ha, N.-C.; Cho, M.-J.; Joo, S.; Kim, K. K.; Choi, K. Y.; Oh, B.-H. *Biochemistry* **1997**, *36*, 14030-14036.

(10) Ha, N.-C.; Kim, M.-S.; Lee, W.; Choi, K. Y.; Oh, B.-H. *J. Biol. Chem.* **2000**, *275*, 41100-41106.

(11) Massiah, M. A.; Abeygunawardana, C.; Gittis, A. G.; Mildvan, A. S. *Biochemistry* **1998**, *37*, 14701-14712.

(12) Cho, H.-S.; Ha, N.-C.; Choi, G.; Kim, H.-J.; Lee, D.; Oh, K. S.; Kim, K. S.; Lee, W.; Choi, K. Y.; Oh, B.-H. *J. Biol. Chem.* **1999**, *274*, 32863-32868.

(13) Oh, K. S.; Cha, S.-S.; Kim, D.-H.; Cho, H.-S.; Ha, N.-C.; Choi, G.; Lee, J. Y.; Tarakeshwar, P.; Son, H. S.; Choi, K. Y.; Oh, B.-H.; Kim, K. S. *Biochemistry* **2000**, *39*, 13891-13896.

(14) Park, H.; Merz, K. M., Jr. *J. Am. Chem. Soc.* **2003**, *125*, 901-911.

(15) Mazumder, D.; Kahn, K.; Bruice, T. C. *J. Am. Chem. Soc.* **2003**, *125*, 7553-7561.

(16) Kim, K. S.; Oh, K. S.; Lee, J. Y. *Proc. Nat. Acad. Sc. USA* **2000**, *97*, 6373-6378.

(17) Pan, Y.; McAllister, M. A. *Journal of Molecular Structure (Theochem)* **2000**, *504*, 29-33.

(18) Feierberg, I.; Aqvist, J. *Biochemistry* **2002**, *41*, 15728-15735.

(19) Feierberg, I.; Aqvist, J. *Theoretical Chemical Accounts* **2002**, *108*, 71-84.

(20) Warshel, A.; Sharma, P. K.; Chu, Z. T.; Aqvist, J. *Biochemistry* **2007**, *46*, 1466-1476.

(21) Wilde, T. C.; Blotny, G.; Pollack, R. M. *J. Am. Chem. Soc.* **2008**, *130*, 6577-6585.

(22) Kraut, D. A.; Sigala, P. A.; Pybus, B.; Liu, C. W.; Ringe, D.; Petsko, G. A.; Herschlag, D. *PLoS Biology* **2006**, *4*, 501-519.

(23) Sigala, P. A.; Fafarman, A. T.; Bogard, P. E.; Boxer, S. G.; Herschlag, D. *J. Am. Chem. Soc.* **2007**, *129*, 12104-12105.

(24) Sigala, P. A.; Kraut, D. A.; Caaveiro, J. M. M.; Pybus, B.; Ruben, E. A.; Ringe, D.; Petsko, G. A.; Herschlag, D. *J. Am. Chem. Soc.* **2008**, *130*, 13696-13708.

(25) Kraut, D. A.; Churchill, M. J.; Dawson, P. E.; Herschlag, D. *ACS Chemical Biology* **2009**, *4*, 269-273.

(26) Sigala, P. A.; Caaveiro, J. M. M.; Ringe, D.; Petsko, G. A.; Herschlag, D. *Biochemistry (in press)*.

(27) Xue, L.; Talalay, P.; Mildvan, A. S. *Biochemistry* **1990**, *29*, 7491-7500.

(28) Warshel, A. *Computer Modeling of Chemical Reactions in Enzymes and Solutions*; John Wiley & Sons, Inc.: New York, 1991.

(29) Batzold, F. H.; Benson, A. M.; Covey, D. F.; Robinson, C. H.; Talalay, P. *Advances in Enzyme Regulation* **1976**, *14*, 243-267.

(30) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. *J. Phys. Chem.* **1993**, *97*, 10269-10280.

(31) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *Journal of the American Chemical Society* **1995**, *117*, 5179-5197.

(32) Wang, J.; Cieplak, P.; Kollman, P. A. *Journal of Computational Chemistry* **2000**, *21*, 1049-1074.

(33) Hawkinson, D. C.; Pollack, R. M.; Ambulos, N. P., Jr. *Biochemistry* **1994**, *33*, 12172-12183.

(34) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.;

Wong, M. W.; Gonzalez, C.; Pople, J. A.; revision C.03 ed.; Gaussian, Inc.: Pittsburgh, PA, 2003.

(35) Ferrenberg, A. M.; Swendsen, R. H. *Phys. Rev. Lett.* **1988**, *61*, 2635-2638.

(36) Ferrenberg, A. M.; Swendsen, R. H. *Phys. Rev. Lett.* **1989**, *63*, 1195-1198.

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

(38) Roux, B. *Comput. Phys. Commun.* **1995**, *91*, 275-282.

(39) Kastner, J.; Thiel, W. *J. Chem. Phys.* **2005**, *123*, 144104.

(40) Kastner, J.; Thiel, W. *J. Chem. Phys.* **2006**, *124*, 234106.

(41) Chakravorty, D. K.; Kumarasiri, M.; Soudackov, A. V.; Hammes-Schiffer, S. *Journal of Chemical Theory and Computation* **2008**, *4*, 1974-1980.

(42) Watney, J. B.; Soudackov, A. V.; Wong, K. F.; Hammes-Schiffer, S. *Chem. Phys. Lett.* **2006**, *418*, 268-271.

(43) Carter, E. A.; Ciccotti, G.; Hynes, J. T.; Kapral, R. *Chem. Phys. Lett.* **1989**, *156*, 472-477.

(44) Schenter, G. K.; Garrett, B. C.; Truhlar, D. G. *J. Chem. Phys.* **2003**, *119*, 5828-5833.

(45) Neria, E.; Karplus, M. *J. Chem. Phys.* **1996**, *105*, 10812-10818.

(46) Bennett, C. H. *Algorithms for Chemical Computation*; American Chemical Society: Washington, DC, 1997.

(47) Keck, J. C. *J. Chem. Phys.* **1960**, *32*, 1035-50.

(48) Anderson, J. B. *J. Chem. Phys.* **1973**, *58*, 4684-92.

(49) Chandler, D. In *Classical and Quantum Dynamics in Condensed Phase Simulations*; Berne, B. J., Ciccotti, G., Coker, D. F., Eds.; World Scientific: Singapore, 1998, p 3-23.

(50) Billeter, S. R.; Webb, S. P.; Agarwal, P. K.; Iordanov, T.; Hammes-Schiffer, S. *J. Am. Chem. Soc.* **2001**, *123*, 11262-272.

(51) Billeter, S. R.; Webb, S. P.; Iordanov, T.; Agarwal, P. K.; Hammes-Schiffer, S. *J. Chem. Phys.* **2001**, *114*, 6925-6936.

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

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

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

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

(56) Xiang, Z.; Honig, B. *J. Mol. Biol.* **2001**, *311*, 421-430.

(57) Xiang, J. Z.; Honig, B.; Columbia University & Howard Hughes Medical Institute: New York, 2002.

(58) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157-1174.

(59) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *Journal of Chemical Physics* **1983**, *79*, 926-935.

(60) Mahoney, M. W.; Jorgensen, W. L. *Journal of Chemical Physics* **2000**, *112*, 8910-8922.

(61) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.

(62) Darden, T.; York, D.; Pedersen, L. *Journal of Chemical Physics* **1993**, *98*, 10089-10092.

(63) Melchionna, S.; Cozzini, S. INFM Udr SISSA, Trieste, Italy, 2001.

(64) Mildvan, A. S.; Massiah, M. A.; Harris, T. K.; Marks, G. T.; Harrison, D. H. T.; Viragh, C.; Reddy, P. M.; Kovach, I. M. *J. Mol. Struct.* **2002**, *615*, 163-175.

(65) Zhao, Q.; Abeygunawardana, C.; Talalay, P.; Mildvan, A. S. *Proc. Nat. Acad. Sci. USA* **1996**, *93*, 8220-8224.

(66) Nose, S. *Molecular Physics* **1984**, *52*, 255-68.

(67) Hoover, W. G. *Physical Review A* **1985**, *31*, 1695-7.

(68) Kim, D.-H.; Jang, D. S.; Nam, G. H.; Choi, G.; Kim, J.-S.; Ha, N.-C.; Kim, M.-S.; Oh, B.-H.; Choi, K. Y. *Biochemistry* **2000**, *39*, 4581-4589.

(69) Kuliopulos, A.; Mildvan, A. S.; Shortle, D.; Talalay, P. *Biochemistry* **1989**, *28*, 149-159.

(70) van der Spoel, D.; Lindahl, E.; Hess, B.; van Buuren, A. R.; Apol, E.; Meulenhoff, P. J.; Tieleman, D. P.; Sijbers, A. L. T. M.; Feenstra, K. A.; van Drunen, R.; Berendsen, H. J. C.; 3.2 ed. 2004.

(71) Humphrey, W.; Dalke, A.; Schulten, K. *Journal of Molecular Graphics* **1996**, *14*, 33-38.

(72) Zhao, Q.; Li, Y.-K.; Mildvan, A. S.; Talalay, P. *Biochemistry* **1995**, *34*, 6562-6572.

(73) Zhao, Q.; Abeygunawardana, C.; Mildvan, A. S. *Biochemistry* **1996**, *35*, 1525-1532.

(74) Holst, M.; Saied, F. *J. Comput. Chem.* **1993**, *14*, 105-113.

(75) DeLano, W. L.; DeLano Scientific: San Carlos, CA, 2002.

(76) Dolinsky, T. J.; Czodrowski, P.; Li, H.; Nielsen, J. E.; Jensen, J. H.; Klebe, G.; Baker, N. A. *Nucleic Acids Res.* **2007**, *35*, W522-5.

(77) Li, Y.-K.; Kuliopulos, A.; Mildvan, A. S.; Talalay, P. *Biochemistry* **1993**, *32*, 1816-1824.

Chapter 4

Impact of Mutation on Proton Transfer Reaction in Ketosteroid Isomerase: Insights from Molecular Dynamics Simulations

Reproduced in part with permission from Dhruva K. Chakravorty, Alexander V. Soudackov, and Sharon Hammes-Schiffer. Submitted for publication in the *Journal of the American Chemical Society*. Unpublished work.

4-1. Introduction

The enzyme Δ^5-3-Ketosteroid isomerase (KSI) catalyzes the isomerization of 3-oxo-Δ^5-steroids to their Δ^4-conjugated isomers by the two-step proton transfer mechanism depicted in Figure 4-1. KSI enzymes from *Pseudomonas putida* and *Commamonas testosteroni* bacteria have been studied extensively. Both experimental[1-9] and computational[10-19] approaches indicate that the catalytic efficiency of KSI is strongly influenced by electrostatic stabilization and hydrogen bonding. As shown in Figure 4-1, this enzyme mechanism involves a dienolate intermediate that is stabilized by hydrogen bonds with Tyr14 and Asp99 in the active site. (In the present paper, the residues are numbered according to *Commamonas testosteroni* KSI.) The catalytic relevance of these residues has been investigated experimentally by kinetic studies on mutant forms of KSI.[2, 3, 20-34] Mutation of these residues has been shown to significantly reduce the catalytic efficiency of KSI. Specifically, mutation of Asp99 to Leu decreases k_{cat} by $10^{2.1}$-fold, mutation of Tyr14 to Phe decreases k_{cat} by $10^{3.4}$-fold, and the corresponding double mutation decreases k_{cat}

by $10^{4.6}$-fold in *Pseudomonas putida* KSI.[3] These trends were also observed experimentally for related single mutants of *Commamonas testosteroni* KSI.[2, 20]

The objective of this paper is to use molecular dynamics (MD) simulations to understand the impact of mutating Tyr14 and Asp99 on the two proton transfer reactions catalyzed by KSI. We perform calculations on both the D99L and Y14F single mutants and the Y14F/D99L double mutant of KSI from *Commamonas testosteroni* bacteria with the Δ^5-androstene-3,17-dione (5-AND) substrate. As in our previous study of wild-type (WT) KSI,[19] the simulations are based on a potential energy surface described by an empirical valence bond (EVB) model.[35] We calculate the free energy profiles along the collective reaction coordinate for both proton transfer steps of all three mutants. We also analyze the hydrogen-bonding patterns, van der Waals and electrostatic interactions, and conformational changes during the proton transfer reactions. Our results are consistent with the experimental data on the mutants and provide insight into the catalytic roles of these hydrogen-bonding residues.

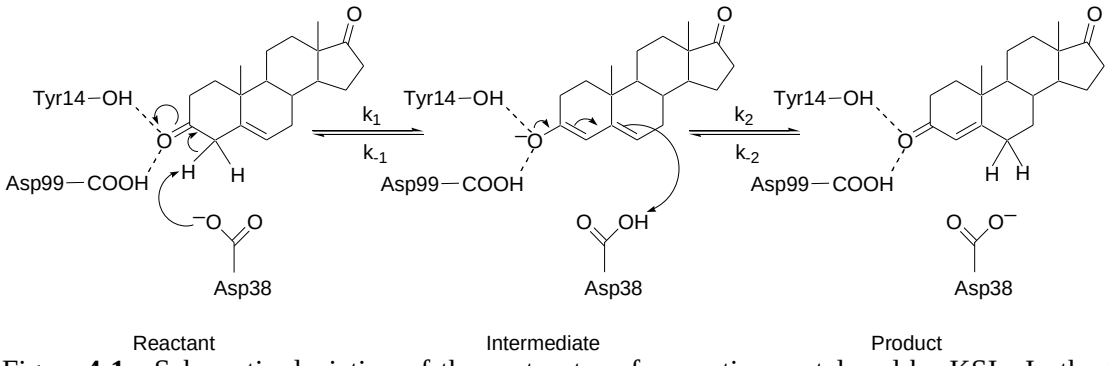


Figure **4-1**: Schematic depiction of the proton transfer reactions catalyzed by KSI. In the first step, the proton transfers from the C4 atom of the substrate to the Asp38 residue. In the second step, the proton transfers from the Asp38 residue to the C6 atom of the substrate. The reactant, intermediate, and product states of the overall reaction are labeled.

This chapter is organized into the following sections. The computational methods used to describe the two proton transfer reactions are described in Section 4-2. The results from the simulations are presented in Section 4-3. The first part of this section analyzes the free energy profiles and relative rate constants, while the second part analyzes the hydrogen bonding patterns, electrostatic interactions, and conformational changes accompanying the proton transfer reactions. Our conclusions are presented in Section 4-4.

4-2.Methods

A. Theory

We used a two-state empirical valence bond (EVB) potential[35] to model the electronic potential energy surface for each proton transfer reaction shown in Figure 4-1. The two EVB states correspond to the Reactant and the Intermediate for the first step and to the Intermediate and the Product for the second step. These two proton transfer reactions are thought to be sequential with negligible coupling between the Reactant and Product states, thereby enabling us to model them with two separate two-state EVB potentials. The ground state electronic potential energy surface for each proton transfer reaction is obtained from the lowest energy eigenvalue of the corresponding 2×2 EVB Hamiltonian. The details of this EVB potential are provided elsewhere.[19]

For each two-state EVB model, the diagonal matrix elements V_{11} and V_{22} of the EVB Hamiltonian were represented by a modified AMBER99 forcefield with a constant energy shift Δ included in V_{22}.[19, 36, 37] The constant energy shift Δ and the coupling V_{12} between the two states were determined by fitting the calculated rates for WT KSI to the experimentally determined rate constants for the forward and backward reactions.[38, 39] Based on the WT KSI data from the present paper, these parameters are Δ = 1.8 kcal/mol and V_{12} = 99.5 kcal/mol for the first step and Δ = –14.5 kcal/mol and V_{12} = 93.1 kcal/mol for the second step. These parameters are similar to

the values obtained from previous MD simulations of WT KSI,[19] and the minor differences are due to typical statistical errors associated with MD. Moreover, these parameters are assumed to be the same for the WT and mutant KSI simulations. Thus, we did not fit any parameters to experimental data for the mutants.

The potential of mean force (PMF) for each proton transfer reaction was calculated as a function of a collective reaction coordinate defined as

$$\Lambda = V_{11} - V_{22} \quad . \tag{1}$$

This energy gap reaction coordinate includes motions of the entire solvated enzyme. The reaction was driven from the reactant to the product state using a series of mapping potentials defined as[35]

$$V_{\text{map}}^{\lambda} = (1-\lambda)V_{11} + \lambda V_{22}, \tag{2}$$

where the mapping parameter λ is varied between zero and unity. The PMF was obtained by propagating a series of classical MD trajectories according to these mapping potentials and combining the results to obtain the complete PMF for the unbiased EVB potential using the weighted histogram analysis method (WHAM)[40-43] or umbrella integration.[44-46]

Within the framework of transition state theory (TST), the rate constant is given by

$$k = \kappa k_{\text{TST}}, \tag{3}$$

where k_{TST} is the TST rate constant and κ is the transmission coefficient accounting for dynamical recrossings of the dividing surface. An expression for the TST rate constant in terms of the PMF along a general reaction coordinate has been derived.[47-49] In this study, we used the simpler form

$$k_{\text{TST}} = \frac{1}{\beta h} e^{-\beta \Delta G^{\ddagger}} \quad , \tag{4}$$

where $\beta = 1/k_B T$, k_B is the Boltzmann constant, and $\Delta G^{\ddagger}$ is the free energy barrier determined directly from the PMF. In our previous study of WT KSI, we showed that the results are similar using the two different forms of the TST rate constant.[19]

Furthermore, we calculated the transmission coefficient κ using the reactive flux approach[50-54] and included the nuclear quantum effects associated with the transferring hydrogen nucleus with the quantized classical path method[55-58] in our previous work.[19] We found that the transmission coefficient decreases the rate and the nuclear quantum effects increase the rate by approximately the same amount in WT KSI, so the inclusion of both effects together does not significantly impact the rate constant.[19] Moreover, we do not expect the trends in the rate constants for the mutant KSI enzymes to be strongly influenced by dynamical barrier recrossings or nuclear quantum effects. Thus, these effects were not included in the calculations of the rate constants for the present study.

B. Simulation Details

The starting structures for the mutant KSI simulations were obtained from our previous study on WT KSI.[19] The initial structure for this previous study was based on the second dimeric unit in the 2.3 Å resolution crystal structure (PDB 1QJG)[59] of *Commonas testosteroni* Δ^5-3-ketosteroid isomerase complexed with equilenin (EQU), which is an analog for the Intermediate state. The equilenin was replaced with the naturally occurring substrate, the mutated residues 38 and 83 were restored to those in WT KSI, and the protonation states were determined from the 1BUQ solution NMR structure.[60] The protein was immersed in a periodically replicated truncated octahedron box of 6837 explicit TIP3P rigid water molecules[61, 62] with 6 Na^+ ions to maintain charge neutrality. The system was equilibrated using a simulated annealing procedure involving MD for an isobaric, isothermal ensemble (NPT) at systematically increasing

temperatures, followed by additional MD for a canonical ensemble (NVT) at 298 K. The details of the system preparation and equilibration are provided elsewhere.[19] For the starting structures in the present study, we chose snapshots obtained after more than 2 ns of NVT MD in the Intermediate state (i.e., λ = 0.95 for the first proton transfer step and λ = 0.05 for the second proton transfer step).

The initial states of the mutant enzymes were prepared from these starting structures as follows. The D99L, Y14F, and Y14F/D99L mutant forms of KSI were created using the profix utility in the JACKAL protein modeling package.[63, 64] The mutated residues were optimized (i.e., the energy was minimized with respect to their coordinates) with the remaining protein and water molecules fixed. Then the active site residues and the substrate were optimized with the remaining protein and water molecules fixed. Subsequently, the water molecules and ions were optimized while keeping the protein fixed, and finally the entire system was optimized. The final stage of equilibration for each mutant consisted of 200 ps of NVT MD at 298 K. All of these equilibration steps were performed in the Intermediate state (i.e., λ = 0.95 for the first proton transfer step and λ = 0.05 for the second proton transfer step).

A modified version of the DLPROTEIN package[65] was used for these simulations. For data collection, the MD trajectories were propagated at 298 K for a canonical ensemble (NVT) with a time step of 1 fs. The hydrogen atoms bonded to heavy atoms were constrained to their equilibrium distances with the SHAKE algorithm.[66] The Smooth Particle Mesh Ewald method[67] was used to calculate the electrostatic interactions. The temperature was maintained with a Nosé -Hoover thermostat.[68, 69] The PMF was generated for WT and the three mutant KSI enzymes using the exact same procedure. The results for WT KSI are qualitatively similar to the data obtained previously for WT KSI.[19] All tables and figures in the present paper were generated from the current data.

To generate the PMF, a series of MD trajectories propagated according to different mapping potentials was started from the Intermediate state for both proton transfer reactions. For the first proton transfer step, the windows with $\lambda = 0.95$ and $\lambda = 0.90$ were started from the corresponding equilibrated structure described above. For the second proton transfer step, the windows with $\lambda = 0.05$ and $\lambda = 0.10$ were started from the corresponding equilibrated structure described above. Each subsequent window was started from the configuration of the previous window after 10 ps of equilibration. For each window, an initial optimization followed by equilibration with 100 ps of MD was performed prior to data collection. A total of 19 different mapping potentials with values of λ = 0.05, 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40, 0.45, 0.50, 0.55, 0.60, 0.65, 0.70, 0.75, 0.80, 0.85, 0.90 and 0.950 was used for each proton transfer step. Two independent sets of trajectories were propagated for each proton transfer step, starting from the same equilibrated configuration but different initial velocities in the first window. We collected 600 ps of data for each mapping potential in both sets of trajectories, leading to a total of 23 ns of data collection for each proton transfer step of each enzyme.

4-3. Results

A. Relative rate constants

As described above, we generated 23 ns of MD data for the WT and the three mutant forms of KSI for both proton transfer steps. We used WHAM[40-43] to generate the PMF curves for the two proton transfer reactions for the WT and mutant KSI enzymes. For comparison, we also used umbrella integration[44-46] to generate the PMF curves and found the barrier height to be the same to within 0.5 kcal/mol for these two methods. The PMF curves for the two proton transfer steps of all three mutants are depicted in Figure 4-2. For each mutant we generated separate PMF curves for the two independent data sets to test convergence and reproducibility. These calculations suggest that the error in the PMF barrier due to numerical procedures is ~1.0

kcal/mol. The corresponding TST rate constants for the mutants are presented in Table 1. As mentioned above, the EVB parameters V_{12} and Δ were determined so that the TST rate constants for WT KSI reproduce the experimentally estimated forward and backward rate constants for each proton transfer step.[38, 39]

Table **4-1**: Rate constants for the two proton transfer reactions catalyzed by WT and mutant forms of KSI.

Rate constants[b]	WT[a]	D99L	Y14F	Y14F/D99L
k_1	1.7×10^5	569.42	43.83	0.904
k_{-1}	5.6×10^5	6.8×10^8	1.7×10^8	8.1×10^8
k_2	2.1×10^5	8.6×10^5	1.1×10^6	2.8×10^7
k_{-2}	40	0.40	6.6×10^{-4}	1.2×10^{-3}

a) The EVB parameters used in all calculations for this table were V_{12} = 99.5 kcal/mol and Δ = 1.8 kcal/mol for the first step and V_{12} = 93.1 kcal/mol and Δ = -14.5 kcal/mol for the second step. These parameters were determining by fitting to the experimental rate constants given in [25-26].

b) All rate constants were calculated with Eq. (4) and are given in s^{-1}.

Figure 4-2 illustrates that the mutations increase the free energy of the Intermediate state relative to the Reactant and Product states. As a result, the free energy barrier increases for the first proton transfer step and decreases for the second proton transfer step upon mutation. This trend is quantified in Table 4-1, which indicates that the forward rate constant decreases for the first proton transfer step and increases for the second proton transfer step upon mutation. For WT KSI, the experimentally measured rate constants imply that the rate constants for the two proton transfer steps are similar.[38, 39] In contrast, our calculations indicate that the rate constant of the first proton transfer step is significantly smaller than the second step for the mutant KSI enzymes because of the destabilized Intermediate state. The calculated forward rate constant for the first

proton transfer step decreases by $10^{2.5}$, $10^{3.6}$, and $10^{5.3}$ for the D99L, Y14F, and Y14F/D99L mutants, respectively, relative to the WT rate constant. This trend is qualitatively consistent with the experimental data showing that k_{cat} is decreased by $10^{2.1}$, $10^{3.4}$, and $10^{4.6}$ for the D99L, Y14F, and Y14F/D99L mutants of *Pseudomonas putida* KSI,[3] and k_{cat} is decreased by $10^{3.7}$ and $10^{4.7}$ for the D99A and Y14F mutants of *Commamonas testosteroni* KSI.[2, 20] Note that k_{cat} depends on other rate constants in the overall enzymatic reaction,[38] so only the qualitative trends can be compared.

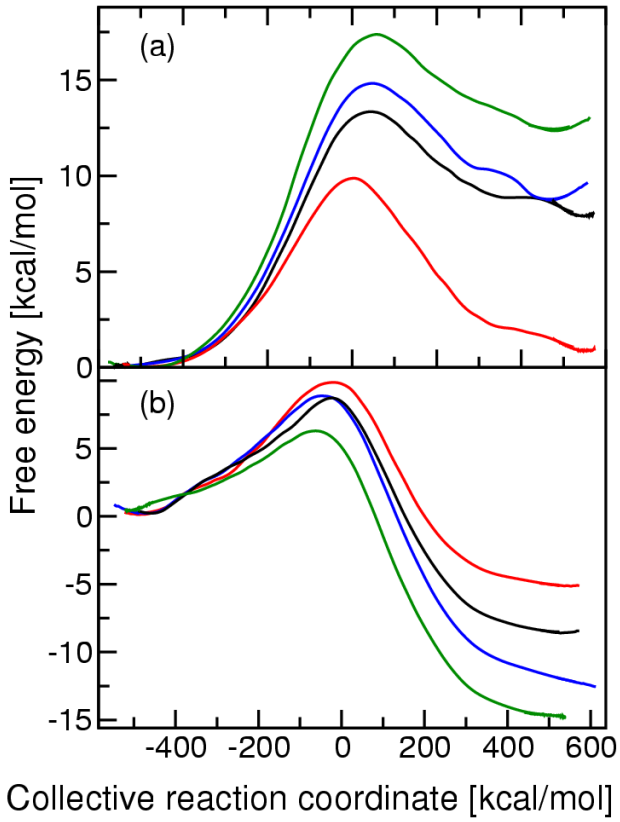


Figure **4-2**: Potential of mean force (PMF) curves for (a) the first proton transfer step and (b) the second proton transfer step catalyzed by KSI. These profiles are depicted for the WT (red), Y14F mutant (blue), D99L mutant (black), and Y14F/D99L double mutant (green) forms of KSI.

We also calculated k_{cat} for the WT and mutant KSIs using the expression given in Ref. [38] . Since this expression also depends on the rate constant of product release, we assumed either that the rate constant for product release is the same in the mutants as in WT KSI or that the rate

constant for product release is much larger than the forward rate constant of the first proton transfer step for the mutants. Using either assumption, the calculated k_{cat} is decreased by $10^{4.6}$, $10^{5.0}$, and $10^{6.0}$ for the D99L, Y14F, and Y14F/D99L mutants, respectively, relative to WT KSI. The qualitative trend in these calculated values of k_{cat} is consistent with the experimental data.

B. Hydrogen bonding in the active site

We examined the impact of the mutations on the catalytically important hydrogen bonds formed in the active site. As discussed above, hydrogen bonds formed by the substrate with Tyr14 and Asp99 play a crucial role in stabilizing the dienolate intermediate. In addition to examining these catalytically important hydrogen bonds, we also analyzed the hydrogen bonds formed by Tyr55 with Tyr14 and, in some cases, with the substrate. These hydrogen bonds are depicted schematically in Figure 4-3. Kinetic experiments indicate that the catalytic effects of mutating the Tyr55 residue are relatively minor,[20] although the crystal structure of the Tyr55Phe mutant suggests that Tyr55 may play a role in positioning the Tyr14 residue.[29] In order to investigate these hydrogen bonds during the proton transfer reactions, we calculated thermally averaged hydrogen bond donor-acceptor distances along the entire range of the collective reaction coordinate for both steps. For this purpose, the thermally averaged distances for the unbiased EVB potential were calculated from the MD trajectories propagated with mapping potentials using methodology related to the WHAM procedure, as described previously.[19]

Figure 4-4 depicts the changes in these hydrogen bond distances during both proton transfer reactions for the WT and three mutant KSI enzymes. These figures were generated from the first data set. The analogous figures generated from the second independent data set are qualitatively similar and are provided in Appendix B. Figure 4-4 also depicts snapshots from the MD trajectories corresponding to the Intermediate state, where the dienolate is stabilized by hydrogen-bonding interactions in the active site. These snapshots illustrate the hydrogen-bonding

patterns for each mutant, as well as the structural rearrangements that occur within the active site upon mutation. We found that a single hydrogen-bonding pattern tends to dominate for each window (i.e., for a trajectory generated with a particular mapping potential), although in some cases we did observe changes in hydrogen bonding within a single window.

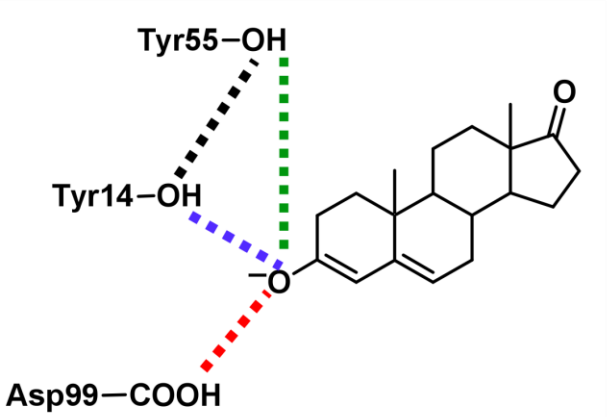


Figure **4-3**: Schematic depiction of the hydrogen-bonding interactions analyzed in Figure 4-4.

In the D99L mutant, the mutation of aspartic acid to leucine eliminates one of the key hydrogen bonds to the substrate. Figure 4-4 indicates that hydrogen bonds between Tyr14 and both the substrate O3 and Tyr55 are present during the first proton transfer step, as observed in WT KSI. In contrast to WT KSI, where the hydrogen bond between the substrate and Tyr14 forms during the first part of this step, the hydrogen bond between the substrate and Tyr14 is present from the beginning of this step in the D99L mutant. For the second proton transfer step, the hydrogen bond between Tyr14 and the substrate is also maintained throughout the entire reaction. In this step, however, Tyr55 alternates between forming a hydrogen bond with either Tyr14 or the substrate O3 for the first part and forms a stable hydrogen bond with Tyr14 for the second part of the reaction.

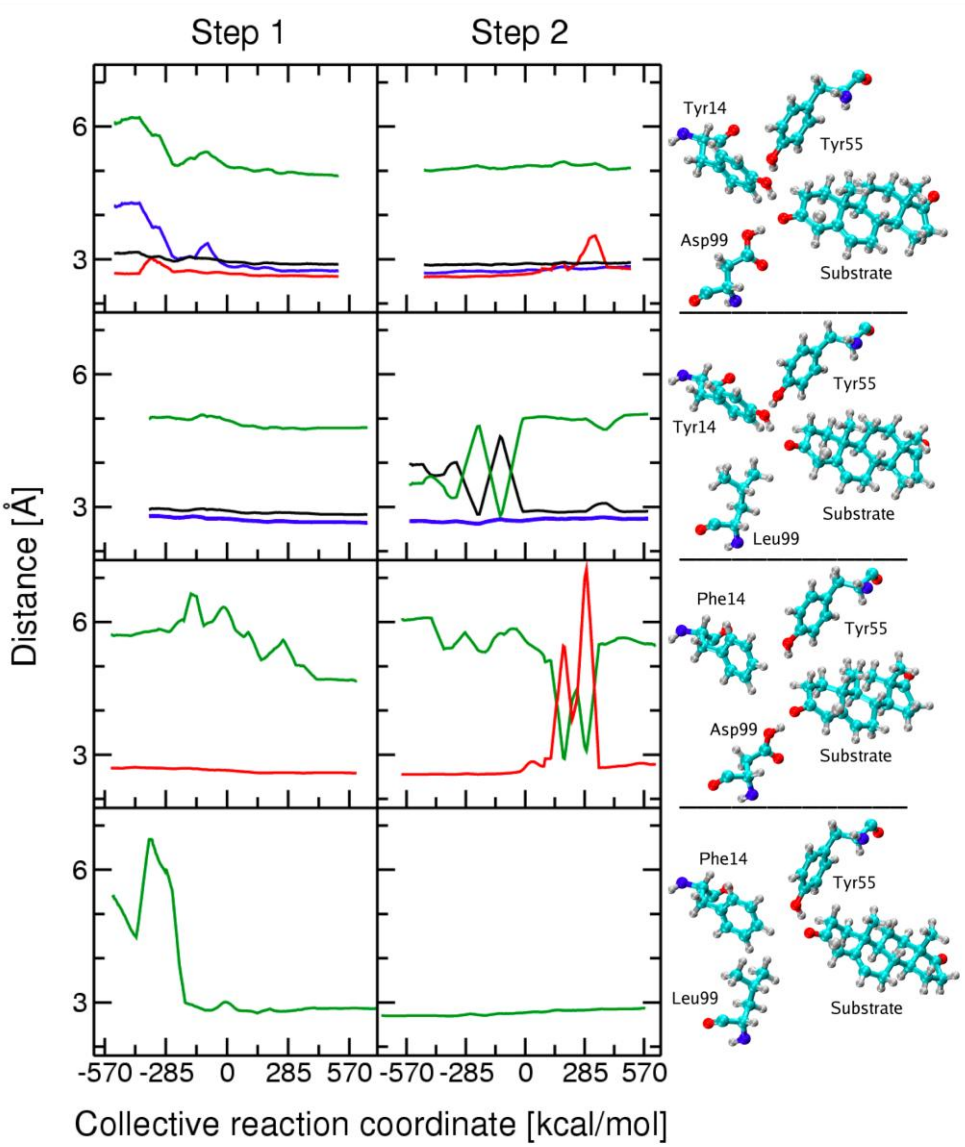


Figure **4-4**: Thermally averaged distances calculated along the collective reaction coordinate for the first (left) and second (right) proton transfer reactions catalyzed by KSI. Snapshots of the hydrogen-bonding pattern in the active site for the Intermediate state are also depicted. The results for the WT, D99L, Y14F, and Y14F/D99L mutants are given from top to bottom. The thermally averaged hydrogen bond donor-acceptor distances between the substrate O3 atom and Tyr14 (blue), between the substrate O3 atom and Asp99 (red), between the substrate O3 atom and Tyr55 (green), and between Tyr14 and Tyr55 (black) are shown. The distances are not shown for the mutated residue. The color coding is depicted in Figure 4-3.

In the Y14F mutant, the mutation of tyrosine to phenylalanine eliminates the other key hydrogen bond to the substrate. Figure 4-4 indicates the presence of a hydrogen bond between the substrate O3 and Asp99 throughout the first proton transfer step and most of the second proton transfer step. Near the end of the second step, this hydrogen bond disappears for short intervals during which the substrate forms a hydrogen bond with Tyr55 instead. In this regime, however, the product is partially dissociating, so the MD results become less reliable.

In the Y14F/D99L double mutant, both critical hydrogen bonds to the substrate are eliminated. In this case, Figure 4-4 illustrates the presence of a new hydrogen bond between Tyr55 and the substrate O3. This hydrogen bond is formed near the beginning of the first proton transfer step and is maintained throughout both proton transfer reactions, although the hydrogen bond distance gradually increases toward the end of the second step. Based on kinetic experiments indicating that this mutant is catalytically active,[3] we conjecture that this hydrogen-bonding interaction partially compensates for the absence of hydrogen bonds formed by the substrate with Tyr14 and Asp99.

In the Y14F/D99L double mutant, both critical hydrogen bonds to the substrate are eliminated. In this case, Figure 4-4 illustrates the presence of a new hydrogen bond between Tyr55 and the substrate O3. This hydrogen bond is formed near the beginning of the first proton transfer step and is maintained throughout both proton transfer reactions, although the hydrogen bond distance gradually increases toward the end of the second step. Based on kinetic experiments indicating that this mutant is catalytically active,[3] we conjecture that this hydrogen-bonding interaction partially compensates for the absence of hydrogen bonds formed by the substrate with Tyr14 and Asp99.

This analysis indicates that when the native hydrogen-bonding interactions are eliminated, structural rearrangements within the active site occur to stabilize the dienolate intermediate by forming new hydrogen-bonding interactions. An example of a new interaction is

the hydrogen bond formed between the substrate and Tyr55 in the Y14F/D99L double mutant and sporadically in the D99L mutant. In the mutants, typically the substrate oxygen atom O3 is hydrogen bonded to only one residue, rather than to both Tyr14 and Asp99 as in WT KSI. As a result, the dienolate intermediate is stabilized enough for the mutants to be catalytically active but not enough to achieve rates as fast as WT KSI. These observations are consistent with the PMF curves depicted in Figure 4-2 and the rate constants given in Table 4-1.

C. Conformational changes and electrostatics

As in our previous simulations of WT KSI,[19] we observed relatively small conformational changes within the active site that facilitate the proton transfer reactions. Specifically, the proton donor-acceptor distance decreases to ~2.67 Å near the transition state of each proton transfer reaction to enable the hydrogen to transfer. In addition, Asp38 exhibits a significant reorientation relative to the substrate between the first and second proton transfer steps to accommodate proton transfer to the C6 atom of the substrate after accepting a proton from the C4 atom of the substrate. These types of conformational changes are essential to the catalytic activity of the enzyme. They are not significantly affected by the mutation of Asp99 and Tyr14, however, and hence are similar for the WT and mutant KSI enzymes.

To investigate the overall structural changes in the protein occurring during the catalytic reaction, we generated thermally averaged structures for the reactant state, transition state, and product state of both proton transfer reactions. These structures were generated from MD trajectories obtained with mapping potentials corresponding to $\lambda = 0.05$, 0.50, and 0.95 for each proton transfer step. The thermally averaged structure for each mapping potential was generated by minimizing the RMSD over all configurations,[70] followed by a weighting and averaging procedure that provides results for the unbiased EVB potential.[19] Note that this process may lead

to non-physical bond distances and angles in the side chains and therefore was used to analyze only the global changes in protein structure.

Table **4-2**: RMSD values [c] calculated from thermally averaged structures for WT and mutant forms of KSI.

Region	WT[b]	D99L	Y14F	Y14F/D99L
Entire enzyme	0.47	0.65	0.68	0.55
Active site[a]	0.21	0.33	0.30	0.27
Tyr88 to Lys92	1.30	1.74	1.62	1.08

a) The active site region consisted of all residues containing an atom within 5Å of the substrate C4 atom or any atom of Tyr14 or Asp99 [25, 26].
b) These values were obtained from Ref. [25, 26].
c) All RMSD values were calculated for backbone atoms using the VMD program Ref. (*25*, *26*) and are given in $Å^{-1}$.

For each mutant, we calculated the average RMSD among the six thermally averaged structures spanning the two proton transfer reactions using the VMD program.[71] The average RMSD values for the backbone atoms of the entire protein, the active site residues, and the mobile loop Tyr88-Lys92 are given in Table 4-2. Here the active site residues are defined to be those containing an atom within 5 Å of the substrate C4 atom or any atom in the Tyr14 or Asp99 residues. The RMSD values for the mutant and WT KSI enzymes are similar, with slightly lower values for WT KSI. The relatively low average RMSD values for the active site are consistent with experimental studies indicating that the active site remains fairly rigid during the proton transfer reactions.[7, 9, 72, 73] The larger average RMSD values for the loop Thr88-Lys92 indicate that this loop is more mobile than the active site. Several other mobile loops that were observed in previous simulations of WT KSI[19] were also observed in the mutant forms of KSI.

Table 4-3 provides the average van der Waals and electrostatic interaction energies between the substrate ligand and the enzyme for the WT and mutant KSI enzymes over the course of the two proton transfer reactions. These interaction energies were averaged over

configurations sampled with 600 ps of MD for the Reactant, Intermediate, and Product, generated using mapping potentials with λ = 0.05 and 0.95 for the first proton transfer step and λ = 0.95 for

Table **4-3**: Electrostatic and van der Waals interaction energies[b] between the substrate and KSI for different states along the reaction pathway.

KSI	Electrostatic			van der Waals		
	Reactant	Intermediate	Product	Reactant	Intermediate	Product
WT[a]	–30	–101s	–27	–26	–24	–26
D99L	–24	–84	–23	–26	–25	–25
Y14F	–27	–91	–23	–26	–25	–27
Y14F/D99L	–19	–83	–21	–26	–24	–25

a) These values were obtained from [25, 26].
b) Interaction energies given in kcal/mol

the second proton transfer step. The virtually constant van der Waals interaction energies imply that no major structural rearrangements occur in the active site during proton transfer. The stronger electrostatic interaction energy for the Intermediate than for the Reactant and Product arises in part from the transfer of the positively charged proton to Asp38 and the resulting redistribution of electronic charge in the substrate. As shown by previous calculations of the hydrogen-bonding interaction energies for WT KSI,[19] this enhanced electrostatic interaction energy is also due to the strengthening of the hydrogen bonds between the substrate oxygen and active site residues (i.e., Asp99, Tyr14, or Tyr55) for the dienolate intermediate. Note that the electrostatic interaction energies are slightly weaker for the mutants than for WT KSI. This difference is consistent with the observation that the mutants exhibit only one hydrogen-bonding interaction, while the WT KSI exhibits two hydrogen-bonding interactions, between the substrate and the active site. Moreover, the electrostatic stabilization of the Intermediate relative to the Reactant is 71 kcal/mol for WT KSI and 60-64 kcal/mol for the mutant forms of KSI. This difference in relative electrostatic interaction energies contributes to the higher free energy of the Intermediate relative to the Reactant in the mutant forms, as depicted in Figure 4-2.

4-4. Conclusions

In this study, we performed MD simulations of the two proton transfer reactions catalyzed by the D99L, Y14F, and Y14F/D99L mutant forms of KSI. The free energy profiles along a collective reaction coordinate illustrate that the mutations destabilize the dienolate intermediate relative to the reactant and product, thereby increasing the free energy barrier for the first proton transfer step and decreasing the free energy barrier for the second proton transfer step. While the rate constants for the two proton transfer steps are similar in WT KSI, our simulations suggest that the rate constant of the first proton transfer step is smaller than the rate constant of the second step in all three mutant forms of KSI. The calculated rate constant was found to decrease along the following series of KSI enzymes: WT, D99L, Y14F, and Y14F/D99L. This trend in the calculated rate constants is qualitatively consistent with kinetic experiments on these mutants. Analysis of the hydrogen-bonding patterns, conformational changes, and van der Waals and electrostatic interactions during the proton transfer reactions provides insight into the physical basis for these trends in the rate constants.

Our simulations illustrated that the mutants typically retain one hydrogen-bonding interaction between the substrate oxygen atom O3 and the active site, while the WT KSI retains two hydrogen-bonding interactions. We observed a new hydrogen bonding interaction between the substrate and Tyr55 in the double mutant and occasionally in the D99L single mutant. Moreover, the electrostatic interaction energy between the substrate and the enzyme for the Intermediate relative to the Reactant is ~8 kcal/mol greater for WT KSI than for the mutant forms of KSI. These observations provide a qualitative explanation for the experimental measurements showing that the mutant KSI enzymes are still catalytically active but exhibit lower catalytic rates. The calculations also predict that a hydrogen bond between Tyr55 and an intermediate analog such as equilenin would be observed in the crystal structure of the double mutant and that

mutation of Tyr55 to Phe would have a greater impact on the proton transfer rates for the double mutant than for WT KSI.

In addition, the simulations illustrated that the van der Waals interactions between the substrate and the enzyme remain relatively constant during the two proton transfer reactions and that the RMSD of the active site among thermally averaged structures spanning both reactions is only ~0.3 Å. These observations suggest that the active site remains relatively rigid during the proton transfer reactions, as also indicated by experimental studies.[7, 9, 72, 73] Nevertheless, as shown previously for WT KSI,[19] changes in the proton donor-acceptor distances and the angles between the substrate and Asp38, which serves as the proton acceptor and donor, must accompany the proton transfer reactions.

The high catalytic efficiency of KSI has been postulated to arise predominantly from a preorganized active site.[18, 74-76] Our simulations indicate that the active site remains relatively rigid during the two proton transfer reactions for WT and the mutant forms of KSI. The calculations also illustrate that mutation of the catalytically important Asp99 and Tyr14 residues leads to structural rearrangements within the active site to retain hydrogen-bonding interactions between the substrate and the enzyme. Thus, these calculations are consistent with the postulate that KSI forms a preorganized active site for both the WT and mutant forms of KSI, but the structure of this preorganized active site is altered upon mutation. Moreover, this preorganized active site still allows for relatively small conformational changes that facilitate the proton transfer reactions by strengthening the hydrogen bonds between the substrate and active site residues, as well as by bringing the proton donor and acceptor closer to each other with the proper orientation for proton transfer.[19] Such conformational changes are due to stochastic thermal motions of the protein and substrate as they sample the multidimensional free energy landscape.[77] Thus, both the concepts of a preorganized active site and conformational sampling are important for this enzyme reaction.

4-5 References:

(1) Zhao, Q.; Abeygunawardana, C.; Talalay, P.; Mildvan, A. S. *Proc. Nat. Acad. Sci. USA* **1996,** *93*, 8220-8224.

(2) Wu, Z. R.; Ebrahimian, S.; Zawrotny, M. E.; Thornburg, L. D.; Perez-Alvarado, G. C.; Brothers, P.; Pollack, R. M.; Summers, M. F. *Science* **1997,** *276*, 415-418.

(3) Choi, G.; Ha, N.-C.; Kim, S. W.; Kim, D.-H.; Park, S.; Oh, B.-H.; Choi, K. Y. *Biochemistry* **2000,** *39*, 903-909.

(4) Kraut, D. A.; Sigala, P. A.; Pybus, B.; Liu, C. W.; Ringe, D.; Petsko, G. A.; Herschlag, D. *PLoS Biology* **2006,** *4*, 501-519.

(5) Mildvan, A. S.; Massiah, M. A.; Harris, T. K.; Marks, G. T.; Harrison, D. H. T.; Viragh, C.; Reddy, P. M.; Kovach, I. M. *J. Mol. Struct.* **2002,** *615*, 163-175.

(6) Sigala, P. A.; Fafarman, A. T.; Bogard, P. E.; Boxer, S. G.; Herschlag, D. *J. Am. Chem. Soc.* **2007,** *129*, 12104-12105.

(7) Sigala, P. A.; Kraut, D. A.; Caaveiro, J. M. M.; Pybus, B.; Ruben, E. A.; Ringe, D.; Petsko, G. A.; Herschlag, D. *J. Am. Chem. Soc.* **2008,** *130*, 13696-13708.

(8) Kraut, D. A.; Churchill, M. J.; Dawson, P. E.; Herschlag, D. *ACS Chemical Biology* **2009,** *4*, 269-273.

(9) Sigala, P. A.; Caaveiro, J. M. M.; Ringe, D.; Petsko, G. A.; Herschlag, D. *Biochemistry* **2009,** *48*, 6932-6939.

(10) Oh, K. S.; Cha, S.-S.; Kim, D.-H.; Cho, H.-S.; Ha, N.-C.; Choi, G.; Lee, J. Y.; Tarakeshwar, P.; Son, H. S.; Choi, K. Y.; Oh, B.-H.; Kim, K. S. *Biochemistry* **2000,** *39*, 13891-13896.

(11) Kim, K. S.; Oh, K. S.; Lee, J. Y. *Proc. Nat. Acad. Sc. USA* **2000,** *97*, 6373-6378.

(12) Pan, Y.; McAllister, M. A. *Journal of Molecular Structure (Theochem)* **2000,** *504*, 29-33.

(13) Park, H.; Merz, K. M., Jr. *J. Am. Chem. Soc.* **2003,** *125*, 901-911.

(14) Feierberg, I.; Aqvist, J. *Biochemistry* **2002,** *41*, 15728-15735.

(15) Feierberg, I.; Aqvist, J. *Theoretical Chemical Accounts* **2002,** *108*, 71-84.

(16) Mazumder, D.; Kahn, K.; Bruice, T. C. *J. Am. Chem. Soc.* **2003,** *125*, 7553-7561.

(17) Warshel, A.; Sharma, P. K.; Chu, Z. T.; Aqvist, J. *Biochemistry* **2007,** *46*, 1466-1476.

(18) Kamerlin, S. C. L.; Sharma, P. K.; Chu, Z. T.; Warshel, A. *Proceedings of the National Academy of Sciences (USA)* **2010,** *107*, 4075-4080.

(19) Chakravorty, D. K.; Soudackov, A. V.; Hammes-Schiffer, S. *Biochemistry* **2009,** *48*, 10608-10619.

(20) Kuliopulos, A.; Mildvan, A. S.; Shortle, D.; Talalay, P. *Biochemistry* **1989,** *28*, 149-159.

(21) Xue, L.; Talalay, P.; Mildvan, A. S. *Biochemistry* **1991,** *30*, 10858-10865.

(22) Austin, J. C.; Sprio, T. G.; Kuliopulos, A.; Mildvan, A. S. *Protein Science* **1992,** *1*, 259-270.

(23) Brothers, P. N.; Blotny, G.; Qi, L.; Pollack, R. M. *Biochemistry* **1995,** *34*, 15453-15458.

(24) Kim, S. W.; Choi, K. Y. *J. Bacteriol.* **1995,** *177*, 2602-2605.

(25) Zhao, Q.; Abeygunawardana, C.; Gittis, A. G.; Mildvan, A. S. *Biochemistry* **1997,** *36*, 14616-14626.

(26) Qi, L.; Pollack, R. M. *Biochemistry* **1998,** *37*, 6760-6766.

(27) Thornburg, L. D.; Henot, F.; Bash, D. P.; Hawkinson, D. C.; Bartel, S. D.; Pollack, R. M. *Biochemistry* **1998,** *37*, 10499-10506.

(28) Pollack, R. M.; Thornburg, L. D.; Wu, Z. R.; Summers, M. F. *Arch. Biochem. Biophys.* **1999,** *370*, 9-15.

(29) Kim, D.-H.; Jang, D. S.; Nam, G. H.; Choi, G.; Kim, J.-S.; Ha, N.-C.; Kim, M.-S.; Oh, B.-H.; Choi, K. Y. *Biochemistry* **2000,** *39*, 4581-4589.

(30) Thornburg, L. D.; Goldfeder, Y. R.; Wilde, T. C.; Pollack, R. M. *J. Am. Chem. Soc.* **2001,** *123*, 9912-9913.

(31) Choi, G.; Ha, N.-C.; Kim, M.-S.; Hong, B.-H.; Oh, B.-H.; Choi, K. Y. *Biochemistry* **2001,** *40*, 6828-6835.

(32) Houck, W. J.; Pollack, R. M. *J. Am. Chem. Soc.* **2004,** *126*, 16416-16425.

(33) Jang, D. S.; Cha, H. J.; Cha, S.-S.; Hong, B. H.; Ha, N.-C.; Lee, J. Y.; Oh, B.-H.; Lee, H.-S.; Choi, K. Y. *Biochem. J.* **2004,** *382*, 967-973.

(34) Kraut, D. A.; Sigala, P. A.; Fenn, T. D.; Herschlag, D. *Proceedings of the National Academy of Sciences (USA)* **2010,** *107*, 1960-1965.

(35) Warshel, A., *Computer Modeling of Chemical Reactions in Enzymes and Solutions*. John Wiley & Sons, Inc.: New York, 1991.

(36) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *Journal of the American Chemical Society* **1995,** *117*, 5179-5197.

(37) Wang, J.; Cieplak, P.; Kollman, P. A. *Journal of Computational Chemistry* **2000,** *21*, 1049-1074.

(38) Hawkinson, D. C.; Eames, T. C. M.; Pollack, R. M. *Biochemistry* **1991,** *30*, 10849-10858.

(39) Hawkinson, D. C.; Pollack, R. M.; Ambulos, N. P., Jr. *Biochemistry* **1994,** *33*, 12172-12183.

(40) Ferrenberg, A. M.; Swendsen, R. H. *Phys. Rev. Lett.* **1988,** *61*, 2635-2638.

(41) Ferrenberg, A. M.; Swendsen, R. H. *Phys. Rev. Lett.* **1989,** *63*, 1195-1198.

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

(43) Roux, B. *Comput. Phys. Commun.* **1995,** *91*, 275-282.

(44) Kastner, J.; Thiel, W. *J. Chem. Phys.* **2005,** *123*, 144104.

(45) Kastner, J.; Thiel, W. *J. Chem. Phys.* **2006,** *124*, 234106.

(46) Chakravorty, D. K.; Kumarasiri, M.; Soudackov, A. V.; Hammes-Schiffer, S. *Journal of Chemical Theory and Computation* **2008,** *4*, 1974-1980.

(47) Watney, J. B.; Soudackov, A. V.; Wong, K. F.; Hammes-Schiffer, S. *Chem. Phys. Lett.* **2006,** *418*, 268-271.

(48) Carter, E. A.; Ciccotti, G.; Hynes, J. T.; Kapral, R. *Chem. Phys. Lett.* **1989,** *156*, 472-477.

(49) Schenter, G. K.; Garrett, B. C.; Truhlar, D. G. *J. Chem. Phys.* **2003,** *119*, 5828-5833.

(50) Neria, E.; Karplus, M. *J. Chem. Phys.* **1996,** *105*, 10812-10818.

(51) Bennett, C. H., *Algorithms for Chemical Computation*. American Chemical Society: Washington, DC, 1997.

(52) Keck, J. C. *J. Chem. Phys.* **1960,** *32*, 1035-50.

(53) Anderson, J. B. *J. Chem. Phys.* **1973,** *58*, 4684-92.

(54) Chandler, D., A story of rare events: From barriers, to electrons, to unknown pathways. In *Classical and Quantum Dynamics in Condensed Phase Simulations*, Berne, B. J.; Ciccotti, G.; Coker, D. F., Eds. World Scientific: Singapore, 1998; pp 3-23.

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

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

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

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

(59) Cho, H.-S.; Ha, N.-C.; Choi, G.; Kim, H.-J.; Lee, D.; Oh, K. S.; Kim, K. S.; Lee, W.; Choi, K. Y.; Oh, B.-H. *J. Biol. Chem.* **1999,** *274*, 32863-32868.

(60) Massiah, M. A.; Abeygunawardana, C.; Gittis, A. G.; Mildvan, A. S. *Biochemistry* **1998,** *37*, 14701-14712.

(61) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *Journal of Chemical Physics* **1983,** *79*, 926-935.

(62) Mahoney, M. W.; Jorgensen, W. L. *Journal of Chemical Physics* **2000,** *112*, 8910-8922.

(63) Xiang, Z.; Honig, B. *J. Mol. Biol.* **2001,** *311*, 421-430.

(64) Xiang, J. Z.; Honig, B. *JACKAL: A Protein Structure Modeling Package*, Columbia University & Howard Hughes Medical Institute: New York, 2002.

(65) Melchionna, S.; Cozzini, S. *DLPROTEIN Version 2.1 Molecular Dynamics Software Package for Macromolecules*, INFM Udr SISSA, Trieste, Italy, 2001.

(66) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977,** *23*, 327.

(67) Darden, T.; York, D.; Pedersen, L. *Journal of Chemical Physics* **1993,** *98*, 10089-10092.

(68) Nose, S. *Molecular Physics* **1984,** *52*, 255-68.

(69) Hoover, W. G. *Physical Review A* **1985,** *31*, 1695-7.

(70) van der Spoel, D.; Lindahl, E.; Hess, B.; van Buuren, A. R.; Apol, E.; Meulenhoff, P. J.; Tieleman, D. P.; Sijbers, A. L. T. M.; Feenstra, K. A.; van Drunen, R.; Berendsen, H. J. C. *GROMACS*, 3.2; 2004.

(71) Humphrey, W.; Dalke, A.; Schulten, K. *Journal of Molecular Graphics* **1996,** *14*, 33-38.

(72) Zhao, Q.; Li, Y.-K.; Mildvan, A. S.; Talalay, P. *Biochemistry* **1995,** *34*, 6562-6572.

(73) Zhao, Q.; Abeygunawardana, C.; Mildvan, A. S. *Biochemistry* **1996,** *35*, 1525-1532.

(74) Rajagopalan, P. T. R.; Benkovic, S. J. *The Chemical Record* **2002,** *2*, 24-36.

(75) Wolfenden, R. *Acc. Chem. Res.* **1972,** *5*, 10-18.

(76) Jencks, W. P., *Catalysis in Chemistry and Enzymology*. 2nd ed.; Dover: New York, 1987.

(77) Benkovic, S. J.; Hammes, G. G.; Hammes-Schiffer, S. *Biochemistry* **2008,** *47*, 3317-3321.

Chapter 5

Conclusions

5-1. Conclusions

This thesis presented a computational study of enzymatic hydrogen transfer reactions in dihydrofolate reductase (DHFR) and Δ^5-3-ketosteroid isomerase (KSI). These studies were performed using a hybrid quantum/classical molecular dynamics approach.[1-3] This methodology allowed for the inclusion of electronic and nuclear quantum effects and incorporated the motion of the protein and the solvent.

The implementation of Umbrella Integration (UI) and WHAM-n methods within the Empirical Valence Bond (EVB) approach was presented.[4] These methods were shown to significantly improve the quality of potential of mean force (PMF) curves obtained from molecular dynamics (MD) simulations of the reaction in DHFR. Unlike the Weighted Histogram Analysis Method (WHAM), UI does not rely upon a binning procedure. This removes statistical errors associated with such a procedure and allows UI to converge faster. Additionally, UI allows for the efficient generation of accurate PMFs with significantly fewer mapping potentials (windows) compared to WHAM. The WHAM-n method was presented as a modification of the WHAM approach. This method worked by providing an analytical form to the distributions built in the WHAM approach. The newly developed WHAM-n approach was found to work more efficiently than WHAM and had fewer associated statistical errors. In a recent study, UI was used to generate the free energy curves of mutant forms of DHFR with minimal sampling.[5] These

advantages show UI and WHAM-n to be promising methods for the generation of free-energy curves.

Hybrid quantum/classical molecular dynamics simulations of the two proton transfer reactions catalyzed by wild-type KSI showed the active site of the enzyme to be pre-organized for the proton transfer reactions.[6] Relatively small conformational changes of the enzyme active site and substrate were shown to work toward strengthening the hydrogen bonds that stabilize the intermediate, thereby facilitating the proton transfer reactions. The van der Waals interaction energy between the substrate and the enzyme was shown to be virtually constant along the reaction pathway, but the electrostatic interaction energy was significantly stronger for the dienolate intermediate than for the reactant and product. Significant structural re-arrangements were observed in the mobile loop regions that were distal to the active site. The active site of the enzyme itself was found to be devoid of any major structural changes in support of the idea of a pre-organized active site.

The impact of mutating the catalytically critical Tyr14 and Asp99 residues on the isomerization reaction catalyzed by KSI was also presented. Our simulations showed new hydrogen bonding interactions within the active sites of the mutant forms of KSI. These interactions were observed over the course of the isomerization reaction, suggesting a pre-organized active site for each mutant. Our simulations of the D99L mutant form of KSI exhibited a hydrogen bonding pattern in which Tyr14 formed a hydrogen bond with the substrate throughout both proton transfer reactions. Tyr55 interchangeably maintained a hydrogen bond with Tyr14 or the substrate over the course of the reactions. Simulations of the Y14F mutant form of KSI showed that Asp99 and Tyr55 alternated in forming a single hydrogen bond with the substrate over the course of both reactions. Simulations of the Y14F/D99L double mutant revealed a novel hydrogen bonding pattern in the active site of the enzyme. A hydrogen bond was found to exist between the substrate and Tyr55 over the course of both proton transfer reactions.

The active site of the enzyme was found to be devoid of significant structural rearrangements during the proton transfer reactions in the mutant forms of KSI. The calculated free-energy profiles for both proton transfer steps for the mutant forms of KSI suggested that the first step of the reaction was predominantly rate-limiting, unlike the case of wild-type KSI where both proton transfer steps are partially rate limiting. These results support the hypothesis that Tyr14 and Asp99 play a vital role in governing the very high catalytic rates of the reaction.

Certain shortcomings exist for the hybrid quantum/classical methodology used in this study. One limitation is that it treats only a few hydrogen atoms quantum mechanically. Other limitations include inaccuracies in the potential energy surface that may arise from the fitting procedure used to parameterize the EVB potential. Methods have been developed that use a functional form of the EVB potential that could describe a more accurate potential energy surface for the reaction. A quantum mechanical/molecular mechanical approach in which the substrate and neighboring residues are treated quantum mechanically with a method such as DFT is another possible direction. In addition, as with any other MD study, potential pitfalls include an inaccurate starting structure and inadequate sampling of phase space.

In conclusion, the fundamental principles governing hydrogen transfer reactions in biological processes can be probed by theoretical chemistry, thereby elucidating the role of the structure and motion of the enzyme for proton transfer reactions. These methodologies have been found to provide useful insight into the mechanisms of enzymatic reactions. Faster processor speeds and improved algorithms have resulted in longer simulation times and more efficient sampling. These improvements have allowed the field to make predictions with higher accuracy. It is certainly a very exciting time to be a computational chemist!

5-3. References

(1) Bileter, S.R.; Webb, S.P.; Iordanov; T., Agarwal; P.K., Hammes-Schiffer, S. *J. Chem. Phys.* **2001,** *114*, 6925.

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

(3) Agarwal, P.K**;** Bileter, S.R; Hammes-Schiffer, S. *J. Phys. Chem. B* **2002,** *106,* 328

(4) Chakravorty, D.K.; Kumarasiri, M.D.; Soudackov, A.; Hammes-Schiffer, S. *J. Chem. Theory Comput.* **2008,** *4*, 1974.

(5) Kumarasiri, M.D.; Baker, G.A.;Soudackov, A.; Hammes-Schiffer, S. *J. Phy. Chem. B.* **2009,** *113*, 3579.

(6) Chakravorty, D.K.; Soudackov, A.; Hammes-Schiffer, S. *Biochemistry* **2009,** *48*, 10608.

Appendix A

Supporting Information For Chapter 3

Table A-1: Rate constants[g] for the two proton transfer reactions catalyzed by KSI.

Rate constants[a]	Experimental[b]	k_{TST}[c]	k_{TST} with QCP[d,f]	k_{tot}[e,f]
k_1	1.7×10^5	1.5×10^5	9.2×10^5	2.4×10^5
k_{-1}	5.6×10^5	7.5×10^5	4.4×10^6	1.1×10^6
k_2	2.1×10^5	2.8×10^5	2.3×10^6	6.4×10^5
k_{-2}	40	30.6	290.6	81.4

a) The EVB parameters used in all calculations for this table were V_{12} = 100.0 kcal/mol and Δ = 4.9 kcal/mol for the first step, and V_{12} = 93.8 kcal/mol and Δ = –15.8 kcal/mol for the second step.
b) The experimental rate constants were obtained from Ref. (*2*, *36*).
c) k_{TST} was calculated in terms of the PMF $W(\Lambda)$ for a general reaction coordinate Λ using the following expression:

$$k_{TST} = \left\{ \left(\frac{Z_\Lambda}{2\pi\beta} \right)^{1/2} \right\}_{\Lambda^\ddagger}^{\text{cond}} \frac{e^{-\beta W(\Lambda^\ddagger)}}{\int_{-\infty}^{\Lambda^\ddagger} d\Lambda e^{-\beta W(\Lambda)}} \quad , \tag{A-1}$$

where

$$Z_\Lambda \equiv \sum_{i=1}^{3N} \frac{1}{M_i} \left(\frac{\partial \Lambda}{\partial R_i} \right)^2 , \tag{A-2}$$

$$e^{-\beta W(\Lambda')} = \frac{\int d\mathbf{R} \delta(\Lambda - \Lambda') e^{-\beta V(\mathbf{R})}}{\int d\mathbf{R} e^{-\beta V(\mathbf{R})}} \tag{A-3}$$

and

$$f\ \mathbf{R}\ \ {}^{\mathrm{cond}}_{\Lambda^{\ddagger}} = \frac{\int d\mathbf{R}\delta\ \ \Lambda - \Lambda^{\ddagger}\ \ e^{-\beta V\ \mathbf{R}}\ f\ \ \mathbf{R}}{\int d\mathbf{R}\delta\ \ \Lambda - \Lambda^{\ddagger}\ \ e^{-\beta V\ \mathbf{R}}}\ . \qquad \text{(A-4)}$$

Here N is the number of nuclei with masses M_i and coordinates R_i, $V\ \mathbf{R}$ is the potential energy of the system, and $\Lambda = \Lambda^{\ddagger}$ at the dividing surface. In our calculations, $\Lambda^{\ddagger} = 0$.

d) The QCP method includes the nuclear quantum effects of the transferring H.
e) k_{tot} was calculated with Eq. (3) and includes nuclear quantum effects and barrier recrossings.
f) QCP corrections and κ values presented in Section III.A were used to calculate 'k_{TST} with QCP' and k_{tot}. Test calculations with different V_{12} and Δ parameters indicated that the QCP corrections and values are not sensitive to these types of minor changes in V_{12} and Δ
g) Rate constants given in s^{-1}.

Figure A-1

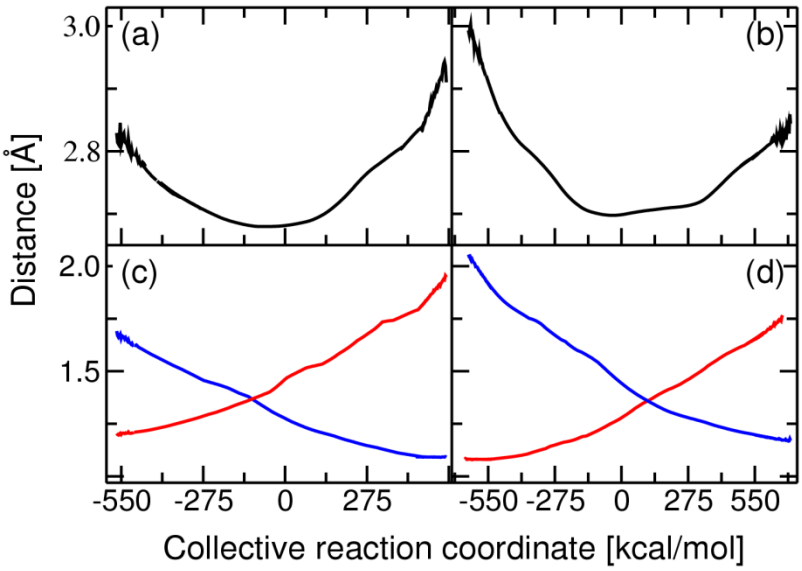


Figure A-1: Thermally averaged distances within the proton transfer interface calculated along the collective reaction coordinate for the two proton transfer reactions catalyzed by KSI for the second data set. For the first proton transfer step, the substrate C4 atom is the donor and the Asp38 OD2 atom is the acceptor. For the second proton transfer step, the Asp38 OD2 atom is the donor and the substrate C6 atom is the acceptor. The donor-acceptor distance is depicted for the first step in (a) and the second step in (b). The donor-hydrogen (red) and acceptor-hydrogen (blue) distances are depicted for the first step in (c) and the second step in (d).

Figure A-2

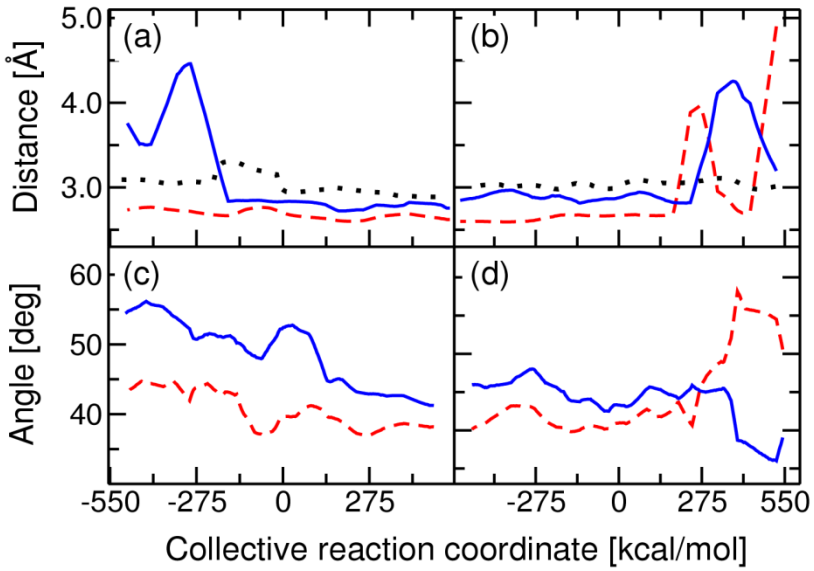


Figure A-2: Thermally averaged distances and angles within the active site calculated along the collective reaction coordinate for the two proton transfer reactions catalyzed by KSI for the second data set. The hydrogen bond donor-acceptor distances between the substrate O3 atom and Tyr14 (solid blue), between the substrate O3 atom and Asp99 (dashed red), and between Tyr14 and Tyr15 (dotted black) are depicted for the first step in (a) and the second step in (b). The angles Tyr14-Asp99-SubstrateO3 (solid blue) and Asp99-Tyr14-SubstrateO3 (dashed red) are depicted for the first step in (c) and the second step in (d). These angles are defined in terms of the heavy atoms involved in the hydrogen bonds between the substrate and both Tyr14 and Asp99.

Figure A-3

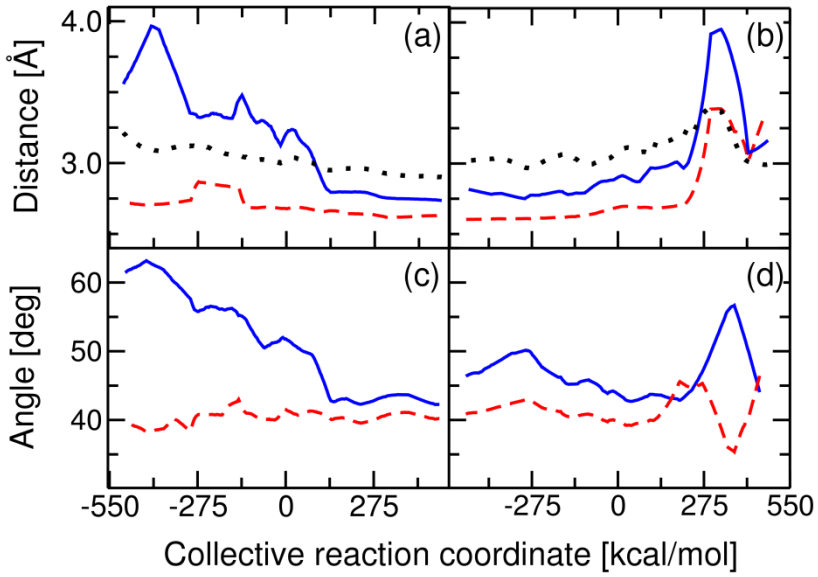


Figure A-3: Thermally averaged distances and angles within the active site calculated along the collective reaction coordinate for the two proton transfer reactions catalyzed by KSI using the set of EVB parameters from k_{TST} calculated from Eq. (A2-1) for the first data set. The hydrogen bond donor-acceptor distances between the substrate O3 atom and Tyr14 (solid blue), between the substrate O3 atom and Asp99 (dashed red), and between Tyr14 and Tyr15 (dotted black) are depicted for the first step in (a) and the second step in (b). The angles Tyr14-Asp99-SubstrateO3 (solid blue) and Asp99-Tyr14-SubstrateO3 (dashed red) are depicted for the first step in (c) and the second step in (d). These angles are defined in terms of the heavy atoms involved in the hydrogen bonds between the substrate and both Tyr14 and Asp99.

Appendix B

Supporting Information For Chapter 4

Table **B-1**: Rate constants obtained for the two proton transfer reactions catalyzed by WT and mutant forms of KSI for the second independent data set.

Region	WT[b]	D99L	Y14F	Y14F/D99L
Entire enzyme	0.47	0.65	0.68	0.55
Active site[a]	0.21	0.33	0.30	0.27
Tyr88 to Lys92	1.30	1.74	1.62	1.08

a) The EVB parameters used in all calculations for this table were V_{12} = 99.5 kcal/mol and Δ = 1.8 kcal/mol for the first step and V_{12} = 93.1 kcal/mol and Δ = -14.5 kcal/mol for the second step. These parameters were determining by fitting to the experimental rate constants.
b) All rate constants were calculated with Eq. (4) and are given in s^{-1}.

Figure B-1

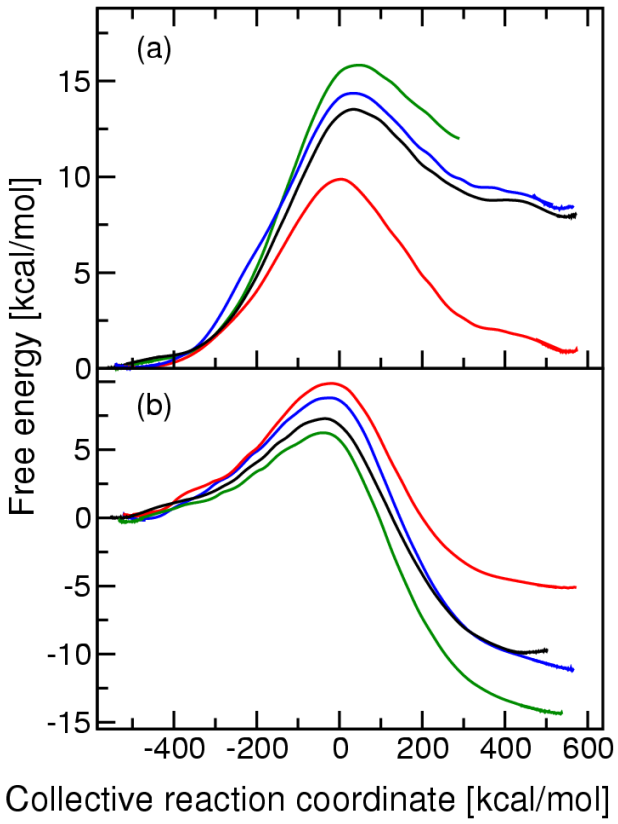


Figure B-1: Potential of mean force (PMF) curves for (a) the first proton transfer step and (b) the second proton transfer step catalyzed by KSI, as obtained from the second independent data set. These profiles are depicted for the WT (red), D99L mutant (black), Y14F mutant (blue), and Y14F/D99L double mutant (green) forms of KSI. For both steps, all curves are shifted so that the reactant is at zero energy, although mechanistically the product of the first step is the same as the reactant of the second step.

Figure B-2

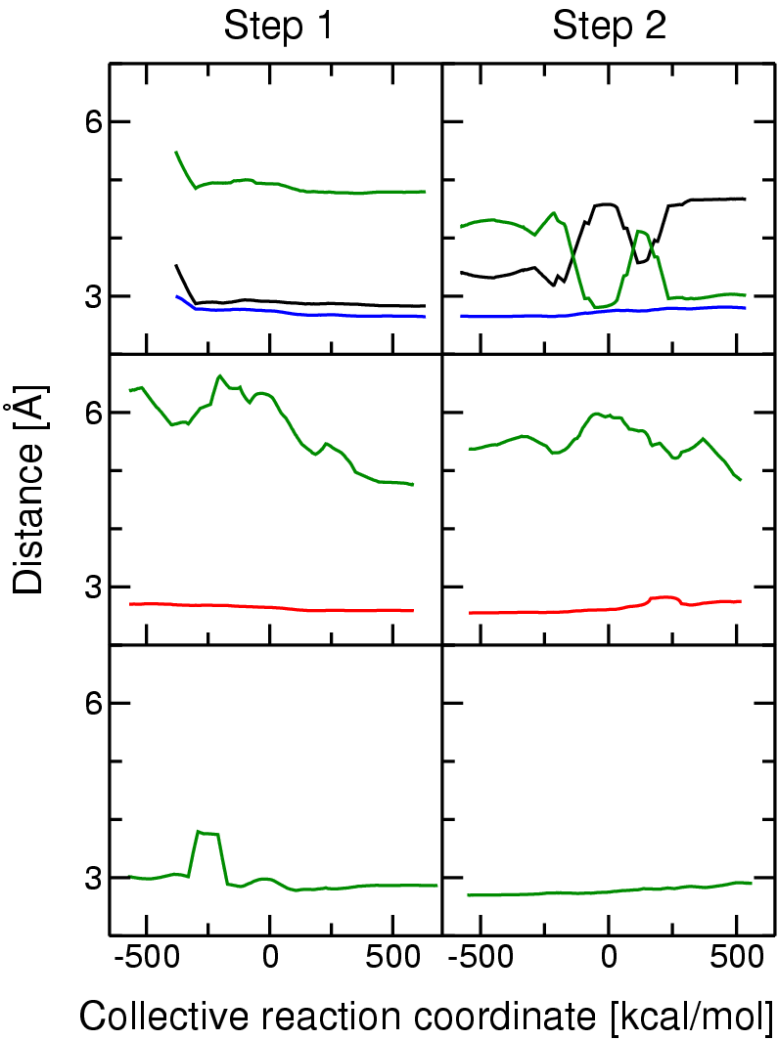


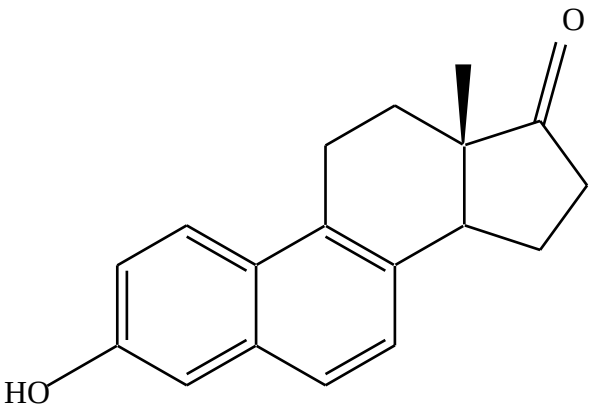
Figure B-2: Thermally averaged distances calculated along the collective reaction coordinate for the first (left) and second (right) proton transfer reactions catalyzed by KSI, as obtained from the second independent data set. The results for the D99L, Y14F, and Y14F/D99L mutants are given from top to bottom. The thermally averaged hydrogen bond donor-acceptor distances between the substrate O3 atom and Tyr14 (blue), between the substrate O3 atom and Asp99 (red), between the substrate O3 atom and Tyr55 (green), and between Tyr14 and Tyr55 (black) are shown. The distances are not shown for the mutated residue. The color coding is the same as depicted in Figure 4-3.

Appendix C

Substrate Analogues

Figure C-1

(a)



(b)

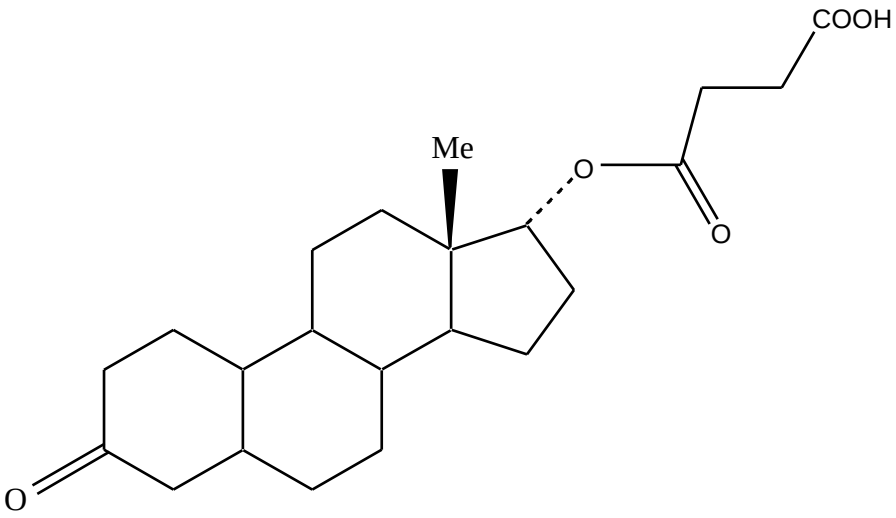


Figure C-1: Substrate analogues used in crystal structures (a) Equilenin and (b) 19-nortestosterone–hemisuccinate.

Appendix D

List of Abbreviations

5-AND: Δ^5-androstene-3,17-dione (substrate)

4-AND: Δ^4-androstene-3,17-dione (product)

DFT: density functional theory

DHFR: dihydrofolate reductase

EQU: equilenin

EVB: empirical valence bond

KSI: ketosteroid isomerase

MD: molecular dynamics

PMF: potential of mean force

QCP: quantized classical path

QM/MM: quantum mechanical/molecular mechanical

RMSD: root mean square deviation

RMSF: root mean square fluctuation

TI: *Commamonas testosteroni*

TST: transition state theory

UI: umbrella integration

VMD: visual molecular dynamics

WHAM: weighted histogram analysis method

WT: wild-type

Vita

Dhruva Kumar Chakravorty

EDUCATION

Ph.D., Physical/Computational Chemistry May 2010

The Pennsylvania State University, University Park, PA

Advisor: Professor Sharon Hammes-Schiffer

M.S., Physical Chemistry Spring 2004

Delhi University, Delhi, India

B.S., Chemistry (with Honors) Spring 2002

St. Stephen's College, Delhi University, Delhi, India

Publications

1. D.K. Chakravorty, M. Kumarasiri, A.V. Soudackov, and S. Hammes-Schiffer, "Implementation of Umbrella Integration within the Framework of the Empirical Valence Bond Approach." *J. Chem. Theory Comput.* **4,** 1974-1980 (2008)
2. D.K. Chakravorty, A.V. Soudackov, and S. Hammes-Schiffer, "Hybrid Quantum/Classical Molecular Dynamics simulations of the Proton Transfer Reactions catalyzed by Ketosteroid Isomerase: Analysis of Hydrogen Bonding, Conformational Motions, and Electrostatics." *Biochemistry* **48,** 10608 (2009).
3. D.K. Chakravorty, and S. Hammes-Schiffer, **"**Impact of Mutation on Proton Transfer Reaction in Ketosteroid Isomerase: Insights from Molecular Dynamics Simulations" *J. of Amer. Chem. Soc.* (Submitted).
