

Proton Transfer and Hydrogen Bonding in Chemical and Biological Systems: A Force Field Approach

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch–Naturwissenschaftlichen Fakultät

der Universität Basel

von

Jing Huang

aus Jiangxi, China

Basel, 2011

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät

auf Antrag von:

Prof. Dr. Markus Meuwly

Prof. Dr. Helmut Grubmüller

Basel, den 21. June 2011

Prof. Dr. M. Spiess (Dekan)

Acknowledgment

First of all I would like to thank Prof. Markus Meuwly. It's a fortune to have you as my PhD supervisor. I have learned a lot - both academically and socially - from you.

I would also like to thank Prof. Helmut Grubmüller who kindly accepted to act as the external examiner.

I would like to thank all the current and past members of the Meuwly group. It has been a pleasure to work in such a friendly group. Special thanks go to Dr. Franziska Schmids and Dr. Stephan Lutz for their kind help when I start my PhD, Dr. Myung Won Lee for keeping me not alone when working in the lab through midnight, Lixian Zhang, Dr. Michael Devereux and Dr. Nuria Plattner for all the stimulating discussions. One central task of my PhD study is to develop MMPT module in CHARMM and to apply it to chemical and biological systems. So I would like to thank all the developers who contribute to the MMPT code: Dr. Sven Lammers, Dr. Stephan Lutz, Dr. Yonggang Yang and Prof. Meuwly.

I am very grateful to everyone who helps me during my PhD study in Basel, especially secretaries in the Department of Physical Chemistry, Ms. Daniela Tischhauser, Ms. Maya Greuter and Ms. Esther Stalder.

Most importantly, I wish to thank my father and mother. You are always there supporting me no matter what.

Above all, this thesis is dedicated to my beloved wife, Cuixian Zhang, for her constant love and support.

Abstract

Proton transfer and hydrogen bonds are fundamental for the function, stability, structure and dynamics of chemically and biologically relevant systems. Hydrogen bonds can be regarded as incipient proton transfer reactions, so theoretically they can be described in unitary way. Here, a molecular mechanics force field approach is pursued. This is a computationally efficient method, so it can be used to study the strong hydrogen bonding and proton transfer processes in large and complex systems.

After reviewing experimental and theoretical methods for proton transfer processes, especially currently available simulation techniques, our method "MMPT" will be presented in detail. The implementation of the code and its validation is discussed. MMPT has been applied to a variety of diverse systems including organic molecules, proteins, and transition metal complexes, and the results will be presented in Chapter 3 - 5. And finally an extension of MMPT to allow proton transport, together with future plans for MMPT development, will be discussed.

Contents

Acknowledgment	i
Abstract	iii
Abbreviations	ix
Units	x
1 Introduction	1
1.1 Background	1
1.2 Current Simulation Methods	4
1.2.1 <i>Ab initio</i> MD Simulations	4
1.2.2 QM/MM Methods	7
1.2.3 Force Field Approaches	8
2 Molecular Mechanics with Proton Transfer	13
2.1 Conceptual Ideas	13
2.1.1 Classical Force Fields	13
2.1.2 Explicit 3D MMPT PES	15
2.1.3 Functional Forms of MMPT Potentials	16
2.1.4 Switching Interactions during PT	20
2.2 Implementation and Validation	22
2.2.1 Implementation into CHARMM	22
2.2.2 Code Structure	22
2.2.3 Validation	23
2.2.4 Limitations	27
2.3 Parameters in the MMPT Potential	28
2.3.1 Zeroth-order “SDM” PES	29
2.3.2 Zeroth-order “LPE” PES	30

3	Applications I: Proton Dynamics in Acetylacetone	37
3.1	Background	37
3.2	Methods	39
3.2.1	Intermolecular Interactions and PES morphing	39
3.2.2	Force Fields for Solvents	41
3.2.3	Generation and Analysis of MD Trajectories	43
3.2.4	Quantum Mechanical Effects	44
3.3	Results	46
3.3.1	Equilibrium Structure	46
3.3.2	IR Spectrum	47
3.3.3	Proton Dynamics and MM Proton Hopping Rate	49
3.3.4	QM Effects of Proton Dynamics	52
3.4	Discussion	54
4	Applications II: Hydrogen Bonds and NMR Couplings in Proteins	57
4.1	Explicit Hydrogen-bond Potentials and their Application to NMR Scalar Couplings in Proteins	57
4.2	Force Field Refinement from NMR Scalar Couplings	68
4.3	Discussion	81
5	Applications III: Molecular Modelling of a Platinum Catalyst	83
5.1	Background	83
5.2	Force Fields for Transition Metal Complexes	85
5.3	Combining MMPT and VALBOND	87
5.3.1	Intermolecular Interactions	87
5.3.2	Force Field Parametrization	88
5.3.3	Computational Methods	89
5.4	Results	91
5.4.1	Equilibrium Structure	91
5.4.2	Potential Energy Surface	91
5.4.3	IR Spectrum	95
5.4.4	Proton Transfer Dynamics	96
5.4.5	UV-Vis Spectrum	98
5.4.6	NMR Spectrum	99
5.5	Discussion	103

6	Extension of MMPT to Study Proton Transport	105
6.1	Basic ideas	105
6.2	Test Systems	106
6.2.1	Proton Transport in Protonated Water Clusters	106
6.2.2	Proton Transfer along Water Chains in Carbon Nanotubes	107
6.3	Schemes and Implementations	108
6.3.1	Identifying the MMPT Motif	108
6.3.2	Smoothing the Dynamics	115
7	Conclusion and Outlook	119
	References	121
	Appendix	153

Abbreviations

CCSD(T)	Coupled-Cluster with single, double and perturbative triple excitations
CHARMM	Chemistry at HARvard Molecular Mechanics
DFT	Density functional theory
ff	force field
IR	infrared
LJ	Lennard-Jones
MD	Molecular Dynamics
MM	Molecular Mechanics
MMPT	Molecular Mechanics with Proton Transfer
MP2	Møller-Plesset second order
NPT	Isothermal-isobaric ensemble
NVE	Microcanonical ensemble
NVT	Canonical ensemble
PT	Proton Transfer
QM	Quantum Mechanics
VdW	Van der Waals

Units

Å	Ångström
nm	Nanometer
e	Elementary charge
Hz	Herz
K	Kelvin
kcal	Kilocalorie
mol	Mole
fs	Femtosecond
ps	Picosecond
ns	Nanosecond
cm^{-1}	Wavenumber
ppm	Parts per million

1 Introduction

"If we were to name the most powerful assumption of all, which leads one on and on in an attempt to understand life, it is that all things are made of atoms, and that everything that living things do can be understood in terms of the jiggings and wiggings of atoms."

Richard Feynman, *The Feynman Lectures on Physics*, 1970

1.1 Background

Water is unique. It has high freezing and boiling points, smaller density in its solid phase (ice) than liquid phase, high heat capacity, and high surface tension. All these properties, which are essential to life on earth, are due to the capability of water molecules to form hydrogen bonds (H-bonds). The electronic structure of water is featured by two "lone pair" electrons, which can interact with the hydrogen atoms in another water molecule, forming O-H \cdots O hydrogen bonds. Unlike ionic bonds constructed by electrostatic interactions, hydrogen bonds are highly directional and prefer near linearity. The strength of hydrogen bonds in water is about 5 kcal/mol, only one order of magnitude greater than the thermal fluctuation k_bT (0.6 kcal/mol) at 300K, so these hydrogen bonds are highly dynamic. If one considers liquid water as a three-dimensional hydrogen-bonded network of H₂O molecules, defects, such as the existence of excess protons, are substantial to this network.

The mobility of protons in water is at least 4.5 times faster than any other cations at room temperature, which can be understood as a consequence of proton transfer along the hydrogen bonds in water. Protons migrate through the H-bond networks as topological defects, which is much faster than the diffusion of individual atoms. This phenomena has been already noticed and explained in similar terms by Grotthuss more than 200 years ago, without even knowing the exact molecular structure of water (he considered water as OH). He wrote "all the molecules of the liquid situated

in this circle would be decomposed and instantly recomposed" in a report entitled "Sur la décomposition de l'eau et des corps qu'elle tient en dissolution à l'aide de l'électricité galvanique" published in Paris in 1806.¹

Although "Grotthuss mechanism" is physically insightful, people still try to understand the detailed molecular mechanism of this fundamental process. In 1925 Hund suggested from semiempirical calculations that the protonated water molecule (H_3O^+) can fit very well into the H-bond network in water, and can be stabilized by such a network.² In 1960's, two elementary structures were proposed by Eigen *et al*³ and Zundel *et al*.⁴, respectively. The Eigen complex, H_9O_4^+ , can be considered as a protonated water molecule (H_3O^+) solvated by accepting three hydrogen-bonded water molecules (first solvation shell), and the proton can "hop" to any of these three water molecules. In the Zundel complex, H_5O_2^+ , one proton is shared equally by two H_2O molecules so the proton can transfer "barrierlessly" between them. These two prototype conformations of protonated water are helpful, but not sufficient for us to understand the PT process in water because more H_2O molecules must be considered. Actually, it has been known for a long time^{5,6} that small changes in the O–O distance can largely impact the proton shuttling along the respective hydrogen bonds. Let's consider Zundel complex H_5O_2^+ , *ab initio* calculations at the CCSD(T)/6-311++G(d,p) level show that in its optimized structure, the proton stays in the middle of the two oxygen atoms, *i.e.*, in a one-dimensional single minimum potential. But if the O–O distances are elongated a little from 2.384 Å to 2.6 Å, then a barrier (0.1 kcal/mol) emerges and the proton dynamics should be described with a double minimum potential. This illustrates that PT process is generally very sensitive to its micro-environment.

After years of investigations and discussions,^{7–23} the current accepted picture of PT in water can be summarized as the follows: the protonated water molecule H_3O_a^+ is surrounded by three hydrogen-bonded water molecules, forming an Eigen-like conformation. But such an Eigen complex is not symmetric, *i.e.*, the three water molecules in the first solvation shell are not equal. One water molecule H_2O_b has a shorter hydrogen bond to the O_a atom than the other two, forming a Zundel-like H_5O_2^+ conformation in a distorted Eigen structure. The identity of H_2O_b changes within the three water molecules in the first solvation shell without actual PT occurrence, which is named "special pair (SP) dance" by Agmon and Voth and happens on an average of 40 fs. The outcome of such SP-dance process is the randomization of the proton hop direction so the proton mobility is diffusive and uncorrelated. The real PT happens

when fluctuations in H-bond network cause a hydrogen bond between O_b and one water molecule in the second solvation shell to break and thus reduce the coordination number of H_2O_b from 4 to 3. The SP is transformed now to a Zundel complex $H_2O_a-H^+-O_bH_2$, in which the proton may rattle many times between O_a and O_b , and finally a new Eigen complex centered at $H_3O_b^+$ may form. The rate-limiting step is the hydrogen bond cleavage between the first and second solvation shell, which happens on the time scale of several picoseconds. During the whole process, neither a large delocalization of proton⁷ nor proton tunneling⁹ is likely to happen.

Hydrogen bonding and proton transfer are not only between water molecules or between oxygen atoms. The donor and acceptor can be any two atoms with large negativity, such as nitrogen or fluorine atoms. Meuwly systematically studied protonated ammonia clusters and protonated ammonia chains where proton transfer happens between nitrogen atoms.²⁴⁻²⁸ Understanding the proton transfer mechanism in polymer materials such as Nafion²⁹ helps design better hydrogen fuel cells.³⁰

Recent advance in molecular biology and structural biology has also highlight the importance of PT in biological systems.³¹ It serves as a fast and common means to transport charges. The proton gradient established by transmembrane proton pumps such as bacteriorhodopsin (bR)^{32,33} and cytochrome c oxidase (CcO),^{34,35} allows energy transduction in living cells. Another example is the highly selective proton-conducting M2 channel in influenza A virus, which plays a key role in the viral life cycle by allowing proton flux from endoplasm into the virion and thus enabling the uncoating and releasing of the viral RNA into the host cell.^{36,37} Proton transfer also participates in numerous enzymatic catalysis - often as an elementary step - in living systems.^{38,39} Examples are the proton transfer in liver alcohol dehydrogenase (LADH)⁴⁰⁻⁴³ and dihydrofolate reductase (DHFR).⁴⁴⁻⁴⁶

Direct investigation of proton transfer dynamics is very difficult experimentally. The most direct evidence for PT to occur are splitting of spectral features.⁴⁷ However, this is only available from high-resolution spectroscopy in the gas phase so only small molecules with intermolecular proton transfer are measured.⁴⁸⁻⁵¹ Another method is to examine the tunneling contribution by proton transfer in the H/D kinetic isotope effect (KIE).⁵²⁻⁵⁴ Nuclear magnetic resonance (NMR) are often used to measure the rate constants and thus to obtain KIEs. With such techniques Limbach *et al* have studied proton transfer thermodynamics and kinetics in a number of complexes,⁵⁵⁻⁵⁸ in which one of the donor or acceptor atoms has to be nitrogen for NMR activity.

Recently, high-resolution vibrational spectra has been recorded for protonated clus-

ters. In 2003 Asmis *et al.* reported the vibrational spectrum of the shared proton in the protonated dimer.⁵⁹ Johnson *et al.* use vibrational predissociation spectroscopy to characterize the spectral signature of cluster ions, such as protonated water clusters $\text{H}^+(\text{H}_2\text{O})_n$ ($n=2-11$),⁶⁰ $\text{OH}^-\cdot(\text{H}_2\text{O})_n$ ($n=1-5$),⁶¹ $[(\text{CO}_2)_n(\text{H}_2\text{O})]^-$ ($n=2-10$)⁶² and $\text{NO}_2^-\cdot\text{H}_2\text{O}$.⁶³ In 2010, Duncan *et al.*⁶⁴ reported photodissociation infrared spectra of water clusters $\text{H}^+(\text{H}_2\text{O})_n$ ($n = 2-5$) with and without Ar tagging. Time-resolved vibrational spectroscopy^{65,66} and photoelectron spectroscopy^{67,68} are also used to study the proton transfer mechanism in liquid water. The analysis of these experimental spectra often relied on the aid of theoretical calculations. The vibrational transitions associated with intermolecular proton transfer or “low-barrier” hydrogen bonds are very sensitive to the chemical environment, and the infrared signatures exhibit a very diffuse character.^{69,70} This make computational simulations, especially molecular dynamics (MD) simulations valuable for assigning spectra and obtaining energetics and dynamics of the studied system.⁷¹⁻⁷³ In the next section, some of these simulation techniques available for proton transfer are reviewed.

1.2 Current Simulation Methods

1.2.1 *Ab initio* MD Simulations

The basic idea of all molecular dynamics simulations is to propagate particles in the system according to the Newton’s second law

$$M_I \ddot{\mathbf{R}}_I(t) = -\nabla_I \Phi(\mathbf{R}_I) \quad (1.1)$$

where M_I and \mathbf{R}_I are the mass and coordinates of particle I. The potential $\Phi(\mathbf{R}_I)$ describe how the particles in the simulation interact with each other and with the environment. In classical molecular dynamics, this potential is calculated by means of a force field:

$$\Phi(\mathbf{R}_I) \approx \sum_{I=1}^N V_1(\mathbf{R}_I) + \sum_{I<J}^N V_2(\mathbf{R}_I, \mathbf{R}_J) + \sum_{I<J<K}^N V_3(\mathbf{R}_I, \mathbf{R}_J, \mathbf{R}_K) + \dots \quad (1.2)$$

, which will be discussed later in more detail .

For *ab initio* molecular dynamics, $\Phi(\mathbf{R}_I)$ is computed from electronic structure calculations that are performed “on-the-fly” as the MD trajectory is generated. Within

the framework of *ab initio* MD, different levels of approximations can be applied. For example, the motion of electrons and motion of nuclei can be solved simultaneously by solving Schrodinger equations 1.3 and Newtonian equations 1.4 at the same time

$$i\hbar \frac{\partial \Psi_0}{\partial t} = H_e \Psi_0 \quad (1.3)$$

$$M_I \ddot{\mathbf{R}}_I(t) = -\nabla_I \langle \Psi_0 | H_e | \Psi_0 \rangle \quad (1.4)$$

where H_e is the electronic Hamiltonian and Ψ_0 represents the ground state wave functions, which leads to the Ehrenfest MD.⁷⁴ However, in order to follow the electrons, time step in Ehrenfest MD simulations has to be set to a very small value (0.01 fs or even less), which largely limits the accessible timescale by this technique. On the other hand, we can also completely decouple the degrees of freedom of electrons and of nuclei based on the Born-Oppenheimer approximation,⁷⁵ and only time-independent Schrodinger equation has to be solved for the electronic degrees of freedom.

$$E_0 \Psi_0 = H_e \Psi_0 \quad (1.5)$$

$$M_I \ddot{\mathbf{R}}_I(t) = -\nabla_I \min_{\psi_0} \langle \Psi_0 | H_e | \Psi_0 \rangle \quad (1.6)$$

where E_0 is the lowest eigenenergy.

In such Born-Oppenheimer MD approaches, a much larger time step (typically 1-2 fs) can be allowed. The CPU-time-consuming electronic structure calculations can be done at different levels of methods, usually density functional theory (DFT),^{24,76-78} but also post-Hartree-Fock (HF) methods such as MP2.⁷⁸ Due to the computational cost of *ab initio* methods, simulations are usually within tens of picosecond. In order to extend MD trajectories to longer time scales, semi-empirical methods such as Self-Consistent Charge Density Functional Tight Binding (SCC-DFTB)⁷⁹ are used. SCC-DFTB is derived from DFT by a second order expansion of the total energy with respect to the charge density, and achieves good balance between accuracy and efficiency.^{24,80,81} SCC-DFTB/MD methods have been used to study the proton transfer in protonated water clusters⁸², amino wires^{24,25} and 2-pyridone·2-hydroxypyridine (2PY2HP) dimer.⁷¹

Born-Oppenheimer MD simulations are carried out strictly on the Born-Oppenheimer surface. To go beyond the Born-Oppenheimer approximation and to effectively couple the nuclear and electronic motions, Car and Parrinello postulated a generalized

purely classical Lagrangian for both electronic and nuclear degrees of freedom⁸³

$$L_{CP} = \sum_I \frac{1}{2} M_I \dot{R}_I^2 + \sum_i \frac{1}{2} \mu_i \langle \dot{\psi}_i | \dot{\psi}_i \rangle - \langle \Psi_0 | H_e | \Psi_0 \rangle + \sum_{ij} \Lambda_{ij} (\langle \psi_i | \psi_j \rangle - \delta_{ij}) \quad (1.7)$$

, where fictitious masses μ_i and fictitious kinetic energies (temperature) $\sum_i \frac{1}{2} \mu_i \langle \dot{\psi}_i | \dot{\psi}_i \rangle$ are associated with the electronic degrees of freedom and Λ_{ij} 's are the Lagrangian multipliers for the wavefunctions' orthonormality $\langle \psi_i | \psi_j \rangle = \delta_{ij}$.

The equations of motion in the Car-Parrinello MD (CPMD) method can be derived from Lagrangian 1.7 as

$$M_I \ddot{\mathbf{R}}_I(t) = -\nabla_I \psi_0 \langle \Psi_0 | H_e | \Psi_0 \rangle \quad (1.8)$$

$$\mu \ddot{\psi}_i(t) = -H_e \psi_i + \sum_j \Lambda_{ij} \psi_j \quad (1.9)$$

The time step used in CPMD is typically 0.1 - 0.25 fs (4 - 10 a.u. approximately), and simulations can reach the timescale of nanoseconds. Shortly after the introduction and implementation of CPMD, it was used to simulate the transport of an "excess" proton in water as well as a OH^- ion which can be considered as a proton hole.¹⁰⁻¹³ Recently, CPMD simulations have been used to study proton transfer in more complex, biologically inspired systems, such as PT between guanine and water⁸⁴ and between nicotine and water.⁸⁵

In the above described *ab initio* simulations, nuclei are treated as pure classical particles with Newtonian dynamics. To consider their quantum nature, the nuclear degrees of freedom can be described using Feynman's path integral formulation of quantum statistical mechanics,⁸⁶ which can be further combined with electronic structure methods in a suitable way.⁸⁷ This leads to the path integral Car-Parrinello molecular dynamics (PICPMD) method,^{87,88} which has been used to study PT in water by Parrinello and collaborators.^{20,21} Tuckerman recently carried out PICPMD simulations to study the proton transport in some common acid membranes employed in hydrogen fuel cells.^{89,90}

One concern regarding *ab initio* MD methods is the accuracy. Since the electronic structure calculations are the most time-consuming part at each time step, they are usually carried out at relatively low levels of theory and with relatively small basis sets. Recently, Lee and Tuckerman reported CPMD simulation of pure liquid water in a complete basis set limit^{91,92} and found substantial difference in the calculated

structural and dynamical properties of water compared with early CPMD simulations with “incomplete” plane-wave basis sets with cutoffs.^{93,94} Although still three fold underestimated, the computed diffusion coefficient with the “complete” basis set was much closer to the experimental value. In general, CPMD simulations assume too structured and too slowly diffusing water, which might be due to the generalized gradient approximations (GGA) employed in the DFT methods of standard CPMD simulations.^{95,96} For semiempirical methods such as SCC-DFTB, the accuracy of the additional parameterizations is also questionable.^{80,97}

Another concern for *ab initio* MD methods is their demanding computational cost. To overcome this drawback, *ab initio* MD and classical MD can be coupled. Such quantum mechanics/molecular mechanics (QM/MM) methods are often used to study chemical reactions in biological systems, such as enzyme catalysis.

1.2.2 QM/MM Methods

In the QM/MM approach,⁹⁸ a localized region, which includes the “active site” where chemical reactions occur, is treated by QM methods and the rest of the system is modeled by MM force fields. Motions of atoms in the QM region can be described by any of the *ab initio* MD methods discussed in section 1.2.1, for example Born-Oppenheimer MD with PM6^{99,100} or SCC-DFTB,^{82,101} as well as CPMD methods.^{102,103} Marx *et al* and Cui *et al* studied the structures, dynamics and infrared spectra of protonated water networks in bacteriorhodopsin with CPMD/MM^{104–106} and SCC-DFTB/MM simulations,^{107,108} respectively. Cui also applied SCC-DFTB/MM methods to investigate the proton transfer processes in carbonic anhydrase (CA)^{109,110}, CcO¹¹¹ and LADH.⁴¹

A technical challenge of QM/MM methods is to correctly deal with the interactions across the boundary between the QM region and the MM region.¹¹² Usually, the separation of atoms into two regions remains constant during a MD simulation. But for proton transfer and transport involving water, this might not be valid anymore. Water is highly diffusive and a water molecule can easily move from the QM region to the MM region and *vice versa*. So a certain water molecule must be able to be switched between a “QM water” and an “MM water”, which is challenging to the definition and treatment of the QM/MM boundary. Besides water molecules, the proton itself might also travel through a relatively large area (consider a proton transport through an entire membrane protein). In currently available QM/MM methods,

the QM region is always fixed as defined at the beginning of the simulation, and cannot move along with the travelling proton. So often a rather large QM region should be assumed, which makes the computational demands of the QM calculations hardly affordable. This limits the applications of QM/MM approaches to systems in which the excess proton is solvated and transported in a fairly small region, which is the case in bR.^{104–108} It must be pointed out that these complications are technical, they can be - and they are being - addressed.^{112–117}

1.2.3 Force Field Approaches

MS-EVB

A well-established force field based approach to simulate proton transfer and transport is multistate empirical valence bond (MS-EVB) method, developed mostly in Voth's group^{17–19,118,119} but also independently in Borgis's group.^{120,121} MS-EVB is in general an extension of the empirical valence bond (EVB) originally proposed by Warshel.^{122,123} The EVB method considers a bond breaking/formation reaction as the system goes from one state to the other, and its potential energy V can be composed from the potential V_{11} of the reactant-like state and the potential V_{22} of the product-like state as

$$V = \frac{1}{2}[V_{11} + V_{22} - \sqrt{(V_{11} - V_{22})^2 + 4V_{12}^2}] \quad (1.10)$$

, which is actually the lowest eigenvalue of the matrix

$$M = \begin{pmatrix} V_{11} & V_{12} \\ V_{12} & V_{22} \end{pmatrix} \quad (1.11)$$

It should be noted that V_{11} , V_{22} and V_{12} are all coordinate-dependent and usually described by empirical MM force fields. The off-diagonal term V_{12} can be considered as the coupling between the two states and needs to be determined "empirically", usually by fitting to *ab initio* PESs. Such standard (two-state) EVB approaches have been employed to develop reactive force fields for proton transfer in water^{124,125} as well as in other more complicated systems such as acetylacetone.¹²⁶

In principle more states can be included in such a valence-bond-inspired picture, which leads to the multistate EVB method.¹¹⁸ The matrix M in 1.11 becomes $N \times N$

instead of 2×2 if there are in total N states included in the EVB calculation.

$$M = \begin{pmatrix} V_{11} & \cdots & V_{1N} \\ \vdots & \ddots & \vdots \\ V_{N1} & \cdots & V_{NN} \end{pmatrix} \quad (1.12)$$

Again, the lowest eigenvalue of M will be the potential energy of the system. For aqueous PT, it is computationally not feasible to go over all the possible states since N will become quite large ($N=20$ even for the protonated water dimer (H_5O_2^+) in gas phase) and matrix diagonalization scales as $O(N^3)$. Fortunately, most of these states will have negligible contributions in a certain configuration. So instead of including the complete set of states, only those “relevant” states are chosen to construct 1.12. However, whether a state is relevant or not is unknown before the construction and diagonalization of M (it can be known afterwards, from the coefficients in the eigenvector). So certain state-selection rules need to be pre-defined, and in each simulation step these rules are applied to determine how many and which states are included in the EVB calculation, *i.e.*, the multiple states will change their identity “on the fly” in the simulations. Voth and his co-workers found out that 22 states are on average required for describing a single proton solvated in the bulk water under ambient conditions, and sometimes $N=40$ must be allowed in MS-EVB MD simulations to reasonably conserve the total energy.¹⁹ This leads to a picture that the proton is extremely delocalized.

The MS-EVB methods have been subsequently improved since its first implementation by Schmitt and Voth in 1998.¹¹⁸ In its original model, charges have to be scaled to achieve quantitative agreement with *ab initio* results, *e.g.*, the net charge of H_3O^+ has to be reduced from $1.00e$ to $0.76e$, which limits its application to PT in more general systems. Such a charge scaling scheme is eliminated by introducing two new repulsive terms in the second generation model (MS-EVB2).¹⁸ These repulsive terms are further improved in the latest MS-EVB3 model¹⁹ to fix an unphysical small artifact in the oxygen-oxygen radial distribution function (RDF). Also in MS-EVB3 the underlying water potential was improved: two specially parameterized flexible water potentials (SPC/Fw¹²⁷ and qSPC/Fw¹²⁸) are included for classical and quantum simulations, respectively. Most importantly, the state-selection algorithm has been refined to provide better total energy conservation - smaller energy drifts - in the MD simulations.

Whether the total energies in NVE simulations are conserved or not is critical to

any deterministic MD methods. Without energy conservation in NVE MD simulations, the ensembles generated by MD simulations and the statistics sampled from MD trajectories will be problematic. However, the energy conservation of MS-EVB methods is not satisfactory. For NVE simulations of one proton solvated in water molecules with a time step of 1 fs, the total energies decrease with a constant rate of 13.1 kcal/(mol⁻¹·ns⁻¹) and 3.4 kcal/(mol⁻¹·ns⁻¹) for the MS-EVB2 and MS-EVB3 models, respectively (see the Figure 16 in Ref. 19). Such drift rates can be reduced by 50% with a 0.5 fs time step but decreasing the time step beyond that value does not lead to significant improvement.¹⁹ It must be pointed out that although energy conservation has been largely improved over generations of models,^{18,19} in theory MS-EVB approaches cannot conserve the total energy due to the simple fact that the states included in EVB calculations are varying during simulations, which means discontinuities in the potential energy surfaces and hence discontinuities in the forces (derivatives of potentials). Such discontinuities can be smoothed by carefully tuned parameters and state-selection algorithms, but cannot be removed completely under the current theoretical frameworks of MS-EVB methods.

Another difficulty of MS-EVB is parametrization. For PT in bulk water, 29 and 13 independent parameters are required for the potential energy in the MS-EVB3 and MS-EVB2 models, respectively. Some of these parameters, especially those appears in the state couplings $V_{ij(i \neq j)}$, are very difficult to fit since they are not physically observables. Moreover, for a particular system, in order to minimize the discontinuities in potential and thus to have reasonable energy conservation, it is very important to tune the parameters in the state-selection algorithm (cutoff radius, cutoff angle, number of solvation shells, and number of EVB states allowed in each solvation shell).¹⁸ This makes setting up MS-EVB simulations very tedious work.

Regardless of these drawbacks, MS-EVB methods achieve enormous success in investigating proton transfer and transport processes in bulk water,^{14-16,18,19,129} in weak acids,¹³⁰⁻¹³² at interfaces,¹³³⁻¹³⁶ and in biomolecular systems such as CA,¹³⁷ CcO,¹³⁸ and the M2 proton channel in influenza A.^{139,140} New techniques are also introduced into the MS-EVB approaches. For example, all the MS-EVB methods described above cannot treat several protons simultaneously, *i.e.*, in these simulations there can be only one excess proton in the systems. To overcome this, a specialized self-consistent iterative multistate EVB (SCI-MS-EVB) is developed to allow multiexcess protons in simulations.¹¹⁹ This technique has been used to simulate acidic solutions at various concentrations¹⁴¹⁻¹⁴³ and is able to, for example, qualitatively reproduce the experi-

mental IR difference spectrum of 1 M HCl solution.¹⁴³

Q-HOP

Q-HOP (quantum hopping) MD, developed in Helms's group^{144–146}, is another classical force field approach to simulate proton transfer in the condensed phase. It allows stochastic proton hopping during the standard MD (Newtonian) evolution of the systems.¹⁴⁵ The probability of proton hopping is determined by coordinate-dependent functions (*e.g.*, depend on the donor-acceptor distance and the energy difference between the minima at the donor and acceptor sides) that have been carefully parameterized against electronic structure calculations.¹⁴⁴ If a proton transfer event takes place (the computed probability is larger than a random number), then the topology of the system is modified before the simulation continues. This is a very efficient method and has been applied to study proton transfer in water as well as in many complex systems such as gated proton channels, proton pumping proteins and fuel cell membranes.^{147–151}

Q-HOP MD has been implemented into the parallel quantum chemistry package NWChem.¹⁵² In principle multi excess protons can be included in Q-HOP simulations, but in practice the system temperature was reported to be unstable if several protons were allowed to transfer simultaneously.¹⁵⁰ So the Q-HOP method is currently restricted to only one hoppable proton. Q-HOP MD is by nature stochastic, meaning that the dynamics are not carried out on a deterministic "underlying" potential energy surface in the form of Newton's equations. This feature makes it difficult to relate actual physical interactions to the generated dynamics, and additional efforts may be needed to interpret the simulation results.¹⁴⁸

There are several other methods to simulate proton transfer and transport based on force field concepts. For example, Keffer *et al* develop a reactive MD model to study proton transport with a series of *ab initio* computed geometry triggers, and once all the triggers are satisfied proton transfer takes place and the system re-equilibrates before the next step of the simulations.¹⁵³ The ReaxFF reactive force field¹⁵⁴ has also been recently extended to study the proton transfer in glycine and its tautomerization of between the neutral form and the zwitterionic form.¹⁵⁵ However, these methods are not as well-developed as the MS-EVB and Q-HOP methods that are reviewed

above, and have only been tested in a few particular applications.

2 Molecular Mechanics with Proton Transfer

"If I had more time, I would have written you a shorter letter."

Blaise Pascal, *Lettres Provinciales*, 1656

In this chapter the MMPT (Molecular Mechanics with Proton Transfer) module implemented in CHARMM is discussed. First we present the basic idea of such a force field for proton transfer reactions. Then different MMPT PESs are introduced. Next, the code structure, implementation and validation of MMPT are analyzed. Finally the parametrization for prototype PT systems are presented.

2.1 Conceptual Ideas

2.1.1 Classical Force Fields

As illustrated in the formula 1.2, explicit exclusion of all electronic effects into a few coordinate-dependent function terms leads to classical force fields. The functional form of a generic force field is

$$V_{\text{tot}} = V_{\text{bond}} + V_{\text{angle}} + V_{\text{dihe}} + V_{\text{elstat}} + V_{\text{vdW}} \quad (2.1)$$

where each term is a function of the atomic coordinates and is separately parametrized

$$V_{\text{bond}} = \sum K_b (r - r_e)^2 \quad (2.2)$$

$$V_{\text{angle}} = \sum K_\theta (\theta - \theta_e)^2 \quad (2.3)$$

$$V_{\text{dihe}} = \sum K_\phi (1 + \cos(n\phi - \delta)) \quad (2.4)$$

$$V_{\text{elstat}} = \frac{1}{4\pi\epsilon_0} \sum \frac{q_i q_j}{r_{ij}} \quad (2.5)$$

$$V_{\text{vdW}} = \sum \epsilon_{ij} \left[\left(\frac{R_{\text{min},ij}}{r_{ij}} \right)^{12} - \left(\frac{R_{\text{min},ij}}{r_{ij}} \right)^6 \right] \quad (2.6)$$

The first three terms represent the internal “bonded” interactions while the latter two describe the “non-bonded” ones. In these expressions, K are the force constants associated with the respective type of interaction, r_e and θ_e are equilibrium values, n is the periodicity of the dihedral and δ is the phase which determines the location of the maximum. The sums for the bonded terms are carried out over all the pre-defined bonds in the system, and all possible valence and dihedral angles for bonded atoms.

Non-bonded interactions include electrostatic and van der Waals terms which sum over all non-bonded atom pairs. The electrostatic terms 2.5 are described with Coulombic interactions between point charges where q_i and q_j are the atomic partial charges of the atoms i and j involved and ϵ_0 is the vacuum dielectric constant. For the van der Waals terms, a Lennard-Jones potential is used with well depth $\epsilon_{ij} = \sqrt{\epsilon_i \epsilon_j}$ and range $R_{\text{min},ij} = (R_{\text{min},i} + R_{\text{min},j})/2$ at the Lennard-Jones minimum. This interaction captures long range dispersion ($\propto -r^{-6}$) and exchange repulsion ($\propto r^{-12}$) where the exponent of the latter is chosen as doubling that of the former for computational convenience.

Eq. 2.1 constitutes a minimal model for force fields, which were developed with the emphasis to carry out studies of the structure and dynamics of macromolecules^{156–163}. These classical force fields, including CHARMM,^{160,164,165} AMBER,¹⁶¹ OPLS¹⁶² and GROMOS,¹⁶³ provide a compromise between accuracy and speed, and are largely applied in characterizing and sampling conformations of extensive molecular structures. They can be further extended by adding additional terms to Eq. 2.1. Taking CHARMM force field for example, the improper dihedral terms and the Urey-Bradley terms can be used to optimize the fit to vibrational spectra and out-of-plane motions, and a numerical correction term named CMAP has been recently introduced^{166,167} to improve the dynamical and structural properties of protein main chains. Also, cross

terms between different internal coordinates (for example the coupling between bond lengths and angles) can be added, constituting the so-called “Class II” force fields¹⁵⁹ that include MMFF¹⁶⁸ and CFF.¹⁵⁸

2.1.2 Replacing the isotropic interaction in standard force fields with explicit 3D MMPT PES $V(R, r, \theta)$

Let’s consider a general proton transfer or hydrogen bonding motif D–H···A where D is the donor, H is the hydrogen and A is the acceptor atom. In a standard force field with all-atom representation and fixed point charges, the interaction within this motif will be considered as the harmonic bond stretching between D and H, plus the non-bonded interactions between D–A and H–A atom pairs. So the only effective interaction that attracts the H atom to the acceptor is the electrostatics between H and A, which is isotropic although the proton transfer reaction or hydrogen bonding are known to be highly directional. Also, the distance-dependence is not correctly modeled with a reciprocal relationship. Proton transfer reactions, *i.e.*, H atom transferring from donor to acceptor is prohibited because of the harmonic constraint between D and H atoms. It should be noted that the simplest model that allows bonds to break is the Morse potential.¹⁶⁹

CHARMM actually contains an explicit hydrogen bonding term

$$E_{\text{hb}} = \sum \left(\frac{A'}{r_{AD}^i} - \frac{B'}{r_{AD}^j} \right) \cos^m(\theta_{A-H-D}) \cos^n(\theta_{AA-A-H}) \quad (2.7)$$

where A' and B' are force field parameters, i and j are positive integers that are usually set to 12 and 10. AA is another atom connected to the acceptor. $m = (0, 2, 4)$ which depends on the type of donor and $n = (0, 2)$ depends on the acceptor’s atom type. In practice the exponents m and n are usually set to zero, which means that only the radial term in Eq. 2.7 is used.

Formula 2.7 was developed in the early times of CHARMM when the extended-atom (united atom) model was assumed in MD simulations. In the extended-atom model, hydrogen atoms are treated as part of the atoms to which they were bonded, which significantly reduces the system size and also allows larger integration step sizes, and the explicit hydrogen bonding term 2.7 is required. However, comparison with *ab initio* calculations show that such additional H-bond terms cannot capture the energy profile of hydrogen bonds.¹⁷⁰ So in later CHARMM force fields such as

“PARAM19”¹⁷¹, the so-called polar hydrogens (OH and NH) are identified separately and treated explicitly, and hydrogen bonds are described by the Lennard-Jones and electrostatic terms alone. With the advance of computational power and the introduction of the SHAKE algorithm^{172,173} that allows larger time steps, all the hydrogen atoms are included explicitly in simulations nowadays. This eliminates the need to include term 2.7 in the current generations of CHARMM force fields such as CHARMM22 and CHARMM27.

We establish a different approach.^{72,174–176} Generally, the interaction within the D–H···A motif in a standard force field is replaced by a parametrized three-dimensional potential energy surface $V(R, r, \theta)$ where R is the distance between donor and acceptor, r is the distance between donor and H atom, and θ is the angle between R and r . For a three-body interaction, 3 internal coordinates ($3 \times 3-6$) lead to a complete description. The PESs should be used together with a standard force field - here, CHARMM force field¹⁶⁴ - and lead to specialized reactive force fields for proton transfer. The method is named Molecular Mechanics with Proton Transfer, or MMPT for short.

To follow the dynamics of proton transfer reactions and hydrogen bonds in chemical and biological systems with MMPT, such 3D PESs $V(R, r, \theta)$ should be given in analytical forms, and are preferable to be implemented into conventional MM and MD codes such as CHARMM. The implementation of MMPT into CHARMM and its validation will be presented in section 2.2, while the functional forms of current available MMPT potentials are discussed below.

2.1.3 Functional Forms of MMPT Potentials

There are five potential types supported in the current MMPT implementation: SSM¹⁷⁶ (symmetric single minimum), SDM¹⁷⁶ (symmetric double minimum), ASM¹⁷⁶ (asymmetric single minimum), NLM¹⁷⁷ (nonlinear hydrogen bond) and LPE (legendre polynomials expansion). To facilitate the parametrization of MMPT potentials, the internal coordinate r (D–H distance) is replaced by a dimensionless coordinate

$$\rho = (r - r_{\min}) / (R - 2r_{\min}) \quad (2.8)$$

where $r_{\min} = 0.8 \text{ \AA}$ is in principle arbitrary but should be sufficiently small to cover the shortest possible D–H separations to avoid ρ becoming negative.

SSM and SDM are symmetric potentials developed according to the prototype sys-

tem $\text{H}_2\text{O}-\text{H}^+ \cdots \text{OH}_2$ (single minimum) and $\text{H}_3\text{N}-\text{H}^+ \cdots \text{NH}_3$ (double minimum), respectively. The function form is

$$\begin{aligned} V(R, \rho, \theta) = & D_{eq}(R)[1 - \exp(-\beta(R)(\rho - \rho_{eq}(R)))]^2 \\ & + D_{eq}(R)[1 - \exp(-\beta(R)(1 - \rho - \rho_{eq}(R)))]^2 \\ & - D_{eq}(R) - c + k\theta^2 \end{aligned} \quad (2.9)$$

and the parameters in Eq. 2.9 are given by:

$$D_{eq}(R) = p_1(1 - \exp(-p_2R - p_3))^2 + p_4 \quad (2.10)$$

$$\beta(R) = p_5 + p_6R \quad (2.11)$$

$$\rho_{eq}(R) = p_7(\exp(-p_8R) + p_9) \quad (2.12)$$

$$k = p_{10} \quad (2.13)$$

$$c = p_{11} \quad (2.14)$$

The radial and angular parts are fully decoupled in Eq. 2.9 as

$$V(R, \rho, \theta) = V_0(R, \rho) + k\theta^2 \quad (2.15)$$

where $V_0(R, \rho)$ can be regarded as the superposition of two morse potentials whose parameters depend on R and the angular dependence is described with a harmonic potential of θ . SSM and SDM potentials have the same function form while the difference in the PES shapes (single minimum or double minimum) is encoded in different parameterizations, so they can in principle be merged as one type. The discrimination into two separate MMPT PES types is due to historical reasons during the development, and is kept in the current MMPT module.

The ASM PES is developed to describe PT in an asymmetric potential such as the PT in $\text{H}_3\text{N}-\text{H}^+ \cdots \text{OH}_2$. The same assumption is adopted that the potential can be decomposed into a two-dimensional $V_0(R, \rho)$ plus a harmonic potential in θ , and the mathematical formula is given by

$$\begin{aligned} V(R, \rho, \theta) = & D_{eq,1}(R) [1 - \exp(-\beta_1(R)(\rho - \rho_{eq,1}(R)))]^2 \\ & + D_{eq,2}(R) [1 - \exp(-\beta_2(R)(\rho_{eq,2}(R) - \rho))]^2 \\ & - c(R) + k\theta^2 \end{aligned} \quad (2.16)$$

where

$$D_{eq,1}(R) = p_1(1 - \exp(-p_2R - p_3))^2 + p_4 \quad (2.17)$$

$$\beta_1(R) = \frac{p_5}{1 - \exp(-p_6(R - p_7))} \quad (2.18)$$

$$\rho_{eq,1}(R) = p_8(1 - \exp(-p_9(R - p_{10})))^2 + p_{11} \quad (2.19)$$

$$D_{eq,2}(R) = p_{12}(1 - \exp(-p_{13}(R - p_{14})))^2 + p_{15} \quad (2.20)$$

$$\beta_2(R) = \frac{p_{16}}{1 - \exp(-p_{17}(R - p_{18}))} \quad (2.21)$$

$$\rho_{eq,2}(R) = p_{19}(1 - \exp(-p_{20}(R - p_{21})))^2 + p_{22} \quad (2.22)$$

$$c(R) = p_{23}(1 - \exp(-p_{24}(R - p_{25})))^2 + p_{26} \quad (2.23)$$

$$k = p_{27} \quad (2.24)$$

MMPT potentials 2.9 and 2.16 assume the predominate linearity of the D–H–A motif and will not be applicable to nonlinear hydrogen bonds. Yang¹⁷⁷ proposed a MMPT potential “NLM” for nonlinear hydrogen bonds such as the intramolecular H-bond in MA. For this, ρ is re-defined as

$$\rho = (r \cos \theta - r_{\min}) / (R - 2r_{\min}) \quad (2.25)$$

and a new variable $d = r \sin \theta$ is introduced as the perpendicular bending displacement. In terms of the new internal coordinates $\{R, \rho, d\}$, the symmetry between D and A atoms can be expressed as

$$V(R, \rho, d) = V(R, 1 - \rho, d) \quad (2.26)$$

The detailed expression of the “NLM” potential is given by:

$$V(R, \rho, d) = V_0(R, \rho) + V_d(R, \rho, d) \quad (2.27)$$

where the isotropic part $V_0(R, \rho)$ is the same as in Eq. 2.9 and 2.15

$$V(R, \rho, \theta) = D_{eq}(R)[1 - \exp(-\beta(R)(\rho - \rho_{eq}(R)))]^2 + D_{eq}(R)[1 - \exp(-\beta(R)(1 - \rho - \rho_{eq}(R)))]^2 - D_{eq}(R) - c \quad (2.28)$$

with the same parametrization as in Eq. 2.10, 2.11, 2.12 and 2.14. The perpendicular

bending part is formulated as an $\{R, \rho\}$ -dependent harmonic potential

$$V_d(R, \rho, d) = \frac{1}{2}[p_{10}V_0(R, \rho) + p_{12}][d - (p_{14} - p_{13}(\rho - 0.5)^2)]^2 \quad (2.29)$$

Table 2.1: Comparison of MMPT potentials

PES type	formula	# params	prototype system
SSM	2.9	11	$\text{H}_2\text{O}-\text{H}^+ \cdots \text{OH}_2$
SDM	2.9	11	$\text{H}_3\text{N}-\text{H}^+ \cdots \text{NH}_3$
ASM	2.16	27	$\text{H}_3\text{N}-\text{H}^+ \cdots \text{OH}_2$
NLM	2.27, 2.28, 2.29	14	$\text{HO}-\text{CH}=\text{CH}-\text{CH}-\text{O}$
LPE	2.30, 2.32, 2.39	114	$\text{H}_2\text{O}-\text{H}^+ \cdots \text{OH}_2$

All of the above MMPT potentials can be regarded as 2D potentials $V_0(R, \rho)$ plus angular revisions. A more sophisticated PES that explicitly couples all the three degrees of freedom has been developed to study proton transfer processes that involve reorientation, such as the proton shuttling in bulk water. This potential is expressed as an expansion in Legendre polynomials and named "LPE"

$$V(R, \rho, \theta) = \sum_{\lambda=0}^n V_\lambda(R, \rho) P_\lambda(\cos\theta) \quad (2.30)$$

where $P_\lambda(x)$ is the λ th-order Legendre polynomial that can be deduced by the recurrence relation

$$(\lambda + 1)P_{\lambda+1}(x) = (2\lambda + 1)xP_\lambda(x) - \lambda P_{\lambda-1}(x) \quad (2.31)$$

and $P_0(x) = 1, P_1(x) = x$.

The zeroth order $V_0(R, \rho)$ is given by

$$\begin{aligned} V_0(R, \rho) = & a_0(R)[1 - \exp(-a_1(R)(\rho - a_2(R)))]^2 \\ & + a_0(R)[1 - \exp(-a_1(R)(1 - \rho - a_2(R)))]^2 \\ & + a_3(R) \exp[-a_4(R)(\rho - 0.5)^2] - a_5(R) \end{aligned} \quad (2.32)$$

where

$$a_0(R) = p_0(1)\{\tanh[p_0(2)(R - p_0(3))] + p_0(4)\} \quad (2.33)$$

$$a_1(R) = p_1(1)\{\tanh[p_1(2)(R - p_1(3))] + p_1(4)\} \quad (2.34)$$

$$a_2(R) = p_2(1)\{\tanh[p_2(2)(R - p_2(3))] + p_2(4)\} \quad (2.35)$$

$$a_3(R) = p_3(1)\{\tanh[p_3(2)(R - p_3(3))] + p_3(4)\} \quad (2.36)$$

$$a_4(R) = p_4(1)\{\tanh[p_4(2)(R - p_4(3))] + p_4(4)\} \quad (2.37)$$

$$a_5(R) = p_5(1)\{\tanh[p_5(2)(R - p_5(3))] + p_5(4)\} \quad (2.38)$$

And for higher orders, $V_{\lambda \neq 0}(R, \rho)$ is expressed as

$$V_{\lambda}(R, \rho) = b_{0,\lambda} + \frac{b_{1,\lambda}(R)}{b_{2,\lambda}(R)[(\rho - 0.5)^2 + b_{1,\lambda}^2(R)]} \quad (2.39)$$

where

$$b_{0,\lambda} = f_0(\lambda) \quad (2.40)$$

$$b_{1,\lambda}(R) = f_1(\lambda)\{\tanh[f_2(\lambda)(R - f_3(\lambda))] + f_4(\lambda)\} \quad (2.41)$$

$$b_{2,\lambda}(R) = f_5(\lambda) + f_6(\lambda)[R - f_7(\lambda)]^2 + f_8(\lambda)[R - f_7(\lambda)]^4 \quad (2.42)$$

The expansion 2.30 is truncated at $n = 10$ for currently implemented MMPT provisions, which leads to in total 114 parameters. In practice not all the terms need to be “activated”. For example, fitting LPE potential 2.30 to an *ab initio* PES scan of rotating a water molecule around H_3O^+ illustrates that taking only $\lambda = 0, 1, 3$ terms already gives a good approximation

$$V(R, \rho, \theta) = V_0(R, \rho) + V_1(R, \rho) \cos \theta + V_3(R, \rho) \left[\frac{5 \cos^3 \theta - 3 \cos \theta}{2} \right] \quad (2.43)$$

, which reduces the number of actual MMPT parameters for “LPE” potential to 44 since all other parameters in Eq. 2.30 are set to zero. See section 2.3.2 for more details.

2.1.4 Switching related interactions on and off during proton transfer

In CHARMM as well as other standard MD codes, no chemical reaction is allowed. Information on all the bonded and non-bonded interactions is specified (for example

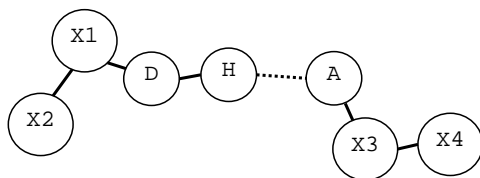


Figure 2.1: Illustration of a general D-H...A motif

in the CHARMM psf file) before the dynamics evolution, and cannot be changed “on the fly”. If two atoms are bonded, they are always bonded. However, proton transfer is a bond breaking/forming process, *i.e.*, the D–H bond breaks and the H–A bond forms. Consequently, the definition of related valence and dihedral angles also need to be revised. To address this problem, MMPT explicitly includes all the interactions on both donor and acceptor side. Bonded and non-bonded terms for the two possible topologies (DH–A or D–HA) are all taken into account in MMPT, and a smooth switching function is applied to tune the interactions on and off (see Table 2.2) depending on the position of the transferring H-atom. The switching function is given by

$$\text{SWF}(R, r, \theta) = \frac{\tanh[2R \cdot (r \cos \theta - 0.5R)] + 1}{2} = \frac{\tanh[2Rr \cos \theta - R^2] + 1}{2} \quad (2.44)$$

which yields 0.5 at the transition state and values close to 1 if H is bound to D and 0 if H is bound to A. For numerical stability, R^2 in Eq. 2.44 is replaced by $0.999997772R^2 + 0.000001203R - 0.000001605$ in the current implementation.

The details of switching force field terms are presented in Table 2.2. X_1 represents all atoms connected to the Donor, and X_2 the atoms connected to X_1 . Similarly, X_3 and X_4 are acceptor antecedents and atoms bonded to the antecedents, as depicted in Fig. 2.1. Three terms in standard CHARMM force field, namely the bonding between D–H and the non-bonded interactions between D–A and H–A atoms, should be removed to avoid double counting of energies because they are already modelled within the 3D MMPT potential $V(R, r, \theta)$. Other interaction terms are either “switched on” or “switched off”. “Switch on” amounts to add this term in force field calculations and to multiply it with $\text{SWF}(R, r, \theta)$, while “switch off” means that this term is already included in standard force field but should be multiplied with $(1 - \text{SWF})$.

Table 2.2: Summary of interactions modified by MMPT. See text for details.

Bond	D-H	remove
Angle	H-D- X_1	switch off
	H-A- X_3	switch on
Dihedral	H-D- X_1 - X_2	switch off
	H-A- X_3 - X_4	switch on
Non-Bonded	D-A	remove
	H-A	remove
	X_1 -H	switch off
	X_2 -H	switch off
	X_3 -H	switch on
	X_4 -H	switch on

2.2 Implementation and Validation

2.2.1 Implementation into CHARMM

MMPT has been implemented as an individual module in the most recent CHARMM version, c36a6 (CHARMM official svn checkout r90). The whole module is encoded as misc/mmpt.src under the CHARMM directory. A documentation file (doc/mmpt.doc) has been provided to instruct the usage of MMPT, while four test cases are prepared and included in test/c36test. Some CHARMM source codes and compilation files have to be slightly modified to allow the inclusion of MMPT energies and forces in MD simulations. The complete list of files added and modified by MMPT is summarized in Table 2.3.

2.2.2 Code Structure

The MMPT module is functionized with two major subroutines: MMPTINIT and EMMPT. MMPTINIT is only called when entering the MMPT module. It reads in the definition of PT motifs (the donor, proton and acceptor atoms, together with the MMPT PES types), and the corresponding parameter files. It also generalizes the list of all interactions that will be removed, added or modified by MMPT. The flowchart of MMPTINIT is presented in Figure 2.2.

EMMPT is called every time a CHARMM ENERGY calculation is carried out. The structure of EMMPT is illustrated in Figure 2.3, where texts in rectangles describe

Table 2.3: Files added or modified by MMPT. The common path is under the CHARMM directory.

Files added by MMPT	Files modified by MMPT
source/misc/mmpt.src	build/UNIX/charmm.mk
doc/mmpt.doc	build/UNIX/energy.mk
test/c36test/mmpt.ma.inp	build/UNIX/misc.mk
test/c36test/mmpt.h5o2p.inp	install.com
test/c36test/mmpt.2py2hp.inp	source/charmm/miscom.src
test/c36test/mmpt.n2h7p.in	source/energy/energy.src
test/data/hbridge.ma.def	source/energy/eutil.src
test/data/mmpt.ma.pdb	source/ltn/energy_ltn.src
test/data/mmpt_nlm_ma.prm	source/misc/mmpt.src
test/data/2py2hp.pdb	
test/data/2py2hp.psf	
test/data/2py2hp.rtf	
test/data/hbridge.2py2hp.def	
test/data/mmpt_sdm_2py2hp_nhn.prm	
test/data/mmpt_sdm_2py2hp_oho.prm	
test/data/h5o2p.rtf	
test/data/h5o2p.par	
test/data/h5o2p.pdb	
test/data/mmpt_ssm_h5o2p.prm	
test/data/hbridge.h5o2p.dat	
test/data/n2h7p.pdb	
test/data/mmpt_sdm_n2h7p.prm	
test/data/hbridge.n2h7p.def	

what the code does while in ellipses subroutines that EMMPT calls are listed. EMMPT returns the modification of potential energies and forces (the derivatives of potential energies) by MMPT.

2.2.3 Validation

As already pointed out in the Introduction, one critical point to validate non-standard force field methods is whether total energy is conserved in NVE simulations. Energy conservation also ensure that the analytical derivatives implemented in the code have the correct functional forms. Here we validate our MMPT module in CHARMM by assessing the energy conservation situations for the following four different MD NVE

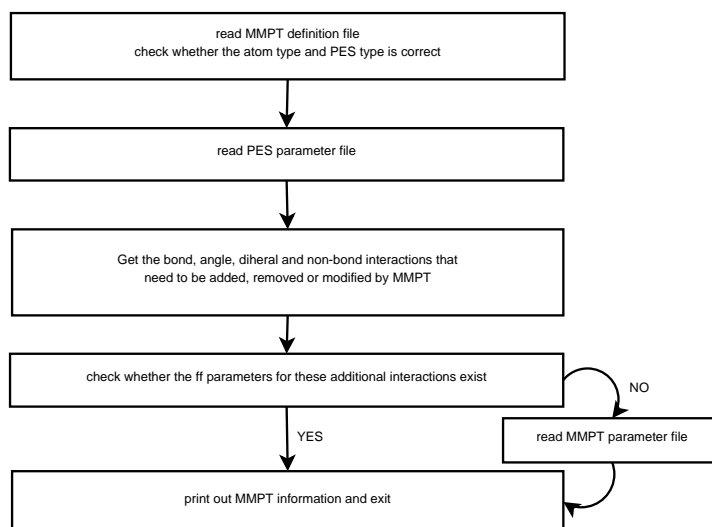


Figure 2.2: code structure of MMPTINIT

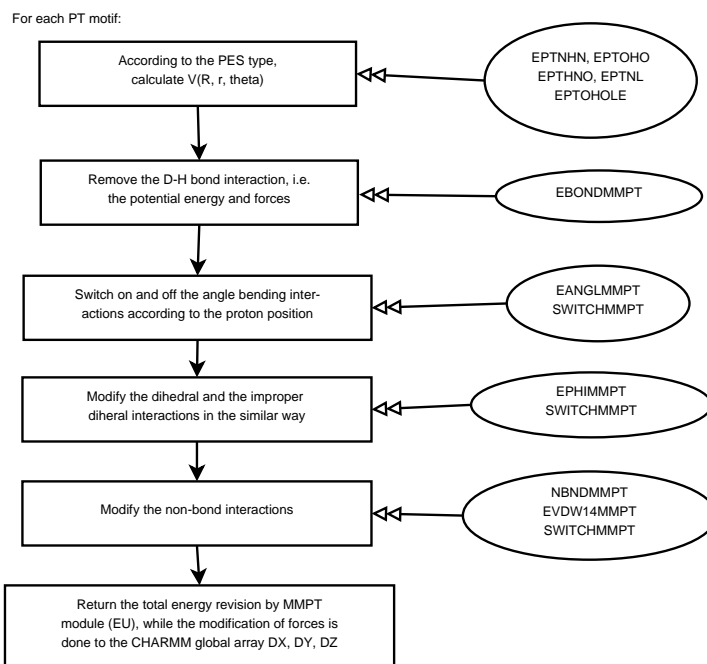


Figure 2.3: code structure of EMMPT

simulations: a) 10 ns simulation of H_5O_2^+ with “SSM” PES* and a time step $\Delta t = 0.1$ fs; b) 10 ns simulation of H_5O_2^+ with “LPE” PES (parameters listed in Table 2.5) and a

*PES parameters: 141.901588, 2.228720, 1.960100, 11.879385, -0.977126, 1.348090, 357.171000, 3.393090, 0.102289, 0.008873, 35.621011.

time step $\Delta t = 0.1$ fs; c) 1 ns simulation of 2PY2HP dimer with two PT motifs treated respectively by “SSM” and “SDM” PES (parameters listed in Table 5.1), and the time step is set to 0.2 fs; and d) 1 ns simulation of ubiquitin (1ubq¹⁷⁸) solvated in a pre-equilibrated TIP3P¹⁷⁹ water box with periodic boundary conditions and a cutoff of 12 Å is applied to the shifted electrostatic and switched Van der Waals interactions, where 29 hydrogen bonding motifs are treated by MMPT with the “ASM” PES[†] and the time step is set to 0.2 fs.

The time evolutions of the total energies in these four MD simulations are plotted in Figure 2.4. Energies are stable and conserved during nanosecond simulations. The fluctuations of total energies along a MD trajectory is computed to characterize quantitatively the energy conservation situations

$$\delta E = E - \langle E \rangle \quad (2.45)$$

where $\langle E \rangle$ is the average of E along the trajectory. Ideally δE should be kept zero during simulations, but numerical errors will render them into a Gaussian distribution. For the four test cases considered here, energy fluctuations in a, b and c obey Gaussian distributions with a width of less than 0.01 kcal/mol (see Figure 2.5), illustrating that the total energies conserve very well. On the other hand, the total energies for the test case d fluctuate within 1 kcal/mol during 1 ns simulation. It should be noted, however, this is a large system containing 16897 atoms. Periodic boundary conditions and cut-off schemes for non-bonded interactions are applied. Standard MD simulations were also carried out by CHARMM with no MMPT module, and resulted in a histogram of δE with similar fluctuation pattern. (see Figure 2.6)

The time steps in above MD simulations are set to 0.1 fs or 0.2 fs. They are shorter than the time steps usually used in standard CHARMM MD simulations (1 fs or 2 fs), but are necessary to follow the proton dynamics. Energy conservation of simulations with larger time steps are also assessed with the test case ‘a’ (10 ns simulation of H_5O_2^+ with “SSM” PES). As shown in Figure 2.7, for larger time step sizes δE still obey Gaussian distributions, but with broader widths. Also, a “long tail” is observed, which skews the distribution to the left. For $\Delta t = 1$ fs, the total energies are stable and conserved well during 10 ns simulation times.

A similar trend is observed if simulation temperature is increased. For all four MD simulations in Figure 2.4 and 2.5, the temperature is set to 300K. This means that be-

[†]PES parameters: 16.079, 1.815, 3.311, 121.346, 4.094, 2.323, 2.815, -0.098, 1.671, 2.743, 0.093, 38.614, 1.355, 3.174, 75.398, 5.500, 1.979, 2.844, 0.08, 1.709, 2.86, 0.769, 63.478, 0.836, 2.913, 60.257, 0.012.

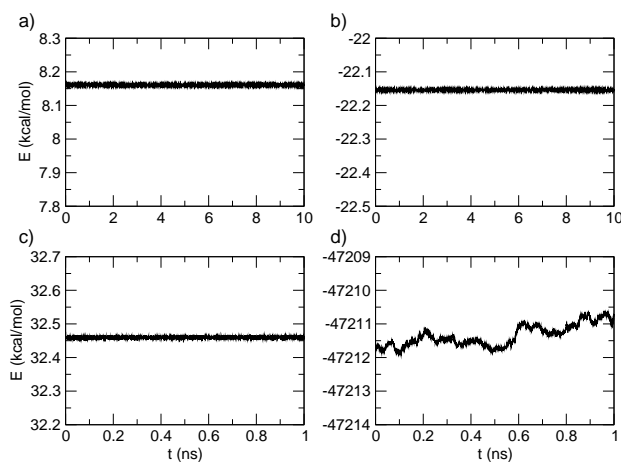


Figure 2.4: Energy conservations of NVE simulations with MMPT. (see text for details)

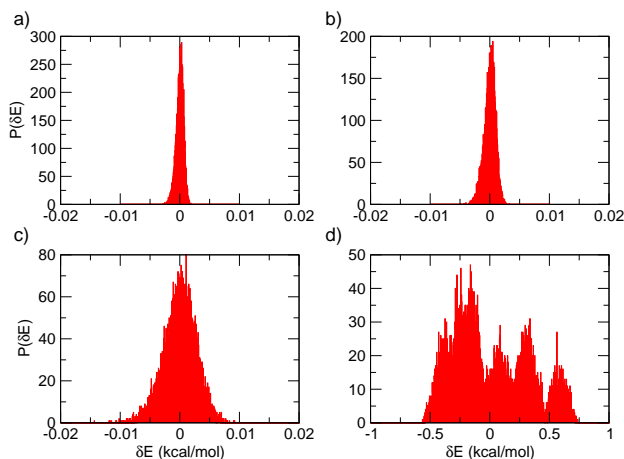


Figure 2.5: Histograms of energy fluctuations δE for NVE simulations with MMPT. (see text for details)

fore free (NVE) dynamics, the systems are first heated to 300K and then equilibrated there for 10^5 steps. The same protocol is applied to the test case b (10 ns simulation of H_5O_2^+ with “LPE” PES) with equilibrated temperature raised to 500, 1000 and 1500 K. The distributions of δE are broadened and slightly distorted Gaussians, as shown

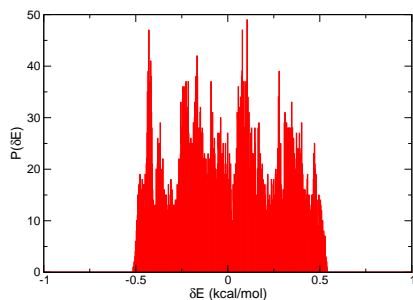


Figure 2.6: Histogram of energy fluctuations δE for NVE simulations of test case (d) with standard CHARMM force field.

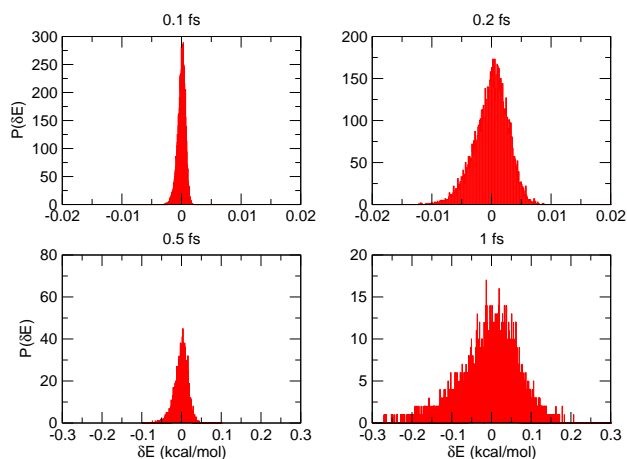


Figure 2.7: Histograms of energy fluctuations δE for NVE simulations of test case (a) with different time step sizes.

in Figure 2.8. In conclusion, MMPT does not deteriorate energy conservation of MD simulations.

2.2.4 Limitations

The current MMPT module in CHARMM is subjected to the following limitations:

- The maximum number of PT motifs treated by MMPT is fixed to 200. MMPT gives an error message if there are more than 200 PT motifs since the corre-

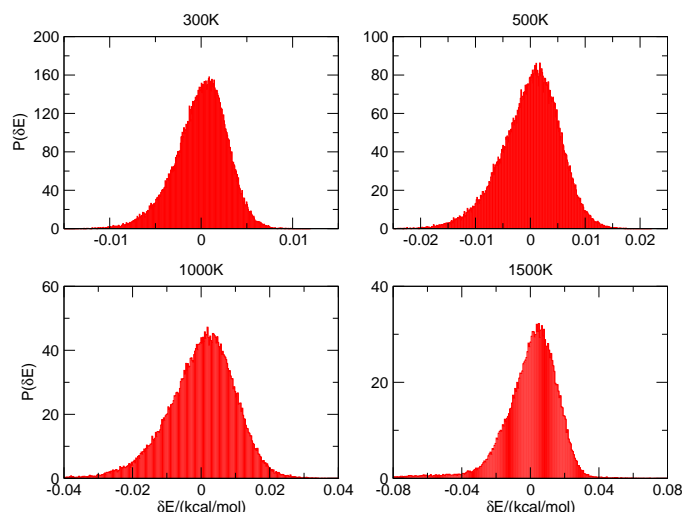


Figure 2.8: Histograms of energy fluctuations δE for NVE simulations of test case (b) under different temperature.

sponding arrays have a fixed size of 200. It would be more convenient to dynamically allocate the corresponding arrays. Alternatively, this can be increased by changing the variable NHBNUM in the subroutine ALLOCFIR if needed.

- Only the cut-off scheme for non-bonded interactions is supported by MMPT. Ewald summation methods¹⁸⁰ such as particle mesh Ewald (PME)^{181,182} is not supported. Also CHARMM allows the electrostatic interactions (Eq. 2.5) to be scaled with different dielectric constants or a distance-dependent dielectric, but this will not be compatible with MMPT.
- Continuous proton transfer, *e.g.* proton shuttling along a water chain, is not possible with current MMPT module. An extension that allows proton transport will be discussed in Chapter 6.

Possibilities to address these limitations, together with plans for future development of MMPT, are discussed in Conclusion and Outlook.

2.3 Parameters in the MMPT Potential

The quality of a force field depends on both its functional form and the corresponding parameters. For MMPT force fields, roughly a dozen parameters are required as listed

in Table 2.1. The general strategy to obtain the MMPT parameters for a particular PT motif is PES “morphing”.^{183–185} This amounts to morphing the zeroth-order PESs of prototype PT systems to adapt their overall shapes to PT or H-bond patterns with similar topology but different energetics. Morphing can be a simple coordinate scaling or a more general coordinate transformation depending on whether the purpose of the study and the experimental data justify such a more elaborate approach.

For this, accurate zeroth-order PESs should be provided. This is done via fitting MMPT PESs to *ab initio* PES of the prototype PT systems that are listed in Table 2.1. An *ab initio* scan is carried out at the MP2/6-311++G(d,p) level with Gaussian03 suite of programs¹⁸⁶, and fitting is performed by I-NoLLS¹⁸⁷, a program for interactive nonlinear least-square fitting. Two examples of parameterizations for MMPT PESs are presented in this section, while the detailed processes of PES morphing will be discussed in the Application chapters.

2.3.1 Zeroth-order “SDM” PES

The prototype system for “SDM” type PES is $\text{H}_3\text{N}-\text{H}^+ \cdots \text{NH}_3$. A set of MMPT parameters was obtained by Lammers^{175,176} by fitting $V_0(R, r)$ to *ab initio* scanning on a regular grid defined by $R \in [2.4 \text{ \AA}, 3.4 \text{ \AA}]$ in increments of 0.1 \AA and $r \in [0.8 \text{ \AA}, R/2]$ in increments of 0.05 \AA . However, on such a PES artificial saddle points exist at $R \sim 2.1 \text{ \AA}$ and $\rho \sim 0.1$ or 0.9 (see Fig. 2.9). In practice this leads to unphysical situations that in MD simulations donor and acceptor atoms can be very close to each other with no hydrogen bonds formed.

A refitting of zeroth-order SDM PES has been performed by taking account of *ab initio* data of $R_{\text{NN}} = 2.0 \text{ \AA}$, $R_{\text{NN}} = 2.2 \text{ \AA}$ and $R_{\text{NN}} = 4.0 \text{ \AA}$. The fitted parameters and corresponding PESs are compared in Table 2.4 and Figure 2.9. The “old set” refers to MMPT parameters reported in Ref. 176 while the “new set” refers to the parameter sets obtained here. The MMPT PES generated with the new parameter set provides an overall high barrier at short D–A distance, which prohibits the system evolving into the region where $R < 2.2 \text{ \AA}$ and $r < 0.85 \text{ \AA}$. The new parameters also provide better a description of the “important” region on the (R, r) space. The standard deviation between MMPT and *ab initio* energies has been computed to evaluate the quality of fitting for 95 such data points ($R \in [2.5 \text{ \AA}, 2.9 \text{ \AA}]$ and $r \in [0.9 \text{ \AA}, R/2]$). The standard deviation is 0.21 kcal/mol for the “new set” compared to 0.33 kcal/mol for the “old set”.

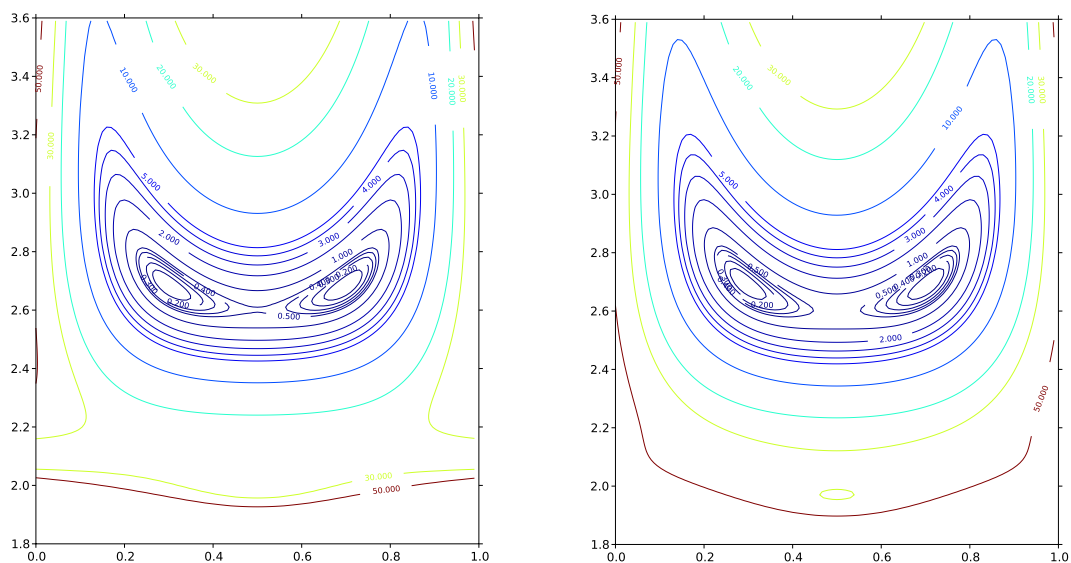


Figure 2.9: Comparison of MMPT SDM PESs generated by the old parameter set (left panel) and the new set (right panel).

	old set	new set
p_1	161.593375	202.404752
p_2	1.936596	1.859491
p_3	2.101121	2.065129
p_4	-0.986537	8.158404
p_5	-0.049717	0.206868
p_6	1.034890	0.863515
p_7	228.382292	210.798809
p_8	2.980991	3.001897
p_9	0.117156	0.109427
p_{10}	0.010000	0.010000
p_{11}	27.446410	40.073220

Table 2.4: Comparison of MMPT parameters

2.3.2 Zeroth-order “LPE” PES

The prototype system for “LPE” type PES is $\text{H}_2\text{O}-\text{H}^+\cdots\text{OH}_2$. *ab initio* data are generated by scanning on a 3D grid defined by $R \in [2.2 \text{ \AA}, 3.2 \text{ \AA}]$ in increments of 0.1 \AA , $r \in [0.8 \text{ \AA}, R/2]$ in increments of 0.05 \AA , and $\theta \in \{11.98^\circ, 27.49^\circ, 43.10^\circ, 58.73^\circ, 74.36^\circ, 90.00^\circ\}$

which are 11th-order Gauss-Legendre quadratures (solution of $P_{11}(\cos \theta) = 0$). For each geometry $\{R, r, \theta\}$ all the remaining degrees of freedom are optimized, which leads to a relaxed 3D *ab initio* PES. With the help of the orthogonality of Legendre polynomials

$$\int_{-1}^1 P_m(x)P_n(x)dx = \frac{2\delta_{mn}}{2n+1} \quad (2.46)$$

, the radial parts $V_\lambda(R, \rho)$ can be effectively decoupled from each other. This allows to first perform fitting to individual $V_\lambda(R, \rho)$ and then use the parameters from 2D fitting as initial values for the overall 3D fitting over all data points. In total there are 2387 *ab initio* data points. All points with a potential energy higher than 35 kcal/mol (respect to the global minimum, the same below) are ignored in the fitting, leaving 1851 points to be included. And those within 15 kcal/mol (824 points) are assigned with large weights in the fitting process. It has been found that the LPE PES can be described very well with three major contributions V_0 , V_1 and V_3 . Including higher order term (V_5 and V_7) does not improve the fitting quality (data not shown). The fitted parameters are listed in Table 2.5. The standard deviation between MMPT and *ab initio* energies is 0.317 kcal/mol for all the data points with potential energy less than 15 kcal/mol, and equals 3.394 kcal/mol and 4.786 kcal/mol for the 1851 points included in fitting and 2387 total *ab initio* points, respectively.

V_0		V_1		V_3			
$p_0(1)$	109.923718	$p_3(1)$	3.684985	$f_0(1)$	1.304358	$f_0(3)$	-2.639399
$p_0(2)$	-1.733356	$p_3(2)$	-6.777204	$f_1(1)$	0.227601	$f_1(3)$	-1.281213
$p_0(3)$	2.192987	$p_3(3)$	2.317678	$f_2(1)$	-0.134110	$f_2(3)$	0.122306
$p_0(4)$	-0.502939	$p_3(4)$	-0.888945	$f_3(1)$	2.964633	$f_3(3)$	-5.412074
$p_1(1)$	1.941748	$p_4(1)$	6.576714	$f_4(1)$	0.959031	$f_4(3)$	-0.928826
$p_1(2)$	0.700997	$p_4(2)$	7.035281	$f_5(1)$	-0.383888	$f_5(3)$	0.172150
$p_1(3)$	2.943756	$p_4(3)$	1.719450	$f_6(1)$	0.090342	$f_6(3)$	-0.047998
$p_1(4)$	1.539175	$p_4(4)$	5.299477	$f_7(1)$	2.747005	$f_7(3)$	2.756022
$p_2(1)$	0.871553	$p_5(1)$	119.454350	$f_8(1)$	-2.713838	$f_8(3)$	1.138408
$p_2(2)$	-1.397980	$p_5(2)$	-1.509945				
$p_2(3)$	1.265474	$p_5(3)$	2.205597				
$p_2(4)$	1.105869	$p_5(4)$	-0.400690				

Table 2.5: Parameters for zeroth-order LPE PES.

In the following the fitted PES $V(R, \rho, \theta)$ is visualized and compared with the *ab initio* calculation results. First we compare $V_0(R, \rho)$, $V_1(R, \rho)$ and $V_3(R, \rho)$ in Figure 2.10, 2.11 and 2.12, respectively. Fitted and *ab initio* computed $V_0(R, \rho)$ are in very

good agreement. Fitted V_1 and V_3 also reproduced qualitatively most feature of the respective *ab initio* scan, with slight differences, for example in the $R = 2.2\text{\AA}$ region. It should be noted that the overall fitted PES only contains V_1 and V_3 components so it is possible that effects from the neglected higher order expansions are included in the fitted $V_1(R, \rho)$ and $V_3(R, \rho)$.

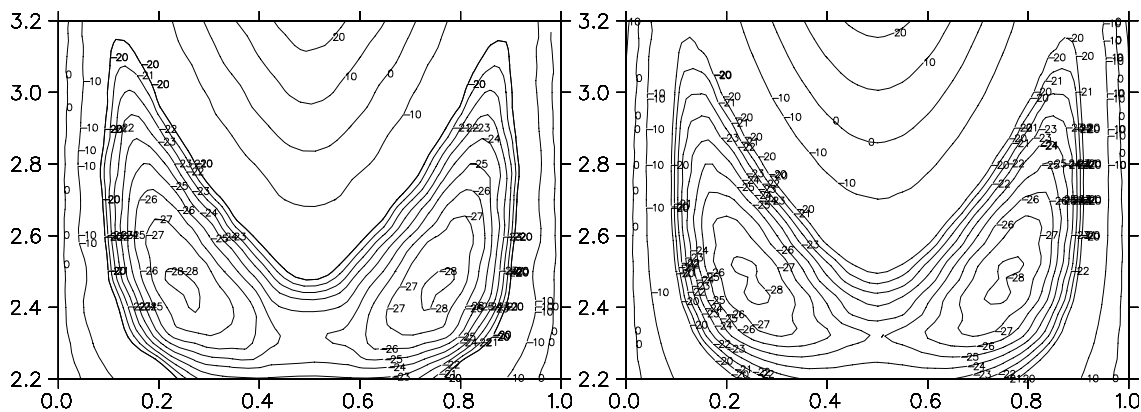


Figure 2.10: Comparison of *ab initio* (left) and fitted (right) $V_0(R, \rho)$.

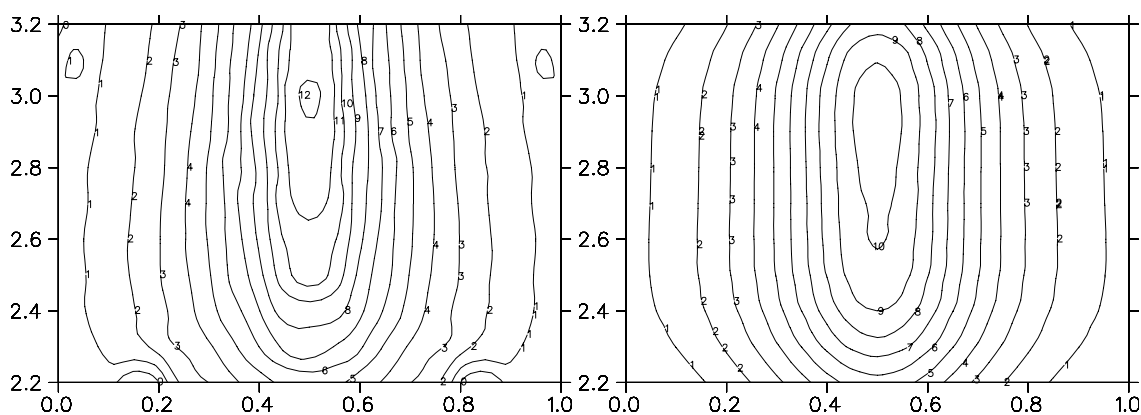


Figure 2.11: Comparison of *ab initio* (left) and fitted (right) $V_1(R, \rho)$.

The major advantage of the “LPE” type of MMPT PES is that the angular dependence is described in an accurate way. Here 2D PESs $V(\rho, \theta)$ are compared between fitted and *ab initio* calculations for given R 's. Results for $R=2.2\text{\AA}$, 2.4\AA , 2.6\AA , 2.8\AA , 3.0\AA and 3.2\AA are plotted in Figure 2.13, 2.14, 2.15, 2.16, 2.17 and 2.18, respectively. In these plots, the X axis represents ρ from 0 to 1 and the Y axis indicates θ from 0 to 180° . All these 2D section slices shows good agreement between fitted and *ab initio* potentials, illustrating that the MMPT potential (Eq. 2.43) combining with the parameter set (Table 2.5) reproduced the three-dimensional interaction within the O–H⁺–O

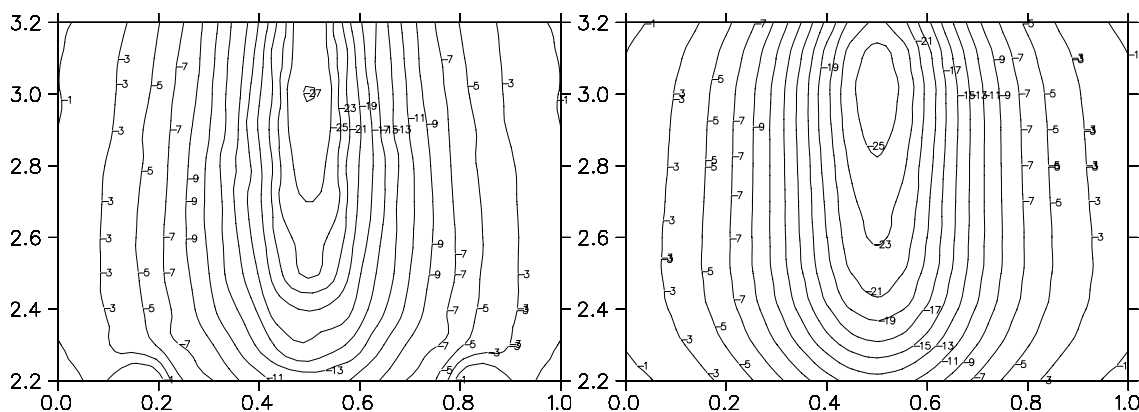


Figure 2.12: Comparison of *ab initio* (left) and fitted (right) $V_3(R, \rho)$.

motif very well.

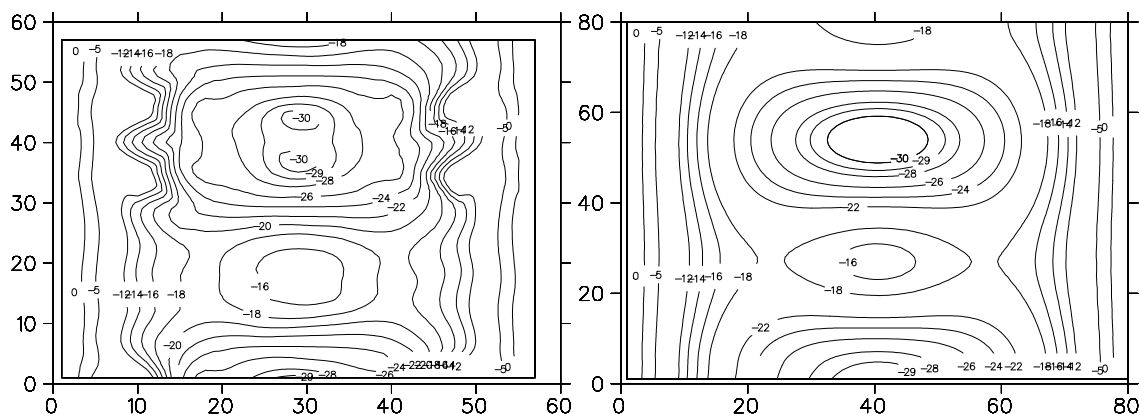


Figure 2.13: Comparison of *ab initio* (left) and fitted (right) $V(\rho, \theta)$ at $R=2.2 \text{ \AA}$ slice. The X axis represents ρ from 0 to 1 and Y axis indicates θ from 0 to 180° .

The PES is fitted to an *ab initio* scan of $R \in [2.2 \text{ \AA}, 3.2 \text{ \AA}]$. It is important to examine its asymptotical properties with the acceptor-donor distance, *i.e.*, whether it behaves correctly when the O–O distance is too large or too small. For this, a potential scan of the O–O distance in H_5O_2^+ was performed by CHARMM with MMPT module, which obtained qualitatively similar results with the *ab initio* scanning (see Fig. 2.19). The global minimum is found at $R_{\text{OO}} = 2.385 \text{ \AA}$ from the fitted MMPT PES, which compares favorably with the O–O equilibrium distance of 2.386 \AA reported by Bowman *et al* at the CCSD(T)/aug-cc-pVTZ level.¹⁸⁸

2 Molecular Mechanics with Proton Transfer

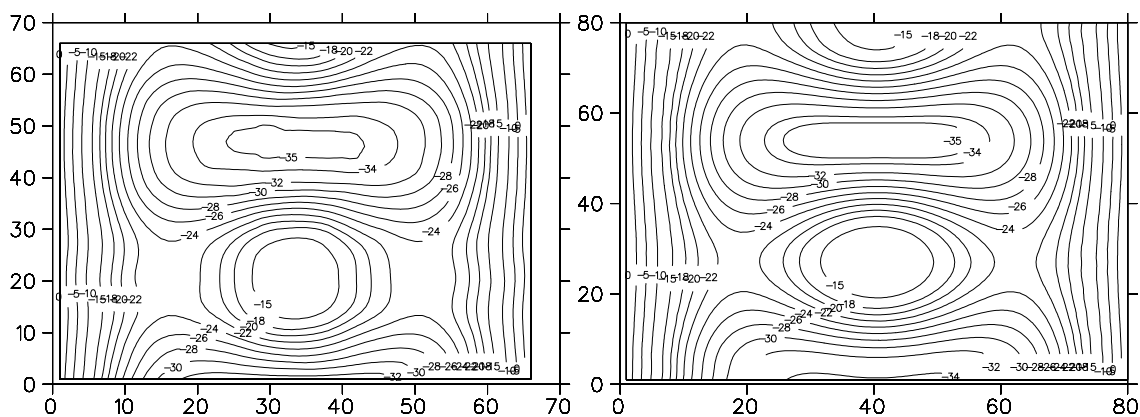


Figure 2.14: Comparison of *ab initio* (left) and fitted (right) $V(\rho, \theta)$ at $R=2.4 \text{ \AA}$ slice. The X axis represents ρ from 0 to 1 and Y axis indicates θ from 0 to 180° .

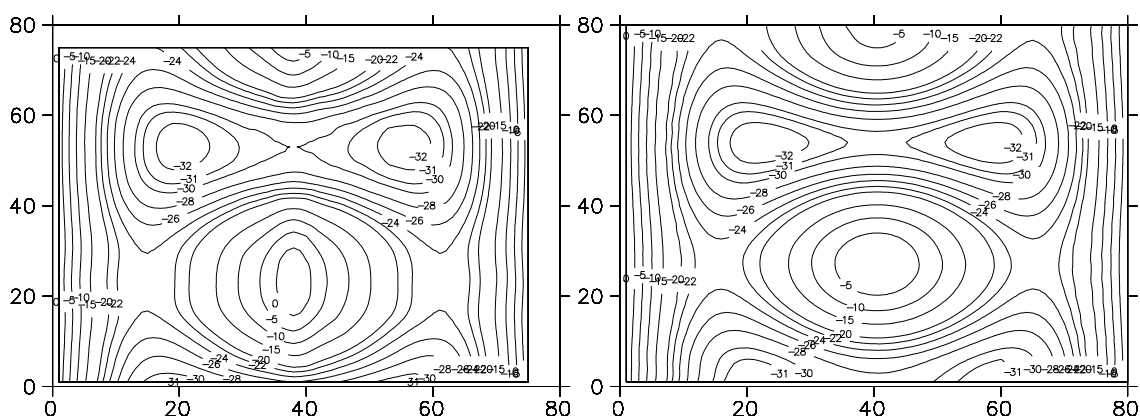


Figure 2.15: Comparison of *ab initio* (left) and fitted (right) $V(\rho, \theta)$ at $R=2.6 \text{ \AA}$ slice. The X axis represents ρ from 0 to 1 and Y axis indicates θ from 0 to 180° .

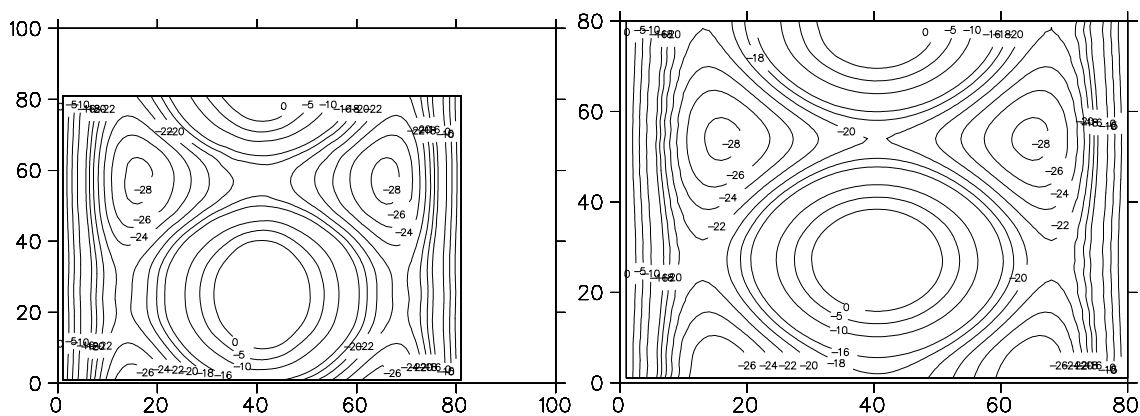


Figure 2.16: Comparison of *ab initio* (left) and fitted (right) $V(\rho, \theta)$ at $R=2.8 \text{ \AA}$ slice. The X axis represents ρ from 0 to 1 and Y axis indicates θ from 0 to 180° .

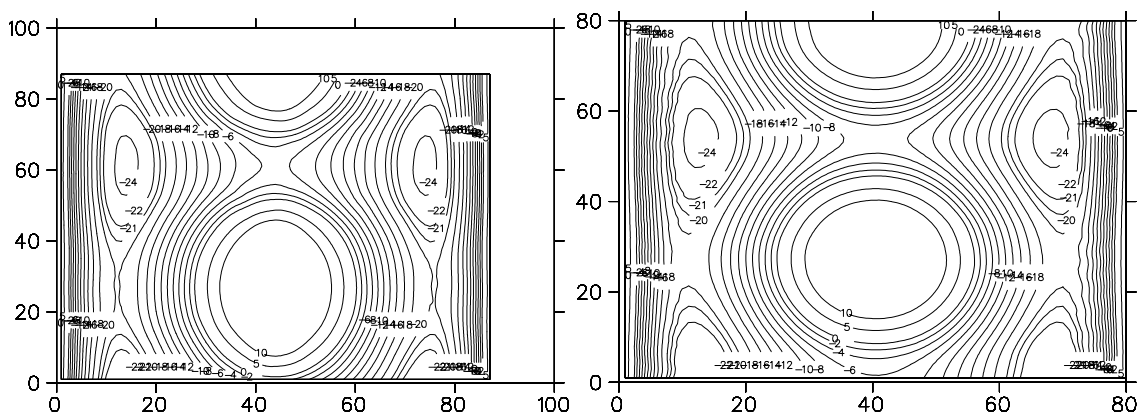


Figure 2.17: Comparison of *ab initio* (left) and fitted (right) $V(\rho, \theta)$ at $R=3.0 \text{ \AA}$ slice. The X axis represents ρ from 0 to 1 and Y axis indicates θ from 0 to 180° .

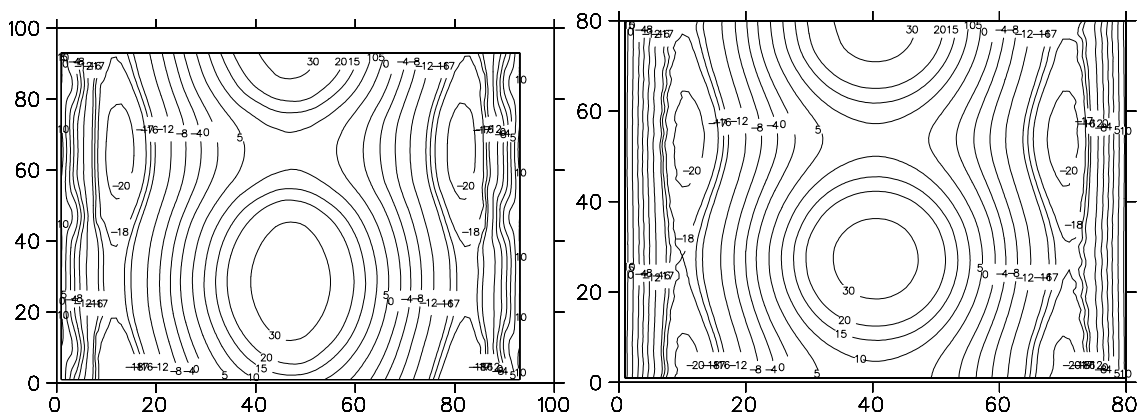


Figure 2.18: Comparison of *ab initio* (left) and fitted (right) $V(\rho, \theta)$ at $R=3.2 \text{ \AA}$ slice. The X axis represents ρ from 0 to 1 and Y axis indicates θ from 0 to 180° .

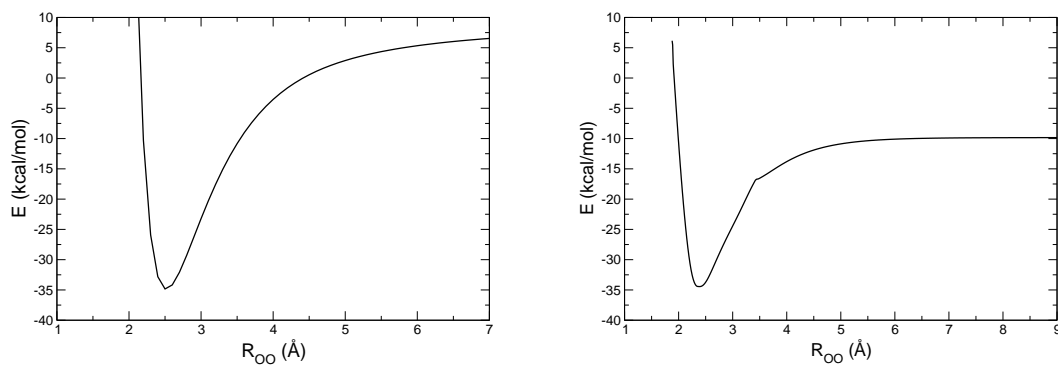


Figure 2.19: PES scan of O–O distance: *ab initio* (left) and fitted (right) results.

3 Applications I: Proton Dynamics in Acetylacetone

"To be fond of something is better than merely to know it, and to find joy in it is better than merely to be fond of it."

Confucius, *Analects*

In the next three chapters, three applications of MMPT will be presented. In this chapter, acetylacetone (AcAc), in particular the proton dynamics in AcAc, is investigated. With this particular example, the practice of generating MMPT potentials via PES morphing will also be illustrated.

3.1 Background

Malonaldehyde (MA) and its derivatives are prototype molecules for studying intramolecular PT processes. Experimentally, a number of infrared spectra¹⁸⁹⁻¹⁹² and tunneling splitting^{48,49} of MA have been measured with high accuracy. Theoretical calculations on different levels have been carried out to assign these spectra and to reproduce the splitting. In general, this requires full dimensional quantum dynamics performed on high level truncated PES^{193,194} such as a CCSD(T) potential with correlations up to seventh order,¹⁹⁵ which is computationally highly expensive. Recently, Yang and Meuwly investigated the proton dynamics in MA using MD simulations with parameterized MMPT force field.¹⁷⁷ Calculated results - including equilibrium structures, infrared spectra, tunneling splitting and proton transfer rates - compare favorably with experimental and previous computational work, which shows that the proton transfer dynamics in MA is very well characterized by MMPT.

The three hydrogen atoms bonded to the carbon atoms in MA can be substituted by different chemical groups. Schaefer *et al* systematically investigated the effect of such

substituent with DFT calculations, to explore the existence of very short hydrogen bonds and to find the derivative with no proton transfer barrier, i.e., a C_{2v} equilibrium structure.¹⁹⁶ The latter one is a difficult problem. Even for acetylacetone, which is malonaldehyde with symmetrically methyl groups substitution, there exists a long debate whether its ground state assumes an asymmetric (C_s) structure or a symmetric (C_{2v}) structure.

Neutron crystallography postulates that the ground state of AcAc has C_s symmetry.¹⁹⁷ Results from electron diffraction experiments are contradictory, suggesting either an asymmetric (C_s) structure^{198,199} or a symmetric (C_{2v}) structure.²⁰⁰ The most recent study performed with ultrafast electron diffraction concluded that the lowest energy form of AcAc has C_s symmetry.²⁰¹ However, based on the high-resolution rotational spectra of AcAc and its singly substituted ^{13}C -isotopologues recorded from microwave and millimeter-wave spectroscopy, enolic AcAc should have a C_{2v} symmetry.²⁰²

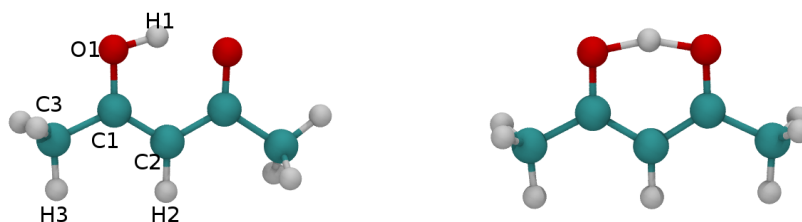


Figure 3.1: Equilibrium structure (left) and PT transition state structure (right) of AcAc, computed by MM force field.

As for MA, the bridging hydrogen atom can be expected to be localized on either of the oxygen atoms, which should lead to a tunneling splitting in a symmetric double well potential. However, contrary to MA where measurements yielded a large ground state splitting (21.6 cm^{-1})⁴⁹, experiments to detect the splitting in AcAc have been unsuccessful so far.^{203–206} Vibrational spectroscopy has also been applied

to characterize the atomic motions, especially the proton motion, in AcAc.^{203,207–212} The IR spectra is featured by an extremely broad peak between 1800 cm^{-1} to 3400 cm^{-1} , which usually indicates the existence of strong hydrogen bonding or proton transfer motifs.^{70–73}

From a theoretical perspective, AcAc is also challenging due to the increased size compared to MA (15 atoms vs. 9 atoms). Most previous computational investigations were limited in predicting the equilibrium geometry and the PT energy barrier of the ground state AcAc with different levels of quantum chemical methods and various basis sets.^{196,213,214} Full dimensional quantum dynamics calculations are in general not feasible, and so far only reduced-dimensionality quantum simulations were reported. For example, a three-dimensional quantum dynamics study was carried out to assign the infrared spectra of acetylacetone in gas the phase.²¹⁵ Mixed quantum-classical MD simulations, where the transferring proton in AcAc is treated quantum mechanically and the remaining degrees of freedom are treated classically with force fields, have been carried out in the past.^{126,212,216–218} Such simulations were based on simulations with a suitable empirical valance bond model with dedicated parameters.^{126,218,219} Roux *et al* parametrized the EVB force field by fitting to an *ab initio* potential surface obtained from HF/4-31G* calculations¹²⁶ and used it to study the potential of mean force and reaction rates for proton transfer in AcAc.²¹⁷ Mavri *et al* fitted the proton transfer force field to *ab initio* calculations at B3LYP/6-311+G(2d,2p) level, and used density matrix evolution (DME) method to compute the infrared spectra of AcAc in the gas phase and in chloroform solution.

3.2 Methods

3.2.1 Intermolecular Interactions and PES morphing

The MMPT potential of AcAc is generated via PES morphing from that of MA,¹⁷⁷ since the PT motifs in the two molecules are chemically similar with each other. The PES type is “NLM” with function form 2.27 and contains 14 parameters (see Table 2.1). $V^{AcAc}(R, \rho, d)$ and $V^{MA}(R, \rho, d)$ can be related through the following simple transformations:

$$(I) \quad V^{AcAc}(R, \rho, d) = \lambda V^{MA}(R, \rho, d) \quad (3.1)$$

$$(II) \quad V^{AcAc}(R, \rho, d) = V^{MA}(R - R_0, \rho, d) \quad (3.2)$$

$$(III) \quad V^{AcAc}(R, \rho, d) = V^{MA}(R, \rho - \rho_0, d) \quad (3.3)$$

$$(IV) \quad V^{AcAc}(R, \rho, d) = V^{MA}(R, \rho, d - d_0) \quad (3.4)$$

There can be the combinations of (I)-(IV), or even more complex transformations. To get the values of the morphing parameters $\lambda, R_0, \rho_0, d_0$, we compare in Table 3.1 the equilibrium structure, the PT transition state structure, and the PT energy barrier of AcAc and MA computed by electronic structure calculations at the same level (MP2) with the same basis set (aug-cc-PVTZ). We can then reasonably assume $\lambda = \Delta E^{AcAc} / \Delta E^{MA} = 0.796$, $R_0 = R_{opt}^{MA} - R_{opt}^{AcAc} = 0.03 \text{ \AA}$, $\rho_0 = \rho_{opt}^{MA} - \rho_{opt}^{AcAc} = 0.02$ and $d_0 = d_{opt}^{MA} - d_{opt}^{AcAc} = 0.02 \text{ \AA}$.

	AcAc		MA	
	opt	ts	opt	ts
$R(\text{\AA})$	2.528	2.362	2.556	2.357
$r(\text{\AA})$	1.0063	1.202	1.0013	1.200
$\theta(^{\circ})$	17.925	10.764	19.382	10.567
ρ	0.1697	0.5	0.1512	0.5
$d(\text{\AA})$	0.310	0.224	0.332	0.220
ΔE (kcal/mol)	2.1822		2.7412	

Table 3.1: Comparison of the electronic structure calculation for AcAc and MA

We summarize in Table 3.2 the results of different morphing transformations. The morphing scheme I+II, *i.e.*, an energy scaling plus a translation of the donor-acceptor distance give good correlation with the reference data (*ab initio* calculation results at the MP2/aug-cc-PVTZ level). So it has been chosen to generate the MMPT PES used in the following section. The detailed parameters in $V(R, \rho, d)$ are listed in Table 3.3.

Other force field parameters for AcAc are taken the same values as in MA,¹⁷⁷ while the substituted methyl groups are treated with standard CHARMM force field parameters.¹⁶⁴

	Ref.	morphed MMPT potential						
		No	I	II	I+II	I+II+III	I+II+IV	I+II+III+IV
$R_{opt}(\text{\AA})$	2.528	2.630	2.616	2.537	2.536	2.546	2.535	2.546
$r_{opt}(\text{\AA})$	1.006	0.989	0.987	1.009	1.009	1.029	1.006	1.029
$\theta_{opt}(\text{\textcircled{C}})$	17.93	18.29	17.88	16.05	15.86	15.52	15.57	15.46
$R_{ts}(\text{\AA})$	2.362	2.359	2.361	2.347	2.350	2.376	2.349	2.375
$r_{ts}(\text{\AA})$	1.202	1.202	1.203	1.196	1.199	1.213	1.196	1.212
$\theta_{ts}(\text{\textcircled{C}})$	10.76	11.07	11.15	11.23	11.27	11.81	10.97	11.58
ΔE (kcal/mol)	2.18	5.09	4.21	2.70	2.35	2.06	2.39	2.08

Table 3.2: Comparison of different morphing schemes for AcAC

	AcAc	MA
p_1	151.75800	190.62641
p_2	1.43800	1.43800
p_3	1.95286	1.99600
p_4	0.40632	0.51039
p_5	0.17015	0.13700
p_6	1.10500	1.10500
p_7	151.97300	151.97300
p_8	3.00819	3.15900
p_9	0.06449	0.06449
p_{10}	14.05000	14.05000
p_{11}	9.88999	12.42305
p_{12}	46.84010	58.83700
p_{13}	0.96300	0.96300
p_{14}	0.21986	0.21986

Table 3.3: Comparison of the parameters for morphed (AcAc) and zeroth-order (MA) MMPT potential

3.2.2 Force Fields for Chloroform and Deuteriochloroform

The explicit solvent molecules used in the MD simulations are chloroform and deuteriochloroform. Force field parameters for them are generated by combining *ab initio* calculation at the MP2/6-311++G** level and values from literature,^{220,221} and they have been written into the following CHARMM str file:

* Topology and Parameter Stream File for Deuteriochloroform

*

! June 2009, Jing Huang

3 Applications I: Proton Dynamics in Acetylacetone

!Parent files that have to be read prior to streaming this file

!top_all22_prot.inp

!par_all22_prot.inp

!The following atom types are unique to the deuteriochloroform and
!have been added to the protein rtf and parameter files. In
!future versions of charmm, if the ability to append MASSes is added
!these atom types should be removed from the original topology and
!the following lines uncommented.

!MASS 93 CLAL 35.453000 CL ! Chlorine Atom

!MASS 18 HD 2.014000 H ! Deuterium Atom

!MASS 22 CTC 12.01100 C ! C in chloroform

read rtf card append

* Topology for deuteriochloroform

*

31 1

!deuteriochloroform, jing huang, 06/09

RESI CDCL 0.00 ! deuteriochloroform

GROUP !

ATOM C1 CTC -0.38 ! D1

ATOM D1 HD 0.26 ! |

ATOM Cl1 CLAL 0.04 !Cl2-C1-Cl3

ATOM Cl2 CLAL 0.04 ! |

ATOM Cl3 CLAL 0.04 ! Cl1

BOND C1 D1 C1 Cl1 C1 Cl2

BOND C1 Cl3

BOND D1 Cl2 D1 Cl3 ! for SHAKE

BOND Cl1 Cl2 Cl2 Cl3 Cl1 Cl3! for SHAKE

ANGLE D1 C1 Cl2 D1 C1 Cl3

ANGLE Cl1 C1 Cl3 Cl1 C1 Cl2

ANGLE Cl2 C1 Cl3

PATCHING FIRST NONE LAST NONE

end

read para card append

* additional parameters for deuteriochloroform

*

BONDS

CTC HD 340.000 1.085

```
CTC CLAL 232.400 1.765
HD CLAL 0.000 2.3345 ! for SHAKE
CLAL CLAL 0.000 2.9139 ! for SHAKE
|*****|
ANGLES
HD CTC CLAL 38.100 107.59
CLAL CTC CLAL 77.700 111.28
|*****|
NONBONDED nbxmod 5 atom cdiel shift vatom vdistance vswitch -
cutnbn 14.0 ctofnb 12.0 ctonnb 10.0 eps 1.0 e14fac 1.0 wmin 1.5
CLAL 0.000000 -0.255000 2.000000 ! rigid
!CLAL 0.000000 -0.030000 1.948000 ! flexible
HD 0.000000 -0.015700 1.187000 ! from H in methane
CTC 0.000000 -0.109400 1.908000 ! from H in methane
end
return
```

The above stream file can be directly included in CHARMM input file for deuteriochloroform (CDCl_3). The topological and parameter files are also valid for chloroform (CHCl_3) as long as the mass of atom type “HD” is changed to 1.008. Such force fields for CHCl_3 and CDCl_3 are also used in Chapter 5.

3.2.3 Generation and Analysis of MD Trajectories

MD Simulations

5 ns MD simulations of AcAc in the gas phase, AcAc solvated in chloroform and deuterated AcAc solvated in deuteriochloroform have been carried out, respectively. For simulations in the condensed phase, an AcAc molecule was solvated in a pre-equilibrated chloroform cubic box with the length of 30.647 \AA , and periodic boundary conditions were applied. A cutoff of 12 \AA was applied to the shifted electrostatic and switched Van der Waals interactions. Before free dynamics simulations, the systems were heated to 300K and then equilibrated there for 1 ns. The time step in all simulations was $\Delta t = 0.1 \text{ fs}$ to follow the rapid proton motion.

Infrared and power spectra

Infrared and power spectra are computed from the time correlation functions from MD simulations. More specifically, the total dipole moment $\vec{M}(t)$ was recorded along

the MD trajectories and correlated over 2^{14} time origins to give $C(t)$

$$C(t) = \langle \vec{M}(t) \cdot \vec{M}(0) \rangle \quad (3.5)$$

To suppress noise, a Blackman filter was used.²²² Then $C(\omega)$, the Fourier-Transform of $C(t)$, was weighted with the Boltzmann factors to give the infrared spectrum $A(\omega)$

$$A(\omega) = \omega \{1 - \exp[-\hbar\omega / (k_B T)]\} C(\omega) \quad (3.6)$$

where k_B is the Boltzmann constant and T is the temperature. The power spectrum corresponding to internal motions can be computed in a similar way, by replacing the dipole-dipole correlation functions $C(t)$ with position-position correlation functions.

Proton Transfer Rate

Hazard plots²²³ are used to analyze the proton transfer rate from MD trajectories.^{28,176} For this, PT events are identified by monitoring the O_1 -H and O_2 -H distances. The proton transition time t_1, t_2, \dots, t_n are sorted in ascending order and the expectation value of the cumulative hazards for the k th transition is given by

$$H_k = \sum_{i=0}^{k-1} \frac{1}{n-i} \quad (3.7)$$

The proton transfer rate can then be computed as the slope of $H_k(t_k)$.

3.2.4 Quantum Mechanical Effects

To consider the quantum mechanical effects and to estimate the tunneling splitting, a suitable Hamiltonian for the quantum dynamics calculation is required. Here we construct a harmonic bath averaged (HBA) Hamiltonian¹⁷⁷

$$H^{\text{HBA}}(p_s, s) = \frac{1}{2} \frac{p_s^2}{1 + \Delta_s} + V_0(s) + \sum_{k=1}^{3N-7} \frac{\omega_k(s)}{2} \left(1 + \frac{\Delta_k(s)}{1 + \Delta_s}\right) \quad (3.8)$$

, where $s = r_1 - r_2$ is set to be the difference between two O-H separation distances. s can be regarded as the reaction path coordinate for proton transfer and p_s is the corresponding momentum. $\Delta_s, V_0(s), \omega_k(s)$ and $\Delta_k(s)$ can be calculated as following:

For each s , the molecule is optimized with all the remaining degrees of freedom

relaxed, and then the Hessian matrix $\mathbf{K}(s)$ is computed respect to the cartesian coordinates. $\mathbf{K}(s)$ is $3N \times 3N$ and for AcAc $N = 15$. The motion along the reaction path, together with the infinitesimal rotations and translations of the molecule system need to be projected out to obtain the projected force constant matrix

$$\mathbf{K}^P(s) = (\mathbf{1} - \mathbf{P}(s)) \cdot \mathbf{K}(s) \cdot (\mathbf{1} - \mathbf{P}(s)) \quad (3.9)$$

The projector \mathbf{P} is also a $3N \times 3N$ matrix given by

$$\mathbf{P} = \sum_{k=3N-6}^{3N} \vec{L}^k \otimes \vec{L}^k \quad (3.10)$$

where the unit vectors \vec{L}^k corresponding to infinitesimal rotations ($k = 3N - 6, 3N - 5, 3N - 4$), infinitesimal translations ($k = 3N - 3, 3N - 2, 3N - 1$) and motion along the reaction path ($k = 3N$) are given, respectively, by

$$L_j^{3N}(s) = -\frac{c}{\sqrt{m_j}} \frac{\partial V(s)}{\partial x_j} \quad (3.11)$$

where j runs from 1 to $3N$, m_j is the mass of the $\lceil \frac{j}{3} \rceil$ th atom and c is the normalization factor;

$$L_{i\gamma}^k(s) = \sqrt{\frac{m_i}{M}} \delta_{\gamma\lambda} \quad (3.12)$$

where $k = 3N - 3, 3N - 2, 3N - 1$ correspond to $\lambda = x, y, z$ respectively, i runs from 1 to N , $\gamma = x, y, z$, m_i is the mass of the i th atom and $M = \sum m_i$;

$$L_{i\gamma}^k(s) = \sum_{\alpha, \beta} (\mathbf{I}_0(s)^{-1/2})_{\lambda\alpha} \epsilon_{\alpha\beta\gamma} a_{i\beta}(s) \quad (3.13)$$

where $k = 3N - 6, 3N - 5, 3N - 4$ correspond to $\lambda = x, y, z$; i runs from 1 to N , $\alpha = x, y, z$; $\beta = x, y, z$; $\gamma = x, y, z$; and 3 dimensional vector $\vec{a}_i(s)$ is given by

$$a_{i\beta}(s) = \sqrt{m_i} [x_{i\beta}(s) - c_\beta(s)] \quad (3.14)$$

where m_i is the mass of the i th atom and $c_\beta(s)$ are the cartesian coordinates of the mass center; and $\mathbf{I}_0(s)$ is the 3×3 inertia tensor on the reaction path given by

$$\mathbf{I}_0(s) = \sum_{i=1}^N [\mathbf{1}\vec{a}_i(s) \cdot \vec{a}_i(s) - \vec{a}_i(s)\vec{a}_i(s)] \quad (3.15)$$

Diagonalizing $\mathbf{P}(s)$ gives $3N$ eigenvalues, of which seven equal zero. Other non-zero eigenvalues $\omega_k(s)$ ($k = 1, \dots, 3N - 7$), and the corresponding eigenvectors $L_{i,k}(s)$ can be used to construct the nonadiabatic coupling parameters²²⁴

$$B_{k,3N} = \sum_{i=1}^{3N} \frac{\partial L_{i,k}(s)}{\partial s} L_{i,3N}(s) \quad (3.16)$$

$$B_{k,l} = \sum_{i=1}^{3N} \frac{\partial L_{i,k}(s)}{\partial s} L_{i,l}(s) \quad (3.17)$$

In formula 3.16 and 3.17 the derivatives of the eigenvector along the reaction path is computed numerically by 4-point centered difference or FFT methods. Once $B_{k,3N}$ and $B_{k,l}$ are available:

$$\Delta_s = \sum_{k=1}^{3N-7} \frac{|B_{k,3N}(s)|^2}{2\omega_k(s)} \quad (3.18)$$

$$\Delta_k = \sum_{l=1}^{3N-7} \frac{|B_{l,k}(s)|^2}{2\omega_l(s)} \quad (3.19)$$

After this one-dimension Hamiltonian 3.8 is constructed, it can be solved discretely with Fourier grid Hamiltonian method.²²⁵

3.3 Results

3.3.1 Equilibrium Structure

The equilibrium and TS structures calculated from force field compare favorably with those from *ab initio* calculation at MP2/6-311G++(d,p) level (see Table 3.4). Also listed in Table 3.4 are the conformationally averaged geometries from 5 ns MD simulations in gas phase. The proton is shared by two oxygen atoms (O₁H equals O₂H) in such an ensemble averaged structure, assuming a C_{2v} symmetry. We further compare the harmonic frequencies of selective modes at the equilibrium and TS geometries in Table 3.5. The correlation is also very good. For example, the OH stretching (proton transfer) mode at TS has an imaginary frequency of 1150i cm⁻¹ from MM force field and 1057i cm⁻¹ from quantum chemistry calculation.

	Min		TS		MD Aver.
	MM	QM	MM	QM	
O ₁ H ₁ (Å)	1.009	1.006	1.199	1.202	1.306
O ₂ H ₁ (Å)	1.590	1.601	1.199	1.202	1.304
O ₁ O ₂ (Å)	2.536	2.528	2.350	2.362	2.518
H ₁ O ₁ O ₂ (°)	15.86	17.93	11.27	10.76	15.03

Table 3.4: Comparison of selective bond length and angles of AcAc minimum and transition state obtained from *ab initio* and force field optimizations, together with MD ensemble averaging values from 5 ns simulations in AcAc.

modes	Min		TS	
	MM	QM	MM	QM
O ₁ H ₁	3363	3013	-1150	-1057
O ₁ O ₂	332, 422	232, 374	270	536
O ₁ H ₁ O ₂ in plane	1257, 1619	1685, 1697	1617, 1404, 1370	1870
O ₁ H ₁ O ₂ out of plane	1118	1004	1259	1295
C ₂ H ₂	3030	3251	3030	3270

Table 3.5: Comparison of selective harmonic frequencies of AcAc minimum and transition state obtained from *ab initio* and force field optimizations. All the frequencies are in cm⁻¹.

3.3.2 IR Spectrum

The computed IR spectrum from the 5 ns MD simulation in the gas phase is shown in Figure 3.2, which reproduce the experimental spectra (see Figure 3.3 and also Fig. 4 in Ref. 212) very well. Spectral features, including both the broad band around 3000 cm⁻¹ and the peaks in the “fingerprint” area (1000 to 1500 cm⁻¹), are reproduced.

One advantage of MD simulations is that power spectra of individual vibrational motions can be calculated, which helps assigning spectra.^{72,73} For example, the strong band at ~ 1620 cm⁻¹ in the IR spectrum of AcAc remained unassigned.^{207–210} In Ref. 211, this peak at 1624 cm⁻¹ is analyzed either in terms of a C=O stretching plus O-H bending or in terms of the free rotation of the methyl group attached to the carbonyl group. The power spectrum of the CO stretching vibration (Figure 3.4) reports a single peak at 1602 cm⁻¹ and clearly relates the experimentally observed band to the CO stretching band. The power spectra of OH stretching spread from 2300 cm⁻¹ to 3300 cm⁻¹, showing a very unharmonic nature of the OH stretching (proton transfer)

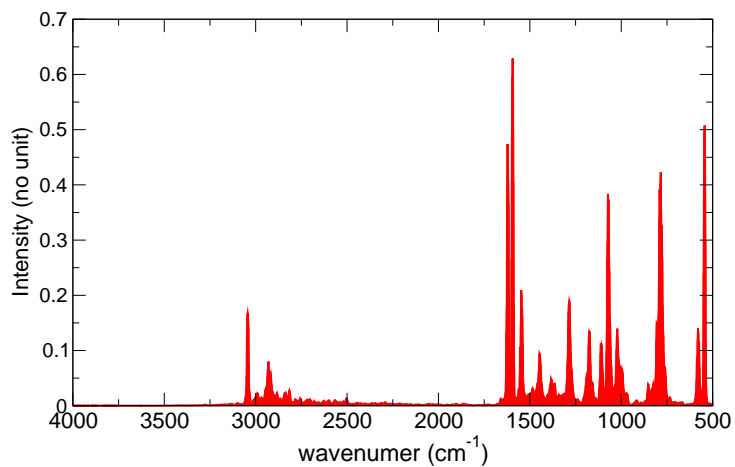


Figure 3.2: IR spectrum of AcAc from 5 ns NVE MD simulations in gas phase under 300K.

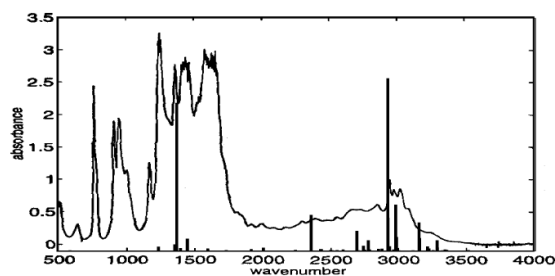


Figure 3.3: Experimental gas-phase spectrum for AcAc²¹² superimposed with IR spectrum computed with quantum dynamics²¹⁵ (Courtesy of I. Matanovic and N. Doslic).²¹⁵

mode. The OH bending modes corresponds to peaks around 1200 cm⁻¹. It is worthwhile to mention that a peak of 520 cm⁻¹ which corresponds to the O–O vibration are found in both OH stretching and OH bending modes, illustrating that these modes

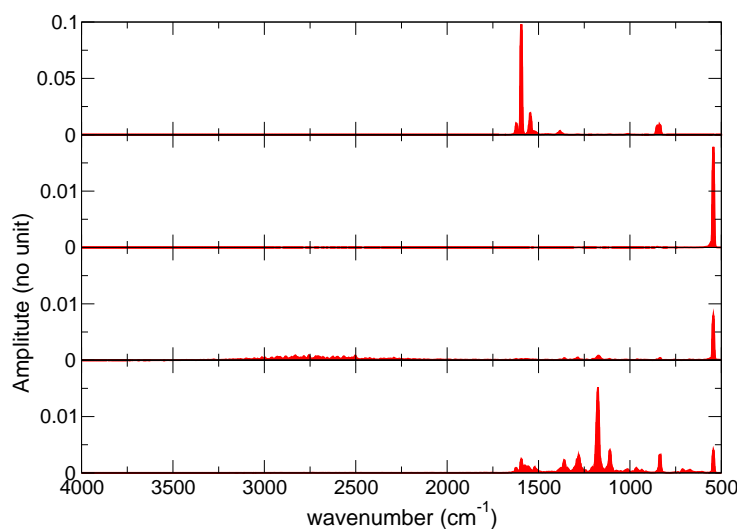


Figure 3.4: Power spectra of AcAc from 5 ns NVE MD simulations in gas phase under 300K. From up to bottom panel: CO stretching; O₁O₂ stretching; O₁H₁ stretching; O₁H₁O₂ bending.

are highly coupled.

The IR spectra of AcAc solvated in CHCl₃ and deuterated AcAc solvated in CDCl₃ are also calculated and shown in Figure 3.5 and 3.6, respectively. The IR spectrum in CHCl₃ is similar to that in gas phase, illustrating that the coupling between the solute and solvent is weak. For the spectra of deuterated AcAc, the PT mode shifts about 800 cm⁻¹ to the range of 2000 to 2250 cm⁻¹. Peaks between 1250 to 1500 cm⁻¹, which corresponds to OH and CH bending, are also missing.

3.3.3 Proton Dynamics and MM Proton Hopping Rate

Apart from the infrared spectra, the MMPT potential can also be used to study the classical over-barrier proton hopping rates. Experimentally, this is a difficult endeavour due to the transient nature of the process. However, having validated the MMPT potential with respect to structure and spectroscopy, a computational investigation of the hopping rates can be attempted. It should be noted that in the following a classical dynamics approach is pursued which should give lower limits for the rates. To include quantum effects, certain flux-flux autocorrelation functions should be used.²²⁶

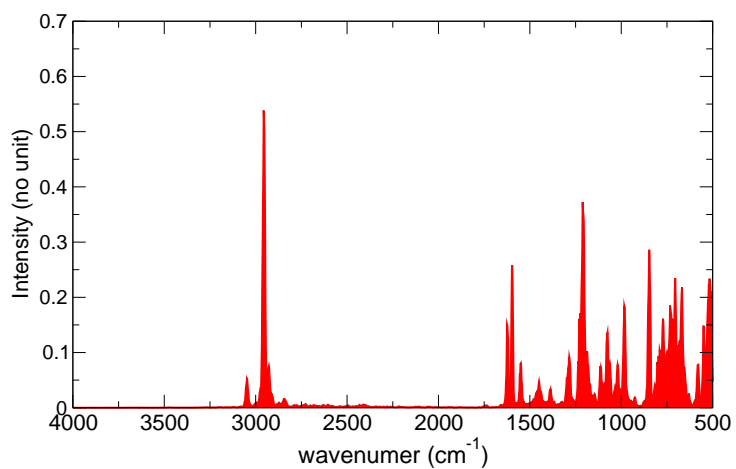


Figure 3.5: IR spectrum of AcAc from 5 ns NVE MD simulations in CHCl₃ under 300K.

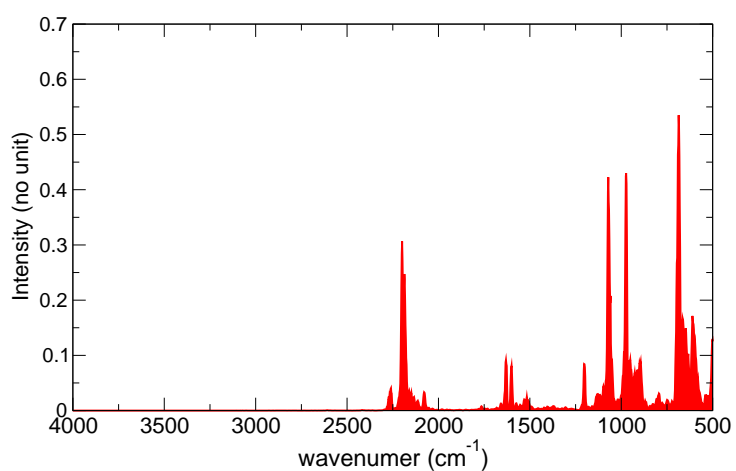


Figure 3.6: IR spectrum of deuterated AcAc from 5 ns NVE MD simulations in CDCl₃ under 300K.

Here only the contribution of the classical part, *i.e.*, the over-barrier hopping processes will be addressed. As has been mentioned in the Method section that the proton hopping rate can be calculated from the slope of hazard plots which is straightforward from MD simulations.

Figure 3.7 shows the proton hopping rates in vacuum and solution, respectively. The proton hopping rate of AcAc in gas phase and in CHCl_3 solvation are estimated to be 0.25/ps. The ignorable difference between rates in gas phase and solution implies a very weak coupling between the transferring proton and the solvent molecules which is consistent with the IR spectra compared above (see Figure 3.2 and 3.5). It should be noted that the real proton transfer in gas-phase AcAc is a non-decaying coherent process, therefore the proton transfer behavior should be better described by a coherent tunneling frequency instead of the hopping rate mentioned above.

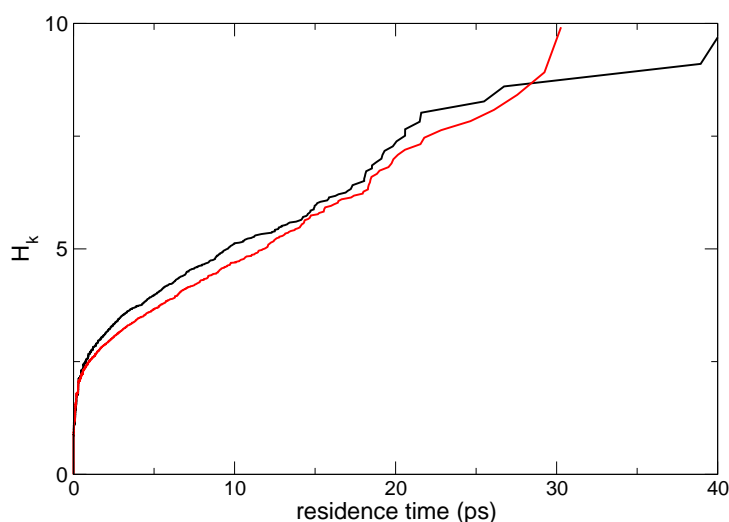


Figure 3.7: Proton hopping rates of AcAc in gas phase (black) and CDCl_3 (red) computed from Hazard plots.

To obtain more insight about the transient behavior of the proton transfer, fluctuations in the O-H and $\text{O} \cdots \text{O}$ distances (r_{OH} and r_{OO}) are investigated. The O-H and $\text{O} \cdots \text{O}$ distances during the first 500 ps in gas phase MD simulations are shown in Figure 3.8 with a “zoom-in” at the 200 - 300 ps time scale. The average $\text{O} \cdots \text{O}$ distance $\langle R \rangle$ is about 2.51 Å and average O-H bond lengths $\langle r_{\text{O}_1\text{H}} \rangle$ and $\langle r_{\text{O}_2\text{H}} \rangle$ equals 1.31 Å and 1.29 Å, respectively. Occasionally the $\text{O} \cdots \text{O}$ distance R can be even shorter than

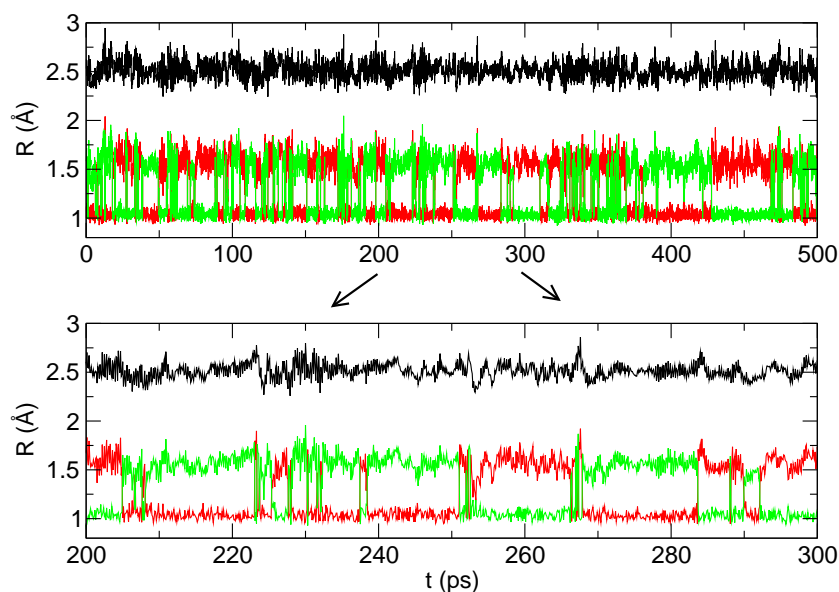


Figure 3.8: Typical time series for donor-acceptor and donor/acceptor-proton distances from the MD trajectory in gas phase at 300K.

2.2 Å or longer than 2.8 Å. The proton transfer occurs each time when the structure gets close to the transition state which is characterized by $r_{O\dots O} \approx 2.35$ Å. This clearly illustrates the role of transition states which agrees with previous investigations.²⁴

3.3.4 QM Effects of Proton Dynamics

Quantum effects are important for systems with fairly strong hydrogen bonds. In particular tunneling rates and quantum correction to the IR absorption spectrum of the PT mode will greatly clarify proton dynamics mechanisms in AcAc. For this we investigate the quantum effects on AcAc in terms of the recently developed HBA Hamiltonian method.

To construct the Hamiltonian 3.8, first the 2D PES of AcAc along r_1 and r_2 is scanned at the MP2/6-311++G** level. The scan is performed on a regular grid from 0.8 Å to 2.4 Å with the interval of 0.01 Å. The calculated 2D PES is presented in Figure 3.9 with the reaction path coordinate $s = r_1 - r_2$ marked with black line. The zeroth-order potential V_0 along the reaction path s is also computed and plotted in Figure 3.10.

Given the structures along the reaction path s , Eq. 3.9 to 3.19 can be used to build the Hamiltonian. $w_k(s)$, which are the non-zero eigenvalues of $\mathbf{K}^P(s)$, constitute a first-order revision to $V_0(s)$ as $V_0(s) + \sum_{k=1}^{3N-7} \omega_k(s)/2$ (see Figure 3.10). One problem is that the curve is not smooth at the minimum states ($s = \pm 0.64 \text{ \AA}$). This is because our definition of $L^{\vec{3}N}(s)$ (the unit vector along proton motion) as the derivatives of the potential (see Eq. 3.11) is not proper at the transition state and minimum states, which leads to large errors in the calculation. A better definition of $L^{\vec{3}N}(s)$ should fix the discontinuity, and related work is in progress. Nevertheless, we can estimate the tunneling splitting from the quantum calculations as follows.

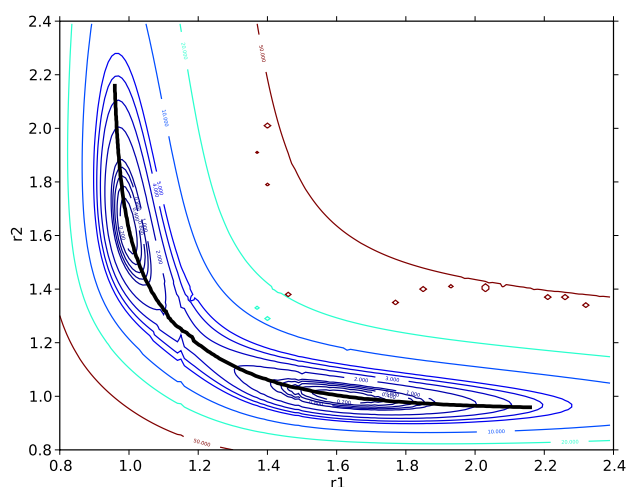


Figure 3.9: calculated 2D PES of AcAc

From Figure 3.10, the PT barrier height is 3.2 kcal/mol at the MP2/6-311++G** level. The zeroth-order Hamiltonian gives a tunneling splitting of 239 cm^{-1} . Though in general zeroth-order Hamiltonian overestimate the tunneling splittings, it can give the correct order of magnitude. For such a high splitting the population of the upper level is limited especially for some low temperature experiments. In this case, experimentally there may be only one of the two IR active transitions which can be observed. The most intense IR active proton transfer bands obtained by the zeroth-order Hamiltonian is 1848 cm^{-1} .

A simple correction to the zeroth-order Hamiltonian is to include zero point energy corrections. By comparing the harmonic frequencies of the minimum and transition state the zero point energy difference between the two structures is found to be 2.1 kcal/mol. Namely, the HBA barrier height should be 5.3 kcal/mol compared to the

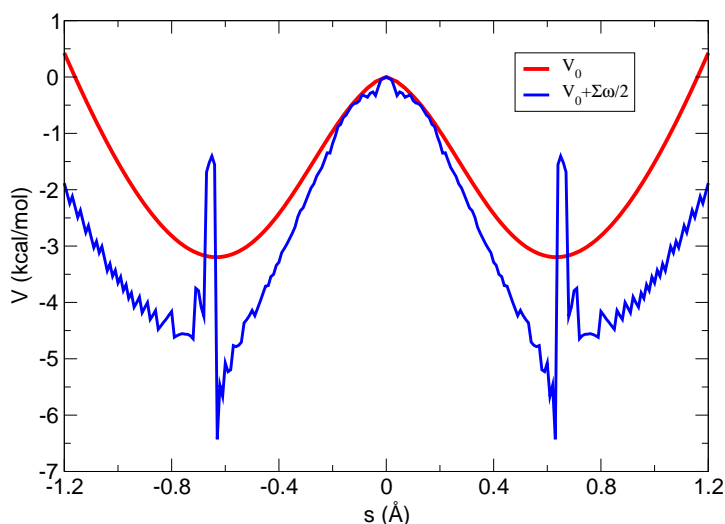


Figure 3.10: V_0 (red) and $V_0 + \sum_{k=1}^{3N-7} \omega_k/2$ (blue) along the reaction path s .

zeroth-order potential barrier of 3.2 kcal/mol. Taking into account of the zero point energy corrections leads to a tunneling splitting of 168 cm^{-1} and a transition frequency of 1974 cm^{-1} for the most intense IR active proton transfer band.

According to Ref. 177 both the potential and the kinetic corrections have been reported to be important. The mass correction Δ_s has been found to be about 0.2 for MA. One can expect the AcAc to have similar mass corrections to MA since the two substitution groups only contribute either some high frequency C-H motions which will not strongly couple to the proton transfer or the two CH_3 torsions which can be separated from the entire Hamiltonian since they have small excitation energies compared to other modes. Therefore the same mass correction of $\Delta_s = 0.2$ has been adopted for the estimation of the tunneling splitting of the AcAc. This leads to the tunneling splitting of 115 cm^{-1} and a transition frequency of 1644 cm^{-1} for the most intense IR active proton transfer band. This is quite close to the two C=O stretches which makes this frequency region much more broadened and structured as can be seen from the experiments.

3.4 Discussion

Here we present the MMPT force field for AcAc. Based on this reactive force field, classical molecular dynamics simulations are carried out to analyze proton transfer dynamics and to assign the IR spectra in gas phase and in solvation. Efforts to con-

struct a one-dimension Hamiltonian based on reaction path Hamiltonian²²⁴ are also presented. This is expected to provide a proper approximation in the quantum effects to the proton dynamics in AcAc.

The global validity and accuracy of the MMPT PES for AcAc is established by comparing to experimental data. Such a PES should be in principle transferable to other β -diketones.¹⁹⁶ An interesting point is that PES morphing can be performed based on limited information of AcAc (desired system) and MA (prototype system) from quantum chemistry calculations. It is possible to systematically map PESs by comparing *ab initio* data (optimized geometries, TS geometries, and energy barriers) between two chemically similar hydrogen bonding motifs, which could generate MMPT parameters for a particular PT system in an accurate and concise way. Work towards this direction is now carried out by a project student in our group.

4 Applications II: Hydrogen Bonds and NMR Couplings in Proteins

"If a man takes no thought about what is distant, he will find sorrow near at hand."

Confucius, *Analects*

In this chapter MMPT is applied to proteins as an explicit hydrogen bond potential with the idea that "hydrogen bond can be regarded as frozen proton transfer".²²⁷ The first section is a published paper "Explicit Hydrogen-bond Potentials and their Application to NMR Scalar Couplings in Proteins" (J. Huang and M. Meuwly, *J. Chem. Theo. Comput.*, 6:467-476, 2010), while the second section includes a manuscript submitted to *Chem. Phys.* (special issue: protein dynamics) entitled "Force Field Refinement from NMR Scalar Couplings". In the last section results are summarized and the future possibilities are discussed.

4.1 Explicit Hydrogen-bond Potentials and their Application to NMR Scalar Couplings in Proteins

Explicit Hydrogen-Bond Potentials and Their Application to NMR Scalar Couplings in Proteins

Jing Huang and Markus Meuwly*

Department of Chemistry, University of Basel, Klingelbergstrasse 80,
4056 Basel, Switzerland

Received October 27, 2009

Abstract: Hydrogen bonds (H bonds) are fundamental for the stability, structure, and dynamics of chemically and biologically relevant systems. One of the direct means to detect H bonds in proteins is NMR spectroscopy. As H bonds are dynamic in nature, atomistic simulations offer a meaningful way to characterize and analyze properties of hydrogen bonds, provided a sufficiently accurate interaction potential is available. Here, we use explicit H-bond potentials to investigate scalar coupling constants ${}^{\text{H}}J_{\text{NC}}$ and characterize the conformational ensemble for increasingly accurate intermolecular potentials. By considering a range of proteins with different overall topology a general procedure to improve the hydrogen-bonding potential (“morphing potentials”) based on experimental information is derived. The robustness of this approach is established through explicit simulations in full solvation and comparison with experimental results. The H-bond potentials used here lead to more directional H bonds than conventional electrostatic representations employed in molecular mechanics potentials. It is found that the optimized potentials lead to H-bond geometries in remarkable agreement with previous *ab initio* and knowledge-based approaches to H bonds in model systems and in proteins. This suggests that, by combining theory, computation, and experimental data, H-bonding potentials can be improved and are potentially useful to better study coupling, energy transfer, and allosteric communication in proteins.

Introduction

Hydrogen bonds are ubiquitous in chemical and biological systems and are essential for the overall structure, function, and dynamics of proteins and other macromolecules.¹ The role of hydrogen bonds in protein folding,² the formation of secondary structural elements,^{3,4} molecular recognition,^{5,6} and catalysis^{7,8} has been established over the past few years. A central feature of H bonds is their directionality, which cannot be easily captured by a superposition of isotropic interactions such as Coulomb interactions, as is done in customary force fields such as CHARMM, AMBER, or OPLS-AA.^{9–11} In small molecules, a hydrogen bond can be characterized spectroscopically. For example, in complexes between simple ions (HCO^+ , HN_2^+) and rare gas atoms (He, Ne, Ar), it is found that the fundamental infrared transitions in the electronic ground state correspond to $\Sigma-\Sigma$ transitions

characteristic for linear molecules.^{12–14} This can be inferred from the structure of the ro-vibrational bands (missing Q branch). Also, fitting of a model Hamiltonian¹⁵ allows for a determination of structural constants, which in turn characterize the average geometry of the molecule. For biological macromolecules, it is more difficult to find direct measures for the directionality of H bonds and to locate the positions of the hydrogen atoms. Direct visualization through recording of a structure is impractical, as H atoms can usually not be seen in X-ray crystallography. Structure determination from NMR data, on the other hand, formulates a search problem in structure space which minimizes a cost function that involves the experimental information (usually nuclear Overhauser data) and additional physical information because experimental data are rarely sufficient to determine the three-dimensional structure of a macromolecule.¹⁶

One experimental signature which recently became more widely available is hydrogen bond scalar couplings, which

* Corresponding author e-mail: m.meuwly@unibas.ch.

can be measured through NMR spectroscopy.^{17,18} Scalar couplings across $N-H\cdots O=C$ H bonds in proteins have a typical range of about -0.2 to -1 Hz, and the measurement errors are usually less than 0.05 Hz.^{19,20} Scalar $^3J_{NC}$ couplings have been observed experimentally in peptides,²¹ nucleic acids,²² and a variety of proteins.^{18,23–28} Together with other NMR parameters such as relaxation times, residual chemical shift anisotropy, and dipolar couplings, $^3J_{NC}$ couplings are important in the identification of conformational dynamics taking place on the NMR time scale.^{29–31} Further interesting and fundamental aspects of $^3J_{NC}$ couplings are their sensitivity to H-bonding network dynamics and cooperativity. Such effects are very difficult to probe directly through experiments, and a combined approach including atomistic simulations may prove advantageous. Earlier work established that the explicit dynamics of the solvated protein have to be taken into account to reliably calculate scalar coupling constants from molecular dynamics (MD) simulations.^{32,33} This naturally paves the way to improve specific terms in empirical force fields to which the observables are sensitive. In the present case, it is the capability of a force field to correctly describe H bonds.

It has been found that $^3J_{NC}$ values can be directly correlated with H-bond geometries. Barfield proposed several empirically parametrized formulas which enable the calculation of scalar couplings from the local $N-H\cdots O=C$ structure.³⁴ As NMR spectroscopy is a time-domain method, measured scalar couplings have to be understood as time averages. From a computational point of view, molecular dynamics simulations are the method of choice for such investigations. In previous work,^{32,33} a good correlation between measured $^3J_{NC}$ couplings and those derived from all-atom simulations was established by carrying out nanosecond MD simulations and averaging $^3J_{NC}$ values over entire trajectories.

Here, we combine a recently developed explicit hydrogen potential (molecular mechanics with proton transfer - MMPT)^{35,36} derived from correlated quantum mechanical calculations with an established force field to characterize $^3J_{NC}$ couplings in a variety of proteins covering different folds (ubiquitin ($\alpha + \beta$), the GB1 domain of protein G ($\alpha + \beta$), cold-shock protein A (all β), apo-calmodulin (all α), holo-calmodulin (all α), and intestinal fatty acid binding protein (all β), see Figure 1). Compared with conventional MD studies, the deviations between calculated and experimental $^3J_{NC}$ values are notably lowered. Next, the topology of the potential energy surfaces for the H-bond potentials is modified through morphing transformations^{37,38} to best describe experimentally determined couplings for three proteins. This approach is then generalized by applying it to the proteins not belonging to the training set, and very good agreement with measured coupling constants is found. Most notably, the approach pursued here leads to an average separation between the hydrogen atom and the acceptor of 1.93 Å, which agrees with a knowledge-based potential

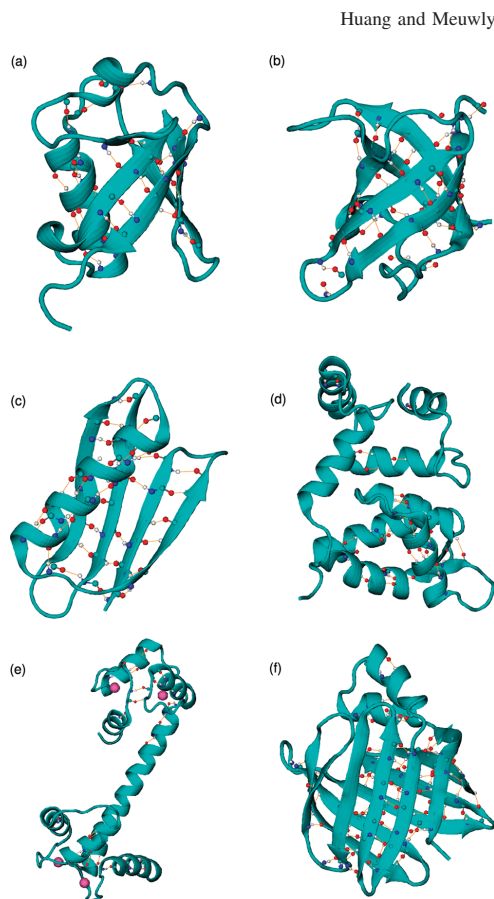


Figure 1. Structure, topology, and H bonds for the six proteins investigated here. (a) Ubiquitin, (b) GB1 domain of protein G, (c) cold-shock protein A, (d) apo-calmodulin, (e) holo-calmodulin, (f) intestinal fatty acid binding protein.

derived from 52 structures and results from electronic structure calculations.^{39,40}

Computational Methods

Molecular Dynamics Simulations. All simulations were carried out with the Charmm program⁴¹ using the CHARMM22 force field⁹ and provisions for MMPT.³⁶ The starting structures were taken from the X-ray structures in the Protein Data Bank⁴² (ubiquitin, 1ubq;⁴³ protein G,^{44b} 2qmt;^{44a} cold-shock protein A (CspA), 1mjc;⁴⁵ apo-calmodulin (apoCAM), 1qx5;⁴⁶ holo-calmodulin (holoCAM), 1cll;⁴⁷ intestinal fatty acid binding protein (IFABP), 1lfc.⁴⁸ Hydrogen atoms were generated with HBUILD,⁴⁹ and the structures were relaxed by 3000 steps of steepest descent minimization. Then, the proteins were solvated in pre-equilibrated water boxes of suitable sizes (1ubq, 65.19 Å \times 52.77 Å \times 49.67 Å; 2qmt, 55.88 Å \times 46.56 Å \times 40.36 Å; 1mjc, 52.77 Å \times 52.77 Å \times 46.56 Å; 1qx5, 71.40 Å \times 58.98 Å \times 58.98 Å; 1cll, 90.03

Explicit Hydrogen-Bond Potentials

$\text{\AA} \times 62.09 \text{\AA} \times 49.67 \text{\AA}$; 1ifc, $65.19 \text{\AA} \times 55.88 \text{\AA} \times 52.77 \text{\AA}$), and periodic boundary conditions were applied. A cutoff of 14\AA was applied to the shifted electrostatic and switched van der Waals interactions. Before free dynamics simulations, the systems were heated to 300 K and then equilibrated for 10^5 time steps.

For conventional MD simulations, all hydrogen atoms were constrained by SHAKE,⁵⁰ whereas for simulations with MMPT, hydrogen atoms involved in $^3J_{\text{NC}}$ couplings were free to move and all other hydrogen atoms were treated with SHAKE. A complete list of H bonds treated by MMPT for all proteins is summarized in Supporting Information S2. In both standard MD and MMPT/MD simulations, the time step was 0.2 fs , and snapshots were taken every 0.02 ps . The hydrogen-bond coordinates were extracted from trajectories and used together with eq 1 to calculate $^3J_{\text{NC}}$ couplings.^{32,34}

$$^3J_{\text{NC}} = (-366 \text{ Hz}) \exp(-3.2r_{\text{HO}'}') [\cos^2 \theta_1 - (0.47 \cos^2 \rho + 0.70 \cos \rho + 0.11) \sin^2 \theta_1] \quad (1)$$

where $r_{\text{HO}'}'$ is the distance between hydrogen and acceptor atoms, while θ_1 and ρ represent the $\text{H} \cdots \text{O}=\text{C}'$ angle and the $\text{H} \cdots \text{O}=\text{C}'-\text{N}'$ dihedral angle, respectively.

A simplified formula (eq 2) is also proposed in ref 34 and was used in a previous work.³³

$$^3J_{\text{NC}} = (-360 \text{ Hz}) \exp(-3.2r_{\text{HO}'}') \cos^2 \theta_1 \quad (2)$$

It captures the dominant effects of scalar couplings, while eq 1 provides a better estimate of $^3J_{\text{NC}}$ couplings in protein G because it accounts for the systematic difference between hydrogen bonds along the α helix and β sheet, respectively, by including a term related to the dihedral angle ρ .³⁴ Equations 1 and 2 can provide the same accuracy as full DFT calculations³² and have been used to calculate $^3J_{\text{NC}}$ couplings in different proteins.^{30,32-34,51} Detailed investigations on small molecules compared the performance of DFT using VWN, BP, or PW91 functionals with results from correlated methods such as the coupled cluster singles and doubles polarization propagator approximation⁵² and found that, with the exception of the HF molecule, the performance of DFT is good and provides almost quantitative spin-spin coupling constants.⁵³ The sensitivity to changes in the parameters of eq 2 has recently been investigated in a systematic fashion.³³ It was found that, overall, a strength factor of $\alpha = -360 \text{ Hz}$ and a decay of $\beta = 3.2 \text{\AA}^{-1}$ provide a good description of most coupling constants. However, for scalar couplings in particular secondary structural elements, the values for α and β could be optimized. As in the present work such aspects are not further pursued, eqs 1 and 2 are used, and the results are virtually identical. Generally, eq 1 leads to slightly smaller deviations between calculated and measured $^3J_{\text{NC}}$ values, so it was used in this work to calculate scalar couplings in all proteins except for CspA, where three backbone-side chain couplings were also included. While distinction between α -helix and β -sheet hydrogen bonds is only relevant for backbone-backbone hydrogen bonds, eq 2 has to be applied for computing these $^3J_{\text{NC}}$ values in CspA.

J. Chem. Theory Comput., Vol. 6, No. 2, 2010 469

The quality of the simulations was assessed by comparing root-mean-square deviations (RMSDs) between calculated and experimental $^3J_{\text{NC}}$ couplings:

$$\text{RMSD} = \sqrt{\frac{1}{N} \sum_{i=1}^N (J_i^{\text{calcd}} - J_i^{\text{exptl}})^2} \quad (3)$$

MMPT Potential and Morphing Transformations. A detailed account of MMPT has been given in ref 36. Briefly, MMPT uses parametrized three-dimensional potential energy surfaces fitted to high-level *ab initio* calculations (MP2/6-311++G(d,p)) to describe the interactions within a general DH-A motif, where D is the donor, H is the hydrogen, and A is the acceptor atom. Together with a standard force field—here, CHARMM⁹ is used—specific rules control how bonded interactions on the donor and acceptor side are switched on and off depending on the position of the transferring H atom (DH-A or D-HA). To adapt the overall shape of the PES to topologically similar, but energetically different, hydrogen bonding patterns—depending on the chemical environment of D and A—the PES can be “morphed”.^{37,38} Morphing can be a simple coordinate scaling or a more general coordinate transformation depending on whether the purpose of the study and the experimental data justify such a more elaborate approach.

For the present case of hydrogen bonds between an amide (NH) group as the donor and the oxygen atom as the acceptor (NH \cdots O), the MMPT potential depends on R (distance between N and O), ρ (relative position of H for a particular value of R), and θ (angle between unit vectors \hat{R} and $\hat{\rho}$). The relationship between ρ and the N-H distance r is given by

$$\rho = (r - r_{\text{min}})/(R - 2r_{\text{min}}) \quad (4)$$

where $r_{\text{min}} = 0.8 \text{\AA}$ is, in principle, arbitrary but should be sufficiently small to cover the shortest D-A separations. The angular dependence of the potential $V(R, \rho, \theta)$ is harmonic, that is, $V(R, \rho, \theta) = V_0(R, \rho) + k\theta^2$, and a typical PES along R and ρ is shown in Figure 2.

As mentioned above, the MMPT potentials are calculated for model systems (zeroth-order potential) and subsequently morphed to describe the situation in the actual chemical environment. Here, the asymmetric zeroth-order potential for $\text{NH}_4^+ \cdots \text{OH}_2$ is morphed to describe the N-H \cdots O=C motif in proteins. Morphing is achieved by modifying the parameters and thus reshaping the MMPT potential. The original potential has a single minimum $\{R_0 = 2.71 \text{\AA}, \rho_0 = 0.23, \theta_0 = 0^\circ\}$ and is mapped to a new one $\{R' = R_0 + \sigma, \rho' = \rho_0 - \delta, \theta_0 = 0^\circ\}$ where σ and δ are positive because hydrogen bonds in proteins are weaker than in a protonated ammonia-water dimer. Since there is a one-to-one correspondence between morphing parameters $\{\sigma, \delta\}$ and PES minima $\{R', \rho'\}$, only one set— $\{R', \rho'\}$ —will be used in the following. The morphed potential has its minimum energy at $\{R', \rho', 0\}$ while maintaining its overall shape, as illustrated in Figure 2.

For most X-ray structures, typically the coordinates of heavy atoms are available since only very rarely can protein crystallography resolve the positions of hydrogen atoms. Therefore, the experimental observable characterizing a

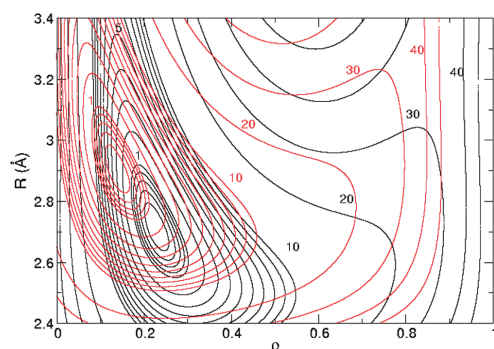


Figure 2. MMPT PES for the NH–O motif and illustration of PES morphing. Black, original PES; red, morphed PES. Contour lines are drawn at intervals of 0.2 kcal/mol for energies below 1 kcal/mol, for energies between 1 and 10 kcal/mol at intervals of 1 kcal/mol, and for higher energies at intervals of 10 kcal/mol. Morphing parameters $\{R', \rho'\} = \{2.92, 0.14\}$.

hydrogen bond is the D–A distance between donor and acceptor. In the Results section, it will be shown that best value for R' corresponds closely to the average D–A distances calculated from the initial X-ray structures. We also establish a relationship between optimized R' and ρ' (eq 5, see below). This leads to the following procedure for optimizing MMPT parameters and calculating ${}^{\text{h}3}\text{J}_{\text{NC}}$ couplings by MMPT/MD simulations:

- (1) From the X-ray/NMR structure, the average distance R' is calculated.
- (2) Compute ρ' by eq 5 (see below).
- (3) Morph the MMPT PES to the minima $\{R', \rho'\}$ by coordinate transformations.
- (4) Carry out MD simulations with the MMPT potential, and calculate the hydrogen-bond scalar couplings according to eq 1 or 2.

Results

Conventional MD as Benchmarks. Standard MD simulations 1 ns in length were first carried out for all six proteins, and the RMSDs between calculated and experimental couplings were computed as benchmarks for comparison. As shown in Supporting Information S3, ${}^{\text{h}3}\text{J}_{\text{NC}}$ couplings converge well within 1 ns. Hence, the RMSD as the average over all ${}^{\text{h}3}\text{J}_{\text{NC}}$ couplings is also stable during our simulation time scale; for example, the RMSDs of CspA calculated from 0.5, 1, and 1.5 ns standard MD trajectories are 0.198, 0.195, and 0.197 Hz, respectively.

We also carried out 500 ps MD simulations with CMAP for ubiquitin and CspA. CMAP is an extension of the CHARMM force field and has recently been shown to obtain a more accurate description of the peptide backbone.⁵⁴ By including grid-based energy correction maps and empirical corrections, this approach yields improved dynamical and structural properties of proteins in various simulations.^{55,56} However, applied to the present simulations of ${}^{\text{h}3}\text{J}_{\text{NC}}$

couplings for ubiquitin and CspA, results are very similar to simulations without CMAP, as illustrated in Supporting Information S4.

MD Simulations with MMPT. The zeroth-order MMPT PES is suitable to describe a N–H \cdots O bond in $\text{NH}_4^+ - \text{H}_2\text{O}$ and will not be directly applicable to hydrogen bonding in proteins. Therefore, it is expected that MD simulations using the unmorphed MMPT potential are unsuited for quantitative work, and large deviations between observed and calculated ${}^{\text{h}3}\text{J}_{\text{NC}}$ couplings should be found, as illustrated in Figure 3. When different morphing parameters are used, the MMPT potentials will have different minimum energy geometries $\{R', \rho'\}$ and lead to different scalar couplings, which is also shown in Figure 3.

The correlation between morphing parameters and RMSDs has been investigated for ubiquitin, CspA, and protein G. First, short (20 ps) test trajectories were run to locate suitable morphing parameters, and then 100 ps MD simulations were carried out on a fine grid ($\Delta = 0.01$ Å) of $\{R', \rho'\}$ and analyzed. For combinations $\{R', \rho'\}$ with low RMSDs, simulations were continued to 500 ps. Longer trajectories (1 ns) were run for ubiquitin (morphing parameters $\{2.92, 0.14\}$), protein G ($\{2.95, 0.16\}$), and CspA ($\{2.96, 0.16\}$), and RMSDs were calculated and are summarized in Table 1, together with results obtained from standard MD simulations. As an illustration, a detailed comparison between measured and calculated ${}^{\text{h}3}\text{J}_{\text{NC}}$ in CspA from standard MD and MMPT/MD simulations, and the squared deviations for each individual hydrogen-bond coupling, are shown in Figure 4. By adopting MMPT PES as the explicit hydrogen-bond potential, the correlation between calculated ${}^{\text{h}3}\text{J}_{\text{NC}}$ couplings and experimental data has been enhanced for most hydrogen bonds, especially those with large deviations ($|J_{\text{calcd}} - J_{\text{exptl}}| > 0.3$ Hz). The range of scalar couplings calculated from MMPT/MD simulations, however, is narrower than that from standard MD simulations. For convergence of most scalar couplings, a total of 500 ps is typically sufficient for MD simulations with the MMPT potential (see Supporting Information S3).

Application of the Morphed Potentials. After establishing that morphed MMPT potentials lead to improved agreement between calculated and experimental scalar coupling constants compared to those of a conventional force field (Table 1), potential morphing is used to further improve scalar coupling constants starting from X-ray and NMR structures. This is done for the three proteins studied in the previous section: ubiquitin, CspA, and protein G. As might be suspected, somewhat different coordinate transformations are most suitable to best describe the scalar couplings in the three different proteins (a summary of the relationship between RMSD and different MMPT PESs is given in the Supporting Information S5). Due to the nonlinearity between parameters (R', ρ') for the H-bond potentials and the calculated RMSDs between calculated and measured ${}^{\text{h}3}\text{J}_{\text{NC}}$ couplings, there is no simple, detectable relationship between the two. However, it is found that deviations are generally small around a certain $\{R', \rho'\}$ combination, and these values are summarized in Table 2. The average N–O distances computed from the initial structure are also reported and are

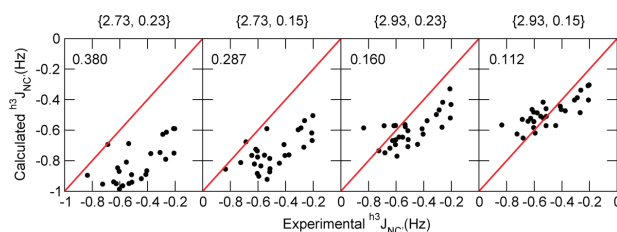


Figure 3. Calculated scalar couplings from 0.2 ns of MMPT/MD simulations compared with experimental data in ubiquitin for different morphing parameters.

Table 1. RMSDs of Ubiquitin, CspA, and Protein G Calculated from 1 ns Trajectory

	ubiquitin	CspA	protein G
standard MD	0.142	0.195	0.134
MD/MMPT	0.118 ^a	0.123 ^b	0.130 ^c

^a Morphing parameters {2.92,0.14}. ^b Morphing parameters {2.96,0.18}. ^c Morphing parameters {2.95,0.16}.

close to R' . With the use of relationship 4, the actual hydrogen bond length r'_{OH} is found to be almost identical for all three proteins, namely, 1.93 Å with an average of $1.931 \text{ \AA} \pm 0.002 \text{ \AA}$. It is worthwhile mentioning that this value is reminiscent of the hydrogen-bond geometry parameter δ_{HA} calculated from a statistical analysis of 52 proteins.⁴⁰ In the following, potential morphing for MMPT PESs is further investigated such that the additional constraint $r'_{OH} = 1.93 \text{ \AA}$ is fulfilled:

$$\rho' = \frac{R' - 2.73}{R' - 1.6} \quad (5)$$

This equation directly relates the two morphing parameters.

To test the procedure, it was applied to apo-CAM, holo-CAM, and IFABP, which were not part of the training set. The 500 ps MMPT/MD simulations were carried out with MMPT PES minima $\{R', \rho'\}$ found above, and scalar couplings were calculated. RMSDs between calculated and experimentally measured $^{13}J_{NC}$ couplings are summarized in Table 3. Compared with results from 1 ns standard MD simulations, considerably better agreement is achieved for all six proteins we investigated.

In all previous MMPT/MD simulations, only hydrogen bonds corresponding to experimentally measured scalar couplings are treated with the explicit hydrogen-bond potential. It would be interesting to test whether $^{13}J_{NC}$ can be predicted by MMPT/MD simulations without knowing which couplings can be observed in E.COSY experiments. Visual Molecular Dynamics (VMD)⁵⁷ has been used to assign hydrogen bonds in ubiquitin and CspA with a distance cutoff of 3.5 Å and an angle cutoff of 40°. In both proteins, more hydrogen bonds are found with this criterion (see Supporting Information S6), but not all of the previously assigned hydrogen bonds are covered. MD simulations with all of these hydrogen bonds treated by MMPT were carried out. $^{13}J_{NC}$ couplings were calculated from 500 ps trajectories and compared to experimental values. The RMSDs (0.122 and 0.160 Hz) are not as good as previous MMPT/MD results

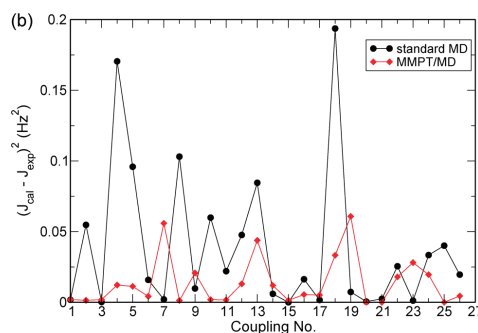
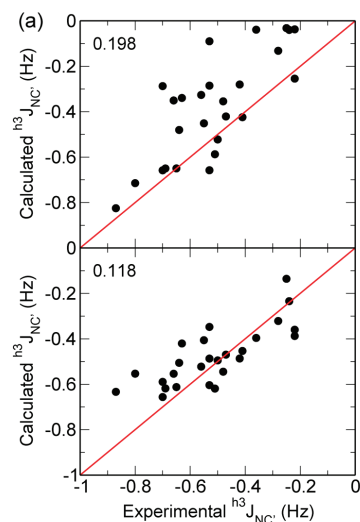


Figure 4. Comparison between scalar couplings calculated by standard MD simulations and MMPT/MD simulations for cold-shock protein A. (a) Comparisons of calculated and experimental $^{13}J_{NC}$ couplings. (b) Squared deviations for hydrogen bond scalar couplings.

(0.116 and 0.140 Hz) but are still significant improvements over standard MD simulations (0.140 and 0.195 Hz).

Characterization of the Conformational Ensemble.

Once suitable morphing parameters are available, MMPT/MD can also be used to characterize the conformational

Table 2. Overview of PES Morphing Parameters for Ubiquitin, CspA, and Protein G

	R' , Å	ρ'	r'_{OH} , Å	$\langle R_{\text{X-ray}} \rangle$, Å
1ubq	2.925	0.145	1.933	2.921
1mjc	2.960	0.170	1.929	2.958
2qmt	2.945	0.160	1.930	2.942

ensemble starting from the X-ray structure. The conformational ensembles generated by MD simulations with and without the MMPT potential are investigated through the distance between hydrogen and acceptor atoms and the angle at the acceptor atom. The respective density distributions (r_{NO} , θ_{HOC}) from 500 ps simulations for 29 hydrogen bonds in ubiquitin are shown in Figures 5 and 6. Using the MMPT potential for the hydrogen bonds in proteins leads to slightly shorter donor–acceptor distances, more pronounced directionality of the H bonds, and significant reductions in the fluctuations of r_{NO} and θ_{HOC} . This is also observed in protein G and CspA (data not shown) and can be explained by the fact that the MMPT potential is stronger and more directional than a conventional superposition of Coulomb terms. Thus, the H bonds are more restricted in the conformational space, which also leads to better stability and convergence of $^{\text{H}}J_{\text{NC}}$ couplings calculated from MMPT/MD compared to standard MD simulations (Supporting Information S3).

The protein dynamics based on using MMPT as an explicit hydrogen-bond potential are also investigated by calculating the root-mean-square fluctuations (or B-factors; Figure 7) and 2D cross-correlation maps (Figure 8). Generally, using the MMPT potential leads to rigidification of the protein, which is consistent with previous efforts to better describe hydrogen bonds in proteins.⁵⁸ The cross-correlation maps of ubiquitin show that most correlated motions are caused by hydrogen-bonding structures in the protein. Cross-correlation maps computed from MMPT/MD and standard MD simulations show similar dynamical features, while the comparison indicates that the MMPT potential enhances the correlations between hydrogen-bonding residues.

Discussion

In this work, we present a general method for deriving quantitative potential energy surfaces for H-bonding motifs and demonstrate their ability to accurately calculate scalar couplings across hydrogen bonds in proteins from atomistic simulations. Compared with standard MD simulations, RMSDs between calculated and experimental $^{\text{H}}J_{\text{NC}}$ couplings have been reduced in all six proteins investigated (Table 3). The $^{\text{H}}J_{\text{NC}}$ couplings can be calculated with an average deviation of 0.14 Hz by MMPT/MD simulation. Better agreement between calculated and experimental values are observed for all different secondary structures (Table 4), while the most significant improvements are found in loop regions. As has been noted previously,⁵⁹ current molecular mechanics force fields perform most poorly in the loop regions in proteins.

Our calculations are based on a force field treating hydrogen bonds explicitly. The MMPT potential, originally developed to investigate proton transfer reactions, has been

shown to be adequate for describing hydrogen bonds in proteins by simple PES morphing techniques. This is consistent with the well-known fact that hydrogen bonds can be regarded as incipient or “frozen stage” proton transfer reactions.¹ It is possible that more sophisticated PES morphing strategies, a more realistic angular dependence (e.g., $V(R, \rho, \theta) = \sum_n V_n(R, \rho) P_n(\cos \theta)$, where P_n are Legendre polynomials), or different MMPT parametrizations for H bonds in different secondary structure elements will lead to additional improvements.

The results presented here are based on an average treatment of H bonds in proteins, which means that the same MMPT potential is used for all hydrogen bonds in a certain protein. This is reflected by the fact that the PES morphing parameter R' corresponds to the average D–A distance from the X-ray structures. However, hydrogen bonds in different chemical environments exhibit different strengths so describing them with environment-specific parametrizations is a possibility for improvement. In fact, this has been previously found to be the case when hydrogen bonds in different secondary structures (α helices, β sheets, and loops) were investigated separately.³³

On the basis of a detailed study of correlation between PES morphing parameters and RMSDs in three proteins (ubiquitin, CspA, and protein G), we propose a generic procedure whereby, starting from X-ray structures, the PES is morphed to a minimum (R'_{NO} , r'_{OH} , θ'_{HNO}). Here, R'_{NO} equals the average N–O distance in the X-ray structure, $r'_{\text{OH}} = 1.93$ Å and $\theta'_{\text{HNO}} = 0$. Such an approach enables us to reliably calculate $^{\text{H}}J_{\text{NC}}$ couplings, and it has been applied to a set of six proteins. Due to the nonlinear relationship between the morphing parameters, the dynamics in proteins, and the calculated RMSDs for scalar couplings, morphing parameters $\{R', \rho'\}$ may not always yield the minimal RMSD between calculated and observed couplings. For example, in CspA, the morphing parameters lead to a RMSD of 0.14 Hz, while the minima $\{2.96, 0.16\}$ yield 0.12 Hz. However, differences are small, and both parameter sets are significant improvements over results from standard MD simulations (0.20 Hz), given that experimental errors are usually smaller than 0.05 Hz.^{19,20}

In previous work relating NMR observables and MD simulations, biased simulations with an additional restraining penalty function have been used.⁵¹ In this approach, $^{\text{H}}J_{\text{NC}}$ couplings were taken as input information and different dynamical ensembles were generated, which enables the determination of accurate geometries and energetics of hydrogen bonds in the native states of proteins. Here, a different approach is pursued. Instead of biasing simulations, the intermolecular interactions are represented more accurately by explicitly including potentials describing H bonds. The dynamical ensemble for the two methods is comparable in that narrower distributions of hydrogen-bond lengths and more restrictions for hydrogen-bond angles are found. Because scalar coupling constants directly characterize the geometries of H bonds, it is tempting to suggest that better quantitative agreement between calculated and experimentally measured $^{\text{H}}J_{\text{NC}}$ couplings also reflects a better description of the conformational ensemble of the protein.

Table 3. Comparison of RMSDs from Conventional MD Simulations and MMPT/MD Simulations with the Morphed Potentials

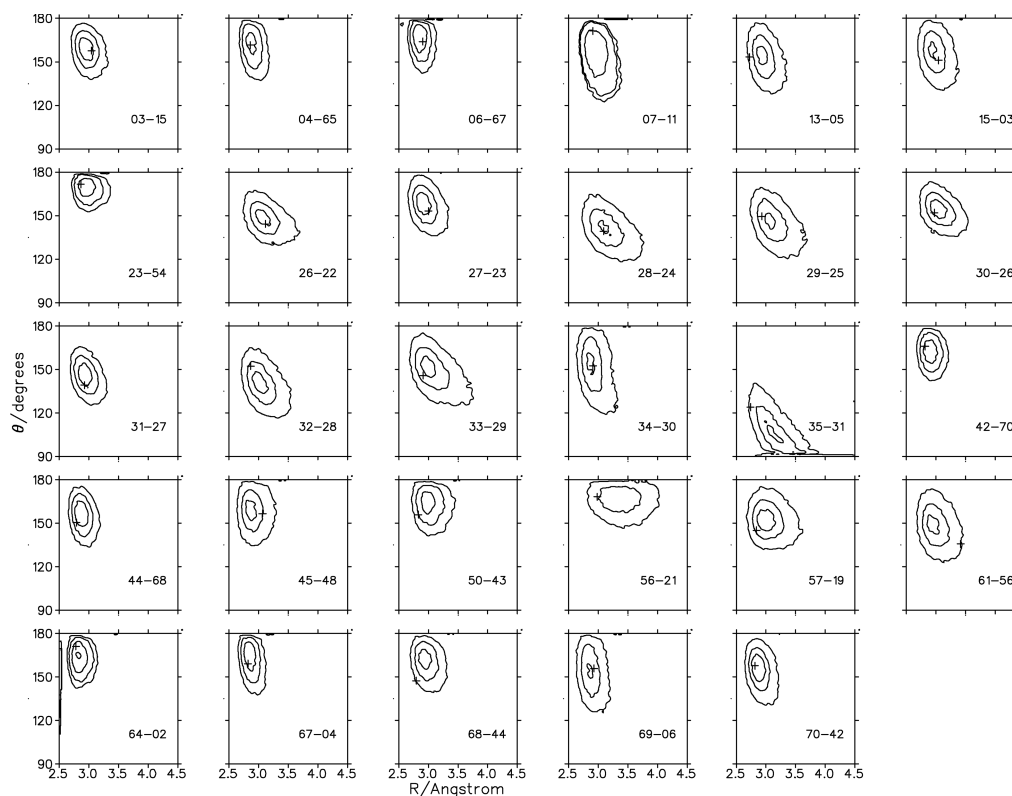
		ubiquitin	CspA	protein G	apoCAM	holoCAM	IFABP
RMSD (standard MD)		0.142	0.195	0.134	0.204	0.203	0.175
morphing parameters	R'	2.921	2.958	2.942	2.986	2.937	2.946
	ρ'	0.144	0.168	0.158	0.185	0.155	0.161
RMSD (MD/MMPT)		0.116	0.140	0.134	0.144	0.142	0.164

As the results show, the procedure pursued here is generally applicable and leads to appreciable improvement for all proteins investigated, and predictions for observables can be attempted. This is, in general, not possible with biased simulations for which the bias introduced is only valid for the particular protein under investigation and is not easily transferred to a different protein.

Hydrogen-bonding dynamics between standard MD and MMPT/MD simulations have been also compared in this work. Stronger hydrogen bonding, shorter hydrogen-bond lengths, and more pronounced directionality have been observed in MMPT/MD simulations. This agrees with a statistical analysis of X-ray structures which yields $\delta_{\text{HA}} = 1.93 \text{ \AA}$, which is identical to the separation found here and close to results from electronic structure calculations (1.94–1.97 \AA). Furthermore, the average NHO angle from

all simulations is 166° , which compares with values between 155° and 162° from electronic structure calculations, and 175° from the knowledge-based potential. Analysis of the protein dynamics shows that the MMPT potential rigidifies the entire protein and leads to stronger correlation between residues coupled by hydrogen bonds. This suggests that using explicit hydrogen-bond potentials shifts the conformational ensemble sampled in MD simulation toward the experimentally measured one.⁴⁰

Here, we showed that an explicit, three-dimensional hydrogen-bond potential leads to—sometimes considerably—improved calculation of hydrogen bond scalar couplings from explicit atomistic simulations in full solvation for six proteins with different folds. A general computational strategy is formulated which employs the coordinates from (high-resolution) X-ray structures and leads to suitably

**Figure 5.** Distributions of hydrogen-bond geometries (r_{NO} , θ_{HOC}) populated during 500 ps standard MD simulation for 29 H-bonds in ubiquitin.

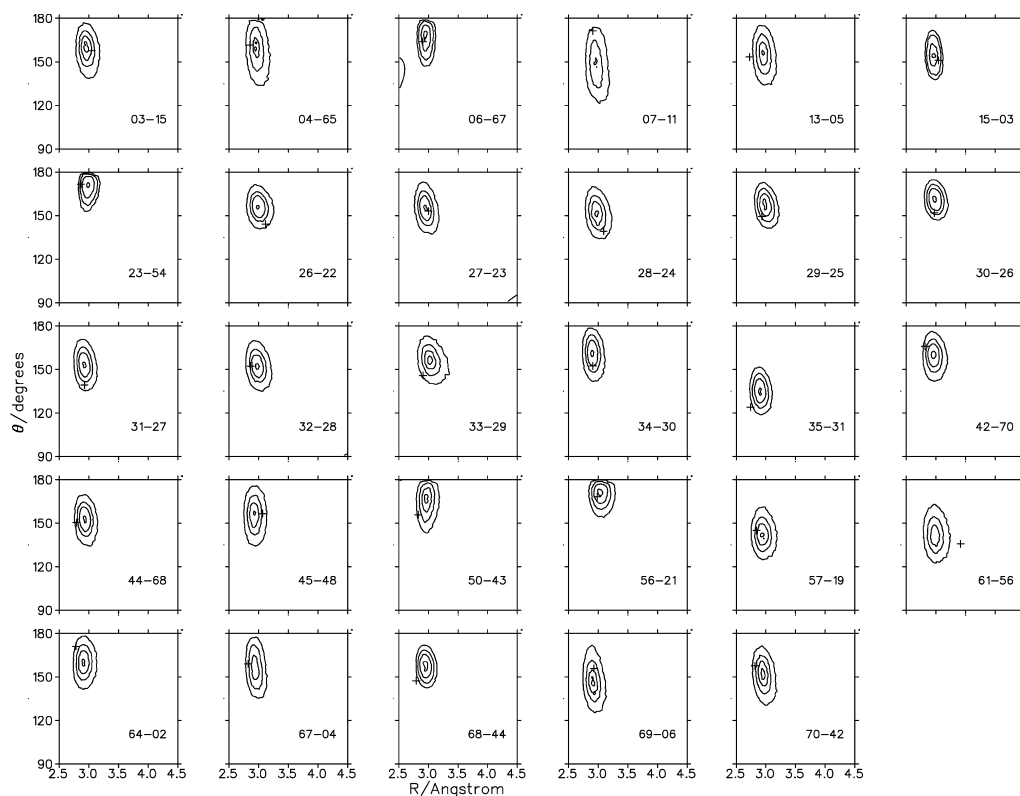


Figure 6. Distributions of hydrogen-bond geometries (r_{NO} , θ_{HOC}) populated during 500 ps MMPT/MD simulation for 29 H bonds in ubiquitin.

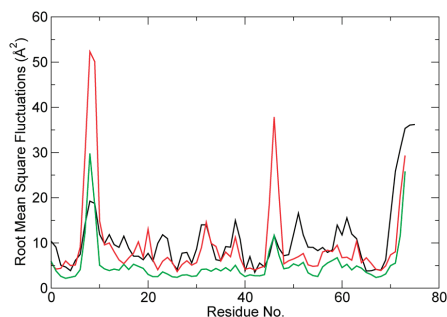


Figure 7. Root mean square fluctuations of backbone atoms calculated from 1 ns standard MD (red) and MMPT/MD (green) simulations of ubiquitin. Experimental B factors are plotted on the black line.

morphed H-bonding potential that can be used to investigate the nuclear dynamics in proteins. It is further illustrated that hydrogen-bonding potentials which lead to better agreement between calculated and measured $^3J_{\text{NC}}$ couplings are those with physically meaningful (PES morphing) parameters. This opens the possibility to further improve force fields by

combining NMR data and atomistic simulations, which is of particular relevance in characterizing conformational ensembles and in studies of signal transduction in proteins. Recently, a detailed analysis of the signaling pathway of rhodopsin led to the proposition that signals in proteins can be conducted through salt bridges and hydrogen bonds because they are more directional and the residues involved can act as molecular switches.⁶⁰ For such studies, which will most likely be intensified in the near future due to the fundamental interest in unraveling the means by which signaling occurs at a molecular level, accurate H-bonding potentials will be particularly important. The additional computational effort involved in using MMPT is minimal because, instead of a few harmonic potentials (conventional force field), the same number of anharmonic (Morse) terms have to be evaluated. What currently limits the standard use of MMPT is the fact that a time step of $\Delta t \approx 0.2$ fs is used to propagate the equations of motion. However, multi-time-step procedures are being considered which will largely circumvent this problem. As has been shown in a recent study on CO relaxation in myoglobin, conventional force fields based on harmonic bonded potentials which accurately describe vibrational spectra can be inappropriate when

Explicit Hydrogen-Bond Potentials

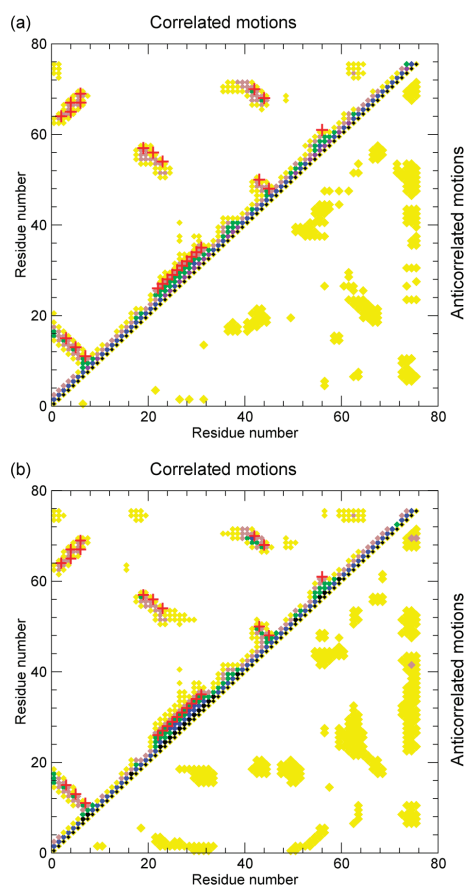


Figure 8. Dynamical cross-correlation maps for ubiquitin calculated from the 1 ns (a) standard MD and (b) MMPT/MD simulations. Positive cross-correlated coefficients are collected in the upper-left triangle and negative ones in the lower-right triangle. Only cross-correlation coefficients C_{ij} larger than 0.20 are shown. The intensity is represented as follows: yellow squares, $0.2 < C_{ij} < 0.35$; brown squares, $0.35 < C_{ij} < 0.5$; green squares, $0.5 < C_{ij} < 0.65$; blue squares, $0.65 < C_{ij} < 0.8$; and black squares, $0.8 < C_{ij} < 1.0$. H bonds with nonvanishing scalar couplings are marked by the red plus.

Table 4. Summary of RMSDs Calculated by Standard MD Simulations and MMPT/MD Simulations for All Couplings and for Couplings in Particular Secondary Structural Elements

	all	α helix	β sheet	loop
number of H bonds	214	86	105	23
standard	0.174	0.168	0.169	0.213
MMPT/MD	0.137	0.128	0.145	0.128

considering energy transfer between vibrational modes with widely separated frequencies.⁶¹ Thus, when energy transfer between modes is studied, details of the interaction potentials may become important. The fundamental role of H bonds,

the sensitivity of $^3J_{\text{NC}}$ couplings to their dynamics, and the possibility to compute couplings from meaningful atomistic simulations provide an ideal stage to further develop and extend the range and applicability of simulations.

Acknowledgment. The authors gratefully acknowledge financial support from the Swiss National Science Foundation under grant 200021-117810.

Supporting Information Available: Tables of all hydrogen bonds, scalar couplings and RMSDs, a comparison of different starting structures for protein G, a comparison of MD simulation with and without CMAP correction, and figures of $^3J_{\text{NC}}$ convergence. This information is available free of charge via the Internet at <http://pubs.acs.org>.

References

- (1) Steiner, T. *Angew. Chem., Int. Ed.* **2002**, *41*, 48–76.
- (2) Dill, K. A. *Biochemistry* **1990**, *29*, 7133–7155.
- (3) Deechongkit, S.; Nguyen, H.; Powers, E. T.; Dawson, P. E.; Gruebele, M.; Kelly, J. W. *Nature* **2004**, *430*, 101–105.
- (4) Wang, M.; Wales, T. E.; Fitzgerald, M. C. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 2600–2604.
- (5) Gray, M.; Cuellar, A. O.; Cooke, G.; Rotello, V. M. *J. Am. Chem. Soc.* **2003**, *125*, 7882–7888.
- (6) Aruksankunwong, O.; Hannongbua, S.; Wolschann, P. *J. Mol. Struct.* **2006**, *790*, 174–182.
- (7) Taylor, M. S.; Jacobsen, E. N. *Angew. Chem., Int. Ed.* **2006**, *45*, 1520–1543.
- (8) Huang, Y.; Unni, A. K.; Thadani, A. N.; Rawal, V. H. *Nature* **2003**, *424*, 146.
- (9) MacKerell, A. D.; et al. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
- (10) Ponder, J.; Case, D. *Adv. Protein Chem.* **2003**, *66*, 27–85.
- (11) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.
- (12) Nizkorodov, S. A.; Dopfer, O.; Ruchti, T.; Meuwly, M.; Maier, J. P.; Bieske, E. J. *J. Phys. Chem.* **1995**, *99*, 17118–17129.
- (13) Meuwly, M.; Nizkorodov, S. A.; Maier, J. P.; Bieske, E. J. *J. Chem. Phys.* **1995**, *104*, 3876–3885.
- (14) Nizkorodov, S.; Dopfer, O.; Meuwly, M.; Bieske, E.; Maier, J. *J. Chem. Phys.* **1996**, *105*, 1770–1777.
- (15) Watson, J. K. G. *Mol. Phys.* **1968**, *15*, 479–490.
- (16) Nilges, M.; Bernard, A.; Bardiaux, B.; Malliavin, T.; Habeck, M.; Rieping, W. *Structure* **2008**, *16*, 1305–1312.
- (17) Dingley, A. J.; Grzesiek, S. *J. Am. Chem. Soc.* **1998**, *120*, 8293–8297.
- (18) Comilescu, G.; Ramirez, B. E.; Frank, M. K.; Clore, G. M.; Gronenborn, A. M.; Bax, A. *J. Am. Chem. Soc.* **1999**, *121*, 6275–6279.
- (19) Grzesieka, S.; Cordiera, F.; Jaravinea, V.; Barfield, M. *Prog. Nucl. Magn. Reson. Spectrosc.* **2004**, *45*, 275–300.
- (20) Alkorta, I.; Elguero, J.; Denisov, G. S. *Magn. Reson. Chem.* **2008**, *46*, 599–624.
- (21) Eberstadt, M.; Mierke, D. F.; Kock, M.; Kessler, H. *Helv. Chim. Acta* **1992**, *75*, 2583–2592.

- (22) Dingley, A. J.; Masse, J. E.; Feigon, J.; Grzesiek, S. *J. Biomol. NMR* **2000**, *16*, 279–289.
- (23) Cordier, F.; Grzesiek, S. *J. Am. Chem. Soc.* **1999**, *121*, 1601–1602.
- (24) Ahn, H.-C.; Juranic, N.; Macura, S.; Markley, J. L. *J. Am. Chem. Soc.* **2006**, *128*, 4398–4404.
- (25) Markwick, P. R. L.; Sprangers, R.; Sattler, M. *J. Am. Chem. Soc.* **2003**, *125*, 644–645.
- (26) Alexandrescu, A. T.; Snyder, D. R.; Abildgaard, F. *Protein Sci.* **2001**, *10*, 1856–1868.
- (27) Juranic, N.; Atanasova, E.; Streiff, J. H.; Macura, S.; Prendergast, F. G. *Protein Sci.* **2007**, *16*, 1329–1337.
- (28) Juranic, N.; Moncrieffe, M. C.; Liki, V. A.; Prendergast, F. G.; Macura, S. *J. Am. Chem. Soc.* **2002**, *124*, 14221–14226.
- (29) Eberstadt, M.; Gemmecker, G.; Mierke, D. F.; Kessler, H. *Angew. Chem., Int. Ed.* **1995**, *34*, 1671–1695.
- (30) Bouvignies, G.; Bernado, P.; Meier, S.; Cho, K.; Grzesiek, S.; Bruschweiler, R.; Blackledge, M. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13885–13890.
- (31) Grzesiek, S.; Sass, H.-J. *Curr. Opin. Struct. Biol.* **2009**, *19*, 585–595.
- (32) Sass, H.-J.; Schmid, F. F.-F.; Grzesiek, S. *J. Am. Chem. Soc.* **2007**, *129*, 5898–5903.
- (33) Schmid, F. F.-F.; Meuwly, M. *J. Chem. Theory Comput.* **2008**, *4*, 1949–1958.
- (34) Barfield, M. *J. Am. Chem. Soc.* **2002**, *124*, 4158–4168.
- (35) Lammers, S.; Meuwly, M. *J. Phys. Chem. A* **2007**, *111*, 1638–1647.
- (36) Lammers, S.; Lutz, S.; Meuwly, M. *J. Comput. Chem.* **2008**, *29*, 1048–1063.
- (37) Bowman, J. M.; Gazdy, B. *J. Chem. Phys.* **1991**, *94*, 816–817.
- (38) Meuwly, M.; Hutson, J. M. *J. Chem. Phys.* **1999**, *110*, 8338–8347.
- (39) Kortemme, T.; Morozov, A. V.; Baker, D. *J. Mol. Biol.* **2003**, *326*, 1239–1259.
- (40) Morozov, A. V.; Kortemme, T.; Tsemekhman, K.; Baker, D. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 6946–6951.
- (41) Brooks, B. R.; Brucoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187–217.
- (42) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235–242.
- (43) Vijay-Kumar, S.; Bugg, C. E.; Cook, W. J. *J. Mol. Biol.* **1987**, *194*, 531–544.
- (44) (a) Schmidt, H. L. F.; Sperling, L. J.; Gao, Y. G.; Wylie, B. J.; Boettcher, J. M.; Wilson, S. R.; Rienstra, C. M. *J. Phys. Chem. B* **2007**, *111*, 14362–14369. (b) 1pga and 2igd are also used as initial structures, and the results are compared and discussed in the Supporting Information S1.
- (45) Schindelin, H.; Jiang, W.; Inouye, M.; Heinemann, U. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 5119–5123.
- (46) Schumacher, M.; Crum, M.; Miller, M. *Structure* **2004**, *12*, 849–860.
- (47) Chattopadhyaya, R.; Meador, W. E.; Means, A. R.; Quijoco, F. A. *J. Mol. Biol.* **1992**, *228*, 1177–1192.
- (48) Scapin, G.; Gordon, J.; Sacchettini, J. *J. Biol. Chem.* **1992**, *267*, 4253–4269.
- (49) Brunger, A. T.; Karplus, M. *Proteins* **1988**, *4*, 148–156.
- (50) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comp. Phys.* **1977**, *23*, 327–341.
- (51) Gsponer, J.; Hopearuoho, H.; Cavalli, C. M.; Dobson, A.; Vendruscolo, M. *J. Am. Chem. Soc.* **2006**, *128*, 15127–15135.
- (52) Raynes, W. T.; Geertsen, J.; Oddershede, J. *Int. J. Quantum Chem.* **1994**, *52*, 153–163.
- (53) Malkin, V. G.; Malkina, O. L.; Salahub, D. R. *Chem. Phys. Lett.* **1994**, *221*, 91–99.
- (54) Alexander, D.; Mackerell, J.; Feig, M.; Charles, L.; Brooks, I. *J. Comput. Chem.* **2004**, *25*, 1400–1415.
- (55) Buck, M.; Bouguet-Bonnet, S.; Pastor, R. W.; Alexander, D.; MacKerell, J. *Biophys. J.* **2006**, *90*, L36–L38.
- (56) Bastug, T.; Kuyucak, S. *Biophys. J.* **2009**, *96*, 4006–4012.
- (57) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33–38.
- (58) Ji, C.; Mei, Y.; Zhang, J. Z. H. *Biophys. J.* **2008**, *95*, 1080–1088.
- (59) Case, D. A. *Acc. Chem. Res.* **2002**, *35*, 325–331.
- (60) Kong, Y.; Karplus, M. *Structure* **2007**, *15*, 611–623.
- (61) Devereux, M.; Meuwly, M. *J. Phys. Chem. B* **2009**, *113*, 13061–13070.

CT9005695

4.2 Force Field Refinement from NMR Scalar Couplings

Force Field Refinement from NMR Scalar Couplings

Jing Huang, Markus Meuwly*

Department of Chemistry, University of Basel,
Klingelbergstrasse 80, 4056 Basel, Switzerland

Abstract

NMR observables contain valuable information about the protein dynamics sampling the high-dimensional potential energy surface. Depending on the observable, the dynamics is sensitive to different time-windows. Scalar coupling constants ${}^{13}\text{C}$ reflect the pico- to nanosecond motions associated with the intermolecular hydrogen bond network. Including an explicit H-bond in the molecular mechanics with proton transfer (MMPT) potential allows us to reproduce experimentally determined ${}^{13}\text{C}$ couplings to within 0.02 Hz at best for ubiquitin and protein G. This is based on taking account of the chemically changing environment by grouping the H-bonds into up to seven classes. However, grouping them into two classes already reduces the RMSD between computed and observed ${}^{13}\text{C}$ couplings by almost 50 %. Thus, using ensemble-averaged data with two classes of H-bonds leads to substantially improved force fields.

Keywords: Force Field Parametrization, NMR, Scalar Coupling, Ensemble Averaging

Introduction

Proteins are highly dynamical systems. Experimentally, nuclear magnetic resonance (NMR) is one of the methods of choice to demonstrate and quantify this. Data from NMR experiments - including spin relaxation data, scalar and residual dipolar coupling constants - contain valuable information about the intermolecular interactions in complex systems. However, using this information in a productive fashion in guiding and improving force fields has been limited due to the considerable conformational averaging that is required to obtain converged results. The - potentially beneficial and useful - interplay between the complementary nature of NMR experiments and molecular dynamics (MD) simulations has been realized from the early days of atomistic simulations for both, fast (relaxation) and slow (chemical exchange) processes, to which NMR is sensitive.[1, 2] However, the development of novel experimental observables (scalar and residual dipolar couplings) and the advent of powerful computer architectures lead the two techniques to go hand-in-hand and to cross-fertilize. Over the past few years a number of reports have appeared which highlight this.[3, 4, 5, 6, 7, 8, 9]

Of particular interest is the fact that NMR data is sensitive to the structural and conformational dynamics of proteins. This information is usually not utilized in current force field development. A primary reason for this is the fact that the already time-consuming and cumbersome fitting process is further complicated by taking ensemble-averaged data into account.[10]

Plenty of efforts are devoted to use MD simulations to analyse, reproduce and predict NMR measurements. For this purpose formulas and algorithms, such as the Lipari-Szabo model-

free formalism [11, 12], have been developed. The reliability of such an approach is based on the fact that NMR is a time-domain method and the assumption that MD simulations can provide a representative ensemble on which the NMR measurements are taken. On the other hand, force fields are a core quantity for MD simulations and their accuracy needs to be continuously improved. Common force fields, and their parameter sets, are mainly developed by inverting experimental data or by fitting to *ab initio* calculations of model systems. Such a methodology dominates force field development for decades and remains very successful, for example the recent development of CMAP [13] is based on fitting to LMP2/cc-pVQZ/MP2/6-31G* calculations of the alanine, glycine, and proline dipeptides.

With easier access to extensive compilations of protein NMR experimental data, force fields can be further improved. Actually, NMR data are already utilized to evaluate, scrutinize and compare the molecular mechanics force fields.[14, 15] In these studies, certain NMR observables in specific proteins are calculated from MD trajectories generated with different existing force fields and the correlations between computed and measured results are compared. Such an approach provides insightful information on the quality of force fields, but usually fails to provide direct suggestions for further force field optimization. Besides, when different kinds of heterogeneous NMR data are used in such assessment of force fields, the results are sometimes contradicting since a particular force field and its parametrization might be good for a certain NMR data but unsuccessful in predicting others.

One example of the recent available NMR observables are scalar couplings across hydrogen bonds (${}^{13}\text{C}$) determined by spin-echo difference techniques and E.Cosy experiments.[16, 17, 18, 19] ${}^{13}\text{C}$ couplings are directly sensitive to orbital overlap between the NH- and OH-groups and provide a direct mea-

*m.meuwly@unibas.ch

sure of N-H \cdots O=C hydrogen bonds in proteins.[20] Barfield correlated the magnitudes with hydrogen bonding geometries [3] and several studies show that ensemble averaging based on MD simulations is essential for realistic coupling calculations.[6, 21] $^{\text{H}}J_{\text{NC}}$ couplings have a typical range of about -0.2 to -1.0 Hz and the measurement errors are below 0.01 Hz [19, 22]. However, scalar couplings across hydrogen bonds are known to be difficult to determine computationally. The RMSDs are typically larger than 0.1 Hz and the correlation coefficients r^2 are usually below 0.5.[8, 14, 21] The reliability of computation for $^{\text{H}}J_{\text{NC}}$ is sometimes so low that the accuracy of experimental measurements of particular couplings is questioned by theoreticians.[14]

In a recent computational study the NH-OC bond was described by an explicit 3-dimensional interaction potential (molecular mechanics with proton transfer - MMPT[23, 24]) as opposed to electrostatic and nonbonded interactions.[8] The rationale for this was that explicit coordinate-dependent potentials provide more control over details in the intermolecular interactions than a superposition of (isotropic) monopolar and van der Waals interactions do. MMPT was demonstrated to provide a conformational ensemble which allows to accurately calculating $^{\text{H}}J_{\text{NC}}$ across hydrogen bonds in a set of six proteins with different folds.[8] The study was based on an average treatment of H-bonds in proteins which means that the same parametrization of the MMPT potential is used for all hydrogen bonds in a given protein. However, hydrogen bonds in different chemical environments exhibit different strength and features. Thus, describing them with environment-specific parametrizations is a possibility for improvement [6], which will be pursued here. In particular, we address three specific questions in this contribution: first, by allowing for different parametrizations of the H-bonds (i.e. classifying H-bonds into different groups), what is the best agreement with experimentally determined values that can be achieved? Second, is it possible to reduce this "maximal model" to one with the fewest number of classes. And third, is it possible to find a quantity which can be determined a priori from structural information and allows one to classify H-bonds into such classes? The latter point is of considerable and general importance for *predictive* instead of *postdictive* calculations of $^{\text{H}}J_{\text{NC}}$. However, for improving force fields answers to questions 1 and 2 above are already of much interest whereas point 3 is important if one wishes to predict scalar coupling constants ab initio.

In this contribution we present force field refinements based on explicit MD simulations using scalar couplings across hydrogen bonds. The MD simulations are based on the MMPT force field and the correlation between calculated and experimental $^{\text{H}}J_{\text{NC}}$ couplings is notably enhanced. Root mean square deviations of scalar couplings between experiment and simulation of 0.07 Hz or better are obtained for both ubiquitin and protein G. Furthermore, we establish a new approach to relate NMR observables and MD simulations. Instead of assessing force fields with NMR data, we directly optimize the force fields.

Methods

Molecular Dynamics Simulations

All simulations were carried out with the Charmm program [25] and provisions for MMPT [24]. The setup of proteins and simulation protocols are those from previous work.[8] Briefly, ubiquitin (lubq [26], (a) in Figure 1) and the B1 domain of protein G (2qmt [27], (b) in Figure 1) were solvated in pre-equilibrated water boxes of suitable sizes and periodic boundary conditions were applied. The system sizes are 16897 atoms for ubiquitin and 10323 for protein G. The systems were first heated to 300K and then equilibrated for 10^5 time steps, using the standard CHARMM force field.[28] Next, the interaction potential was switched to MMPT/MD (see below) with different H-bond partitions (see below) and morphing parameters, the systems were equilibrated for another 10^4 time steps before free dynamics simulations were carried out. Since the MD simulations are the most time-consuming step in the fitting cycles, the length of *NVE* simulations is limited to 100 ps. The convergence of RMSDs between calculated and experimental couplings from 100 ps MD trajectories is already good for MMPT/MD simulations (see Figure S3 in Ref. 8). Simulations with MMPT parameters with relatively low RMSDs between calculated and experimental $^{\text{H}}J_{\text{NC}}$ couplings are further extended to 500 ps to establish convergence. The time step in all simulations is $\Delta t = 0.2$ fs.

Intermolecular Interactions

The interactions within the hydrogen bonding motifs N-H \cdots O are described by the MMPT potential. A detailed account of MMPT has been given in Ref. 24. Briefly, MMPT uses parametrized three-dimensional potential energy surfaces (PESs) fitted to high level *ab initio* calculations (MP2/6-311++G(d,p)) to describe the interactions within a general DH-A motif where D is the donor, H is the hydrogen and A is the acceptor atom. Together with a standard force field - here, CHARMM [28] is used - specific rules control how bonded interactions on the donor and acceptor side are switched on and off depending on the position of the transferring H-atom (DH-A or D-HA).

The MMPT potentials are parametrized for prototype systems (zeroth-order potential) and subsequently "morphed" [29, 30] to topologically similar, but energetically different hydrogen bonding patterns. Morphing can be a simple coordinate scaling or a more general coordinate transformation depending on whether the purpose of the study and the experimental data justify such a more elaborate approach. Here, the asymmetric PES for $\text{NH}_4^+ \cdots \text{OH}_2$ is the zeroth-order potential $V^{(0)}$ and morphed as following to describe the N-H \cdots O=C motif in proteins

$$V(R, \rho, \theta) = V^{(0)}(R + \sigma, \rho - \delta, \theta) \quad (1)$$

, i.e., the minimum of the PES is shifted from $\{R_0 = 2.71 \text{ \AA}, \rho_0 = 0.23, \theta_0 = 0^\circ\}$ to $\{R' = R_0 + \sigma, \rho' = \rho_0 - \delta, \theta' = 0^\circ\}$. Other morphing strategies such as energy scaling ($V = \lambda V^{(0)}$ where λ is the morphing parameter) have also been tested, but do not improve the correlation between calculated and experimental $^{\text{H}}J_{\text{NC}}$ couplings (data not shown), and they are not used

in this work. Since there is a one-to-one correspondence between morphing parameters $\{\sigma, \delta\}$ and PES minima $\{R', \rho'\}$ and in order to be consistent with our previous study [8], only one set - $\{R', \rho'\}$ - will be reported.

Force Field Optimization

It is known that scalar coupling constants can be calculated from the H-bond geometry (see below) and are therefore sensitive to the parametrization of the force field.[6, 8, 14] In the present case this amounts to a dependence of ${}^{\text{H}}J_{\text{NC}'}$ on morphing parameters $\{R', \rho'\}$. The general problem is, however, that calculation of an expectation value of ${}^{\text{H}}J_{\text{NC}'}$ which can be compared with experimental data requires conformational sampling. Therefore, for every parameter set individual MD simulations have to be run. This leads to a procedure which can be automated by using a dedicated fitting environment. For this purpose, an interface [10] between CHARMM [25] and I-NoLLS [31] is used to fit the morphing parameters $\{R', \rho'\}$. User supervision of the fitting process is allowed within this interface, which provides control to maintain parameter values physically meaningful and makes the fitting of parameters for MD simulations more efficient.

Because each H-bonded motif is embedded in a slightly different chemical environment provided by the amino acids surrounding the motif, it is expected that the shape of the proton transfer potential is somewhat altered. This is reflected by different morphing parameters $\{R', \rho'\}$ which leads to a potentially large number of parameters and therefore to a high-dimensional fit. The aim is, however, to determine the smallest necessary number of different H-bonded groups required to reliably determine scalar coupling constants ${}^{\text{H}}J_{\text{NC}'}$.

In the following, hydrogen bonds are partitioned into groups to be described by different MMPT potentials. This amounts to introducing "meta-atom types" in the sense that we ask for a) the minimal number of parametrized H-bonding motifs necessary and b) the actual parameters describing a specific type of H-bonding motif. This can be compared with coarse-grained potentials from protein structures which also determine effective interactions between different atom types. The same logic is used in deriving models for coarse-grain simulations [32, 33, 34] or knowledge-based protein structure prediction.[35]

To generate a new H-bond partitioning from an existing one, three schemes are used: splitting, clustering and shuffling. For splitting, we calculate the deviations between the computed and experimental ${}^{\text{H}}J_{\text{NC}'}$ couplings in one H-bond group, and divide it into (up to) three groups according to whether the deviation $\Delta J = {}^{\text{H}}J_{\text{NC}'}^{\text{calc}} - {}^{\text{H}}J_{\text{NC}'}^{\text{exp}}$ for a particular H-bond is positive, negative or negligible ($|\Delta J| < 0.01$ Hz). The number of H-bond groups is increased by splitting while it can be also decreased by clustering, in which H-bond groups with close MMPT parameters (generally $\Delta R' < 0.03$ Å and $\Delta \rho' < 0.02$) are collected. Finally, those H-bonds with large deviations (for example, those with $|\Delta J| > 0.1$ Hz) can be selected and placed into another group, and this process to generate new H-bond partitions is named shuffling.

These three schemes are manually chosen during parametrization processes in the current study. It should be noted that the

time-limiting factor is the MD simulation itself, whereas parameter selection and decisions concerning splitting, clustering and shuffling are rapid and should be done manually for the current purpose. However, given that these decisions can be cast in a quantitative fashion, the entire process could be potentially automated if fitting of a larger data set should be attempted. For each H-bond partitioning (i.e. assigning a particular H-bond to a particular group characterized by morphing parameters $\{R', \rho'\}$), morphing parameters are optimized by fitting to experimental scalar couplings using the CHARMM-INoLLS interface [10], with the aim to lower the RMSD between ${}^{\text{H}}J_{\text{NC}'}^{\text{calc}}$ and ${}^{\text{H}}J_{\text{NC}'}^{\text{exp}}$ as much as possible (see Table 3). In general about five fitting cycles are attempted, which might not be sufficient to establish full convergence. However, the RMSDs and corresponding parameters we obtained and reported can be considered to be an upper-limit estimate for the lowest RMSD possible for a particular H-bond partition.

To analyze the scalar coupling constants, snapshots were taken every 0.02 ps from the *NVE* trajectories. The hydrogen bond coordinates were extracted from these snapshots and equation 2 was evaluated to calculate ${}^{\text{H}}J_{\text{NC}'}$ [3, 6]:

$${}^{\text{H}}J_{\text{NC}'} = (-360\text{Hz}) \exp(-3.2r_{\text{HO}'}) \cos^2 \theta_1 \quad (2)$$

where $r_{\text{HO}'}$ is the distance between the hydrogen and the acceptor atom and θ_1 is the $\text{H} \cdots \text{O}=\text{C}'$ angle. The computed scalar couplings are then averaged along the MD trajectory. The quality of the simulations was assessed by comparing root mean square deviations (RMSDs) between calculated and experimental ${}^{\text{H}}J_{\text{NC}'}$ couplings:

$$\text{RMSD} = \sqrt{\frac{1}{N} \sum_{i=1}^N (J_i^{\text{calc}} - J_i^{\text{exp}})^2} \quad (3)$$

Results

Secondary Structure as Indicator for Grouping

Considering the two proteins, ubiquitin and protein G, it is tempting to group H-bonding motifs depending on the secondary structural element in which they appear, i.e., helices, β -sheets, or loops. This would allow to directly infer a partitioning from structure alone and provide one possible and generic way to improve simulations based on a well defined procedure. This was pursued in an earlier attempt to better reproduce the ${}^{\text{H}}J_{\text{NC}'}$ couplings by scaling partial atomic charges in the H-bond motifs in MD simulations with the standard CHARMM force field.[6] It has been found that increased polarity of the hydrogen bond improves the correlation between computed and experimentally measured ${}^{\text{H}}J_{\text{NC}'}$ couplings, and H-bonds in different secondary structures are better represented by different magnitudes of electrostatic interactions which are directly related to the partial charges on the H-bonding motif.

A similar procedure was attempted with MMPT. The 29 backbone hydrogen bonds in ubiquitin are partitioned into three groups, according to their secondary structure being α -helix (10 H-bonds), β -sheet (15 H-bond) and loop region (4 H-bonds).

The 32 hydrogen bonds in protein G are partitioned into two groups: α -helix and β -sheet, each containing 13 and 19 H-bonds, respectively. This amounts to 6 morphing parameters ($R_\alpha, \rho_\alpha, R_\beta, \rho_\beta, R_{\text{loop}}$ and ρ_{loop}) for ubiquitin and 4 morphing parameters ($R_\alpha, \rho_\alpha, R_\beta$ and ρ_β) for protein G. These parameters are determined via fitting the computed ${}^{13}\text{J}_{\text{NC}'}$ couplings to experimental data by the CHARMM-INoLLS interface and are listed in Table 1.

The correlation between calculated and experimental scalar couplings has been improved for both H-bonds in α -helix and β -sheet in both proteins (see Table 2 and Figure 2) compared to previous work which treated all H-bonds equally.[8] The difference of H-bonds in different secondary structural elements is mainly in the difference of morphing parameter ρ . But the improvement is minor, and for H-bonds in the loop region in ubiquitin no improvement is found. The RMSDs decrease from 0.118 Hz to 0.115 Hz for ubiquitin and from 0.128 Hz to 0.125 Hz for protein G. As a comparison, other H-bond partition schemes with the same number of H-bond groups yields RMSDs of 0.055 Hz for ubiquitin and 0.070 Hz for protein G (see below). Also, the parameter $\{R_\alpha, \rho_\alpha\}$ and $\{R_\beta, \rho_\beta\}$ optimized with scalar couplings in ubiquitin differ from those in protein G, suggesting that one set of parameters for H-bonds in different secondary-structure of proteins might not be sufficient.

Number of Groupings and Overall RMSD

Because a straightforward grouping by secondary structural elements did not provide satisfactory improvement of the RMSD, a more systematic approach was pursued. Taking ubiquitin for example, all H-bonds are treated equally at the beginning.[8] In such a partitioning, all H-bonds are considered as one group and with the morphing parameters $\{2.92, 0.14\}$ a RMSD between ${}^{13}\text{J}_{\text{NC}'}$ and ${}^{13}\text{J}_{\text{NC}'}$ of 0.12 Hz is obtained. By comparing computed ${}^{13}\text{J}_{\text{NC}'}$ couplings with experimental values, the total 29 H-bonds were assigned into three groups, with 13, 4 and 12 H-bonds, respectively. The couplings were assigned to a particular group by clustering them into those that over- or underestimate the experimental ${}^{13}\text{J}_{\text{NC}'}$ coupling by more than 0.01 Hz and those that were within 0.01 Hz. In this way a new H-bond partitioning with 3 H-bond groups is generated by the “splitting” scheme (see Figure 3). Then the morphing parameters $\{R_1, \rho_1\}$, $\{R_2, \rho_2\}$ and $\{R_3, \rho_3\}$ for these three groups are determined by fitting. The initial values of these parameters in such a fitting process are $\{2.90, 0.16\}$, $\{2.92, 0.14\}$ and $\{2.94, 0.12\}$. The initial guess is based on the optimized parameters from the last H-bond partitioning ($\{2.92, 0.14\}$) and the experience that larger R and smaller ρ typically correspond to smaller scalar couplings. A good initial guess for the parameters can potentially reduce the number of fitting cycles required and thus reduce the overall computational cost, but it is not mandatory for the fitting process. After six fitting cycles, the RMSD has been reduced by 40% and the fitting results are summarized in Table 3. The optimized morphing parameters for this H-bond partition are $\{2.868, 0.181\}$, $\{2.922, 0.14\}$ and $\{2.975, 0.099\}$, corresponding an RMSD of 0.055 Hz.

Similarly, a new H-bond partitioning, *i.e.*, another way of assigning H-bonds into groups, can be generated by either “clus-

tering” or “shuffling” from a given H-bond partition, as is described in the method section (see Figure 3). The procedure continues with a new H-bond partitioning and for each of them optimized morphing parameters and corresponding RMSDs are obtained from fitting with the CHARMM-INoLLS interface. The relationship between the lowest RMSDs we obtain during our fitting and the numbers of groups which we partition the hydrogen bonds into in ubiquitin is shown in Figure 4 (see also Table S3 in the supporting information), together with the results for protein G. It must be pointed out that such a process can in principle go through an arbitrarily large number of cycles. Not all of them are explored here as the aim of the procedure we established is to produce meaningful data sets to address the questions we asked in the Introduction.

One relevant question concerns the trajectory length from which the RMSDs are computed. The 100 ps used so far may not be sufficient to fully converge the expectation value of ${}^{13}\text{J}_{\text{NC}'}$ and for the cycle described in Table 3 additional data from 500 ps trajectories was accumulated. This was done for cycles No. 2, 5 and 6. In cycles 2 and 6 the RMSDs are reduced by 0.003 Hz after extending the simulation time to 500 ps, while in cycle 5 the RMSD increases from 0.046 Hz to 0.068 Hz. Given that the RMSDs appear quite well converged it was decided to restrict the trajectories to 100 ps.

As illustrated in Figure 4, applying different hydrogen bonding potentials to different hydrogen bonds in proteins effectively enhances the correlation between measured and calculated scalar couplings. Partitioning hydrogen bonds into two groups already leads to very good correlation between computed and measured ${}^{13}\text{J}_{\text{NC}'}$ couplings. The correlation coefficients are $r^2 = 0.91$ and $r^2 = 0.93$ for ubiquitin and protein G, respectively (see Figure 5 and Table S1 in the SI). The RMSD obtained from partitioning into two groups for both proteins is 0.07 Hz, which is a reduction of 40% compared to previous results.[8] The maximal absolute error ($|{}^{13}\text{J}_{\text{NC}'}$ - ${}^{13}\text{J}_{\text{NC}'}$) and absolute relative error ($(|{}^{13}\text{J}_{\text{NC}'}$ - ${}^{13}\text{J}_{\text{NC}'})/|{}^{13}\text{J}_{\text{NC}'}$) have been reduced from 0.28 Hz (H-bond No. 21) and 104% (H-bond No. 10) to 0.14 Hz (H-bond No. 15) and 65% (H-bond No. 15) in ubiquitin, and from 0.25 Hz (H-bond No. 3 and No. 4) and 194% (H-bond No. 3) to 0.12 Hz (H-bond No. 19) and 91% (H-bond No. 13) in protein G. Groot *et al.* also calculated the ${}^{13}\text{J}_{\text{NC}'}$ couplings in ubiquitin and protein G from sub-microsecond MD simulations, and used the results to assess the agreement between experiment and simulations.[14] They report three scalar couplings across H-bonds, namely H-bond No. 5 and No. 17 in ubiquitin and H-bond No. 15 in protein G, to be outliers. For all six atomistic force fields tested there (OPLS/AA, CHARMM22, GROMOS96-43a1, GROMOS96-53a6, AMBER99sb, and AMBER03), the absolute errors ($|{}^{13}\text{J}_{\text{NC}'}$ - ${}^{13}\text{J}_{\text{NC}'}$) of all these three H-bonds are larger than 0.25 Hz.[14] In our simulations with two H-bond groups, the errors are 0.03 Hz, 0.05 Hz and 0.1 Hz, respectively, illustrating that these ${}^{13}\text{J}_{\text{NC}'}$ couplings can be reproduced by MD simulations with the MMPT force field and its parametrization for a partitioning into two groups.

Going beyond two groups can further improve RMSD and correlation as shown in Figure 4. However, it should be pointed

out that sometimes increasing the number of groups does not affect the RMSDs. For example in ubiquitin with 3 H-bond groups a RMSD of 0.055 Hz is obtained whereas with 4 H-bond groups 0.056 Hz is found. This suggests that the fitting cycles may not be fully converged. Increasing the number of groups to seven yields a RMSD of 0.026 Hz in ubiquitin. This is much lower than the 0.12 Hz with averaging MMPT treatment [8] or 0.06 Hz from biased simulations.[36] However, it should be mentioned that partitioning into more than two or three groups is only of interest in that they demonstrate how closely one can reproduce experimental coupling constants. In practice it will be much more important to have a balance between the number of parameters (as small as possible) and the ability to assign particular H-bonds to a specific group. This point will be further addressed in the Discussion.

Discussion and Conclusions

In this work we showed that very good agreement with experimentally measured scalar couplings can be achieved with MMPT/MD simulations by allowing different parametrizations for different kinds of H-bonds in proteins. The RMSDs between computed and experimental $^3J_{\text{NC}}$ couplings can be as low as 0.03 Hz, comparable to the experimental errors which are usually about 0.005 - 0.01 Hz [19, 22]. We also show that partitioning H-bonds into two groups gives a good correlation and is sufficient to reproduce experimentally measured $^3J_{\text{NC}}$ couplings to within 0.07 Hz. Such parametrizations can now be used to generate improved conformational ensembles from extensive MD simulations as has been previously done.[6, 8] It is important to point out that the force field refinement discussed in the present work includes both, the nuclear dynamics (through conformational averaging) and the influence of solvent (as all MD simulations are carried out with explicit solvent). An interesting point will be to determine the robustness of such parametrizations in view of water models other than TIP3P [37] which was used here, such as SPC/E [38] or TIP4P.[37]

An important next step after determining optimal force field parameters for H-bonds is the question whether it is possible to assign H-bonds in an arbitrary protein only from its structure. In other words, a way to group H-bonds from structural data alone was sought. One possibility is to use secondary structure. This possibility was already pursued above and it was found that this was not sufficiently robust. We illustrate this further in the following by considering a few couplings in ubiquitin. Given the partitioning from fitting $^3J_{\text{NC}}$ in ubiquitin (see Figure 1(c)) it is observed that coupling No. 16 at the end of the α -helix belongs to a different class than all other couplings in the helix. Conversely, couplings 4 and 21 are better described by α -helix parametrizations although they appear in β -sheets. Assigning coupling 16 to the group of α -helix couplings and couplings 4 and 21 to that for β -sheet and refitting the parameters leads to an RMSD of 0.11 Hz. This is an increase by 40% compared to the original assignment which demonstrates that secondary structure may not be the most robust criterion for classifying H-bond motifs. On the other hand it should be noted that these three couplings appear at special locations along the

protein structure, e.g. at one end of the α -helix (coupling 16) or at the end of a short β -sheet (coupling 21).

Another useful measure that is readily available from protein structure is the solvent exposure of hydrogen bonds. For this we consider the solvent accessible surface area (ASA) of H-bonded motifs in ubiquitin and protein G (see Figure 6). ASAs of a particular backbone H-bond motifs N-H...O=C were calculated by CHARMM with a probe radius of 1.4 Å, [39] and averaged along 10 ns MD trajectories with the standard CHARMM force field. 10 ns is sufficient for converging the ensemble averaged ASAs, as shown in Figure S1 in the SI. Again, H-bonds in the two different groups are discriminated by red and blue marks. In ubiquitin, the ASA appears to serve as a very good indicator. As shown in Figure 6(a), an ASA smaller or larger than 225 Å² separates the H-bonds into two groups and leads to an RMSD for scalar couplings of 0.070 Hz. However, using the same criterion in protein G (see Figure 6(b)) only yields an RMSD of 0.122 Hz. Therefore, ASA is also not a universally indicative criterion.

It should be emphasized that the current approach to improve force fields by fitting to ensemble-averaged data from simulations in explicit solvent is a computationally rather expensive procedure. To reduce this, it may be useful to combine this with a recently proposed scheme which retains the conformational ensemble and only re-evaluates the observables from a new parametrization.[9] This is of particular interest for the fitting cycles in which the number of groups in a partitioning is maintained. Another aspect is the combined fit of J-couplings from a library of proteins. It is hoped that this will yield robust and transferable parametrizations across classes of proteins.

Acknowledgement

We are grateful for the financial support granted by the Schweizerischer Nationalfonds through project Nr. 200021-117810 and through the NCCR MUST.

References

- [1] R. M. Levy, M. Karplus, P. G. Wolynes, Nmr relaxation parameters with interanl motion: Exact langevin trajectory results compared with simplified relaxation models, *J. Am. Chem. Soc.* 103 (1981) 5998–5011.
- [2] G. Lipari, A. Szabo, R. M. Levy, Protein dynamics and nmr relaxation: Comparison of simulations with experiment, *Nature* 300 (1982) 197–198.
- [3] M. Barfield, Structural dependencies of interresidue scalar coupling $^3J_{\text{NC}}$ and donor 1h chemical shifts in the hydrogen bonding regions of proteins, *J. Am. Chem. Soc.* 124 (2002) 4158–4168.
- [4] J. Gsponer, H. Hopearuoho, A. Cavalli, C. M. Dobson, M. Vendruscolo, Geometry, energetics, and dynamics of hydrogen bonds in proteins: Structural information derived from NMR scalar couplings, *J. Am. Chem. Soc.* 128 (2006) 15127–15135.
- [5] H.-J. Sass, F.-F. Schmid, S. Grzesiek, Correlation of protein structure and dynamics to scalar couplings across hydrogen bonds, *J. Am. Chem. Soc.* 129 (2007) 5898–5903.
- [6] F. F.-F. Schmid, M. Meuwly, Direct comparison of experimental and calculated nmr scalar coupling constants for force field validation and adaptation, *J. Chem. Theory Comput.* 4 (2008) 1949–1958.
- [7] S. A. Showalter, R. Bruschweiler, Validation of molecular dynamics simulations of biomolecules using NMR spin relaxation as benchmarks: Application to the AMBER99SB force field, *J. Chem. Theo. Comp.* 3 (2007) 961–975.

- [8] J. Huang, M. Meuwly, Explicit hydrogen-bond potentials and their application to nmr scalar couplings in proteins, *J. Chem. Theo. Comp.* 6 (2010) 467–476.
- [9] D.-W. Li, R. Bruschweiler, Nmr-based protein potentials, *Angew. Chem. Intern. Ed.* 49 (2010) 6778–6780.
- [10] M. Devereux, M. Meuwly, Force field optimization using dynamics and ensemble-averaged data: Vibrational spectra and relaxation in bound mbco, *J. Chem. Inf. Model.* 50 (2010) 349–357.
- [11] G. Lipari, A. Szabo, Model-free approach to the interpretation of nuclear magnetic-resonance relaxation in macromolecules i. theory and range of validity, *J. Am. Chem. Soc.* 104 (1982) 4546–4559.
- [12] G. Lipari, A. Szabo, Model-free approach to the interpretation of nuclear magnetic-resonance relaxation in macromolecules ii. analysis of experimental results, *J. Am. Chem. Soc.* 104 (1982) 4559–4570.
- [13] J. Alexander D. Mackerell, M. Feig, I. Charles L. Brooks, Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations, *J. Comp. Chem.* 25 (2004) 1400–1415.
- [14] O. F. Lange, D. van der Spoel, B. L. de Groot, Scrutinizing molecular mechanics force fields on the microsecond timescale with nmr data, *Biophys. J.* 99 (2010) 647–655.
- [15] L. Wickstrom, A. Okur, C. Simmerling, Evaluating the performance of the ff99sb force field based on nmr scalar coupling data, *Biophys. J.* 97 (2009) 853–856.
- [16] A. J. Dingley, S. Grzesiek, Direct observation of hydrogen bonds in nucleic acid base pairs by internucleotide $^2J_{\text{nm}}$ couplings, *J. Am. Chem. Soc.* 120 (1998) 8293–8297.
- [17] G. Cornilescu, B. E. Ramirez, M. K. Frank, G. M. Clore, A. M. Gronenborn, A. Bax, Correlation between $^3J_{\text{H}^{\alpha}\text{H}^{\beta}}$ and hydrogen bond length in proteins, *J. Am. Chem. Soc.* 121 (1999) 6275–6279.
- [18] F. Cordier, S. Grzesiek, Direct observation of hydrogen bonds in proteins by interresidue $^3J_{\text{H}^{\alpha}\text{H}^{\beta}}$ scalar couplings, *J. Am. Chem. Soc.* 121 (1999) 1601–1602.
- [19] S. Grzesieka, F. Cordiera, V. Jaravinea, M. Barfield, Insights into biomolecular hydrogen bonds from hydrogen bond scalar couplings, *Prog. Nucl. Magn. Reson. Spectrosc.* 45 (2004) 275–300.
- [20] K. Pervushin, A. Ono, C. Fernandez, T. Szyperski, M. Kainosho, K. Wüthrich, *Proc. Natl. Acad. Sci.* 95 (1998) 14147–14151.
- [21] H.-J. Sass, F. F.-F. Schmid, S. Grzesiek, Correlation of protein structure and dynamics to scalar couplings across hydrogen bonds, *J. Am. Chem. Soc.* 129 (2007) 5898–5903.
- [22] I. Alkorta, J. Elguero, G. S. Denisov, A review with comprehensive data on experimental indirect scalar nmr spin-spin coupling constants across hydrogen bonds, *Magn. Reson. Chem.* 46 (2008) 599–624.
- [23] S. Lammers, M. Meuwly, On the relationship between infrared spectra of shared protons in different chemical environments: A comparison of protonated diglyme and protonated water dimer, *J. Phys. Chem. A* 111 (2007) 1638–1647.
- [24] S. Lammers, S. Lutz, M. Meuwly, Reactive force fields for proton transfer dynamics, *J. Comput. Chem.* 29 (2008) 1048–1063.
- [25] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, M. Karplus, Charmm: A program for macromolecular energy, minimization, and dynamics calculations, *J. Comp. Chem.* 4 (1983) 187–217.
- [26] S. Vijay-Kumar, C. E. Bugg, W. J. Cook, Structure of ubiquitin refined at 1.8 Å resolution, *J. Mol. Biol.* 194 (1987) 531–544.
- [27] H. L. F. Schmidt, L. J. Sperling, Y. G. Gao, B. J. Wylie, J. M. Boettcher, S. R. Wilson, C. M. Rienstra, Crystal polymorphism of protein gb1 examined by solid-state nmr spectroscopy and x-ray diffraction, *J. Phys. Chem. B* 111 (2007) 14362–14369.
- [28] A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkerich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wirkiewicz-Kuczera, D. Yin, M. Karplus, All-atom empirical potential for molecular modeling and dynamics studies of proteins, *J. Phys. Chem. B* 102 (1998) 3586–3616.
- [29] J. M. Bowman, B. Gazdy, A simple method to adjust potential energy surfaces: Application to hco, *J. Chem. Phys.* 94 (1991) 816–817.
- [30] M. Meuwly, J. M. Hutson, Morphing *ab initio* potentials: A systematic study of nehf, *J. Chem. Phys.* 110 (1999) 8338–8347.
- [31] M. Law, J. Hutson, I-nolls: a program for interactive nonlinear least-squares fitting of the parameters of physical models, *Comput. Phys. Commun.* 102 (1997) 252–268.
- [32] J. C. Shelley, M. Y. Shelley, R. C. Reeder, S. Bandyopadhyay, M. L. Klein, A coarse grain model for phospholipid simulations, *J. Phys. Chem. B* 105 (2001) 4464–4470.
- [33] S. Izvekov, M. Parrinello, C. J. Burnham, G. A. Voth, Effective force fields for condensed phase systems from *ab initio* molecular dynamics simulation: A new method for force-matching, *J. Chem. Phys.* 120 (2004) 10896–10913.
- [34] S. Izvekov, G. A. Voth, A multiscale coarse-graining method for biomolecular systems, *J. Phys. Chem. B* 109 (2005) 2469–2473.
- [35] T. Kortemme, A. V. Morozov, D. Baker, An orientation-dependent hydrogen bonding potential improves prediction of specificity and structure for proteins and protein-protein complexes, *J. Mol. Biol.* 326 (2003) 1239–1259.
- [36] J. Gsponer, H. Hopearuo, C. M. Cavalli, A.; Dobson, M. Vendruscolo, Geometry, energetics, and dynamics of hydrogen bonds in proteins: Structural information derived from nmr scalar couplings, *J. Am. Chem. Soc.* 128 (2006) 15127–15135.
- [37] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, Comparison of simple potential functions for simulating liquid water, *J. Chem. Phys.* 79 (1983) 926–935.
- [38] H. J. C. Berendsen, J. R. Grigera, T. P. Straatsma, The missing term in effective pair potentials, *J. Phys. Chem.* 91 (1987) 6269–6271.
- [39] B. Lee, F. Richards, The interpretation of protein structures: Estimation of static accessibility, *J. Mol. Biol.* 55 (1971) 379 – 400.

Figure Legends

Figure 1:

Ubiquitin (panels a, c) and protein G (panels b, d). In (a) and (b), H-bonds are highlighted by CPK atoms and yellow or black lines, which corresponds to the partition into two H-bond groups which yields RMSDs of 0.07 Hz. (See text for details) The 3D structures (a) and (b) are projected into 2D schemes (c) and (d), in which the secondary structures are shown. H-bonds are marked by arrows pointing from donor to acceptor residues, and numbered with red or blue for the discrimination of two H-bond groups as above.

Figure 2:

Comparison of experimental and computed $^{13}J_{NC'}$ coupling constants from MMPT/MD simulations for ubiquitin with all H-bonds treated equally (1ubq-equal), for ubiquitin with H-bonds in different secondary structures treated separately (1ubq-SS), and for protein G with all H-bonds treated equally (2qmt-equal) and for protein G with H-bonds in different secondary structures treated separately (2qmt-SS). The red filled circles, green filled squares and blue empty circles correspond to H-bonds in α -helix, β -sheets and loop regions, respectively.

Figure 3:

Illustration of the fitting procedure of the morphing parameters for ubiquitin. Four H-bond partitionings are shown. The “splitting”, “clustering” and “shuffling” schemes to generate them are illustrated with blue dotted lines. For each H-bond partition, the H-bond groups, the optimized morphing parameters $\{R, \rho\}$ and the lowest RMSD we found are listed.

Figure 4:

Correlation between lowest RMSDs obtained from the fitting procedure and the numbers of H-bond groups for ubiquitin (black) and protein G (red).

Figure 5:

Comparison of experimental and computed $^{13}J_{NC'}$ coupling constants from MMPT/MD simulations with two H-bonds groups (marked as red circles and blue squares) described with different morphing parameters. For ubiquitin (panel a), the morphing parameters found are $\{2.871, 0.182\}$ and $\{2.979, 0.100\}$, which leads to a RMSD of 0.07 Hz. For protein G (panel b), the morphing parameters used are $\{2.905, 0.167\}$ and $\{2.979, 0.079\}$, which also leads to a RMSD of 0.07 Hz.

Figure 6:

The solvent accessible surface areas of hydrogen bonding motifs in ubiquitin (a) and protein G (b). The two groups of H-bonds are marked with red and blue circles.

4 Applications II: Hydrogen Bonds and NMR Couplings in Proteins

Table 1: Summary of morphing parameters and RMSDs calculated by MMPT/MD simulations with separate treatment of H-bonds in different secondary structural elements.

	α -helix $\{R_\alpha, \rho_\alpha\}$	β -sheet $\{R_\beta, \rho_\beta\}$	loop $\{R_{loop}, \rho_{loop}\}$
ubiquitin	{2.929, 0.148}	{2.923, 0.160}	{2.932, 0.142}
protein G	{2.991, 0.157}	{2.951, 0.147}	

Table 2: Comparison of RMSDs between experimental and computed $^{\text{h}^3}J_{\text{NC}}$ coupling from MMPT/MD simulations with all H-bonds treated equally (equal) and with H-bonds in different secondary structures treated separately (SS) in ubiquitin and protein G. RMSDs calculated for all couplings as well as for couplings in particular secondary structures (α -helix, β -sheet and loop region) are reported.

		α -helix	β -sheet	loop	total
ubiquitin	equal	0.119	0.126	0.087	0.118
	SS	0.117	0.121	0.088	0.115
protein G	equal	0.165	0.112		0.128
	SS	0.156	0.110		0.125

Table 3: Fitting cycles for ubiquitin with H-bonds separated into three groups with corresponding morphing parameters $\{R_1, \rho_1\}$, $\{R_2, \rho_2\}$ and $\{R_3, \rho_3\}$. RMSD_{100} is calculated from 100 ps simulations while RMSD_{500} is from 500 ps simulations (see text for details).

cycle No.	R_1 (Å)	ρ_1	R_2 (Å)	ρ_2	R_3 (Å)	ρ_3	RMSD_{100} (Hz)	RMSD_{500} (Hz)
cycle 1	2.900	0.160	2.920	0.140	2.940	0.120	0.088	-
cycle 2	2.900	0.165	2.922	0.140	2.955	0.115	0.074	0.071
cycle 3	2.901	0.165	2.920	0.141	2.955	0.115	0.079	-
cycle 4	2.866	0.181	2.925	0.141	2.976	0.100	0.062	-
cycle 5	2.866	0.180	2.922	0.140	2.975	0.100	0.046	0.068
cycle 6	2.868	0.181	2.922	0.140	2.975	0.099	0.055	0.052

4 Applications II: Hydrogen Bonds and NMR Couplings in Proteins

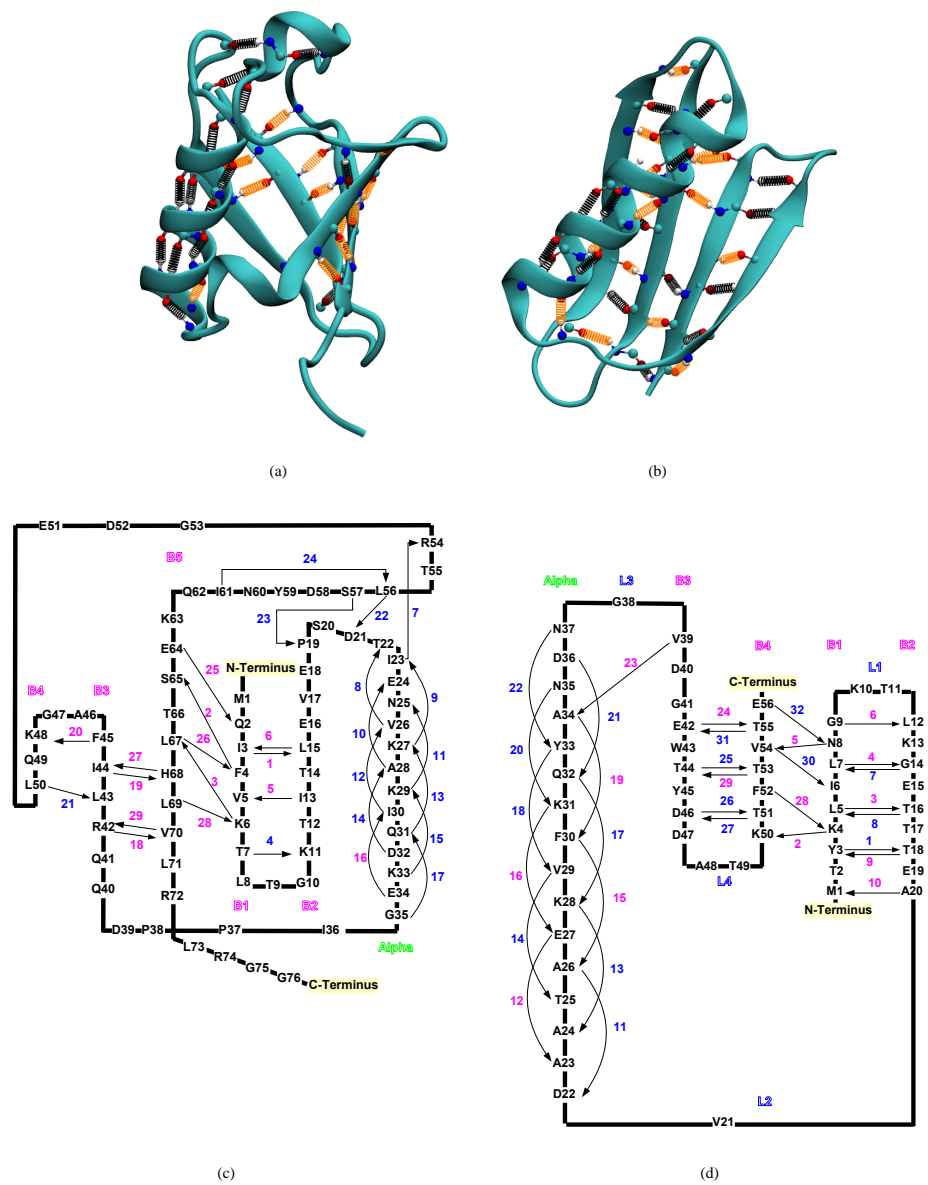


Figure 1: Huang et al

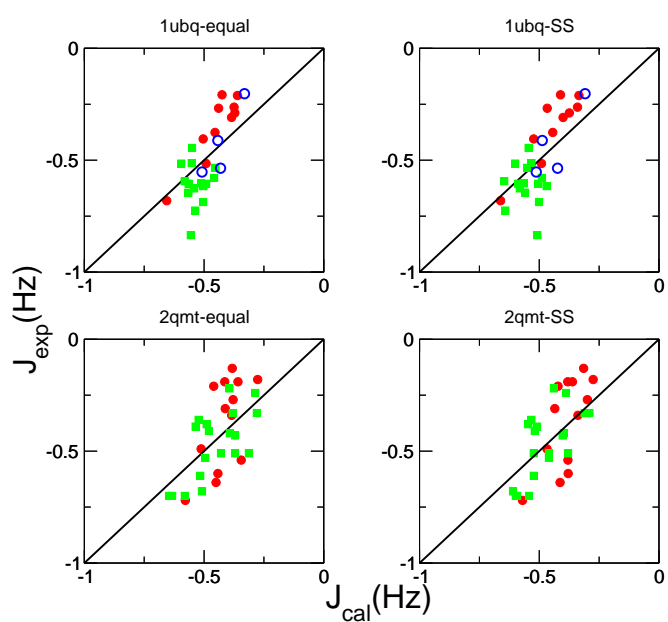


Figure 2: Huang et al

4 Applications II: Hydrogen Bonds and NMR Couplings in Proteins

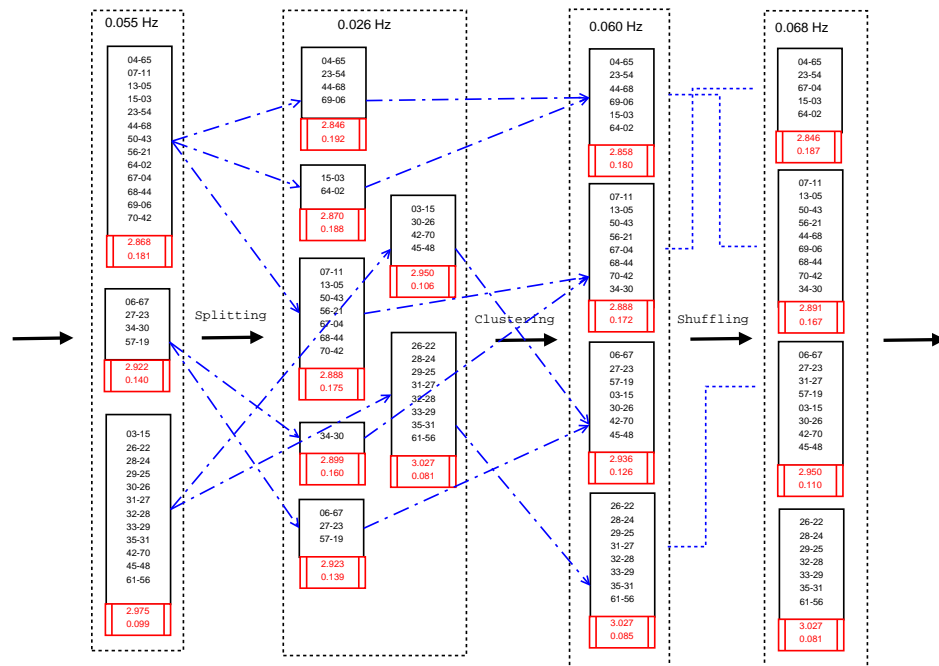


Figure 3: Huang et al

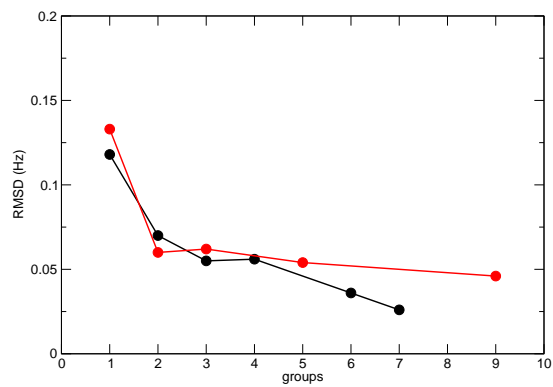


Figure 4: Huang et al

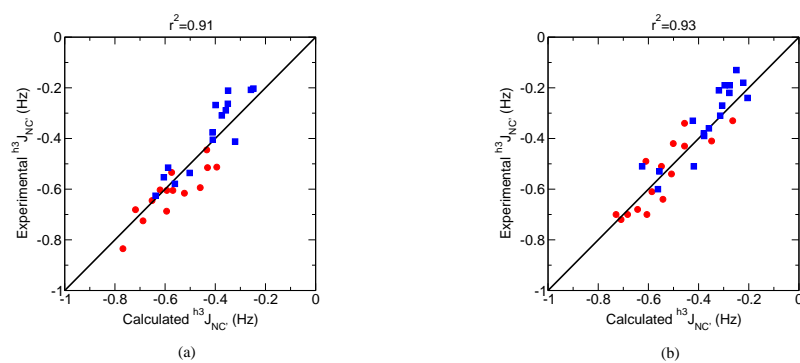
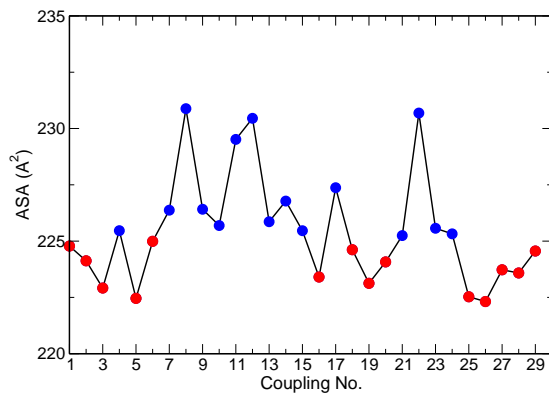
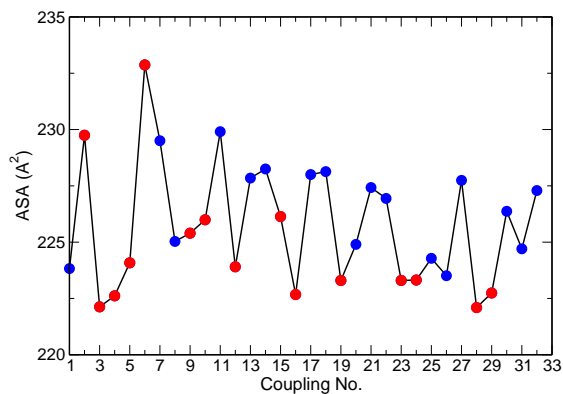


Figure 5: Huang et al



(a)



(b)

Figure 6: Huang et al

4.3 Discussion

Recent advances in NMR methodologies^{228,229} and atomistic simulations^{230,231} make protein dynamics accessible at atomic resolution on a wide range of time scales. However, direct characterization of H-bonds in proteins is still very difficult both experimentally and computationally. One recently available NMR observable is the hydrogen bond scalar couplings that can be measured through NMR spectroscopy.^{232,233} Scalar couplings across N–H···O=C H-bonds in proteins have a typical range of about -0.2 to -1 Hz and the measurement errors are usually less than 0.01 Hz.^{234,235} On the other hand, ${}^{\text{h}^3}J_{\text{NC}'}$ scalar couplings are known to be very difficult to be determined computationally. The RMSDs between experimental and back-calculated scalar couplings are typically larger than 0.15 Hz and the correlation coefficients r^2 are usually below 0.5.^{236,237} The reliability of computation for ${}^{\text{h}^3}J_{\text{NC}'}$ is sometimes so low that the accuracy of experimental measurements of particular couplings is questioned by theoreticians.²³⁷ Here, MMPT is used as an explicit coordinate-dependent potential for H-bonds in proteins, which leads to better correlations between computationally predicted and experimentally measured scalar couplings across hydrogen bonds.

In section 4.1, a general procedure to improve the hydrogen bonding potential via “PES morphing” based on experimental information is derived by considering six proteins with different overall topology. Hydrogen bonding dynamics between standard MD and MMPT/MD simulations have been also compared. Stronger hydrogen bonding, shorter hydrogen-bond length and more pronounced directionality have been observed in MD simulations with MMPT. Analysis of the protein dynamics shows that the MMPT potentials rigidify the entire protein and lead to stronger correlation between residues coupled by H-bonds. It has also been found out that the optimized potentials lead to H-bond geometries in remarkable agreement with a statistical analysis of 52 high-resolution protein structures.²³⁸ This suggests that using explicit hydrogen bond potentials shifts the conformational ensemble sampled in MD simulation toward the experimentally measured one.²³⁸

Results in section 4.1 are based on an average treatment of H-bonds in proteins which means that the same parametrization of the MMPT potential is used for all hydrogen bonds in a given protein. Better agreement with experimentally measured scalar couplings can be achieved with MMPT/MD simulations by allowing different parametrizations for different kinds of H-bonds in proteins. For this, a more system-

atical way to determine PES morphing parameters has been established with the help of an interface²³⁹ between CHARMM and I-NoLLS, which allows us to reproduce experimentally determined $^{\text{h}3}J_{\text{NC}'}$ couplings to within 0.02 Hz at best for ubiquitin. This is based on taking account of the chemically changing environment by grouping the H-bonds into up to seven classes. However, grouping them into two classes already reduces the RMSD between computed and observed $^{\text{h}3}J_{\text{NC}'}$ couplings to 0.07 Hz in both ubiquitin and protein G. Thus, using ensemble-averaged data with two classes of H-bonds leads to substantially improved force fields.

An important next step is to provide a generic procedure on H-bond partitions in proteins. The question is whether it is possible to assign H-bonds in an arbitrary protein only from its structure. Two possible measures, secondary structure and solvent accessible surface area, have been tested in section 4.2, but none of them seem to be a universal and robust indicator on how to group H-bonds. This is still under investigation.

The procedure we established in section 4.2 is not limited to optimize MMPT parameters, but can in principle be applicable to any force field parameters. Heterogeneous experimental data can be used in such an approach to improve force fields by fitting to ensemble-averaged data from MD simulations in explicit solvent. This opens the possibility to include more NMR observables, such as order parameters²⁴⁰ and residual dipole couplings (RDCs)²⁴¹⁻²⁴³ in the force field parameterization and optimization.

It would also be interesting to improve the standard force field with the knowledge gained from MD simulation with MMPT, for example the dynamical ensembles and the H-bond partition schemes. The atomic partial charges²⁴⁴ and VdW parameters of the hydrogen bonding atoms could be reparameterized, possibly also according to the different chemical environments. This would make the results here useful for conventional atomistic simulations and interesting to a wider readership.

5 Applications III: Molecular Modelling of a Platinum Catalyst

"The gentleman is satisfied and composed; the mean man is always full of distress."

Confucius, *Analects*

In this chapter the molecular mechanics modelling of a novel hydrogen-bonded self-assembly bidentate Platinum catalyst is presented. We will briefly introduce force field approaches for transition metal complexes, in particular VALBOND methods. Then we will show how to combine MMPT with VALBOND to investigate such a catalytic system. MD simulations have been carried out, which provide effective ensemble sampling of the complex and enhance our understanding of the atomistic detail of motions.

5.1 Background

Hydroformylation is the largest homogeneous catalytic process in industry and the selectivity to linear or branched products is important. While bidentate ligand systems such as Ruthenium/BINAP²⁴⁵ and Rhodium/XANTPHOS²⁴⁶ can provide unique selectivity in homogeneous catalysis (see Scheme 5.1), the synthesis of these ligands is difficult and expensive. As an alternative to classical bidentate ligands in which the two donor sites are connected by a covalent connection, Breit *et al* proposed a new concept of generating bidentate ligands by self-assembly of monodentate ligands via complementary hydrogen bonding.²⁴⁷⁻²⁴⁹ A typical example is the self-assembly of 6-diphenylphosphanylpyridone (6-DPPon) in the presence of a transition metal such as platinum(II) and rhodium(I), as illustrated in Scheme 5.1. Such a complex can provide high activity and regioselectivity in the hydroformylation of a

range of functionalized terminal alkenes.²⁴⁷ The approach is intrinsically combinatorial and thus self-assembly ligand library have been prepared and tested in various catalytic transformations.²⁵⁰

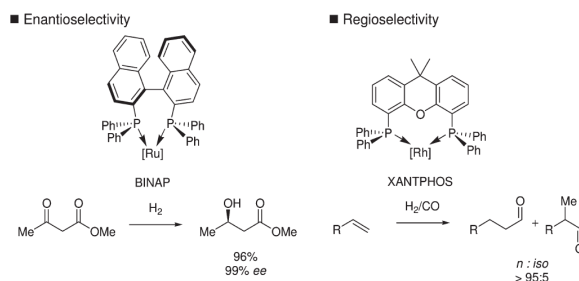


Figure 5.1: Enantio- and regioselectivity control with selected metal catalysts derived from bidentate ligands (courtesy of U. Gellrich, J. Huang, W. Seiche, M. Keller, M. Meuwly and B. Breit).²⁵¹

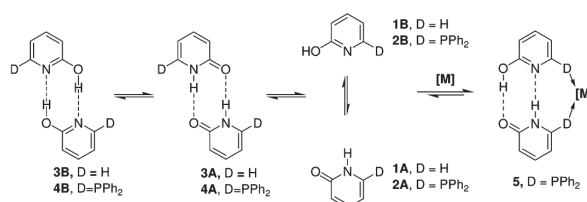


Figure 5.2: Self-Assembling of the 6-DPPon Ligand in the Presence of a Transition Metal Center (courtesy of U. Gellrich, J. Huang, W. Seiche, M. Keller, M. Meuwly and B. Breit).²⁵¹

Previously, the cis -[Cl₂Pt(6-DPPon)₂] complex has been characterized in detail using a range of spectroscopic methods (UV-Vis, IR, NMR, X-ray) and DFT calculations were used to rationalize a number of observations.²⁵¹ However, as the complex is rather dynamic, in particular around the double-hydrogen-bond motif between the pyridone and hydroxypyridine rings, and solvent effects are expected to play a role, methods beyond electronic structure calculations are needed. To this end, a combination of MMPT and force fields suitable for transition metal complexes are employed together with atomistic simulations.

5.2 Force Fields for Transition Metal Complexes

Computational chemistry can be of considerable value for our understanding of catalysis. For example reliable modeling and prediction of catalytically active transition metal complexes, can make the design and optimization of catalysts more efficient and targeted.²⁵² Such *in silico* screening approach remains a challenge despite the increasing computer power, because of the concurrent requirement of both computational efficiency and accuracy. DFT methods provide a good balance between the two and has been successful in building descriptors and revealing mechanisms.^{253,254} DFT scales cubically with system size, which often limits its application to practical catalytic systems. Furthermore, the effect of solvation is neglected or only described implicitly, and environmental factors including temperature and pressure are not considered at all although they may be critical to real catalytic processes.

Therefore, having a much faster, better scaling, even if approximate method is desirable. MM force fields offer computationally efficient means to investigate structural and dynamic properties of molecules in gas phase or condensed phase, and are routinely used for studying biomolecular systems with up to millions of atoms.²⁵⁵. However, the application of general force fields such as CHARMM¹⁶⁴, AMBER¹⁶¹ and OPLS¹⁶² to transition metal complexes have been limited. Difficulties encountered include: metals can have a variety of coordination numbers, π -binding ligands can bind in various ways, and electronic effects such as the Jahn-Teller distortion occur.^{256–259}

To address these problems, methods have been developed and improved over the past decade. Ligand Field Molecular Mechanics (LFMM)^{260–262} is based on ligand field theory of metal-ligand bonding interaction. By introducing an additional ligand field stabilization energy (LFSE) term, it has been successfully applied to “first-row” transition metal complexes^{263,264}. Another method is the “Sum of Interactions Between Fragments *Ab Initio*” (SIBFA),^{265,266} which is based on fragment-based energy decomposition and has been widely applied to organometallic systems from small complexes to metallo-proteins.^{267–269} The LFSE term can also be included and leads to the SIBFA-LF force field.²⁷⁰

An alternative approach to understand ligand-metal interactions in complexes is valence bond theory²⁷¹. VALBOND^{272–276} is a force field that computes the angle bending energies based on this theory. To capture bending energies over a wide range of angular distortions VALBOND replaces the harmonic angle bending potential in

conventional force field with a more sophisticated functional form²⁷²⁻²⁷⁴

$$E_i = k_i(S_i^{max} - S_i(\alpha_{ij})) \quad (5.1)$$

which is applicable for every ligand i . Here, k_i is an empirical scaling factor, S_i^{max} is the maximum strength for a particular hybrid orbital and $S(\alpha_{ij})$ is the reduction in maximal strength due to the particular angular arrangement of the two hybrid orbitals involved on two different ligands i and j . For two $sp^m d^n$ hybrid orbitals at an angle α the expression for S^{max} is

$$S^{max} = \sqrt{\frac{1}{1+m+n}(1 + \sqrt{3m} + \sqrt{5n})} \quad (5.2)$$

and $S(\alpha)$ is in general given by

$$S(\alpha) = S^{max} \sqrt{1 - \frac{1 - \sqrt{1 - \Delta^2}}{2}}. \quad (5.3)$$

The overlap Δ for two $sp^m d^n$ hybrid orbitals is

$$\Delta = \frac{1}{1+m+n} (|1s\rangle + m \cos \alpha |p_z\rangle + \frac{n}{2} (3 \cos^2 \alpha - 1) |d_{z^2}\rangle) \quad (5.4)$$

Combining equations 5.1 to 5.4 constitutes the core of VALBOND methods.

VALBOND is also applicable to hypervalent compounds by means of a 3-center-4-electron bonding model. It should be pointed out that VALBOND uses a 12-electron rule so all transition metals are considered hypervalent.²⁷⁴⁻²⁷⁶ Thus VALBOND can describe various geometries found in transition metal complexes, such as octahedral, trigonal bipyramidal, and square planar. For more detailed explanation and formula of nonhypervalent and hypervalent VALBOND, the reader is referred to the original literature. Recently, we have extended VALBOND by accounting for the trans influence²⁷⁷ which allows to carry out molecular dynamics (MD) simulations of metal-containing complexes. The MM force field methods, as Comba pointed out,²⁵⁹ can provide the same, or even better accuracy compared to DFT calculations, if a suitably parametrized force field is available.

5.3 Combining MMPT and VALBOND

5.3.1 Intermolecular Interactions

One advantage of force field methods is their extensibility. To model a chemically complicated compound, different and separately developed force field methods can be combined. A more familiar example might be the QM/MM approach that treats a localized region by QM methods and the rest of the system by MM force fields. Similarly, we can combine MMPT and VALBOND to study $\text{cis-}[\text{Cl}_2\text{Pt}(\text{6-DPPon})_2]$, a transition metal complexes whose ligands interact via strong hydrogen bonds.

As illustrated in 5.3, the $\text{cis-}[\text{Cl}_2\text{Pt}(\text{6-DPPon})_2]$ complex is partitioned into three parts, coordination sphere, hydrogen bonding motifs and the remaining atoms, which are treated with different force fields. The coordination sphere includes the center Pt atom and the ligand atoms P and Cl, and these five atoms are treated with VALBOND. The atoms involved in the hydrogen bonding between the pyridone (PY) and hydroxypyridine (HPY) rings are described by MMPT. The remaining atoms - aromatic carbon and hydrogen atoms - are described by the standard CHARMM force field.¹⁶⁴ Dealing with the interactions across the boundary of different MM regions is relatively simple since force field is additive. Both VALBOND and MMPT were developed together with CHARMM22,^{176,272} so mixing these force fields is meaningful.

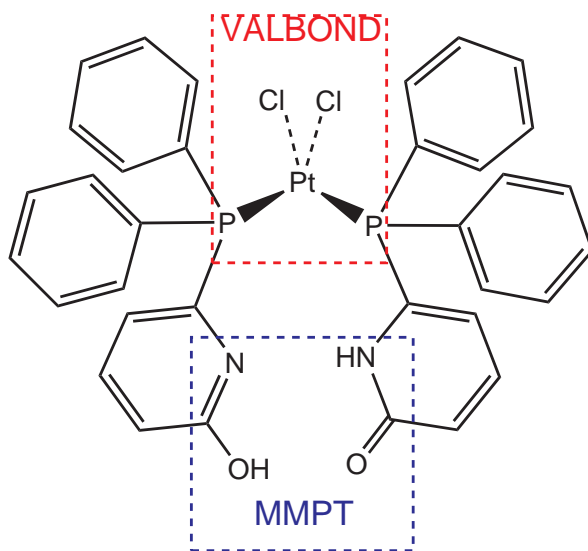


Figure 5.3: $\text{Cis-}[\text{Cl}_2\text{Pt}(\text{6-DPPon})_2]$, and the force field decomposition.

5.3.2 Force Field Parametrization

The reliability and accuracy of a force field depends on its parameters, and parametrization is usually done through careful and sometimes tedious fitting to experimental structure data and/or high level electronic structure calculations. In this chapter, however, parameters from each force field are directly carried over and already leads to good accuracy as shown in next section. For example, the VALBOND parameters for Platinum, Phosphorus and Chlorine are taken from Landis's original work^{272,274} and the corresponding supplementary materials.

For the MMPT part, "SDM" PESs in the previous study of 2PY2HP dimer¹⁷⁶ are considered as the zeroth order potentials.* The double hydrogen bonding motif in the Pt complex is chemically similar to the free 2PY2HP dimer but is embedded in a rigid framework due to the coordination. The steric and electronic effects of such a constraint can be taken into account by the following PES morphing²⁷⁸

$$\begin{aligned} V^{\text{NHN}}(R', \rho', \theta') &= \lambda_E \cdot V_0^{\text{NHN}}(R - R_1, \rho, \theta) \\ V^{\text{OHO}}(R', \rho', \theta') &= \lambda_E \cdot V_0^{\text{OHO}}(R - R_2, \rho, \theta) \end{aligned} \quad (5.5)$$

, where the morphing variables $\lambda_E = 0.8$, $R_1 = 0.2\text{\AA}$ and $R_2 = 0.08\text{\AA}$ have been chosen to reproduce the N-N and O-O distances measured by x-ray crystallography.²⁴⁷ MMPT parameters for the O-H-O and N-H-N motifs before and after PES morphing are listed in Table 5.1. Shape modification of the PESs is necessary to capture effects of the chemical environment.

The CHARMM22 parameter set¹⁶⁴ is used for all other force field terms required. When CHARMM parameters were missing, those from analogous CHARMM atom types or from relevant literature^{279,280} are taken. To help assign atomic partial charges, natural population analysis (NPA)²⁸¹ has been carried out and MM charges are set to the half of the NPA charges²⁷⁷ since NPA generally overestimates partial atomic charges.²⁸²

*Parameter p_{11} in Eq. 2.9 that is related to angular part of MMPT force field has been reparametrized. For isolated 2PY·2HP system, both hydrogen bonding motifs are almost linear, so the angular dependence is not important in our previous study and a heuristic 0.01 kcal/(mol·degree²) is used. Here a more carefully parametrized angular potential is needed, so this parameter is refitted and has been determined to be 0.007 kcal/(mol·degree²) for N-H-N motif and 0.009 kcal/(mol·degree²) for O-H-O motif, respectively.

2PY2HP		[Cl ₂ Pt(6-DPPon) ₂]	
O-H-O	N-H-N	O-H-O	N-H-N
161.593375	161.593375	129.2747	129.2747
2.160022	1.913428	2.160022	1.913428
2.016843	2.192999	1.936843	1.992999
-0.986537	-0.986537	-0.7892296	-0.7892296
0.534987	-0.101133	0.60618148	0.1023868
0.889931	1.017599	0.889931	1.017599
229.064043	198.052047	178.81829673	112.944422314
3.095393	2.808170	3.095393	2.80817
0.113593	0.114284	0.113593	0.114284
0.010000	0.010000	0.0072	0.0048
23.510782	27.482766	18.8086256	21.9862128

Table 5.1: MMPT PES parameters for the 2PY2HP dimer and the cis-[Cl₂Pt(6-DPPon)₂] complex.

5.3.3 Computational Methods

Molecular Dynamics Simulations

With rapid evaluation of energies and gradients, molecular dynamics (MD) simulations can be carried out. For MD simulations in the gas phase, the complex was first heated to 300 K by 6000 steps and equilibrated at that temperature for 100 ps. Then a 10 ns *NVE* trajectory was generated by free dynamics. The time step was $\Delta t = 0.1$ fs to follow the rapid proton motion. The total energies conserve well such MD simulations with the combination of VALBOND and MMPT force fields, as illustrated in Figure 5.4 where energies fluctuate according to a Gaussian distribution with a width of less than 0.01 kcal/mol.

MD simulations have also been carried out in explicit solvent including deuteriochloroform (CDCl₃), water (H₂O), toluene (C₆H₅CH₃), dichloromethane (CH₂Cl₂), deuterodichloromethane (CD₂Cl₂) and tetrachloromethane (CCl₄). Topologies and force field parameters are taken from Refs.^{283–285}. For these simulations, cubic solvent boxes with suitable sizes were used (CDCl₃: (45.97 × 45.97 × 40.86) Å³; H₂O: (34.15 × 31.04 × 31.04) Å³; toluene: (50.48 × 50.48 × 50.48) Å³; CH₂Cl₂ & CD₂Cl₂: (42.64 × 42.64 × 37.90) Å³; CCl₄: (48.96 × 48.96 × 48.96) Å³). The [Cl₂Pt(6-DPPon)₂] complex was then solvated and the entire system was heated to the desired temperature. Periodic boundary conditions and a cutoff of 12 Å was applied to the shifted electrostatic and switched Van der Waals interactions. *NVT* simulations were

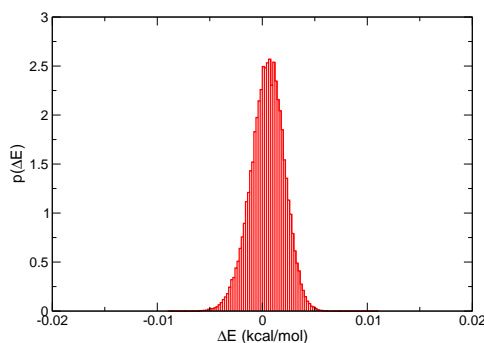


Figure 5.4: The fluctuations of total energies ΔE along a 10 ns MD trajectory with VALBOND and MMPT force fields for cis-Pt[Cl₂(6-DPPon)₂] in gas phase at 300K

carried out using a Nose-Hoover thermostat^{286,287} with a coupling constant of 50 kcal·mol⁻¹·ps². Before free dynamics simulations, the system was equilibrated for 100 ps.

Electronic Structure Calculations

DFT calculations were carried out using the B3LYP functional^{288,289} with the all-electron 6-31G(d,p) basis set for C, H, N, O, P atoms and a LANL2DZ effective core potential²⁹⁰ for Pt and Cl atoms. For the NMR and UV-Vis spectra, coordinates of the complex were recorded from MD snapshots at regular intervals and used in Gauge-Independent Atomic Orbital (GIAO)²⁹¹ or time dependent DFT (TD-DFT) calculations with a non-equilibrium polarized continuum model (PCM) solvation. To compare computed NMR shielding tensors with experimentally measured proton chemical shifts, the isotropic shielding constants $\sigma_{\text{iso}}(r)$ were referenced to TMS calculated at the same level of theory, *i.e.*, $\delta_{\text{iso}}(r) = \sigma_{\text{ref}} - \sigma_{\text{iso}}(r)$ where $\sigma_{\text{ref}} = 31.756$ ppm. To compute the UV-Vis spectrum, snapshots were extracted every 100 ps from a 5 ns NVE simulation trajectory of [Cl₂Pt(6-DPPon)₂] solvated in explicit CH₂Cl₂. Transition wavelengths and strength were computed with the CAM-B3LYP functional²⁹², convoluted using Gaussian curves by the cclib package²⁹³ and averaged over 50 frames. All electronic structure calculations were carried out using Gaussian03 with the grid=ultrafine option.¹⁸⁶

Infrared Spectrum

Infrared and power spectra are computed from the time correlation functions, as described in section 3.2.3.

5.4 Results

5.4.1 Equilibrium Structure

The calculated equilibrium structure of $\text{cis-}[\text{Cl}_2\text{Pt}(\text{6-DPPon})_2]$ from the force field (MM structure) compares favorably with structures from X-ray crystallography²⁴⁷ (X-ray structure) and DFT optimization (DFT structure) (see Figure 5.5). This is quantified in Table 5.2, where selected bond lengths and angles are reported. The natural bite angle ($\angle\text{P1-Pt-P2}$)²⁹⁴, a popular descriptor for chelating ligands, is reproduced by the MM structure compared to the X-ray structure. Another useful steric descriptor for P-Donor ligands is $S4'$ ²⁹⁵ constructed from angles at the coordinated phosphines. For these angles, the X-ray and DFT structures give the same values, which differ from those found from the force field calculations by 4° . Further improvement would be possible via the reparametrization of phosphine in VALBOND force field, but will not be pursued in the current work, as the difference is rather small. The N-H and O-H bond lengths reported in the MM structure are consistent with the DFT structure, and both are larger than those from X-ray experiments. Such a difference can be expected since locating hydrogen atoms in X-ray crystallography is difficult because of their low scattering power and asymmetrical electron density, and the refined structure depends on the computational algorithms used.

5.4.2 Potential Energy Surface

The PES for the hydrogen bonding motif in the $\text{cis-}[\text{Cl}_2\text{Pt}(\text{6-DPPon})_2]$ complex is further studied in detail. For this, the high-dimensional PES is projected onto a two-dimensional map $V(\rho_1, \rho_2)$, following the investigation of similar double proton transfer (DPT) systems such as 2PY2HP dimer⁷¹ and formic acid dimer²⁹⁶. The progression coordinate $\rho_1 := \frac{1}{2}(r_{\text{NH}} - \frac{1}{2}r_{\text{NN}} + r_{\text{OH}} - \frac{1}{2}r_{\text{OO}})$ describes mainly the proton motion and the coordinate $\rho_2 := \frac{1}{2}(r_{\text{NN}} + r_{\text{OO}})$ characterizes the intermonomer vibrations, where r_{NN} , r_{OO} , r_{NH} and r_{OH} are the separations between N-N, O-O, N-H and O-H, respectively. All the remaining degrees of freedom are relaxed. The PES

	X-ray	DFT	MM
Pt-P1 (Å)	2.256	2.313	2.312
Pt-P2 (Å)	2.236	2.306	2.302
∠P1-Pt-P2 (°)	97.58	105.44	97.80
∠Cl-Pt-Cl (°)	87.17	88.18	90.80
Pt-Cl* (Å)	2.339	2.462	2.406
N-N (Å)	3.070	2.966	2.967
N-H (Å)	0.84	1.033	1.032
∠N-H-N (°)	136.9	156.3	139.0
O-O (Å)	2.684	2.631	2.662
O-H (Å)	0.86	1.003	1.001
∠O-H-O (°)	164.3	173.3	178.6
P-C* (Å)	1.825	1.841	1.833
∠Pt-P-C* (°)	114.42	114.42	110.09
∠C-P-C* (°)	104.67	104.15	107.21

Table 5.2: Comparison of selected bond lengths and angles from the X-ray structure, the DFT structure and the MM structure. (*: average values.)

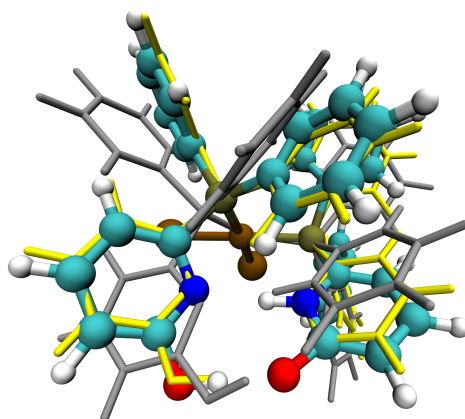


Figure 5.5: Equilibrium structure of $\text{cis-}[\text{Cl}_2\text{Pt}(\text{6-DPPon})_2]$ deduced from optimization with force field. The X-ray structure and DFT structure are also plotted with yellow and silver licorice, respectively.

is characterized by two minimum energy configurations corresponding to the two tautomeric forms of 2PY2HP. The first observation from the computed PES is that a

second minimum (marked by cross in Figure 5.6) lies 0.8 kcal/mol above the global minimum (marked by plus). This is also confirmed qualitatively by DFT calculations, which find an energy difference of 0.4 kcal/mol between these two minima. Although the global minima computed by MM and DFT overlay with each other, the predicted position of this second minimum from MM force field differs from that from DFT method by $\approx 0.1 \text{ \AA}$ in both ρ_1 and ρ_2 . Such an asymmetry, which is not observed in prototype systems such as the free 2PY2HP dimer, is expected since the hydrogen bonding motif is embedded into a complicated chemical environment and can be related to the difference in π -stacking effects between the aromatic rings in the complex. A transition state for proton transfer is identified via DFT calculations (marked by dot in Figure 5.6) and lies in the same transition state region observed from the MM computed PES. However, the MM force field predicts a PT energy barrier of 5 kcal/mol, compared to a barrier of 7.9 kcal/mol found from DFT calculations. A coordinate-depending energy scaling maybe used to improve this.

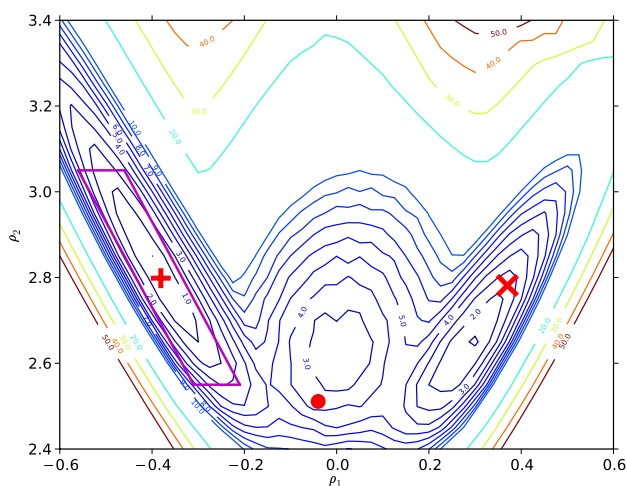


Figure 5.6: Potential energy surface for the hydrogen bonding motif in cis-[Cl₂Pt(6-DPPon)₂]. DFT determined minima and transition states are marked with red plus, cross and dot, respectively. Magenta parallelogram is drawn to cover the points to compare MM and DFT energies, see text for details.

One difficulty of the present force field is the disability to correctly describe ion pair states in which two hydrogen atoms are bonded to the same PY/HPY ring. Compared to DFT calculations, such conformations are overstabilized. DFT calculations

find such states at 13 kcal/mol above the global minimum whereas optimization of such a conformation with the force field leads to a local minimum around $\rho_1 = 0$. However, the stabilization is marginal and does not affect the dynamics simulations because trajectories seldom sample this region of configuration space, as shown in Figure 5.7. This figure reports projections of configurations onto the ρ_1 , ρ_2 and PES from 10 ns simulations in the gas phase and 5 ns in CDCl_3 .

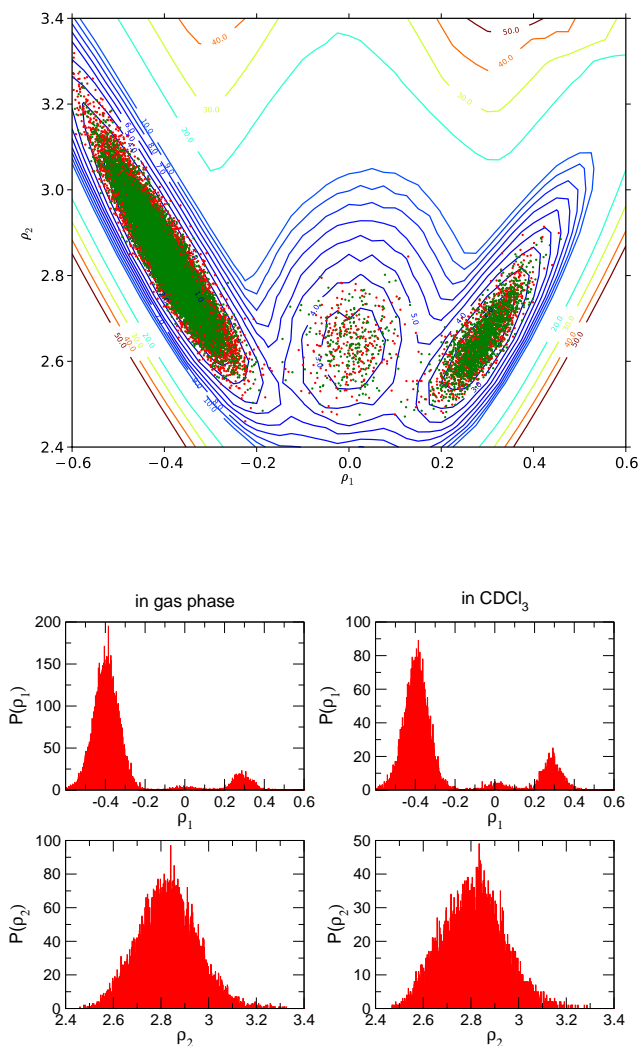


Figure 5.7: Up panel: The projection of a 10 ns MD trajectory in gas phase (red) and a 5 ns MD trajectory in explicit CDCl_3 solvation (green) onto the potential energy surface computed by MM force field. Bottom panel: The histograms of both projection onto ρ_1 and ρ_2 .

Further assessment is carried out by comparing a series of single point energy calculations with the N-N distance R_{NN} , the O-O distance R_{OO} , the N-H bond length r_{NH} and the O-H bond length r_{OH} fixed and all the remaining degrees of freedom fully relaxed. Such constrained optimizations are carried out with both DFT and MM force field on a regular four-dimensional grid of 108 points. These points are shown in the magenta parallelogram in Figure 5.6, and the resulting two sets of energies are compared in Figure 5.8. The correlation coefficient r^2 between the MM and DFT energies is 0.92, the mean absolute error (MAE) equals 0.31 kcal/mol, and the root mean standard deviation (RMSD) is 0.41 kcal/mol.

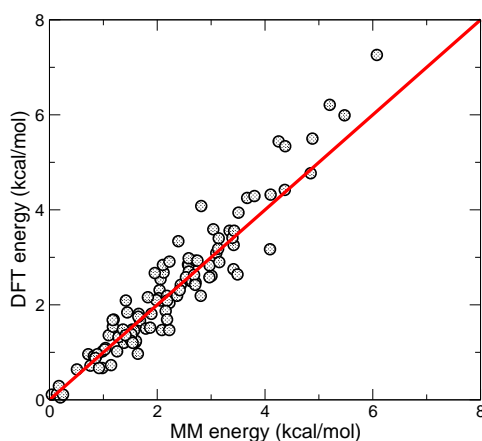


Figure 5.8: The comparison of MM and DFT energy profiles.

5.4.3 IR Spectrum

The experimental infrared spectrum of $\text{cis}[\text{Cl}_2\text{Pt}(\text{6-DPPon})_2]$ ²⁵¹ is characterized by an unusually broad peak between 2500 cm^{-1} and 3500 cm^{-1} . DFT calculations on the equilibrium structure did not find any spectral signature in this region. However by perturbing the equilibrium structure towards the DPT transition state, the harmonic vibrational analysis yielded the characteristic spectral signatures. This suggests that the infrared signatures between 2500 cm^{-1} and 3500 cm^{-1} have a dynamical origin. To better characterize and assign these features, the IR and power spectra for $\text{cis}[\text{Cl}_2\text{Pt}(\text{6-DPPon})_2]$ are calculated directly from MD simulations. One advantage of this approach is that power spectra of individual vibrational motions can be calculated, which helps assigning spectra.⁷²

Computed IR spectra from the 10 ns MD simulation in gas phase are shown in Figure 5.9, together with the experimental and DFT calculated ones. Peaks at the red end ($1000\text{ cm}^{-1} - 1700\text{ cm}^{-1}$) are consistent with the experimental measurements. The power spectrum of the HPY OH-bending (5.10 a) and the CO stretching vibration in both PY and HPY rings (5.10 b) enable us to assign the measured 1653 cm^{-1} peak to the CO-symmetric and asymmetric stretching vibration and the 1584 cm^{-1} peak to OH rocking vibration. Furthermore, one finds that these two motions are strongly coupled. The power spectrum of the N-N distance reports a single peak at 1200 cm^{-1} . The N-N separation motions can be regarded as the relative vibration between PY and HPY rings in the complex and correspond to the observed peak at 1240 cm^{-1} , which cannot be assigned with DFT calculations due to the considerable anharmonicity. It is worthwhile to mention that this peak occurs in the IR spectra of the $[\text{Cl}_2\text{Pt}(\text{6-DPPon})_2]$ complex but is missing in the spectrum of the 6-DPPon ligand itself. (See the Figure 11 in Ref. 251.) Finally, power spectra spread from 3000 cm^{-1} to 3500 cm^{-1} for OH bond stretching and from 2500 cm^{-1} to 3000 cm^{-1} for NH bond stretching, which assign the very broad peak between 2500 cm^{-1} and 3500 cm^{-1} to the proton motions in the hydrogen bonding motif.

The experimental spectrum is recorded in CCl_4 with $c = 10^{-3}\text{ M}$.²⁵¹ Although the solution is dilute, it is not identical to the situation in the gas phase for which the above DFT and force field calculations were carried out. Therefore, the IR spectrum from nanosecond MD simulations in explicit CCl_4 solution was also calculated, together with other organic solutes. Spectra of the complex in CCl_4 , CDCl_3 and CH_2Cl_2 share similar features. With the presence of these organic solutes in the simulations, the band between 2700 cm^{-1} and 3400 cm^{-1} , which corresponds to the hydrogen bonding in the complex, is now visibly broad. (See Figure 5.11.)

5.4.4 Proton Transfer Dynamics

Direct investigation of PT dynamics is very difficult experimentally. The most direct evidence for PT to occur are splitting of spectral features.⁴⁷ However, this is only available from high-resolution spectroscopy in the gas phase. Another method is to examine the tunneling contribution by proton transfer in the H/D KIE.⁵²⁻⁵⁴ Both experimental methods can not be applied to the system studied here. However, with a force field capable of forming and breaking hydrogen bonds, this is possible from MD simulations. The proton transfer dynamics in the complex is studied in detail and

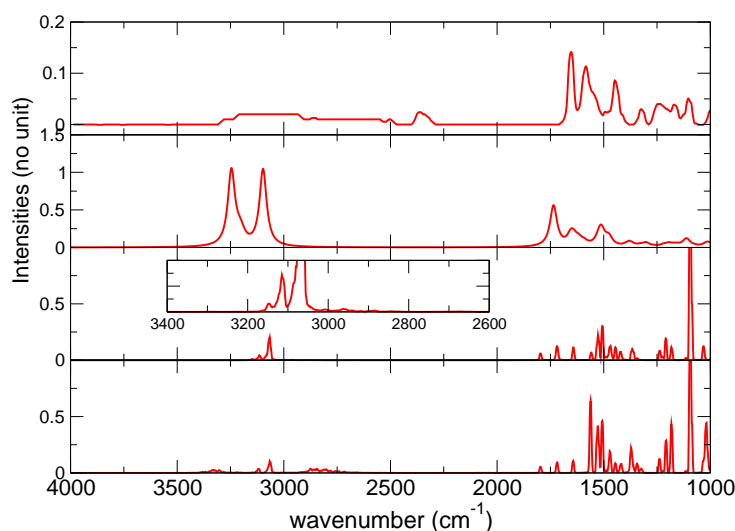


Figure 5.9: IR spectrum of $\text{cis-}[\text{Cl}_2\text{Pt}(\text{6-DPPon})_2]$. Upper panel: Experimental measurement; Middle upper panel: DFT calculation of the equilibrium structure; Middle bottom panel: computation from 10ns NVE MD simulations in gas phase; Bottom panel: computation from 2ns NVE MD simulations in CCl_4 .

compared with the PT in the topological similar but chemically simpler 2PY2HP. Figure 5.12 shows the time series of N-N, O-O, N-H and O-H distances for the complex during a 5 ns molecular dynamics simulation with explicit CDCl_3 solvation at 300K. Several PT events are observed and they are clearly correlated with the decrease of N-N separation distances. Also PT along the N-H-N and O-H-O motif is concerted. This agrees with the situation in the 2PY2HP dimer. Since the present force field underestimates the PT barrier by ≈ 3 kcal/mol, the actual PT rate is probably overestimated (up to three orders of magnitudes). However PT is likely to occur on a microsecond timescale. This agrees with results from NMR experiments which provide evidence for PT: in the ^{31}P -NMR spectra²⁵¹ only one signal is observed which indicates that the phosphorus bonded to the PY ring and that bonded to the HPY ring are magnetically equivalent, *i.e.*, tautomerization via double proton transfer is likely to take place on NMR timescales (see below).

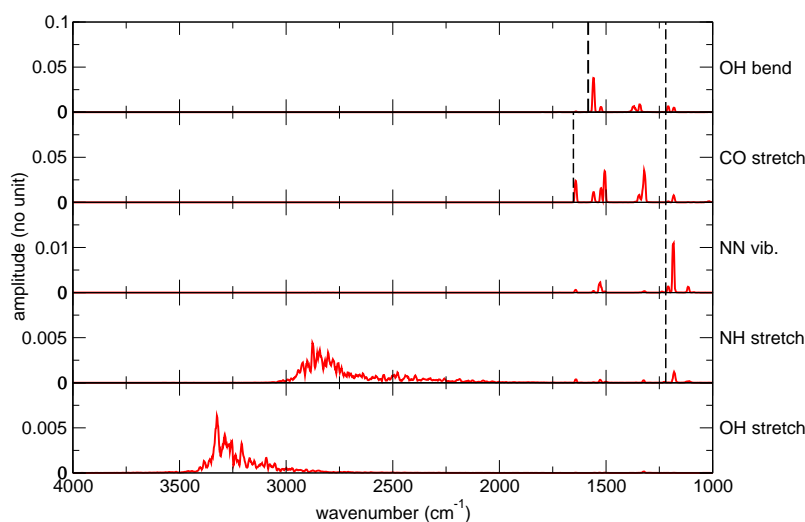


Figure 5.10: Power spectrum of $\text{cis-}[\text{Cl}_2\text{Pt}(\text{6-DPPon})_2]$ computed from 10 ns NVE MD simulations in gas phase. From up to bottom panel: OH bending; CO stretching; NN vibration; H vibration in the N–H···N motif; H vibration in the O–H···O motif. The dash lines from left to right correspond to 1653 cm^{-1} , 1584 cm^{-1} and 1240 cm^{-1} , respectively.

5.4.5 UV-Vis Spectrum

In an earlier work the UV-Vis spectrum of the complex was reported.²⁵¹ Comparison with the UV-Vis spectra of “fixed tautomer” model systems such as $[\text{Cl}_2\text{Pt}(\text{6-diphenylphosphanyl-1-methyl-pyridine-2-one})_2]$ and $[\text{Cl}_2\text{Pt}(\text{2-methoxy-6-diphenylphosphanylpyridine})_2]$ complexes concludes that the UV-Vis spectrum of the complex appears as the superposition of multi-tautomeric forms. This cannot be reproduced by a single TD-DFT calculation on the equilibrium structure. However, superposition of UV-Vis spectra from combined TD-DFT/MD calculations correctly reproduce the peaks around 325 nm and around 280 nm as shown in Figure 5.13. In fact, the ground to excited state transitions differ dramatically from frame to frame, while a realistic absorption profile is obtained by the ensemble average over 50 frames.

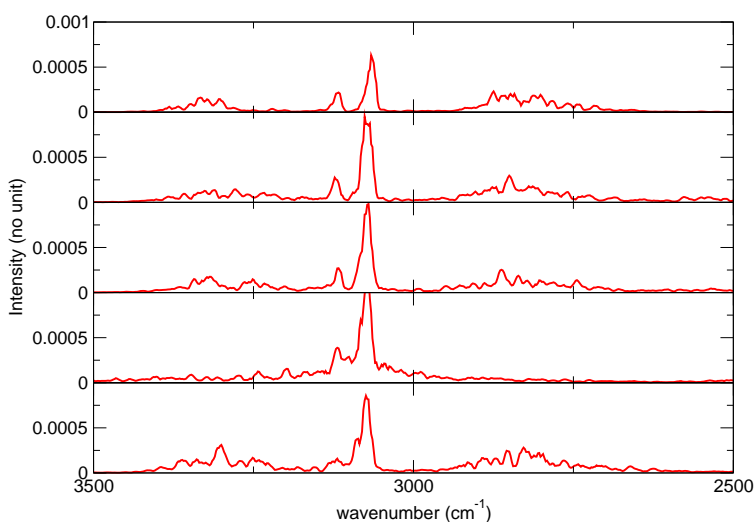


Figure 5.11: IR spectrum of $\text{cis-}[\text{Cl}_2\text{Pt}(\text{6-DPPon})_2]$ in different solvents. From up to bottom panel, 2 ns MD in CCl_4 solvation; 5 ns MD in CDCl_3 solvation; 5 ns MD in CH_2Cl_2 solvation; 5 ns MD in water; 2 ns MD in toluene.

5.4.6 NMR Spectrum

The proton chemical shifts can provide useful information on the structure and dynamics of the hydrogen atoms in the complex due to their sensitivity to the chemical environment. Here the ^1H chemical shift spectra for $[\text{Cl}_2\text{Pt}(\text{6-DPPon})_2]$ were simulated from MD simulations carried out at four different temperatures (200K, 250K, 300K and 350K) and in two different solvents (CDCl_3 and CD_2Cl_2). Coordinates, generated every 10 ps from individual 2 ns MD trajectories, were used in GIAO calculations and then averaged over 200 snapshots. The distribution of a typical proton chemical shift along a particular MD trajectory can extend over several ppm (see Figure 5.14 and Figure 5.15). Such broad distributions are also reported in the literature²⁹⁷⁻²⁹⁹. The ensemble averaged chemical shifts are summarized in Table 5.3, together with the single point calculation at the DFT optimized structure.

As illustrated in Table 5.3, computed chemical shifts for aromatic protons are temperature- and solvent-independent. The distributions of ^1H chemical shifts for a particular aromatic proton calculated from MD snapshots at different temperatures in CDCl_3 are reported in Figure 5.14. The distributions indicate that the MD ensem-

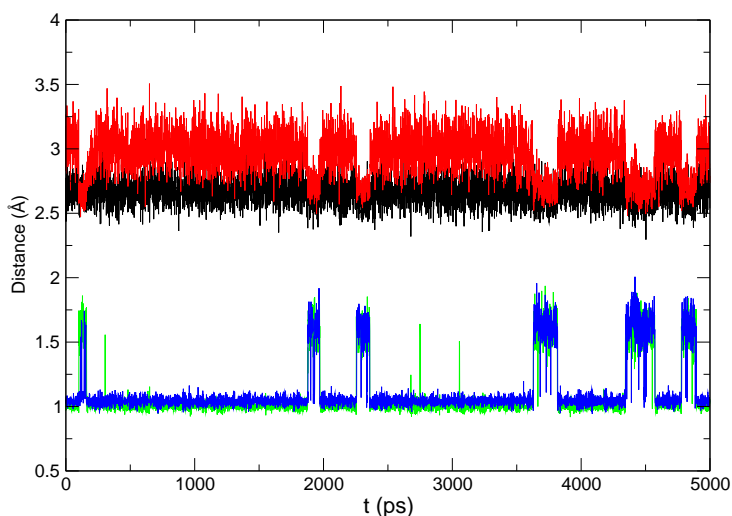


Figure 5.12: Time series of a 5 ns molecular dynamics trajectory for $\text{cis-}[\text{Cl}_2\text{Pt}(\text{6-DPPon})_2]$ in CDCl_3 solvation at 300K showing the N-N distances (red), the O-O distances (black), the N-H distances (blue) and the O-H distances (green).

		O-H-O	N-H-N	8H	6H	8H	2H	2H
single point	DFT opt	13.33	13.04	8.08	7.36	7.19	6.62	6.49
MD in CDCl_3	200K	11.88	14.45	7.63	7.33	7.17	6.70	6.43
	250K	11.71	14.63	7.66	7.35	7.18	6.69	6.48
	300K	12.07	15.18	7.67	7.34	7.18	6.76	6.46
	350K	12.59	15.41	7.69	7.38	7.19	6.74	6.45
MD in CD_2Cl_2	200K	12.30	15.37	7.63	7.34	7.15	6.75	6.43
	250K	11.80	14.65	7.65	7.35	7.17	6.75	6.47
	300K	11.64	14.72	7.63	7.34	7.20	6.72	6.42
	350K	12.23	15.71	7.68	7.34	7.20	6.77	6.46

Table 5.3: Calculated ^1H NMR chemical shifts (in ppm).

bles reflect the different conditions under which they were generated. However, the averaged chemical shifts are almost identical, as one would expect from NMR spectroscopy.³⁰⁰ This also shows that 200 snapshots might be sufficient to obtain statistically meaningful results. For protons involved in the double hydrogen-bonding motif, chemical shifts computed from MD ensemble are sensitive to the temperature.

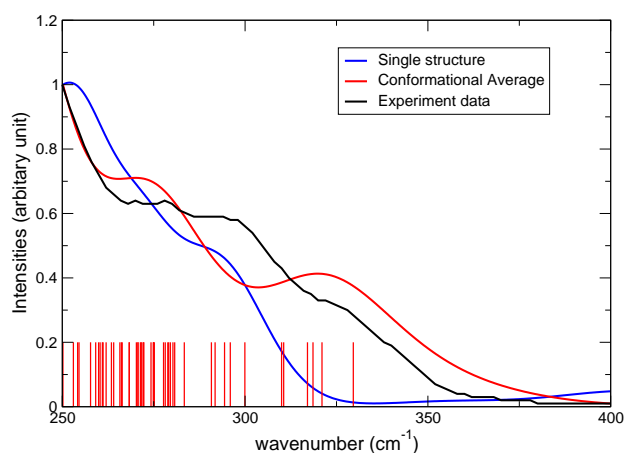


Figure 5.13: UV-Vis spectrum of $\text{cis-}[\text{Cl}_2\text{Pt}(\text{6-DPPon})_2]$. Black: experimental data; Blue: calculation on the optimized structure; Red: ensemble average along MD trajectories in CH_2Cl_2 ; Drop lines: the maximum peak position obtained from TD-DFT calculation of each single conformation.

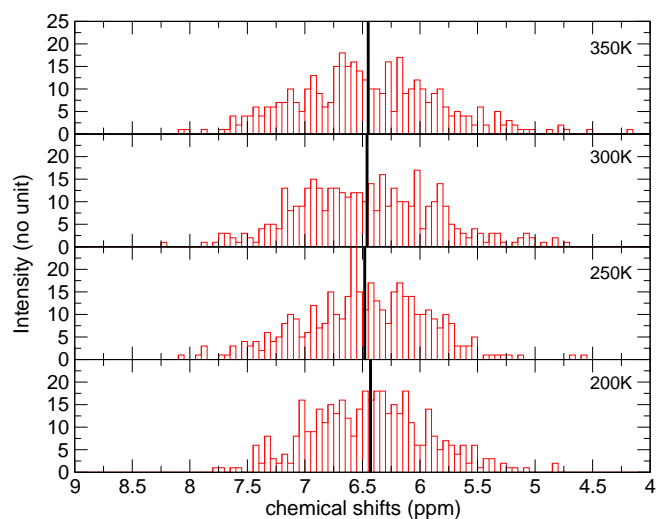


Figure 5.14: The distribution of computed proton chemical shifts from 2 ns MD simulations under different temperatures (200K, 250K, 300K, 350K) in CDCl_3 .

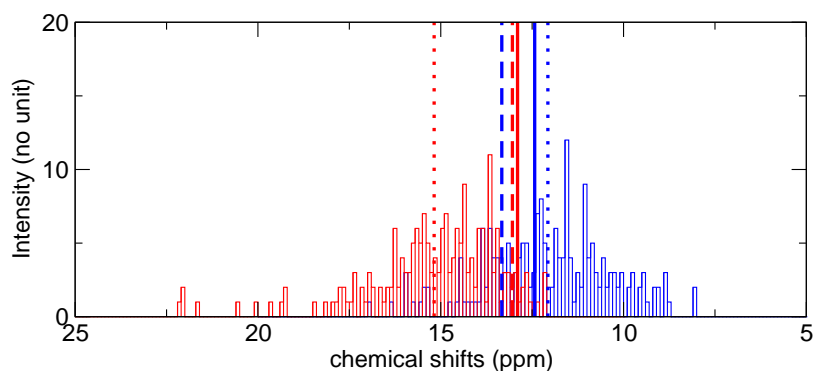


Figure 5.15: The distribution of computed proton chemical shifts in the DPT motif. The solid lines indicate the experimentally measured ones. Break lines corresponds to the calculation on the DFT optimized structure. Histograms show the distribution of computed chemical shifts calculated from 300K MD simulations in CDCl_3 . Red is for the proton in N-H-N motif while blue for that in O-H-O motif.

However, the dependence does not reflect the one found in earlier experiments.²⁵¹ In particular, the experimentally observed coalescence of two signals at 325K is not reproduced by the current approach. This suggests that the process that is reflected in the coalescence is not double proton transfer along the H-bonding motif. In fact, the rotation and tautomerization process that renders the two protons indistinguishable, as we conclude in our previous work, should compete with DPT and is likely to be the process that is responsible for such an NMR signal.²⁵¹

It is worth pointing out that calculation from MD ensemble averaging indicates that the proton involved in the N-H-N motif has a larger chemical shift than the proton in the O-H-O motif. This is at variance with the NMR-calculation on the DFT optimized structure as well as our assignment in the previous work.²⁵¹ Usually O-H-O protons exhibit larger chemical shifts than N-H-N protons because the electronegativity of

oxygen is larger than that of nitrogen. The result from the MD simulations reported above has, however, been confirmed by measuring the ^1H - ^{31}P couplings with 2D NMR spectroscopy. A direct coupling between the ^{31}P and the ^1H signal at 12.90 ppm has been detected by our collaborators, which clearly assigns the relatively larger (12.90 ppm) chemical shift to the proton in the N-H-N motif and the smaller one (12.43 ppm) to the proton in the O-H-O motif. This means that the amine proton is involved in stronger hydrogen bonding than the hydroxyl proton. This is supported by the fact that DPT is always triggered by the transferring of proton between two nitrogen atoms (see Figure 5.12).

5.5 Discussion

In this chapter we present a MM force field approach for a particular Pt complex featured by a double proton transfer motif. The force field is validated by comparing structures and energies with experimental data and DFT calculation results. MD simulations are carried out in gas phase and in different solvation, IR, UV-Vis and NMR spectra are computed and compared with experimental results, and hydrogen bonding patterns are analyzed.

The $\text{cis-}[\text{Cl}_2\text{Pt}(\text{6-DPPon})_2]$ complex studied here can be considered as a model system. In Breit's group, series of transition metal complexes (Pt, Rh and Ir as the center metal) with the self-assembly of structurally simple monodentate ligands - usually based on 2PY/2HP or aminopyridine/isoquinolone - are synthesised, tested and analyzed. This eventually leads to a library of hydrogen-bonded bidentate catalysts,²⁵⁰ with the system sizes vary from about 70 atoms to 300 atoms. The same combination of MMPT and VALBOND force field can be applicable to the compounds in this library, and the transferability of parameters can be expected since no particular parameter fitting is used in our investigation of the model $[\text{Cl}_2\text{Pt}(\text{6-DPPon})_2]$ complex. Based on a reasonable molecular modelling, virtual screening on the library of such catalysts would be a possibility.

The complex is also of interest in view of investigating the influence of the chemical environment on proton transfer (PT) processes. The parent system of the ligands is the free 2PY2HP dimer, a paradigm system for double proton transfer (DPT) that has been investigated both experimentally^{47,301} and through simulations.⁷¹ It is interesting to find out how the energetic and dynamics of DPT is affected by the complexation to a relatively rigid framework such as platinum coordination, and this will also

enhance our understanding of the concept of self-assembly via hydrogen bonding.

6 Extension of MMPT to Study Proton Transport

"Simplicity does not precede complexity, but follows it."

Alan Perlis, *Epigrams on Programming*, 1982

In the last three chapters, MMPT was applied successfully to the proton transfer and hydrogen bonding in complicated chemical and biological systems. However, the standard MMPT method presented in Chapter 2 cannot correctly describe continuous proton transfer. The proton transfer motifs that are treated with MMPT PESs should be defined at the beginning of MD simulations, and cannot be changed through the dynamical evolution. This prohibits the usage of MMPT in many applications, for example the step-wise proton transfer in LADH or the proton transport through transmembrane channels.

In this chapter, attempts to extend the MMPT method to study proton transport processes are presented. This involves identifying PT motifs in the system, and always moving the MMPT potential spatially to follow such PT motifs.

6.1 Basic ideas

Let's consider applying MMPT to the simplest (and perhaps also the most complicated) system: proton transport in bulk water. First the excess proton should be located, which can be done by finding out the oxygen atom with three hydrogen atoms bonded to it, *i.e.*, the H_3O^+ cation. The second step will be to find out in its first solvation shell the nearest water molecule (oxygen atom) and applying MMPT to the Zundel complex (H_5O_2^+) formed by them. The dynamics of this Zundel complex can be described with very high accuracy, given that MMPT provides suitable and well-parameterized PESs. During the time evolution of systems by MD simulations, the

oxygen atoms (potential proton acceptors) will rearrange. There exist two possibilities: a) another water molecule (oxygen atom) in the first solvation shell approaches the H_3O^+ , and becomes more appropriate to be considered in the Zundel complex treated by MMPT. For such a “special pair dance”,^{8,302} the MMPT potential applied to the original $\text{O}-\text{H}^+ \cdots \text{O}$ motif should be turned off and be applied to this new $\text{O}-\text{H}^+ \cdots \text{O}$ motif; b) the proton rattles between the donor and the acceptor atoms and transfers eventually to the acceptor oxygen atom, which is usually initiated by the reorganization of oxygen atoms in the first and second solvation shells. For this, the definition of donor and acceptor in MMPT motif should be swapped, and a new hydronium ion H_3O^+ is formed. Then dynamics continues from the second step as mentioned above.

The purpose of extending MMPT as “Molecular Mechanics with Proton Transfer” to “Molecular Mechanics with Proton Transport” is not for studying PT in water, but the proton transport processes in much larger biological and chemical systems, which is difficult to characterize using high-level *ab initio* MD simulations. The same logic can be used, though. First the correct D-H-A motifs treated with MMPT have to be located. Second, the methodology to change a particular part of the system from standard force field to MMPT force field, or *vice versa*, has to be developed. This can be compared with exchanging a water molecule between QM water and MM water in QM/MM methods or exchanging the representation of a particle between atomistic one and coarse-grained one in multiscale simulations.^{303,304} Here the description of certain atoms is changing between a general force field and a specialized force field with high accuracy (MMPT). Algorithms for this have been proposed and implemented into the MMPT module within CHARMM, and two test systems were setup for validation.

6.2 Test Systems

6.2.1 Proton Transport in Protonated Water Clusters

The first test system is the protonated water cluster $\text{H}_{2n+1}\text{O}_n^+$, which has been studied extensively from both experimental^{60,64} and computational perspectives.^{77,82,305} Here we use a small cluster with $n = 4$ as the test case, and set the initial conformation as a water chain, as plotted in Figure 6.1.

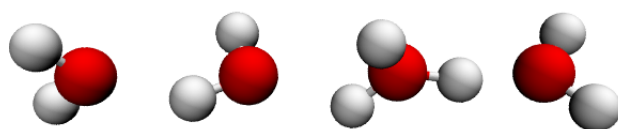


Figure 6.1: Test case for MMPT with proton transport: H_9O_4^+

6.2.2 Proton Transfer along Water Chains in Carbon Nanotubes

A more realistic system would be protonated water chains in carbon nanotubes. The confinement of protons into such narrow channels creates proton wires having up to 40 times faster proton transport than in bulk water.^{99,306–309} Also, this is a less complicated system than PT in water, because the carbon nanotube provides a semirigid environment that largely restricts the lateral and translational degrees of freedom.

Here, a carbon nanotube with a length of ~ 46 Å and diameter of 7.6 Å is constructed. The diameter was chosen such as to prevent reorganization (interchange of oxygen atoms) of the water molecules while allowing for some lateral freedom, so the water chain can be considered as a highly ordered one dimensional system. The proton wire inside the nanotube contains 18 water molecules with an excess proton, and in total the system is made up of 475 atoms (see Figure 6.2). A similar system setup was used to study the proton transfer along ammonia chains in carbon nanotubes with SCC-DFTB/MD method by Zoete and Meuwly.²⁵

With these two test systems, three questions are asked: a) whether the proton can transport. This amounts to whether the identities of PT motifs, *i.e.*, the definitions of donor, proton and acceptor atoms can be changed “on the fly” during simulations, and whether MMPT potentials can correctly follow the PT motifs; b) whether the dynamics is carried out correctly, which can be assessed by the energy conservation situation during NVE simulations. Also strange (energetically highly unfavorable) conformations, for example two O-H bonds pointing towards each other with very close H-H separation, should not be frequently observed in the simulations. c) how the physical observables computed from MMPT/MD simulations compare with other simulation methods. These can be structural information such as the radial distribution function (RDF) of the protonated water cluster, or dynamical properties such as the time (in a statistical sense) it takes for H_9O_4^+ to evolve from an extended proton

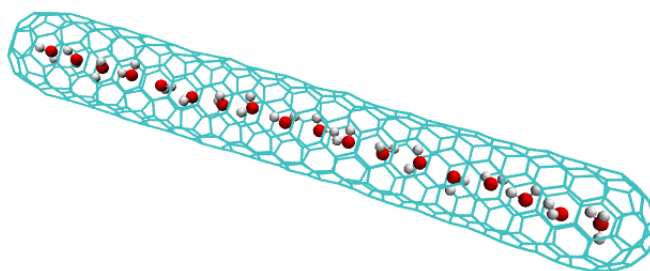


Figure 6.2: Test case for MMPT with proton transport: 1D water wire ($\text{H}_3\text{7O}_{18}^+$) confined in a carbon nanotube.

water wire to the conformation of an Eigen cation, or in a similar way the time for a proton to transport from one end of the nanotube to the other end or from the middle of the nanotube to its ends. SCC-DFTB/MD simulations will be carried out as a benchmark, while MS-EVB or CPMD methods could also be used for comparison.

We will try to answer these questions with the methods as described in the next section.

6.3 Schemes and Implementations

6.3.1 Identifying the MMPT Motif

As briefly mentioned above, the first step to identify a MMPT motif is to find the donor atom. Two criteria can be used: a) for each oxygen atom, find the three nearest hydrogen atoms H_1 , H_2 and H_3 , then compare all the $r_{\text{OH}_1} + r_{\text{OH}_2} + r_{\text{OH}_3}$'s and the smallest one corresponds to the donor atom; b) based on the assumption that PT is step-wise, the current donor atom is either the donor or the acceptor in the previous step, thus by directly comparing the D–H and A–H distances we can pick up the donor atom as the smaller one. Both criteria are implemented and they are found to give the same results, so both criteria a) and b) are examined every step to locate the donor. Whenever they contradict (which is very rare), the criteria 'b' is used while a warning message is given by MMPT.

The next step is to locate all potential acceptors in the system. This is done with

a geometry cutoff r_{cut} , and all possible acceptor atoms within the radius r_{cut} of the donor atom will be included. Here we set $r_{\text{cut}} = 4 \text{ \AA}$.

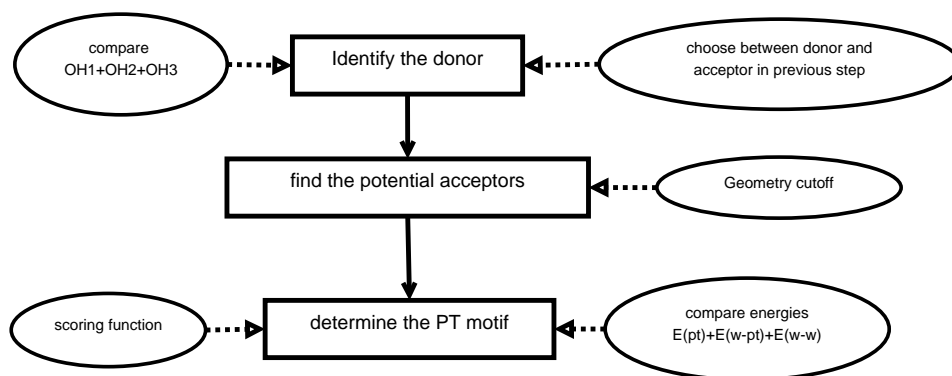


Figure 6.3: Flowchart for identifying MMPT motifs. See text for details.

Now for each donor atom there exist three potential protons and several possible acceptors, which leads to a number of possible D-H-A motifs. One D-H-A (the PT motif) should be chosen to be described with the MMPT PES. Two schemes have been proposed for this purpose. The first one is based on a scoring function that depends on internal coordinates. Scores are computed for each possible D-H-A motif, and the motif with the smallest score is treated by MMPT. The scoring function could be a simple function such as the distance between the hydrogen and the acceptor atom

$$\text{Score} = R_{HA} \quad (6.1)$$

or more complex ones such as

$$\text{Score} = 13(R_{DA} - 2.39)^2 + 1.5[1 + \cos(\theta_{HDA})] \quad (6.2)$$

, where R_{DA} is the D–A distance and θ_{HDA} is the angle between D–H and D–A. Eq. 6.2 can be considered as a harmonic approximation to our MMPT potential.

Such scoring functions can determine which D-H-A combination is the MMPT motif, however, directly changing the identity of the MMPT motif leads to large energy jumps. This is caused by the inconsistency between the MMPT potentials and standard force field terms (the harmonic bonding between possible donor and protons, plus several non-bonded terms), and will cause instability in MD simulations. To avoid such spurious energy jumps from instantaneous switching of potential functions, two PT motifs D-H₁-A₁ and D-H₂-A₂ are chosen with the smallest score S_1 and

the second smallest score S_2 , respectively. Energies computed with either D-H₁-A₁ motif treated by MMPT (V_1) or D-H₂-A₂ motif treated by MMPT (V_2) are mixed.

One way of mixing is to use an EVB-like formula

$$E = \frac{V_1 + V_2}{2} - \sqrt{\left(\frac{V_1 - V_2}{2}\right)^2 + \epsilon^2} \quad (6.3)$$

where ϵ is set to be a constant such as 1 kcal/mol. Another possibility is to use the already computed scores S_1 and S_2 :

$$E = V_1 f(S_1, S_2) + V_2 [1 - f(S_1, S_2)] \quad (6.4)$$

with the switching function

$$f(S_1, S_2) = \frac{1}{2} \left\{ \tanh\left(\frac{S_1 - S_2}{2}\right) + 1 \right\} \quad (6.5)$$

Forces are then computed by taking the derivatives of 6.3 and 6.4. The general idea is that when V_1 and V_2 differ a lot (V_1 much smaller than V_2), $E \sim V_1$ and the changing of D-H₂-A₂ motifs, *i.e.*, the discontinuity brought by V_2 , is neglectable. When V_2 is comparable to V_1 , then the potential energies are mixed and the D-H₂-A₂ motif will gradually become the MMPT motif, and eventually when V_2 is much smaller than V_1 , the changing of the identity of the MMPT motif (MMPT transition) finishes. Both scoring functions 6.1 and 6.2, and both mixing schemes 6.3 and 6.4, have been implemented and tested. None of them lead to energy conservation, however. Energy jumps of 5 to 20 kcal/mol can be observed during the transition of MMPT motifs. This is probably caused by the fact that the MMPT transition happens within only a few steps so the total energy remains discontinuous although a mixing scheme has been applied.

During implementation of the algorithms to mix V_1 and V_2 , it was noticed that most force field terms in V_1 and V_2 are the same. The only difference is the MMPT PES $V(R, \rho, \theta)$ and a few bonded and non-bonded terms within the D-H₁-A₁ and D-H₂-A₂ motifs. Since force fields are additive, all the energetic terms related to the D-H₁-A₁ and D-H₂-A₂ motifs can be isolated. This leads to the idea to directly use energy, or more precisely, parts of the potential energy as the scoring function

$$E^{(i)} = E(pt) + E(w - pt) + E(w - w) \quad (6.6)$$

where $E(pt)$ refers to the interactions within the particular D-H_i-A_i motif as described by the MMPT PES, $E(w - pt)$ refers to all interactions between the D-H_i-A_i motif and other possible motifs, and $E(w - w)$ is interactions within all the other possible motifs. The MMPT PES is always applied to the motif with the lowest $E^{(i)}$. Thus MMPT transitions only take place when $E^{(1)}(t) < E^{(2)}(t)$ and $E^{(1)}(t + \Delta t) > E^{(2)}(t + \Delta t)$. $E^{(1)}$ is used as the potential at t and $E^{(2)}$ is used at $t + \Delta t$, and if Δt is small enough the resulting potential energy should be continuous.

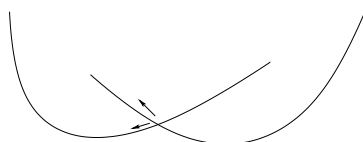


Figure 6.4: Illustration of the crossing of two potential energy surfaces. The potential energies at the point of intersection are the same for each surface, but their derivatives are different.

The whole process is summarized in the flowchart 6.5. UPTPSF is a subroutine that updates the CHARMM psf file since when the donor changes, the D-H bond should break and the corresponding information must be re-written into CHARMM. MMPTINIT and EMMPT are subroutines that initialize MMPT and that compute the modification of potential energies and forces by MMPT, respectively, as described in section 2.2.2.

The algorithm has been implemented and tested with the two test cases in section 6.2. For protonated water cluster $H_9O_4^+$, proton transport events have been observed with picosecond simulations. Four snapshots from a typical MD trajectory are presented in Figure 6.6. The conformational changes are shown, but it is not indicated in the Figure 6.6 whether proton transfer has taken place. since all protons are plotted in the same way as Van der Waals spheres with the VDW drawing method in VMD.³¹⁰ The snapshots have been replotted in Figure 6.7 with the CPK drawing method, in which the initial bonds between protons and oxygen atoms are always connected. It is shown that protons transfer from one oxygen atom to another. At 30 ps protons in the system are completely rearranged. This illustrates that the algorithm we implemented allows proton transport to happen periodically during a MD simulation. The total energy is in general stable during proton transfer and transport processes as plotted in Figure 6.8, but it increases rapidly after ~ 37 ps. This is mainly due to the quick increase of kinetic energy (see bottom panel in Figure 6.8). The reason is

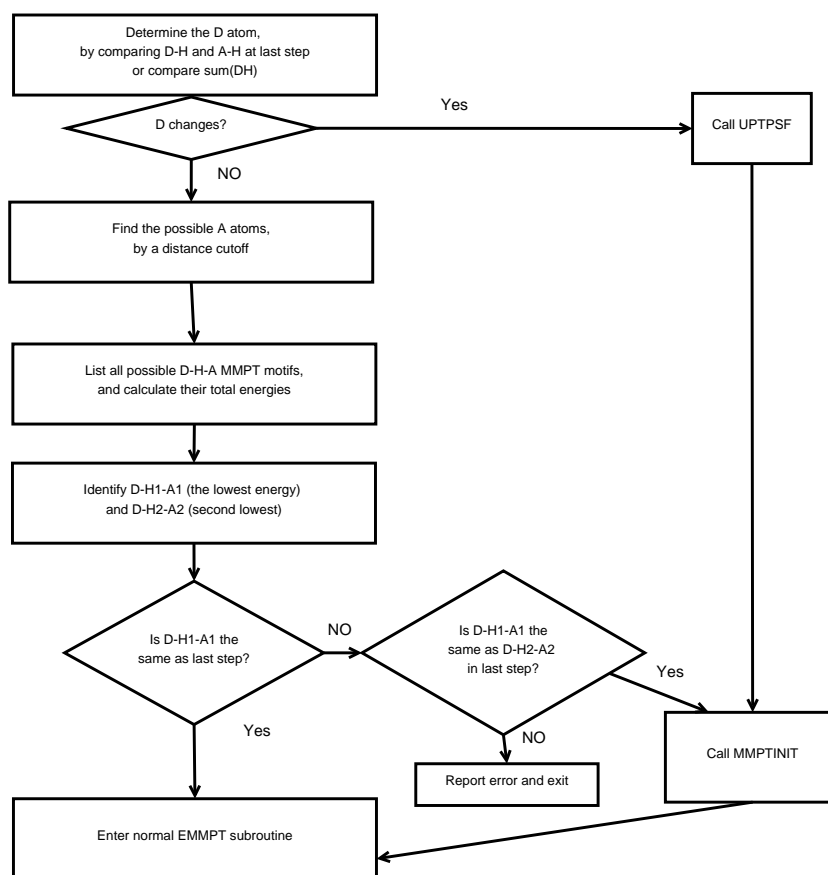


Figure 6.5: Code structure of MMPT with proton transport.

that when we switch MMPT motifs though the potential is continuous, the gradient of the potential is not (see Figure 6.4). This means that the velocities and forces of certain atoms are not compatible at the time of MMPT transition, and it is possible that velocities (kinetic energies) are not able to relax, which makes the total energy unstable and affects dynamics. This doesn't happen in every simulation with MMPT. For example, the total energy of another MD trajectory is stable during the 100 ps simulation time as illustrated in Figure 6.9. Energy conservation becomes very good after ~ 60 ps, but this is because the Zundel complex is far away from other water molecules, *i.e.*, the system falls apart.

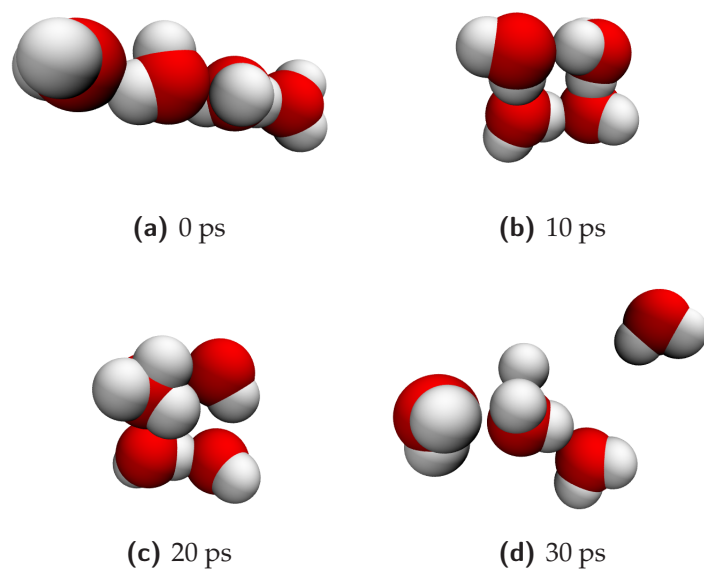


Figure 6.6: Snapshots from a MD trajectory of H_9O_4^+ where proton transport takes place.

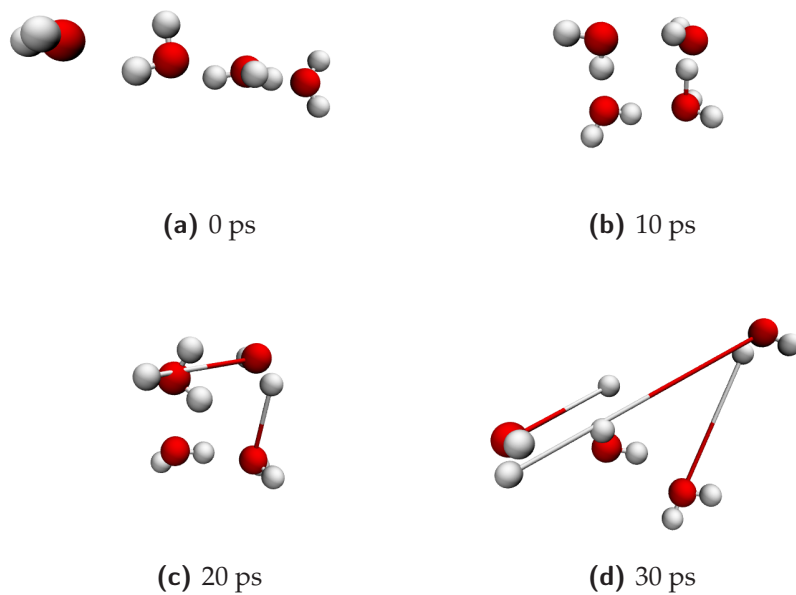


Figure 6.7: Snapshots from a MD trajectory of H_9O_4^+ where proton transport takes place. The initial bonds between protons and oxygen atoms are shown.

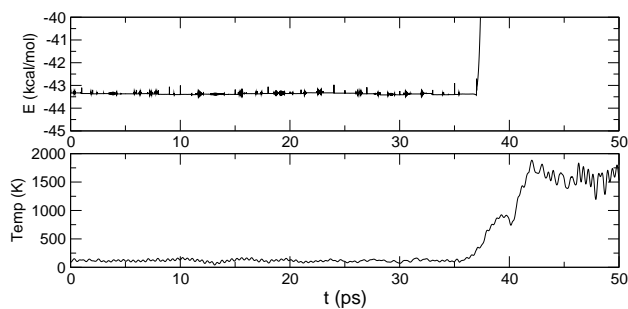


Figure 6.8: Total energy, and the kinetic energy (temperature) along a MD trajectory with MMPT that allows proton transport.

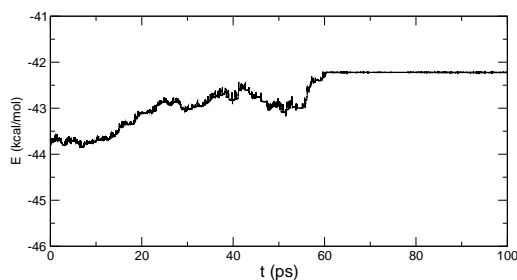


Figure 6.9: Total energy along a MD trajectory with MMPT that allows proton transport.

For the proton wire confined in a carbon nanotube, five independent 1 ns MD simulations have been carried out at 300K but no proton transport events have been detected. Longer simulation time or higher temperature might be necessary. Besides, in the current setup the proton is placed initially at one end of the nanotube (see Figure 6.2). To enhance the probability of observing proton transport we should start simulations with the conformation of proton in the middle of the carbon nanotube.

6.3.2 Smoothing the Dynamics

In order to improve energy conservation, we propose here to mix the two PESs $E^{(1)}$ and $E^{(2)}$ in time domain, as has been done in the adiabatic reactive molecular dynamics (ARMD) method.^{311–313}

ARMD is a conceptually simple surface crossing technique developed in our group, and has been successfully applied to study reactive processes such as NO rebinding to myoglobin³¹² and NO dioxygenation in truncated hemoglobin.³¹⁴ Briefly, the reactant and product are treated in ARMD with individually parametrized force fields (potential surfaces) which differ in a small number of terms. An energy criterion is used to decide whether a crossing should occur. Once a crossing is detected, the mixing of surfaces will be introduced *a posteriori* to create the adiabatic surface. This is achieved by reconstructing the system at the time point $t_{\text{cross}} - t_{\text{mix}}/2$ and gradually mixing the surface V_R and V_P by a time dependent switching function

$$f(t) = \frac{\tanh[a(t - t_{\text{cross}})] + 1}{2} \quad (6.7)$$

where t_{cross} is the time at which the crossing occurs, t_{mix} is the total mixing time which usually equals 40 time steps, and a is a constant which is usually set to 0.06. The mixed surface is given by

$$V_{\text{mix}} = V_R f(t) + V_P (1 - f(t)) \quad (6.8)$$

After the mixing of two states finishes at time $t_{\text{cross}} + t_{\text{mix}}/2$, the simulation continues on the product surface.

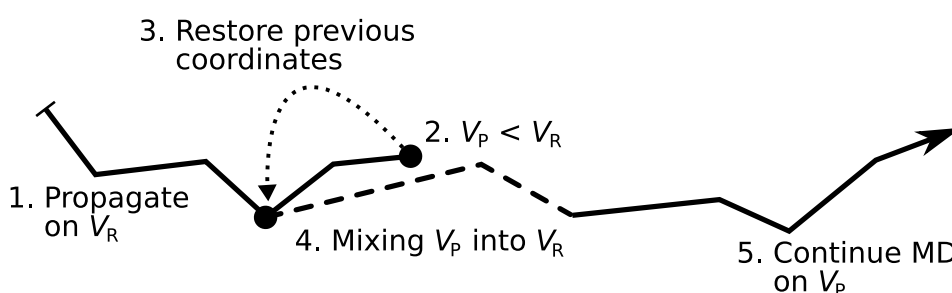


Figure 6.10: Diagram showing how ARMD works during a MD simulation involving a transition between two potential surfaces V_R and V_P along a fictitious trajectory. The dashed line describes the time t_{mix} during which V_R is slowly transformed into V_P (courtesy of S. Lutz).

Based on ARMD, an algorithm is proposed as depicted in Figure 6.11. If a MMPT transition is detected at t_{cross} by finding that the identity of the lowest energy D-H-A motif changes, a logical flag Q_{mix} is set, and the coordinates and velocities at the $t_{\text{cross}} - t_{\text{mix}}/2$ time point are restored, and the mixing of two potential starts, which is done in the subroutine EMMPTMIX as described in Eq. 6.7 and 6.8. The variable N_{mixstep} represents the current mixing steps. If it is the first step in mixing, subroutine MMPTMIX that initializes MMPT should be called, and if it is the last step then some logical flags, such as Q_{da} that defines whether the donor atoms need be changed and Q_{mix} , are set. The rewinding and mixing process finishes at $t_{\text{cross}} + t_{\text{mix}}/2$, and the standard subroutine EMMPTNOR as described in section 2.2.2 is then used to compute the MMPT energies and forces. No MMPT transition will be detected until time ' $t_{\text{cross}} + t_{\text{mix}}$ ' (as controlled by Q_{ext}) to avoid complexity that two mixing processes might overlap.

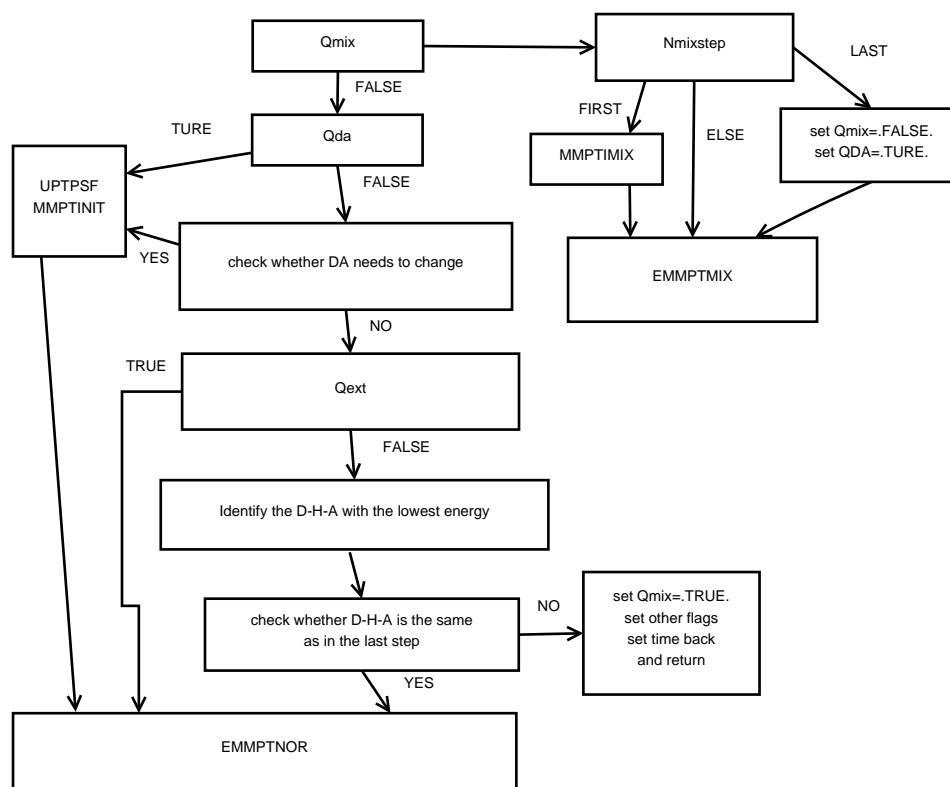


Figure 6.11: Proposed code structure of MMPT with proton transport. See text for details.

The implementation of such an algorithm into MMPT in CHARMM is still under-way. Based on the success of the ARMD method, it is hoped that improved dynamics

will result.

7 Conclusion and Outlook

"The way ahead is long and endless, yet high and low I'll explore with my will unbending."

Qu Yuan, *The Lament*

The methodology, implementation and applications of MMPT are presented in this thesis. MMPT uses parametrized three-dimensional potential energy surfaces fitted to high level *ab initio* calculations (MP2/6-311++G(d,p)) to describe the interactions within a general DH–A motif where D is the donor, H is the hydrogen and A is the acceptor atom. Together with a standard force field - here, the CHARMM force field is used - specific rules control how bonded interactions on the donor and acceptor side are switched on and off depending on the position of the transferring H-atom (DH–A or D–HA). The MMPT method provides an efficient and accurate way to study proton transfer reactions and hydrogen bonding interactions in complex chemical and biological systems.

Three applications of MMPT have been presented. The first one is to study proton dynamics in AcAc. Here MMPT parameters for AcAc are generated from those of MA by PES morphing, and lead to good agreement between computational results and experimental data including equilibrium structures and infrared spectra. In general, MMPT PESs can be “morphed” to adapt their overall shapes to topologically similar, but energetically different hydrogen bonding patterns, depending on the chemical environment. This is a concise and accurate way to obtain MMPT potentials.

In the second application,^{315,316} MMPT is applied as an explicit hydrogen bond potential in proteins and leads to better correlations between computationally predicted and experimentally measured NMR scalar couplings. We also develop a general procedure to refine and optimize force fields with NMR observables, which is not limited to MMPT PESs but can be applicable to any force field parameters. The last application is a novel hydrogen-bonded self-assembly bidentate catalyst,^{73,251} and with MMPT we are able to reproduce and explain a variety of experimental spectra.

We also demonstrate the possibility to combine MMPT with other sophisticated force field methods with this example.

The extension of MMPT to study proton transport processes is discussed. With the implemented algorithm, proton transport can take place during MD simulations, and the total energy is stable at most times. However, the dynamics becomes incorrect due to the switching of PESs when changing the identity of MMPT motifs. To solve this problem, the mixing schemes used in the ARMD method can be incorporated into MMPT. Corresponding algorithms have been proposed and are now being implemented.

Another future development of MMPT will be to incorporate a fluctuating charge model, which enables atomic partial charges of the D–H···A motif to freely adapt to the H-bond geometries during MD simulations. The importance of charge fluctuation has been shown in a recent study of the water-assisted proton transfer in Ferredoxin I.³¹⁷ It has also been found that revision of partial charges for hydrogen bonding atoms leads to differences in computed $^{\text{h3}}J_{\text{NC}'}$ couplings in proteins.²⁴⁴ An updated module with fluctuating charges will thus greatly enhance the accuracy and utility of MMPT.

Bibliography

1. C. J. T. de Grotthuss. Sur la décomposition de l'eau et des corps qu'elle tient en dissolution á l'aide de l'électricité galvanique. *Ann. Chim. (Paris)*, LVIII:54–74, 1806.
2. F. Hund. Die Gestalt mehratomiger polarer Molekeln. II. Molekeln, die aus einem negativen Ion und aus Wasserstoffkernen bestehen. *Z. Physik A Hadrons and Nuclei*, 32:1–18, 1925.
3. M. Eigen. Proton transfer, acid-base catalysis, and enzymatic hydrolysis. part i: Elementary processes. *Angew. Chem. Int. Ed.*, 3:1–19, 1964.
4. G Zundel and H Metzger. Energy bands of excess tunneling protons in fluid acids. ir spectroscopy of h_5o_2^+ groups. *Z. Phys. Chem. (N. F. Frankfurt)*, 58:225–245, 1968.
5. Maurice L. Huggins. The role of hydrogen bonds in conduction by hydrogen and hydroxyl ions. *J. Am. Chem. Soc.*, 53:3190–3191, 1931.
6. Maurice L. Huggins. Hydrogen bridges in ice and liquid water. *J. Phys. Chem.*, 40:723–731, 1936.
7. Noam Agmon. The grotthuss mechanism. *Chem. Phys. Lett.*, 244:456 – 462, 1995.
8. Omer Markovitch, Hanning Chen, Sergei Izvekov, Francesco Paesani, Gregory A. Voth, and Noam Agmon. Special pair dance and partner selection: Elementary steps in proton transport in liquid water. *J. Phys. Chem. B*, 112:9456–9466, 2008.
9. Dominik Marx. Proton transfer 200 years after von grotthuss: Insights from ab initio simulations. *ChemPhysChem*, 7:1848–1870, 2006.
10. M E Tuckerman, K Laasonen, M Sprik, and M Parrinello. Ab initio simulations of water and water ions. *Journal of Physics: Condensed Matter*, 6:A93, 1994.

11. Mark Tuckerman, Kari Laasonen, Michiel Sprik, and Michele Parrinello. Ab initio molecular dynamics simulation of the solvation and transport of h_3o^+ and oh^- ions in water. *J. Phys. Chem.*, 99:5749–5752, 1995.
12. M. Tuckerman, K. Laasonen, M. Sprik, and M. Parrinello. Ab initio molecular dynamics simulation of the solvation and transport of hydronium and hydroxyl ions in water. *J. Chem. Phys.*, 103:150–161, 1995.
13. Mark E. Tuckerman, Dominik Marx, Michael L. Klein, and Michele Parrinello. On the quantum nature of the shared proton in hydrogen bonds. *Science*, 275:817–820, 1997.
14. Tyler J. F. Day, Udo W. Schmitt, and Gregory A. Voth. The mechanism of hydrated proton transport in water. *J. Am. Chem. Soc.*, 122:12027–12028, 2000.
15. Jeongho Kim, Udo W. Schmitt, Julie A. Gruetzmacher, Gregory A. Voth, and Norbert E. Scherer. The vibrational spectrum of the hydrated proton: Comparison of experiment, simulation, and normal mode analysis. *J. Chem. Phys.*, 116:737–746, 2002.
16. U. W. Schmitt and G. A. Voth. The isotope substitution effect on the hydrated proton. *Chem. Phys. Lett.*, 329:36–41, 2000.
17. Jessica M. J. Swanson, C. Mark Maupin, Hanning Chen, Matt K. Petersen, Jiancong Xu, Yujie Wu, and Gregory A. Voth. Proton solvation and transport in aqueous and biomolecular systems: Insights from computer simulations. *J. Phys. Chem. B*, 111:4300–4314, 2007.
18. Tyler J. F. Day, Alexander V. Soudackov, Martin Cuma, Udo W. Schmitt, and Gregory A. Voth. A second generation multistate empirical valence bond model for proton transport in aqueous systems. *J. Chem. Phys.*, 117:5839–5849, 2002.
19. Yujie Wu, Hanning Chen, Feng Wang, Francesco Paesani, and Gregory A. Voth. An improved multistate empirical valence bond model for aqueous proton solvation and transport. *J. Phys. Chem. B*, 112:467–482, 2008.
20. M E Tuckerman, D Marx, ML Klein, and M Parrinello. On the quantum nature of the shared proton in hydrogen bonds. *Science*, 275:817–820, 1997.

21. D Marx, M E Tuckerman, J Hutter, and M Parrinello. The nature of the hydrated excess proton in water. *Nature*, 397:601–604, 1999.
22. JT Hynes. The protean proton in water. *Nature*, 397:565, 1999.
23. Timothy C. Berkelbach, Hee-Seung Lee, and Mark E. Tuckerman. Concerted hydrogen-bond dynamics in the transport mechanism of the hydrated proton: A first-principles molecular dynamics study. *Phys. Rev. Lett.*, 103:238302, 2009.
24. M. Meuwly and M. Karlpus. Simulation of proton transfer along ammonia wires: An ab initio and semiempirical density functional comparison of potentials and classical molecular dynamics. *J. Chem. Phys.*, 116:2572–2585, 2002.
25. V. Zoete and M. Meuwly. On the influence of semirigid environments on proton transfer along molecular chains. *J. Chem. Phys.*, 120:7085–7094, 2004.
26. Stéphane Coussan, Markus Meuwly, and Samuel Leutwyler. Proton transfer and tautomerization in 7-hydroxyquinoline · (nh₃)_n clusters: Structure and energetics at the self-consistent field level. *J. Chem. Phys.*, 114:3524–3534, 2001.
27. Markus Meuwly, Andreas Bach, and Samuel Leutwyler. Grotthus-type and diffusive proton transfer in 7-hydroxyquinoline·(nh₃)_n clusters. *J. Am. Chem. Soc.*, 123:11446–11453, 2001.
28. A. Fouqueau and M. Meuwly. Structures and dynamics of protonated ammonia clusters. *J. Chem. Phys.*, 123:244308, 2005.
29. Kenneth A. Mauritz and Robert B. Moore. State of understanding of nafion. *Chem. Rev.*, 104:4535–4586, 2004.
30. Klaus-Dieter Kreuer, Stephen J. Paddison, Eckhard Spohr, and Michael Schuster. Transport in proton conductors for fuel-cell applications: Simulations, elementary reactions, and phenomenology. *Chem. Rev.*, 104:4637–4678, 2004.
31. Colin A. Wraight. Chance and design—proton transfer in water, channels and bioenergetic proteins. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, 1757:886–912, 2006.
32. Ulrich Haupts, Joerg Tittor, and Dieter Oesterhelt. Closing in on bacteriorhodopsin: Progress in understanding the molecule. *Ann. Rev. Biophys. Biomole. Struct.*, 28:367–399.

33. Janos K. Lanyi. Bacteriorhodopsin. *Ann. Rev. Physiol.*, 66:665–688, 2004.
34. Marten Wikstrom. Proton pump coupled to cytochrome c oxidase in mitochondria. *Nature*, 266:271 – 273, 1997.
35. Marten Wikstrom. Cytochrome c oxidase: 25 years of the elusive proton pump. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, 1655:241 – 247, 2004.
36. Ari Helenius. Unpacking the incoming influenza virus. *Cell*, 69:577 – 578, 1992.
37. Jun Hu, Riqiang Fu, Katsuyuki Nishimura, Li Zhang, Huan-Xiang Zhou, David D. Busath, Viksita Vijayvergiya, and Timothy A. Cross. Histidines, heart of the hydrogen ion channel from influenza a virus: Toward an understanding of conductance and proton selectivity. *Proc. Natl. Acad. Sci.*, 103:6865–6870, 2006.
38. Sharon Hammes-Schiffer. Hydrogen tunneling and protein motion in enzyme reactions. *Acc. Chem. Res.*, 39:93–100, 2006.
39. Zachary D. Nagel and Judith P. Klinman. Tunneling and dynamics in enzymatic hydride transfer. *Chem. Rev.*, 106:3095–3118, 2006.
40. Salomon R. Billeter, Simon P. Webb, Pratul K. Agarwal, Tzvetelin Iordanov, and Sharon Hammes-Schiffer. Hydride transfer in liver alcohol dehydrogenase: Quantum dynamics, kinetic isotope effects, and role of enzyme motion. *J. Am. Chem. Soc.*, 123:11262–11272, 2001.
41. Qiang Cui, Marcus Elstner, and Martin Karplus. A theoretical analysis of the proton and hydride transfer in liver alcohol dehydrogenase (ladh). *J. Phys. Chem. B*, 106:2721–2740, 2002.
42. Laura Masgrau, Anna Roujeinikova, Linus O. Johannissen, Parvinder Hothi, Jaswir Basran, Kara E. Ranaghan, Adrian J. Mulholland, Michael J. Sutcliffe, Nigel S. Scrutton, and David Leys. Atomic description of an enzyme reaction dominated by proton tunneling. *Science*, 312:237–241, 2006.
43. Jigar N. Bandaria, Christopher M. Cheatum, and Amnon Kohen. Examination of enzymatic h-tunneling through kinetics and dynamics. *J. Am. Chem. Soc.*, 131:10151–10155, 2009.

44. Pratul K. Agarwal, Salomon R. Billeter, P. T. Ravi Rajagopalan, Stephen J. Benkovic, and Sharon Hammes-Schiffer. Network of coupled promoting motions in enzyme catalysis. *Proc. Natl. Acad. Sci.*, 99:2794–2799, 2002.
45. Mireia Garcia-Viloca, Jiali Gao, Martin Karplus, and Donald G. Truhlar. How enzymes work: Analysis by modern rate theory and computer simulations. *Science*, 303:186–195, 2004.
46. Sharon Hammes-Schiffer and James B Watney. Hydride transfer catalysed by *Escherichia coli* and *Bacillus subtilis* dihydrofolate reductase: coupled motions and distal mutations. *Phil. Trans. R. Soc. B*, 361:1365–1373, 2006.
47. David R. Borst, Joseph R. Rosciolia, David W. Pratt, Gina M. Florio, Timothy S. Zwier, Andreas Mueller, and Samuel Leutwyler. Hydrogen bonding and tunneling in the 2-pyridone-2-hydroxypyridine dimer. effect of electronic excitation. *Chem. Phys.*, 112:3717–3726, 2002.
48. S. L. Baughcum, R. W. Duerst, W. F. Rowe, Z. Smith, and E. B. Wilson. Microwave spectroscopic study of malonaldehyde (3-hydroxy-2-propenal) .2. structure, dipole-moment, and tunneling. *J. Am. Chem. Soc.*, 103:6296–6303, 1981.
49. D. W. Firth, K. Beyer, M. A. Dvorak, S. W. Reeve, A. Grushow, and K. R. Leopold. Tunable far infrared spectroscopy of malonaldehyde. *J. Chem. Phys.*, 94:1812–1819, 1991.
50. Keiichi Tanaka, Masaaki Toshimitsu, Kensuke Harada, and Takehiko Tanaka. Determination of the proton tunneling splitting of the vinyl radical in the ground state by millimeter-wave spectroscopy combined with supersonic jet expansion and ultraviolet photolysis. *J. Chem. Phys.*, 120:3604–3618, 2004.
51. Adam M. Daly, P. R. Bunker, and Stephen G. Kukolich. Evidence for proton tunneling from the microwave spectrum of the formic acid-propionic acid dimer. *J. Chem. Phys.*, 132:201101, 2010.
52. Edward Sheldon. Lewis and Lance. Funderburk. Rates and isotope effects in the proton transfers from 2-nitropropane to pyridine bases. *J. Am. Chem. Soc.*, 89:2322–2327, 1967.

53. William von E. Doering and Xin Zhao. Effect on kinetics by deuterium in the 1,5-hydrogen shift of a cisoid-locked 1,3(z)-pentadiene, 2-methyl-10-methylenebicyclo[4.4.0]dec-1-ene: Evidence for tunneling? *J. Am. Chem. Soc.*, 128:9080–9085, 2006.
54. Thomas C. Wilde, Grzegorz Blotny, and Ralph M. Pollack. Experimental evidence for enzyme-enhanced coupled motion/quantum mechanical hydrogen tunneling by ketosteroid isomerase. *J. Am. Chem. Soc.*, 130:6577–6585, 2008.
55. Detlef Gerritzen and Hans Heinrich Limbach. Kinetic isotope effects and tunneling in cyclic double and triple proton transfer between acetic acid and methanol in tetrahydrofuran studied by dynamic proton and deuterium nmr spectroscopy. *J. Am. Chem. Soc.*, 106:869–879, 1984.
56. Uwe Langer, Christof Hoelger, Bernd Wehrle, Lidia Latanowicz, Emanuel Vogel, and Hans-Heinrich Limbach. ^{15}N nmr study of proton localization and proton transfer thermodynamics and kinetics in polycrystalline porphycene. *J. Phys. Org. Chem.*, 13:23–34, 2000.
57. Shasad Sharif, Gleb S. Denisov, Michael D. Toney, and Hans-Heinrich Limbach. Nmr studies of solvent-assisted proton transfer in a biologically relevant schiff base: toward a distinction of geometric and equilibrium h-bond isotope effects. *J. Am. Chem. Soc.*, 128:3375–3387, 2006.
58. Juan Miguel Lopez, Ferdinand Mandle, Iwona Wawer, Gerd Buntkowsky, and Hans-Heinrich Limbach. Nmr studies of double proton transfer in hydrogen bonded cyclic n,n[prime or minute]-diarylformamidinium dimers: conformational control, kinetic hh/hd/dd isotope effects and tunneling. *Phys. Chem. Chem. Phys.*, 9:4498–4513, 2007.
59. K. R. Asmis, N. L. Pivonka, G. Santambrogio, M. Bruemmer, C. Kaposta, D. M. Neumark, and L. Woeste. Gas phase infrared spectrum of the protonated water dimer. *Science*, 299:1375–1377, 2003.
60. Jeffrey M. Headrick, Eric G. Diken, Richard S. Walters, Nathan I. Hammer, Richard A. Christie, Jun Cui, Evgeniy M. M yshakin, Michael A. Duncan, Mark A. Johnson, and Kenneth D. Jordan. Spectral signatures of hydrated proton vibrations in water clusters. *Science*, 308:1765–1769, 2005.

61. William H. Robertson, Eric G. Diken, Erica A. Price, Joong-Won Shin, and Mark A. Johnson. Spectroscopic determination of the oh^- solvation shell in the $\text{oh}^- \cdot (\text{h}_2\text{o})_n$ clusters. *Science*, 299:1367–1372, 2003.
62. Azusa Muraoka, Yoshiya Inokuchi, Nathan I. Hammer, Joong-Won Shin, Mark A. Johnson, and Takashi Nagata. Structural evolution of the $[(\text{co}_2)_n(\text{h}_2\text{o})]^-$ cluster anions: Quantifying the effect of hydration on the excess charge accommodation motif. *J. Phys. Chem. A*, 113:8942–8948, 2009.
63. Rachael A. Relph, Ben M. Elliott, Gary H. Weddle, Mark A. Johnson, Jing Ding, and Kenneth D. Jordan. Vibrationally induced interconversion of h-bonded $\text{no}_2^- \cdot \text{h}_2\text{o}$ isomers within $\text{no}_2^- \cdot \text{h}_2\text{o} \cdot \text{ar}_m$ clusters using ir-ir pump-probe through the oh and no stretching vibrations. *J. Phys. Chem. A*, 113:975–981, 2009.
64. G. E. Douberly, R. S. Walters, J. Cui, K. D. Jordan, and M. A. Duncan. Infrared spectroscopy of small protonated water clusters, $\text{h}^+(\text{h}_2\text{o})_n$ ($n = 2-5$): Isomers, argon tagging, and deuteration. *J. Phys. Chem. A*, 114:4570–4579, 2010.
65. Sander Woutersen and Huib J. Bakker. Ultrafast vibrational and structural dynamics of the proton in liquid water. *Phys. Rev. Lett.*, 96:138305, 2006.
66. Wafa Amir, Guilhem Gallot, François Hache, S. Bratos, J.-C. Leicknam, and R. Vuilleumier. Time-resolved observation of the eigen cation in liquid water. *J. Chem. Phys.*, 126:034511, 2007.
67. Bernd Winter, Manfred Faubel, Ingolf V. Hertel, Christian Pettenkofer, Stephen E. Bradforth, Barbara Jagoda-Cwiklik, Lukasz Cwiklik, and Pavel Jungwirth. Electron binding energies of hydrated h_3o^+ and oh^- : Photoelectron spectroscopy of aqueous acid and base solutions combined with electronic structure calculations. *J. Am. Chem. Soc.*, 128:3864–3865, 2006.
68. Barbara Kirchner. Eigen or zundel ion: News from calculated and experimental photoelectron spectroscopy. *ChemPhysChem*, 8:41–43, 2007.
69. Oriol Vendrell, Fabien Gatti, and Hans-Dieter Meyer. Dynamics and infrared spectroscopy of the protonated water dimer. *Angew. Chem. Int. Ed.*, 46:6918–6921, 2007.
70. J. R. Roscioli, L. R. McCunn, and M. A. Johnson. Quantum structure of the intermolecular proton bond. *Science*, 316:249–254, 2007.

71. M. Meuwly, A. Müller, and S. Leutwyler. Energetics, dynamics and infrared spectra of the dna base-pair analogue 2-pyridone · 2-hydroxypyridine. *PCCP*, 5:2663–2672, 2003.
72. S. Lammers and M. Meuwly. Investigating the relationship between infrared spectra of shared protons in different chemical environments: A comparison of protonated Diglyme and protonated water dimer. *J. Phys. Chem. A*, 111:1638–1647, 2007.
73. Jing Huang, Urs Gellrich, Wolfgang Seiche, Bernhard Breit, and Markus Meuwly. Energetics, dynamics and spectra of a hydrogen-bonded bidentate catalyst: a computational approach. *in preparation*.
74. P. Ehrenfest. Bemerkung ueber die angenaeherte gueltigkeit der klassischen mechanik innerhalb der quantenmechanik. *Z. Phys.*, 45:455–457, 1927.
75. M. Born and R. Oppenheimer. Zur quantentheorie der molekeln. *Ann. Phys. (Leipzig)*, 84:457, 1927.
76. N. Galamba and B. J. Costa Cabral. Born-oppenheimer molecular dynamics of the hydration of na^+ in a water cluster. *J. Phys. Chem. B*, 113:16151–16158, 2009.
77. Martina Kaledin and Christopher A. Wood. Ab initio studies of structural and vibrational properties of protonated water cluster h_7o_3^+ and its deuterium isotopologues: An application of driven molecular dynamics. *J. Chem. Theo. Comp.*, 6:2525–2535, 2010.
78. Martina Kaledin, John M. Moffitt, Craig R. Clark, and Fareeha Rizvi. Ab initio molecular dynamics simulations of the infrared spectra of h_3o_2^- and d_3o_2^- . *J. Chem. Theo. Comp.*, 5:1328–1336, 2009.
79. M. Elstner, D. Porezag, G. Jungnickel, J. Elsner, M. Haugk, T. Frauenheim, S. Suhai, and G. Seifert. Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties. *Phys. Rev. B*, 58:7260–7268, 1998.
80. H. Y. Zhou, E. Tajkhorshid, T. Frauenheim, S. Suhai, and M. Elstner. Performance of the am1, pm3, and scc-dftb methods in the study of conjugated schiff base molecules. *Chem. Phys.*, 277:91–103, 2002.

81. M. Elstner. The scc-dftb method and its application to biological systems. *Theor. Chim. Acta*, 116:316–325, 2006.
82. Puja Goyal, Marcus Elstner, and Qiang Cui. Application of the scc-dftb method to neutral and protonated water clusters and bulk water. *J. Phys. Chem. B*, 115:6790–6805, 2011.
83. R. Car and M. Parrinello. Unified approach for molecular dynamics and density-functional theory. *Phys. Rev. Lett.*, 55:2471–2474, 1985.
84. Holger Langer, Nikos L. Doltsinis, and Dominik Marx. Excited-state dynamics and coupled proton-electron transfer of guanine. *ChemPhysChem*, 6:1734–1737, 2005.
85. Marie-Pierre Gageot, Alvaro Cimas, Mahamadou Seydou, Ju-Young Kim, Sungyul Lee, and Jean-Pierre Schermann. Proton transfer from the inactive gas-phase nicotine structure to the bioactive aqueous-phase structure. *J. Am. Chem. Soc.*, 132:18067–18077, 2010.
86. R. P. Feynman. Space-time approach to non-relativistic quantum mechanics. *Rev. Mod. Phys.*, 20:367–387, 1948.
87. Dominik Marx and Michele Parrinello. Ab initio path integral molecular dynamics: Basic ideas. *J. Chem. Phys.*, 104:4077–4082, 1996.
88. Mark E. Tuckerman, Dominik Marx, Michael L. Klein, and Michele Parrinello. Efficient and general algorithms for path integral car-parrinello molecular dynamics. *J. Chem. Phys.*, 104:5579–5588, 1996.
89. Robin L. Hayes, Stephen J. Paddison, and Mark E. Tuckerman. Proton transport in triflic acid hydrates studied via path integral car-parrinello molecular dynamics. *J. Phys. Chem. B*, 113:16574–16589, 2009.
90. Robin L. Hayes, Stephen J. Paddison, and Mark E. Tuckerman. Proton transport in triflic acid pentahydrate studied via ab initio path integral molecular dynamics. *J. Phys. Chem. A*, in press, 2011.
91. Hee-Seung Lee and Mark E. Tuckerman. Structure of liquid water at ambient temperature from ab initio molecular dynamics performed in the complete basis set limit. *J. Chem. Phys.*, 125:154507, 2006.

92. Hee-Seung Lee and Mark E. Tuckerman. Dynamical properties of liquid water from ab initio molecular dynamics performed in the complete basis set limit. *J. Chem. Phys.*, 126:164501, 2007.
93. K. Laasonen, M. Sprik, M. Parrinello, and R. Car. “ab initio” liquid water. *J. Chem. Phys.*, 99:9080–9089, 1993.
94. Michiel Sprik, Jürg Hutter, and Michele Parrinello. Ab initio molecular dynamics simulation of liquid water: Comparison of three gradient-corrected density functionals. *J. Chem. Phys.*, 105:1142–1152, 1996.
95. Sergei Izvekov and Gregory A. Voth. Ab initio molecular-dynamics simulation of aqueous proton solvation and transport revisited. *J. Chem. Phys.*, 123:044505, 2005.
96. Joost VandeVondele, Fawzi Mohamed, Matthias Krack, Jürg Hutter, Michiel Sprik, and Michele Parrinello. The influence of temperature and density functional models in ab initio molecular dynamics simulation of liquid water. *J. Chem. Phys.*, 122, 2005.
97. Kurt W. Sattelmeyer, Julian Tirado-Rives, and William L. Jorgensen. Comparison of scc-dftb and nndo-based semiempirical molecular orbital methods for organic molecules. *J. Phys. Chem. A*, 110:13551–13559, 2006.
98. A. Warshel and M. Levitt. Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.*, 103:227–249, 1976.
99. Regis Pomes and Benoit Roux. Free energy profiles for h⁺ conduction along hydrogen-bonded chains of water molecules. *Biophys. J.*, 75:33 – 40, 1998.
100. Regis Pomes and Benoit Roux. Molecular mechanism of h⁺ conduction in the single-file water chain of the gramicidin channel. *Biophys. J.*, 82:2304 – 2316, 2002.
101. Qiang Cui, Marcus Elstner, Efthimios Kaxiras, Thomas Frauenheim, and Martin Karplus. A qm/mm implementation of the self-consistent charge density functional tight binding (scc-dftb) method. *J. Phys. Chem. B*, 105:569–585, 2001.

102. Tom K. Woo, Peter M. Margl, Peter E. Bloechl, and Tom Ziegler. A combined car-parrinello qm/mm implementation for ab initio molecular dynamics simulations of extended systems: Application to transition metal catalysis. *J. Phys. Chem. B*, 101:7877–7880, 1997.
103. Stefano Piana, Daniel Sebastiani, Paolo Carloni, and Michele Parrinello. Ab initio molecular dynamics-based assignment of the protonation state of pepstatin a/hiv-1 protease cleavage site. *J. Am. Chem. Soc.*, 123:8730–8737, 2001.
104. Roger Rousseau, Volker Kleinschmidt, Udo W. Schmitt, and Dominik Marx. Modeling protonated water networks in bacteriorhodopsin. *Phys. Chem. Chem. Phys.*, 6:1848–1859, 2004.
105. Roger Rousseau, Volker Kleinschmidt, Udo W. Schmitt, and Dominik Marx. Assigning protonation patterns in water networks in bacteriorhodopsin based on computed ir spectra. *Angew. Chem. Int. Ed.*, 43:4804–4807, 2004.
106. Gerald Mathias and Dominik Marx. Structures and spectral signatures of protonated water networks in bacteriorhodopsin. *Proc. Natl. Acad. Sci.*, 104:6980–6985, 2007.
107. Prasad Phatak, Nilanjan Ghosh, Haibo Yu, Qiang Cui, and Marcus Elstner. Amino acids with an intermolecular proton bond as proton storage site in bacteriorhodopsin. *Proc. Natl. Acad. Sci.*, 105:19672–19677, 2008.
108. M. Clemens, P. Phatak, Q. Cui, A. N. Bondar, and M. Elstner. Role of arg82 in the early steps of the bacteriorhodopsin proton-pumping cycle. *The Journal of Physical Chemistry B*, 115:7129–7135, 2011.
109. Zorka Smedarchina, Willem Siebrand, Antonio Fernandez-Ramos, and Qiang Cui. Kinetic isotope effects for concerted multiple proton transfer: A direct dynamics study of an active-site model of carbonic anhydrase ii. *J. Am. Chem. Soc.*, 125:243–251, 2003.
110. P. H. Konig, N. Ghosh, M. Hoffmann, M. Elstner, E. Tajkhorshid, T. Frauenheim, and Q. Cui. Toward theoretical analysis of long-range proton transfer kinetics in biomolecular pumps. *J. Phys. Chem. A*, 110:548–563, 2006.

111. Nilanjan Ghosh, Xavier Prat-Resina, M. R. Gunner, and Qiang Cui. Microscopic pka analysis of glu286 in cytochrome c oxidase (rhodobacter sphaeroides): Toward a calibrated molecular model. *Biochemistry*, 48:2468–2485, 2009.
112. Hai Lin and Donald G. Truhlar. Qm/mm: what have we learned, where are we, and where do we go from here? *Theor. Chim. Acta*, 117:185–199, 2007.
113. A. V. Nemukhin, B. L. Grigorenko, I. A. Topol, and S. K. Burt. Flexible effective fragment qm/mm method: Validation through the challenging tests. *J. Comp. Chem.*, 24:1410–1420, 2003.
114. Teodoro Laino, Fawzi Mohamed, Alessandro Laio, and Michele Parrinello. An efficient linear-scaling electrostatic coupling for treating periodic boundary conditions in qm/mm simulations. *J. Chem. Theo. Comp.*, 2:1370–1378, 2006.
115. Yan Zhang and Hai Lin. Flexible-boundary quantum-mechanical/molecular-mechanical calculations: Partial charge transfer between the quantum-mechanical and molecular-mechanical subsystems. *J. Chem. Theo. Comp.*, 4:414–425, 2008.
116. Erin E. Dahlke and Donald G. Truhlar. Electrostatically embedded many-body expansion for simulations. *J. Chem. Theo. Comp.*, 4:1–6, 2008.
117. Yan Zhang and Hai Lin. Flexible-boundary qm/mm calculations: Ii. partial charge transfer across the qm/mm boundary that passes through a covalent bond. *Theor. Chim. Acta*, 126:315–322, 2010.
118. Udo W. Schmitt and Gregory A. Voth. The computer simulation of proton transport in water. *J. Chem. Phys.*, 111:9361–9381, 1999.
119. Feng Wang and Gregory A. Voth. A linear-scaling self-consistent generalization of the multistate empirical valence bond method for multiple excess protons in aqueous systems. *J. Chem. Phys.*, 122:144105, 2005.
120. Rodolphe Vuilleumier and Daniel Borgis. Quantum dynamics of an excess proton in water using an extended empirical valence-bond hamiltonian. *J. Phys. Chem. B*, 102:4261–4264, 1998.
121. Rodolphe Vuilleumier and Daniel Borgis. Transport and spectroscopy of the hydrated proton: A molecular dynamics study. *J. Chem. Phys.*, 11:4251–4267, 1999.

122. A. Warshel and R. M. Weiss. An empirical valence bond approach for comparing reactions in solutions and in enzymes. *J. Am. Chem. Soc.*, 102:6218–6226, 1980.
123. S. C. L. Kargerlin, J. Cao, E. Rosta, and A. Warshel. On unjustifiably misrepresenting the evb approach while simultaneously adopting it. *J. Phys. Chem. B*, 113:10905–10915, 2009.
124. J. Lobaugh and Gregory A. Voth. The quantum dynamics of an excess proton in water. *J. Chem. Phys.*, 104:2056–2070, 1996.
125. Diane E. Sagnella and Mark E. Tuckerman. An empirical valence bond model for proton transfer in water. *J. Chem. Phys.*, 108:2073–2084, 1998.
126. Konrad Hinsen and Benoit Roux. A potential function for computer simulation studies of proton transfer in acetylacetone. *J. Comp. Chem.*, 18:368–380, 1997.
127. Yujie Wu, Harald L. Tepper, and Gregory A. Voth. Flexible simple point-charge water model with improved liquid-state properties. *J. Chem. Phys.*, 124:024503, 2006.
128. Francesco Paesani, Wei Zhang, David A. Case, Thomas E. Cheatham, and Gregory A. Voth. An accurate and simple quantum model for liquid water. *J. Chem. Phys.*, 125:184507, 2006.
129. Daniel Laria, Jordi Marti, and Elvira Guardia. Protons in supercritical water: A multistate empirical valence bond study. *J. Am. Chem. Soc.*, 126:2125–2134, 2004.
130. Martin Cuma, Udo W. Schmitt, and Gregory A. Voth. A multi-state empirical valence bond model for weak acid dissociation in aqueous solution. *J. Phys. Chem. A*, 105:2814–2823, 2001.
131. Martin Cuma, Udo W. Schmitt, and Gregory A. Voth. A multi-state empirical valence bond model for acid-base chemistry in aqueous solution. *Chem. Phys.*, 258:187–199, 2000.
132. C. Mark Maupin, Kim F. Wong, Alexander V. Soudackov, Sun Kim, and Gregory A. Voth. A multistate empirical valence bond description of protonatable amino acids. *J. Phys. Chem. A*, 110:631–639, 2006.

133. Alexander M. Smondyrev and Gregory A. Voth. Molecular dynamics simulation of proton transport near the surface of a phospholipid membrane. *Biophys. J.*, 82:1460–1468, 2002.
134. Matt K. Petersen, Srinivasan S. Iyengar, Tyler J. F. Day, and Gregory A. Voth. The hydrated proton at the water liquid/vapor interface. *J. Phys. Chem. B*, 108:14804–14806, 2004.
135. Harald L. Tepper and Gregory A. Voth. Protons may leak through pure lipid bilayers via a concerted mechanism. *Biophys. J.*, 88:3095–3108, 2005.
136. Harald L. Tepper and Gregory A. Voth. Mechanisms of passive ion permeation through lipid bilayers: Insights from simulations. *J. Phys. Chem. B*, 110:21327–21337, 2006.
137. C. Mark Maupin, Robert McKenna, David N. Silverman, and Gregory A. Voth. Elucidation of the proton transport mechanism in human carbonic anhydrase ii. *J. Am. Chem. Soc.*, 131:7598–7608, 2009.
138. Jiancong Xu and Gregory A. Voth. Computer simulation of explicit proton translocation in cytochrome c oxidase: The d-pathway. *Proc. Natl. Acad. Sci.*, 102:6795–6800, 2005.
139. Alexander M. Smondyrev and Gregory A. Voth. Molecular dynamics simulation of proton transport through the influenza a virus m2 channel. *Biophys. J.*, 83:1987–1996, 2002.
140. Yujie Wu and Gregory A. Voth. A computational study of the closed and open states of the influenza a m2 proton channel. *Biophys. J.*, 89:2402–2411, 2005.
141. Matt K. Petersen and Gregory A. Voth. Characterization of the solvation and transport of the hydrated proton in the perfluorosulfonic acid membrane nafion. *J. Phys. Chem. B*, 110:18594–18600, 2006.
142. Hanning Chen, Jianqing Xu, and Gregory A. Voth. Unusual hydrophobic interactions in acidic aqueous solutions. *J. Phys. Chem. B*, 113:7291–7297, 2009.
143. Jianqing Xu, Yong Zhang, and Gregory A. Voth. Infrared spectrum of the hydrated proton in water. *J. Phys. Chem. Lett.*, 2:81–86, 2011.

144. Markus A. Lill and Volkhard Helms. Reaction rates for proton transfer over small barriers and connection to transition state theory. *J. Chem. Phys.*, 115:7985–7992, 2001.
145. Markus A. Lill and Volkhard Helms. Molecular dynamics simulation of proton transport with quantum mechanically derived proton hopping rates (q-hop md). *J. Chem. Phys.*, 115:7993–8005, 2001.
146. Elena Herzog, Tomaso Frigato, Volkhard Helms, and C. Roy D. Lancaster. Energy barriers of proton transfer reactions between amino acid side chain analogs and water from ab initio calculations. *J. Comp. Chem.*, 27:1534–1547, 2006.
147. Markus A. Lill and Volkhard Helms. Proton shuttle in green fluorescent protein studied by dynamic simulations. *Proc. Natl. Acad. Sci.*, 99:2778–2781, 2002.
148. Wei Gu, Tomaso Frigato, Tjerk Straatsma, and Volkhard Helms. Dynamic protonation equilibrium of solvated acetic acid. *Angew. Chem. Int. Ed.*, 46:2939–2943, 2007.
149. Wei Gu and Volkhard Helms. Tightly connected water wires facilitate fast proton uptake at the proton entrance of proton pumping proteins. *J. Am. Chem. Soc.*, 131:2080–2081, 2009.
150. Ram Devanathan, Arun Venkatnathan, Roger Rousseau, Michel Dupuis, Tomaso Frigato, Wei Gu, and Volkhard Helms. Atomistic simulation of water percolation and proton hopping in nafion fuel cell membrane. *J. Phys. Chem. B*, 114:13681–13690, 2010.
151. Wei Gu, Bo Zhou, Tihamer Geyer, Michael Hutter, Haiping Fang, and Volkhard Helms. Design of a gated molecular proton channel. *Angew. Chem. Int. Ed.*, 50:768–771, 2011.
152. Ricky A. Kendall, Edoardo Apra, David E. Bernholdt, Eric J. Bylaska, Michel Dupuis, George I. Fann, Robert J. Harrison, Jialin Ju, Jeffrey A. Nichols, Jarek Nieplocha, T.P. Straatsma, Theresa L. Windus, and Adrian T. Wong. High performance computational chemistry: An overview of nwchem a distributed parallel application. *Comput. Phys. Commun.*, 128:260–283, 2000.

153. Myvizhi Esai Selvan, David J. Keffer, Shengting Cui, and Stephen J. Paddison. A reactive molecular dynamics algorithm for proton transport in aqueous systems. *J. Phys. Chem. C*, 114:11965–11976, 2010.
154. A. C. T. van Duin, S. Dasgupta, F. Lorant, and W. A. Goddard III. Reaxff: A reactive force field for hydrocarbons. *J. Phys. Chem. A*, 105:9396–9409, 2001.
155. Obaidur Rahaman, Adri C. T. van Duin, William A. Goddard, and Douglas J. Doren. Development of a reaxff reactive force field for glycine and application to solvent effect and tautomerization. *J. Phys. Chem. B*, 115:249–261, 2011.
156. S. Lifson and A. Warshel. Consistent force field for calculations of conformations vibrational spectra and enthalpies of cycloalkane and n-alkane molecules. *J. Chem. Phys.*, 49:5116–5129, 1968.
157. M. Levitt and S. Lifson. Refinement of protein conformations using a macromolecular energy minimization procedure. *J. Mol. Biol.*, 46:269–279, 1969.
158. M. J. Hwang, T. P. Stockfish, and A. T. Hagler. Derivation of class ii force fields: 2. derivation and characterization of a class ii force field, cff93, for the alkyl functional group and alkane molecules. *J. Am. Chem. Soc.*, 116:2515–2525, 1994.
159. J. R. Maple, M. J. Hwang, T. P. Stockfish, U. Dinur, M. Waldman, C. S. Ewig, and A. T. Hagler. Derivation of class-ii force-fields: 1. methodology and quantum force-field for the alkyl functional-group and alkane molecules. *J. Comp. Chem.*, 15:162–182, 1994.
160. B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, and M. Karplus. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.*, 4:18–217, 1983.
161. S. J. Weiner, P. A. Kollman, D. A. Case, U.C. Singh, C. Ghio, G. Alagona, S. Profeta Jr, and P. Weiner. A new force-field for molecular mechanical simulation of nucleic-acids and proteins. *J. Am. Chem. Soc.*, 106:765–784, 1984.
162. W. L. Jorgensen and J. Tirado-Rives. The opls potential functions for proteins - energy minimizations for crystals of cyclic-peptides and crambin. *J. Am. Chem. Soc.*, 110:1657–1666, 1988.

163. J. Hermans, H. J. C. Berendsen, W. F van Gunsteren, and J. P. M. Postma. A consistent empirical potential for water-protein interactions. *Biopol.*, 23:1513–1518, 1984.
164. Jr. A. D. MacKerell, D. Bashford, M. Bellott, Jr. R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, III W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorcikiewicz-Kuczera, D. Yin, and M. Karplus. All atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B*, 102:3586–3616, 1998.
165. B. R. Brooks, C. L. Brooks, A. D. Mackerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, and M. Karplus. Charmm: The biomolecular simulation program. *J. Comp. Chem.*, 30:1545–1614, 2009.
166. Jr. Alexander D. Mackerell, Michael Feig, and Charles L. Brooks III. Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comp. Chem.*, 25:1400–1415, 2004.
167. Matthias Buck, Sabine Bouguet-Bonnet, Richard W. Pastor, and Jr. Alexander D. MacKerell. Importance of the cmap correction to the charmm22 protein force field: Dynamics of hen lysozyme. *Biophys J.*, 90:L36–L38, 2006.
168. Thomas A. Halgren. Merck molecular force field. i. basis, form, scope, parameterization, and performance of mmff94. *J. Comp. Chem.*, 17:490–519, 1998.
169. Philip M. Morse. Diatomic molecules according to the wave mechanics. ii. vibrational levels. *Phys. Rev.*, 34:57–64, 1929.
170. W. E. Reiher III. *Theoretical Studies of Hydrogen Bonding*. PhD thesis, Harvard University, 1985.
171. Eyal Neria, Stefan Fischer, and Martin Karplus. Simulation of activation free energies in molecular systems. *J. Chem. Phys.*, 105:1902–1921, 1996.

172. W.F. van Gunsteren and H.J.C. Berendsen. Algorithms for macromolecular dynamics and constraint dynamics. *Mol. Phys.*, 34:1311–1327, 1977.
173. Jean-Paul Ryckaert, Giovanni Ciccotti, and Herman J. C. Berendsen. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comp. Phys.*, 23:327–341, 1977.
174. S. Lammers and M. Meuwly. Double proton transfer using dissociable force fields. *Aust. J. Chem.*, 57:1223–1228, 2004.
175. S. Lammers. *Simulations of proton transfer processes using reactive force fields*. PhD thesis, University of Basel, 2006.
176. Sven Lammers, Stephan Lutz, and Markus Meuwly. Reactive force fields for proton transfer dynamics. *J. Comput. Chem.*, 29:1048–1063, 2008.
177. Yonggang Yang and Markus Meuwly. A generalized reactive force field for non-linear hydrogen bonds: Hydrogen dynamics and transfer in malonaldehyde. *J. Chem. Phys.*, 133:064503, 2010.
178. S. Vijay-Kumar, C. E. Bugg, and W. J. Cook. Structure of ubiquitin refined at 1.8 a resolution. *J. Mol. Biol.*, 194:531–544, 1987.
179. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.*, 79:926–935, 1983.
180. P. P. Ewald. Die berechnung optischer und elektrostatischer gitterpotentiale. *Annalen der Physik*, 369:253–287, 1921.
181. Tom Darden, Darrin York, and Lee Pedersen. Particle mesh ewald: An n·log(n) method for ewald sums in large systems. *J. Chem. Phys.*, 98:10089–10092, 1993.
182. Tom Darden, Lalith Perera, Leping Li, and Lee Pedersen. New tricks for modelers from the crystallography toolkit: the particle mesh ewald algorithm and its use in nucleic acid simulations. *Structure*, 7:R55 – R60, 1999.
183. Joel M. Bowman and Bela Gazdy. A simple method to adjust potential energy surfaces: Application to hco. *J. Chem. Phys.*, 94:816–817, 1991.

184. Jeremy M. Hutson. Vibrational dependence of the anisotropic intermolecular potential of ar-hf. *J. Chem. Phys.*, 96:6752–6767, 1992.
185. Markus Meuwly and Jeremy M. Hutson. Morphing *ab initio* potentials: A systematic study of ne-hf. *J. Chem. Phys.*, 110:8338–8347, 1999.
186. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, Jr. J. A. Montgomery, T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople. Gaussian 03, Revision C.02. Gaussian, Inc., Wallingford, CT, 2004.
187. M. M. Law and J. M. Hutson. I-nolls: a program for interactive nonlinear least-square fitting of the parameters of physical models. *Comp. Phys. Comm.*, 102:252–268, 1997.
188. S. Huang, B. J. Braams, and J. M. Bowman. Ab initio potential energy and dipole moment surfaces for H_5O_2^+ . *J. Chem. Phys.*, 122:044308, 2005.
189. Z. Smith and E. B. Wilson. The infrared spectrum of gaseous malonaldehyde (3-hydroxy-2-propenal). *Spectrochim. Acta A*, 39:1117–1129, 1983.
190. D. W. Firth, P. F. Barbara, and H. P. Trommsdorf. Matrix induced localization of proton tunneling in malonaldehyde. *Chem. Phys.*, 136:349–360, 1989.
191. T. Chiavassa, R. Roubin, L. Piazzala, P. Verlaque, A. Allouche, and F. Marinelli. Experimental and theoretical studies of malonaldehyde - vibrational analysis

- of a strongly intramolecularly hydrogen-bonded compound. *J. Phys. Chem.*, 96:10659–10665, 1992.
192. C. Duan and D. Luckhaus. High resolution ir-diode laser jet spectroscopy of malonaldehyde. *Chem. Phys. Lett.*, 391:129–133, 2004.
193. Alexandra Viel, Mauricio D. Coutinho-Neto, and Uwe Manthe. The ground state tunneling splitting and the zero point energy of malonaldehyde: A quantum monte carlo determination. *J. Chem. Phys.*, 126:024308, 2007.
194. Anirban Hazra, Jonathan H. Skone, and Sharon Hammes-Schiffer. Combining the nuclear-electronic orbital approach with vibronic coupling theory: Calculation of the tunneling splitting for malonaldehyde. *J. Chem. Phys.*, 130:054108, 2009.
195. Yimin Wang, Bastiaan J. Braams, Joel M. Bowman, Stuart Carter, and David P. Tew. Full-dimensional quantum calculations of ground-state tunneling splitting of malonaldehyde using an accurate *ab initio* potential energy surface. *J. Chem. Phys.*, 128:224314, 2008.
196. Jacqueline C. Hargis, Francesco A. Evangelista, Justin B. Ingels, and Henry F. Schaefer III. Short intramolecular hydrogen bonds: Derivatives of malonaldehyde with symmetrical substituents. *J. Am. Chem. Soc.*, 130:17471–17478, 2008.
197. M. R. Johnson, N. H. Jones, A. Geis, A. J. Horsewill, , and H. P. Trommsdorff. Structure and dynamics of the keto and enol forms of acetylacetone in the solid state. *J. Chem. Phys.*, 116:5694–5700, 2002.
198. A. H. Lowrey, C. George, P. D'Antonio, and J. Karle. Structure of acetylacetone by electron diffraction. *J. Am. Chem. Soc.*, 93:6399–6403, 1971.
199. A. L. Andreassen and S. H. Bauer. The structures of acetylacetone, trifluoroacetyl-acetone and trifluoroacetone. *J. Mol. Spectrosc.*, 12:381–403, 1972.
200. Kinya Iijima, Atsushi Ohnogi, and Shuzo Shibata. The molecular structure of acetylacetone as studied by gas-phase electron diffraction. *J. Mol. Spectrosc.*, 156:111–118, 1987.
201. R. Srinivasan, J. S. Feenstra, S. T. Park, S. Xu, and A. H. Zewail. Direct determination of hydrogen-bonded structures in resonant and tautomeric reactions using ultrafast electron diffraction. *J. Am. Chem. Soc.*, 126:2266, 2004.

-
202. Walther Caminati and Jens-Uwe Grabow. The c_{2v} structure of enolic acetylacetone. *J. Am. Chem. Soc.*, 128:854–857, 2006.
203. A. Trivella, P. Roubin, P. Theule, M. Rajzmann, and S. Coussan. Uv and ir photoisomerization of acetylacetone trapped in a nitrogen matrix. *J. Phys. Chem. A*, 111:3074–3081, 2007.
204. Daniel W. Firth, Paul F. Barbara, and H. Peter Trommsdorff. Matrix induced localization of proton tunneling in malonaldehyd. *Chem. Phys.*, 136:349–360, 1989.
205. T. Chiavassa, P. Roubin, L. Pizzala, P. Verlaque, A. Allouche, and F. Marinelli. Experimental and theoretical studies of malonaldehyde: vibrational analysis of a strongly intramolecularly hydrogen bonded compound. *J. Phys. Chem.*, 96:10659–10665, 1992.
206. L. Pizzala A. Allouche P. Roubin T. Chiavassa, P. Verlaque. Experimental and theoretical studies of the photoisomerization of malonaldehyde isolated in rare gas matrixes. *J. Phys. Chem.*, 97:5917–5925, 1993.
207. S. Bratoz, D. Hadzi, and G. Rossmy. The infra-red absorption bands associated with the chelate ring in some unsaturated hydroxycarbonyl compounds. *Trans. Faraday Soc.*, 52:464–470, 1956.
208. Hisanobu Ogoshi and Kazuo Nakamoto. Normal-coordinate analyses of hydrogen-bonded compounds. v. the enol forms of acetylacetone and hexafluoroacetylacetone. *J. Chem. Phys.*, 45:3113–3120, 1966.
209. S. F. Tayyari, Th. Zeegers-Huyskens, and J. L. Wood. Spectroscopic study of hydrogen bonding in the enol form of β -diketones-i. vibrational assignment and strength of the bond. *Spectrochim. Acta A*, 39:1265–1272, 1979.
210. S. F. Tayyari, Th. Zeegers-Huyskens, and J. L. Wood. Spectroscopic study of hydrogen bonding in the enol form of β -diketones-ii. symmetry of the hydrogen bond. *Spectrochim. Acta A*, 39:1289–1295, 1979.
211. S. F. Tayyari and F. Milani-nejad. Vibrational assignment of acetylacetone. *Spectrochim. Acta A*, 56:2679–2691, 2000.
212. Janez Mavri and Joze Grdadolnik. Proton transfer dynamics in acetylacetone: A mixed quantum-classical simulation of vibrational spectra. *J. Phys. Chem. A*, 105:2045–2051, 2001.

213. Stacy A. Broadbent, Lori A. Burns, Chandrima Chatterjee, and Patrick H. Vaccaro. Investigation of electronic structure and proton transfer in ground state acetylacetone. *Chem. Phys. Lett.*, 434:31–37, 2007.
214. Sebastian Schlund, Eline M. Basilio Janke, Klaus Weisz, and Bernd Engels. Predicting the tautomeric equilibrium of acetylacetone in solution. i. the right answer for the wrong reason? *J. Comp. Chem.*, 31:665–670, 2010.
215. I. Matanovic and N. Doslic. Infrared spectroscopy of the intramolecular hydrogen bond in acetylacetone: A computational approach. *J. Phys. Chem. A*, 109:4185–4194, 2005.
216. O Sharafeddin, K Hinsen, T Carrington, and B. Roux. Mixing quantum-classical molecular dynamics methods applied to intramolecular proton transfer in acetylacetone. *J. Comp. Chem.*, 18:1760–1772, 1997.
217. Konrad Hinsen and Benoit Roux. Potential of mean force and reaction rates for proton transfer in acetylacetone. *J. Chem. Phys.*, 106:3567–3577, 1997.
218. Janez Mavri and Joze Grdadolnik. Proton potential in acetylacetone. *J. Phys. Chem. A*, 105:2039–2044, 2001.
219. A. Warshel and M. Levitt. Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.*, 103:227–249, 1976.
220. Thomas Fox and Peter A. Kollman. Application of the resp methodology in the parametrization of organic solvents. *J. Phys. Chem. B*, 102:8070–8079, 1998.
221. Gunter Kab, Christian Schroder, and Dirk Schwarzer. Intramolecular vibrational redistribution and energy relaxation in solution: A molecular dynamics approach. *PCCP*, 4:271–278, 2002.
222. M. P. Allen and D. J. Tildesley. *Computer Simulation of Liquids*. Clarendon Press, Oxford, 1987.
223. E. Helfand. Brownian dynamics study of transitions in a polymer chain of bistable oscillators. *J. Chem. Phys.*, 69:1010–1018, 1978.
224. W. H. Miller, N. C. Handy, and J. E. Adams. Reaction path hamiltonian for polyatomic molecules. *J. Chem. Phys.*, 72:99–112, 1980.

-
225. C. Clay Marston and Gabriel G. Balint-Kurti. The fourier grid hamiltonian method for bound state eigenvalues and eigenfunctions. *J. Chem. Phys.*, 91:3571–3577, 1989.
226. W. Miller. Quantum mechanical transition state theory and a new semiclassical model for reaction rate constants. *J. Chem. Phys.*, 61:1823, 1974.
227. Thomas Steiner. The hydrogen bond in the solid state. *Angew. Chem. Int. Ed.*, 41:48–76, 2002.
228. Anthony Mittermaier and Lewis E. Kay. New tools provide new insights in nmr studies of protein dynamics. *Science*, 312:224–228, 2006.
229. Stephan Grzesiek and Hans-Jurgen Sass. From biomolecular structure to functional understanding: new nmr developments narrow the gap. *Curr. Opin. Struct. Bio.*, 19:585–595, 2009.
230. Kevin J. Bowers, Ron O. Dror, and David E. Shaw. The midpoint method for parallelization of particle simulations. *J. Chem. Phys.*, 124:184109, 2006.
231. Abhinav Bhatele, Laxmikant V. Kale, and Sameer Kumar. Dynamic topology aware load balancing algorithms for molecular dynamics applications. *Proceedings of 23rd ACM International Conference on Supercomputing*, 2009.
232. Andrew J. Dingley and Stephan Grzesiek. Direct observation of hydrogen bonds in nucleic acid base pairs by internucleotide $^2j_{nm}$ couplings. *J. Am. Chem. Soc.*, 120(33):8293–8297, 1998.
233. Gabriel Cornilescu, Benjamin E. Ramirez, M. Kirsten Frank, G. Marius Clore, Angela M. Gronenborn, and Ad Bax. Correlation between $^3j_{nc'}$ and hydrogen bond length in proteins. *J. Am. Chem. Soc.*, 121:6275–6279, 1999.
234. Stephan Grzesieka, Florence Cordiera, Victor Jaravinea, and Michael Barfield. Insights into biomolecular hydrogen bonds from hydrogen bond scalar couplings. *Prog. Nucl. Magn. Reson. Spectrosc.*, 45:275–300, 2004.
235. Ibon Alkorta, Jose Elguero, and Gleb S. Denisov. A review with comprehensive data on experimental indirect scalar nmr spin-spin coupling constants across hydrogen bonds. *Magn. Reson. Chem.*, 46:599–624, 2008.

236. Hans-Jurgen Sass, Franziska Fang-Fang Schmid, and Stephan Grzesiek. Correlation of protein structure and dynamics to scalar couplings across hydrogen bonds. *J. Am. Chem. Soc.*, 129:5898–5903, 2007.
237. Oliver F. Lange, David van der Spoel, and Bert L. de Groot. Scrutinizing molecular mechanics force fields on the microsecond timescale with nmr data. *Biophys. J.*, 99:647–655, 2010.
238. Alexandre V. Morozov, Tanja Kortemme, Kiril Tsemekhman, and David Baker. Close agreement between the orientation dependence of hydrogen bonds observed in protein structures and quantum mechanical calculations. *Proc. Natl. Acad. Sci.*, 101:6946–6951, 2004.
239. M. Devereux and M. Meuwly. Force field optimization using dynamics and ensemble-averaged data: Vibrational spectra and relaxation in bound mbco. *J. Chem. Inf. Model.*, 50:349–357, 2010.
240. D. A. Case. Molecular dynamics and nmr spin relaxation in proteins. *Acc. Chem. Res.*, 35:325–331, 2002.
241. Guillaume Bouvignies, Pau Bernado, Sebastian Meier, Kyuil Cho, Stephan Grzesiek, Rafael Bruschweiler, and Martin Blackledge. Identification of slow correlated motions in proteins using residual dipolar and hydrogen-bond scalar couplings. *Proc. Natl. Acad. Sci.*, 102:13885–13890, 2005.
242. Oliver F. Lange, Nils-Alexander Lakomek, Christophe Fares, Gunnar F. Schroder, Korvin F. A. Walter, Stefan Becker, Jens Meiler, Helmut Grubmuller, Christian Griesinger, and Bert de Groot. Recognition dynamics up to microseconds revealed from an rdc-derived ubiquitin ensemble in solution. *Science*, 320:1471–1475, 2008.
243. Nils-Alexander Lakomek, Korvin Walter, Christophe Fares, Oliver Lange, Bert de Groot, Helmut Grubmuller, Rafael Bruschweiler, Axel Munk, Stefan Becker, Jens Meiler, and Christian Griesinger. Self-consistent residual dipolar coupling based model-free analysis for the robust determination of nanosecond to microsecond protein dynamics. *J. Biomol. NMR*, 41:139–155, 2008.
244. Franziska F.-F. Schmid and Markus Meuwly. Direct comparison of experimental and calculated nmr scalar coupling constants for force field validation and adaptation. *J. Chem. Theory Comput.*, 4:1949–1958, 2008.

245. A. Miyashita, A. Yasuda, H. Takaya, K. Toriumi, T. Ito, T. Souchi, and R. Noyori. Synthesis of 2,2'-bis(diphenylphosphino)-1,1'-binaphthyl (binap), an atropisomeric chiral bis(triaryl)phosphine, and its use in the rhodium(i)-catalyzed asymmetric hydrogenation of α -(acylamino)acrylic acids. *J. Am. Chem. Soc.*, 102:7932–7934, 1980.
246. Mirko Kranenburg, Yuri E. M. van der Burgt, Paul C. J. Kamer, Piet W. N. M. van Leeuwen, Kees Goubitz, and Jan Fraanje. New diphosphine ligands based on heterocyclic aromatics inducing very high regioselectivity in rhodium-catalyzed hydroformylation: Effect of the bite angle. *Organometallics*, 14:3081–3089, 1995.
247. Bernhard Breit and Wolfgang Seiche. Hydrogen bonding as a construction element for bidentate donor ligands in homogeneous catalysis: regioselective hydroformylation of terminal alkenes. *J. Am. Chem. Soc.*, 125:6608–6609, 2003.
248. Wolfgang Seiche, Alexander Schuschkowski, and Bernhard Breit. Bidentate ligands by self-assembly through hydrogen bonding: A general room temperature/ambient pressure regioselective hydroformylation of terminal alkenes. *Adv. Synth. Catal.*, 347:1488–1494, 2005.
249. Floris Chevallier and Bernhard Breit. Self-assembled bidentate ligands for ruthenium-catalyzed anti-markovnikov hydration of terminal alkynes. *Angew. Chem. Int. Ed.*, 45:1599–1602, 2006.
250. J. Wieland and B. Breit. A combinatorial approach to the identification of self-assembled ligands for rhodium-catalyzed asymmetric hydrogenation. *Nat. Chem.*, 2:832–837, 2010.
251. Urs Gellrich, Jing Huang, Wolfgang Seiche, Manfred Kellera, Markus Meuwly, and Bernhard Breit. Origin of ligand self-assembling through complementary hydrogen-bonding in the coordination sphere of a transition metal center: The 6-diphenylphosphanylpyridin-2(1h)-one system. *J. Am. Chem. Soc.*, 133:964–975, 2011.
252. K. N. Houk and Paul Ha-Yeon Cheong. Computational prediction of small-molecule catalysts. *Nature*, 455:309–313, 2008.
253. Brian V. Popp, Johanna E. Wendlandt, Clark R. Landis, and Shannon S. Stahl. Reaction of molecular oxygen with an nhc-coordinated pd^0 complex: Computa-

- tional insights and experimental implications. *Angew. Chem. Int. Ed.*, 46:601–604, 2007.
254. Clement Mazet, Sebastian P. Smidt, Markus Meuwly, and Andreas Pfaltz. A combined experimental and computational study of dihydrido(phosphinooxazoline)iridium complexes. *J. Am. Chem. Soc.*, 126:14176–14181, 2004.
255. Peter L. Freddolino, Anton S. Arkhipov, Steven B. Larson, Alexander McPherson, and Klaus Schulten. Molecular dynamics simulations of the complete satellite tobacco mosaic virus. *Structure*, 14:437–449, 2006.
256. P. Comba. The relation between ligand structures, coordination stereochemistry, and electronic and thermodynamic properties. *Cord. Chem. Rev.*, 123:1–48, 1993.
257. M. Zimmer. Bioinorganic molecular mechanics. *Chem. Rev.*, 95:2629–2649, 1995.
258. Peter Comba and R. Remenyi. Inorganic and bioinorganic molecular mechanics modeling - the problem of the force field parameterization. *Cord. Chem. Rev.*, 238-239:9–20, 2003.
259. Peter Comba and Marion Kerscher. Computation of structures and properties of transition metal compounds. *Cord. Chem. Rev.*, 253:564–574, 2009.
260. V. J. Burton, R. J. Deeth, C. M. Kemp, and P. J. Gilbert. Molecular mechanics for coordination complexes: The impact of adding d-electron stabilization energies. *J. Am. Chem. Soc.*, 117:8407–8415, 1995.
261. Robert J. Deeth, Natalie Fey, and Benjamin Williams-Hubbard. Dommimoe: An implementation of ligand field molecular mechanics in the molecular operating environment. *J. Comp. Chem.*, 26:123–130, 2005.
262. R.J. Deeth, A. Anastasi, C. Diedrich, and K. Randell. Molecular modelling for transition metal complexes: Dealing with d-electron effects. *Coord. Chem. Rev.*, 253:795–816, 2009.
263. Robert J. Deeth. Comprehensive molecular mechanics model for oxidized type i copper proteins: Active site structures, strain energies, and entatic bulging. *Inorg. Chem.*, 46:4492–4503, 2007.

264. Robert J. Deeth. General molecular mechanics method for transition metal carboxylates and its application to the multiple coordination modes in mono- and dinuclear mn(ii) complexes. *Inorg. Chem.*, 47:6711–6725, 2008.
265. N. Gresh, P. Claverie, and A. Pullman. Theoretical studies of molecular conformation. derivation of an additive procedure for the computation of intramolecular interaction energies. comparison with *ab initio* scf computations. *Theor. Chim. Acta*, 66:1–20, 1984.
266. N. Gresh. Energetics of zn binding to a series of biologically relevant ligands: A molecular mechanics investigation grounded on *ab initio* self-consistent field supermolecule calculations. *J. Comp. Chem.*, 16:856–882, 1995.
267. N. Gresh, C. Policar, and C. Giessner-Prettre. Modeling copper(i) complexes: Sibfa molecular mechanics versus *ab initio* energetics and geometrical arrangements. *J. Phys. Chem. A*, 106:5660–5670, 2002.
268. N. Gresh, J.-P. Piquemal, and M. Krauss. Representation of zn(ii) complexes in polarizable molecular mechanics. further refinements of the electrostatic and short-range contributions. comparison with *ab initio* computations. *J. Comp. Chem.*, 26:1113–1130, 2005.
269. B. de Courcy, J.-P. Piquemal, and N. Gresh. Energy analysis of zn polycoordination in a metalloprotein environment and of the role of a neighboring aromatic residue. what is the impact of polarization? *J. Chem. Theo. Comp.*, 4:1659–1668, 2008.
270. Jean-Philip Piquemal, Benjamin Williams-Hubbard, Natalie Fey, Robert J. Deeth, Nohad Gresh, and Claude Giessner-Prettre. Inclusion of the ligand field contribution in a polarizable molecular mechanics: Sibfa-lf. *J. Comp. Chem.*, 24:1963–1970, 2003.
271. L. Pauling. *The nature of the chemical bond*. Cornell University Press, Ithaca, NY, 1960.
272. D. M. Root, C. R. Landis, and T. Cleveland. Valence bond concepts applied to the molecular mechanics description of molecular shapes. 1. application to nonhypervalent molecules of the p-block. *J. Am. Chem. Soc.*, 115:4201–4209, 1993.

273. T. Cleveland and C. G. Landis. Valence bond concepts applied to the molecular mechanics description of molecular shapes. 2. applications to hypervalent molecules of the p-block. *J. Am. Chem. Soc.*, 118:6020–6030, 1996.
274. C. R. Landis, T. Cleveland, and T. K. Firman. Valence bond concepts applied to the molecular mechanics description of molecular shapes. 3. applications to transition metal alkyls and hydrides. *J. Am. Chem. Soc.*, 120:2641–2649, 1998.
275. C. R. Landis, T. K. Firman, D. M. Root, and T. Cleveland. A valence bond perspective on the molecular shapes of simple metal alkyls and hydrides. *J. Am. Chem. Soc.*, 120:1842–1854, 1998.
276. T. K. Firman and C. R. Landis. Valence bond concepts applied to the molecular mechanics description of molecular shapes. 4. transition metals with pi-bonds. *J. Am. Chem. Soc.*, 123:11728–11742, 2001.
277. I. Tubert-Brohman, M. Schmid, and M. Meuwly. A molecular mechanics force field for octahedral organometallic compounds with inclusion of the trans influence. *J. Chem. Theo. Comp.*, 5:530–539, 2009.
278. M. Meuwly and J. Hutson. Morphing ab initio potentials: A systematic study of Ne-HF. *J. Chem. Phys.*, 110:8338–8347, 1999.
279. Jonas Sjoqvist, Mathieu Linares, and Patrick Norman. Platinum(ii) and phosphorus mm³ force field parametrization for chromophore absorption spectra at room temperature. *J. Phys. Chem. A*, 114:4981–4987, 2010.
280. Achim Lienke, Gnter Klatt, David J. Robinson, Klaus R. Koch, and Kevin J. Naidoo. Modeling platinum group metal complexes in aqueous solution. *Inorg. Chem.*, 40:2352–2357, 2001.
281. E. D. Glendening, A. E. Reed, J. E. Carpenter, and F. Weinhold. Nbo version 3.1. *NBO Version 3.1*.
282. Célia Fonseca Guerra, Jan-Willem Handgraaf, Evert Jan Baerends, and F. Matthias Bickelhaupt. Voronoi deformation density (vdd) charges: Assessment of the mulliken, bader, hirshfeld, weinhold, and vdd methods for charge analysis. *J. Comp. Chem.*, 25:189–210, 2004.

-
283. K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, and A. D. Mackerell. Charmm general force field: A force field for drug-like molecules compatible with the charmm all-atom additive biological force fields. *J. Comp. Chem.*, 31:671–690, 2009.
284. Thomas Fox and Peter A. Kollman. Application of the resp methodology in the parametrization of organic solvents. *J. Phys. Chem. B*, 102:8070–8079, 1998.
285. Gunter Kab, Christian Schroder, and Dirk Schwarzer. Intramolecular vibrational redistribution and energy relaxation in solution: A molecular dynamics approach. *PCCP*, 4:271–278, 2002.
286. S. Nose. A unified formulation of the constant temperature molecular-dynamics methods. *J. Chem. Phys.*, 81:511–519, 1984.
287. William G Hoover. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A*, 31:1695–1697, 1985.
288. C. Lee, R. Yang, and W. G. Parr. Development of the colle-salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B*, 37:785, 1988.
289. Axel D. Becke. Density-functional thermochemistry. iii. the role of exact exchange. *J. Chem. Phys.*, 98:5648–5652, 1993.
290. P. J. Hay and W. R. Wadt. Ab initio effective core potential for molecular calculation. *J. Chem. Phys.*, 82:270–283, 1985.
291. Krzysztof Wolinski, James F. Hinton, and Peter Pulay. Efficient implementation of the gauge-independent atomic orbital method for nmr chemical shift calculations. *J. Am. Chem. Soc.*, 112:8251–8260, 1990.
292. Takeshi Yanai, David P. Tew, and Nicholas C. Handy. A new hybrid exchange-correlation functional using the coulomb-attenuating method. *Chem. Phys. Lett.*, 393:51–57, 2004.
293. N. M. O’Boyle, A. L. Tenderholt, and K. M. Langner. cclib: A library for package-independent computational chemistry algorithms. *J. Comp. Chem.*, 29:839–845, 2008.

294. Zoraida Freixa and Piet W. N. M. van Leeuwen. Bite angle effects in diphosphine metal catalysts: steric or electronic? *Dalton Trans.*, 2003:1890–1901, 2003.
295. Barry J. Dunne, Richard B. Morris, and A. Guy Orpen. Structural systematics. part 3. geometry deformations in triphenylphosphine fragments: a test of bonding theories in phosphine complexes. *J. Chem. Soc. Dalton Trans.*, 1991:653–661, 1991.
296. Shinichi Miura, Mark E. Tuckerman, and Michael L. Klein. An ab initio path integral molecular dynamics study of double proton transfer in the formic acid dimer. *J. Chem. Phys.*, 109:5290–5299, 1998.
297. Jeffrey C. Hoch, Christopher M. Dobson, and Martin Karplus. Fluctuations and averaging of proton chemical shifts in the bovine pancreatic trypsin inhibitor. *Biochem.*, 21:1118–1125, 1982.
298. Motoyuki Shiga, Kimichi Suzuki, and Masanori Tachikawa. The chemical shift of deprotonated water dimer: Ab initio path integral simulation. *J. Chem. Phys.*, 132:114104, 2010.
299. Itzam De Gortari, Guillem Portella, Xavier Salvatella, Vikram S. Bajaj, Patrick C. A. van der Wel, Jonathan R. Yates, Matthew D. Segall, Chris J. Pickard, Mike C. Payne, and Michele Vendruscolo. Time averaging of nmr chemical shifts in the mlf peptide in the solid state. *J. Am. Chem. Soc.*, 132:5993–6000, 2010.
300. James Keeler. *Understanding NMR Spectroscopy*. Wiley, 2005.
301. Andreas Mueller, Francis Talbot, and Samuel Leutwyler. Intermolecular vibrations of jet-cooled (2-pyridone)₂: A model for the uracil dimer. *J. Chem. Phys.*, 283:341–354, 2000.
302. Jessica M. J. Swanson and Jack Simons. Role of charge transfer in the structure and dynamics of the hydrated proton. *J. Phys. Chem. B*, 113:5149–5161, 2009.
303. Bernd Ensing, Steven O. Nielsen, Preston B. Moore, Michael L. Klein, and Michele Parrinello. Energy conservation in adaptive hybrid atomistic/coarse-grain molecular dynamics. *J. Chem. Theo. Comp.*, 3:1100–1105, 2007.
304. Shina C.L. Kamerlin, Spyridon Vicatos, Anatoly Dryga, and Arieh Warshel. Coarse-grained (multiscale) simulations in studies of biophysical and chemical systems. *Ann. Rev. Phys. Chem.*, 62:41–64, 2011.

-
305. Haibo Yu and Qiang Cui. The vibrational spectra of protonated water clusters: A benchmark for self-consistent-charge density-functional tight binding. *J. Chem. Phys.*, 127:234504, 2007.
306. Mark L. Brewer, Udo W. Schmitt, and Gregory A. Voth. The formation and dynamics of proton wires in channel environments. *Biophys. J.*, 80:1691 – 1702, 2001.
307. Christoph Dellago, Mor M. Naor, and Gerhard Hummer. Proton transport through water-filled carbon nanotubes. *Phys. Rev. Lett.*, 90:105902, 2003.
308. Christoph Dellago, Mor M. Naor, and Gerhard Hummer. Erratum: Proton transport through water-filled carbon nanotubes [phys. rev. lett. 90, 105902 (2003)]. *Phys. Rev. Lett.*, 91:139902, 2003.
309. Zhen Cao, Yuxing Peng, Tianying Yan, Shu Li, Ailin Li, and Gregory A. Voth. Mechanism of fast proton transport along one-dimensional water chains confined in carbon nanotubes. *J. Am. Chem. Soc.*, 132:11395–11397, 2010.
310. William Humphrey, Andrew Dalke, and Klaus Schulten. VMD – Visual Molecular Dynamics. *J. Mol. Graph.*, 14:33–38, 1996.
311. D. Nutt and M. Meuwly. CO migration in native and mutant myoglobin: Atomistic simulations for the understanding of protein function. *Proc. Natl. Acad. Sci.*, 101:5998–6002, 2004.
312. J. Danielsson and M. Meuwly. Atomistic simulation of adiabatic reactive processes based on multi-state potential energy surfaces. *J. Chem. Theo. Comp.*, 4:1083–1093, 2008.
313. S. Lutz. *Computational Investigation of Conformational and Spectroscopic Substates in Neuroglobin*. PhD thesis, University of Basel, 2011.
314. Sabyashachi Mishra and Markus Meuwly. Atomistic simulation of no dioxygenation in group i truncated hemoglobin. *J. Am. Chem. Soc.*, 132:2968–2982, 2010.
315. Jing Huang and Markus Meuwly. Explicit hydrogen-bond potentials and their application to nmr scalar couplings in proteins. *J. Chem. Theo. Comp.*, 6:467–476, 2010.

Bibliography

316. Jing Huang and Markus Meuwly. Force field refinement from nmr scalar couplings. *Chem. Phys.*, submitted, 2011.
317. Stephan Lutz, Ivan Tubert-Brohman, Yonggang Yang, and Markus Meuwly. Water-assisted proton transfer in ferredoxin i. *J. Biol. Chem.*, in press, 2011.

Appendix

Curriculum vitae

Personal Data

Address: Jing Huang
St. Johannis-Vorsdadt 70
4056 Basel

Email: jing.huang@unibas.ch

Nationality: China

Education

08/2007 – 07/2011 **PhD Studies, Department of Chemistry, University of Basel**
Thesis: Proton Transfer and Hydrogen Bonding in Chemical and Biological Systems: A Force Field Approach
Supervisor: Prof. Dr. Markus Meuwly
External examiner: Prof. Dr. Helmut Grubmüller

09/2005 – 07/2007 **M. Sc. in Physics, Tsinghua University**
Thesis: Dynamical Potential Approach to Molecular Highly Excited Vibration
Supervisor: Prof. Dr. Guozhen Wu

09/2001 – 07/2005 **B. Sc. in Physics, Tsinghua University**

Honors and Awards

- Camille- und Henry-Dreyfus Stipendien, *University of Basel*, 2011
- "Friends of Tsinghua" scholarship, *Tsinghua University*, 2002 (first prize) and 2003 (second prize)

Publications

- U. Gellrich, J. Huang, W. Seiche, M. Keller, M. Meuwly and B. Breit, "**Ligand Self-Assembling through Complementary Hydrogen-Bonding in the Coordination Sphere of a Transition Metal Center: The 6-Diphenylphosphanylpyridin-2(1H)-one system**", *J. Am. Chem. Soc.*, 133, 964 (2011)
- J. Huang and M. Meuwly, "**Explicit Hydrogen-bond Potentials and their Application to NMR Scalar Couplings in Proteins**", *J. Chem. Theo. Comput.*, 6, 467 (2010)
- J. Huang and G. Wu, "**Dynamical Potential Approach to DCO Highly Excited Vibration**", *Chem. Phys. Lett.*, 439, 231 (2007)

In press or submitted:

- J. Huang, M. Devereux, F. Hofmann and M. Meuwly
Computational Organometallic Chemistry with Force Fields
book chapter in "Computational Organometallic Chemistry", Ed., O. Wiest, Springer, in press
- J. Huang and M. Meuwly
Force Field Refinement from NMR Scalar Couplings
submitted

Presentations

- **Efficient Computational Methods for Transition Metal Complexes: Energetics and Dynamics of a Hydrogen-Bonded Bidentate Catalyst**
4th Workshop of the International Research Training Group "Catalysts and Catalytic Reactions for Organic Synthesis", June 10th, 2010, Basel, Switzerland
- **Efficient Computational Methods for Transition Metal Complexes: Computational Characterization of a Hydrogen-Bonded Bidentate Catalyst**
25th Molecular Modelling Workshop, April 5th, 2011, Erlangen, German