การจำลองพลวัตเชิงโมเลกุลที่ผสมผสานกลศาสตร์ควอนตัมและ

กลศาสตร์โมเลกุล (QM/MM MD) ของระบบไอออนฟอร์เมต (HCOO⁻)

และอะซีเตท (CH₃COO⁻) ในสารละลายน้ำ

นายอภิรักษ์ พยัคฆา

AB INITIO QM/MM MD SIMULATIONS OF

FORMATE (HCOO⁻) AND ACETATE (CH₃COO⁻) IONS

IN AQUEOUS SOLUTION

Apirak Payaka

A Thesis Submitted in Partial Fulfillment of the Requirements for

the Degree of Doctor of Philosophy in Chemistry

Suranaree University of Technology

Academic Year 2009

AB INITIO QM/MM MD SIMULATIONS OF FORMATE (HCOO⁻)

AND ACETATE (CH₃COO⁻) IONS IN AQUEOUS SOLUTION

Suranaree University of Technology has approved this thesis submitted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy.

Thesis Examining Committee

_____

(Assoc. Prof. Dr. Albert Schulte)

Chairperson

_____

(Assoc. Prof. Dr. Anan Tongraar)

Member (Thesis Advisor)

_____

(Prof. Dr. Jiali Gao)

Member

_____

(Prof. Dr. Kritsana Sagarik)

Member

_____

(Asst. Prof. Dr. Kunwadee Rangsriwatananon)

Member

_____   _____

(Prof. Dr. Sukit Limpijumnong)   (Assoc. Prof. Dr. Prapun Manyum)

Vice Rector for Academic Affairs   Dean of Institute of Science

อภิรักษ์ พยัคฆา : การจำลองพลวัตเชิงโมเลกุลที่ผสมผสานกลศาสตร์ควอนตัมและกลศาสตร์โมเลกุล (QM/MM MD) ของระบบไอออนฟอร์เมต (HCOO⁻) และอะซีเตท (CH₃COO⁻) ในสารละลายน้ำ (*AB INITIO* QM/MM MD SIMULATIONS OF FORMATE (HCOO⁻) AND ACETATE (CH₃COO⁻) IONS IN AQUEOUS SOLUTION) อาจารย์ที่ปรึกษา : รองศาสตราจารย์ ดร. อนันต์ ทองระอา, 150 หน้า.

การจำลองพลวัตเชิงโมเลกุลที่ผสมผสานกลศาสตร์ควอนตัมและกลศาสตร์โมเลกุล 2 แบบ ได้แก่ HF/MM   และ   B3LYP/MM ได้ถูกดำเนินการเพื่อคำนวณคุณลักษณะของพันธะไฮโดรเจนระหว่างไอออนฟอร์เมต (HCOO⁻) และไอออนอะซีเตท (CH₃COO⁻) กับน้ำ โดยเทคนิค QM/MM ที่นำมาใช้นี้ส่วนของระบบที่เป็นไอออนกับน้ำบริเวณรอบ ๆ ไอออนจะถูกดำเนินการโดยวิธีฮาร์ทรี-ฟอร์ก (HF) และวิธี B3LYP โดยใช้เบซีสเซตชนิด DZV+   ในขณะที่ส่วนที่เหลือของระบบอธิบายโดยใช้ฟังก์ชันศักย์คู่ ผลที่ได้จากการจำลองแสดงให้เห็นความแข็งแรงของพันธะไฮโดรเจนระหว่าง HCOO⁻ กับน้ำ และ CH₃COO⁻ กับน้ำ โดยเฉพาะเมื่อเปรียบเทียบกับพันธะไฮโดรเจนของตัวทำละลายน้ำ จากข้อมูลวิถีโคจรที่ได้จากการจำลองพบว่าน้ำที่อยู่ในซอลเวชันชั้นแรกสามารถที่จะยึดเหนี่ยวกับอะตอมออกซิเจนของหมู่คาร์บอนิลของ HCOO⁻ และ CH₃COO⁻ ได้ในลักษณะหลวม ๆ และในลักษณะที่เหนียวแน่น ส่งผลให้ได้ค่าเลขไฮเดรชันที่แต่ละอะตอมออกซิเจนของหมู่คาร์บอนิลของ HCOO⁻ และ CH₃COO⁻ ที่เบี่ยงเบนสูง โดยมีค่าได้ตั้งแต่ 1 ถึง 6 โดยมีค่าที่มีความถี่สูงสุดคือ 3 เมื่อเปรียบเทียบพันธะไฮโดรเจนระหว่าง HCOO⁻ กับน้ำ และ CH₃COO⁻ กับน้ำ พบว่าพันธะไฮโดรเจนของ CH₃COO⁻ กับน้ำ มีความแข็งแรงมากกว่าของ HCOO⁻ กับน้ำ ซึ่งสามารถอธิบายได้จากการที่หมู่เมททีลของ CH₃COO⁻ สามารถให้อิเล็กตรอนกับหมู่คาร์บอนิลของ CH₃COO⁻ ส่งผลให้หมู่คาร์บอนิลของ CH₃COO⁻ สามารถเกิดพันธะไฮโดรเจนกับน้ำที่บริเวณรอบ ๆ ได้แข็งแรงมากขึ้น ในส่วนของการเปรียบเทียบการใช้งานวิธีฮาร์ทรี-ฟอร์กกับวิธี B3LYP พบว่าวิธีฮาร์ทรี-ฟอร์กให้ผลการศึกษาเกี่ยวกับพันธะไฮโดรเจนนี้อ่อนกว่าการคำนวณ B3LYP เล็กน้อย และทำนายระยะพันธะไฮโดรเจนที่ยาวกว่า แต่อย่างไรก็ตามการคำนวณ B3LYP ให้ผลการคำนวณในส่วนของสมบัติเชิงพลวัตที่ค่อนข้างต่ำกว่าผลการทดลองมาก

สาขาวิชาเคมี   ลายมือชื่อนักศึกษา_____

ปีการศึกษา 2552   ลายมือชื่ออาจารย์ที่ปรึกษา_____

APIRAK PAYAKA : *AB INITIO* QM/MM MD SIMULATIONS OF FORMATE (HCOO⁻) AND ACETATE (CH$_3$COO⁻) IONS IN AQUEOUS SOLUTION. THESIS ADVISOR : ASSOC. PROF. ANAN TONGRAAR, Ph.D. 150 PP.

FORMATE/ACETATE/HYDROGEN BOND/WATER EXCHANGE

Two combined quantum mechanics/molecular mechanics (QM/MM) molecular dynamics simulations, namely HF/MM and B3LYP/MM, have been performed to investigate the characteristics of HCOO⁻-water and CH$_3$COO⁻-water hydrogen bonds in aqueous solution. By the QM/MM technique, the ion and its surrounding water molecules were treated at HF and B3LYP levels of accuracy, respectively, using DZV+ basis set, while the rest of the system was described by means of classical pair potentials. In this work, the results obtained by the HF/MM and B3LYP/MM simulations clearly indicate strong HCOO⁻-water and CH$_3$COO⁻-water hydrogen bonds, especially when compare to that of water-water hydrogen bonds in bulk. Regarding to the HF/MM and B3LYP/MM trajectories, it is observed that first-shell waters are either "loosely" or "tightly" bound to their respective HCOO⁻ and CH$_3$COO⁻ oxygen atoms. Consequently, this leads to large fluctuations in the hydration number, varying from 1 to 6, with the prevalent value of 3. Comparing between the HCOO⁻-water and CH$_3$COO⁻-water hydrogen bonds, the latter one is found to be stronger as a consequence of electronic effects in CH$_3$COO⁻, *i.e.*, the methyl group of CH$_3$COO⁻ can donate electron to CH$_3$COO⁻ carbonyl group. With regard to the HF and B3LYP methods for the description of QM treated region, the

first one leads to slightly too weak and thus longer hydrogen bonds, while the latter predicts them stronger with the too slow dynamical data.

School of Chemistry   Student's Signature_____

Academic Year 2009   Advisor's Signature _____

ACKNOWLEDGMENTS

Apirak Payaka

CONTENTS

CONTENTS (Continued)

CONTENTS (Continued)

CONTENTS (Continued)

LIST OF TABLES

LIST OF TABLES (Continued)

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF ABBREVIATIONS

XRD   =   X-ray diffraction

ND   =   Neutron diffraction

NMR   =   Nuclear magnetic resonance spectroscopy

XAS   =   X-ray absorption spectroscopy

MC   =   Monte Carlo

MD   =   Molecular dynamics

CPMD   =   Car-Parrinello molecular dynamics

QM/MM   =   Quantum mechanics/molecular mechanics

QM   =   Quantum mechanics

MM   =   Molecular mechanics

$RCOO^-$ =   Carboxylate ions

$HCOO^-$ =   Formate ion

CH_3COO^- =   Acetate ion

mole%   =   Mole percent

OPLS   =   Optimized potentials for liquid simulations

SPC/E   =   Simple point charge effective pair water model

MC-OPLS   =   Monte Carlo simulation with optimized potentials for liquid simulations

MD-SPC/E   =   Molecular dynamics with simple point charge effective pair water model

LIST OF ABBREVIATIONS (Continued)

BLYP	=	Becke hybrid functional combined with Lee-Yang-Parr correlation function
HF	=	Hartree-Fock
B3LYP	=	Beck three-parameters hybrid functional combined with Lee-Yang-Parr correlation function
DFT	=	Density functional theory
DZV+	=	Double zeta valence with diffuse function
CPU	=	Central processing unit
BJH	=	Flexible water model developed by Bopp, Jancsó and Heinzinger
CF2	=	Central force model version 2
RDF	=	Radial distribution function
ADF	=	Angular distribution function
MRT	=	Mean residence times
λ	=	wavelength
ψ	=	wavefunction
m	=	Mass
V	=	Potential field
H	=	Hamiltonian operator
$\hbar$	=	Plank's constant
m_e	=	Mass of electron

LIST OF ABBREVIATIONS (Continued)

m_k	=	Mass of nucleus
e	=	Electron charge
Z	=	Atomic number
r_{ab}	=	Distance between particles a and b
∇^2	=	Laplacian operator
Φ	=	Trial function
V_N	=	The nuclear-nuclear repulsion energy
E_{el}	=	Electronic energy
$\left|\psi_i(r)\right|^2$	=	Spatial distribution of an electron
$\chi(x)$	=	Spin orbital
etc	=	et cetera
LCAO-MO	=	Linear combination of atomic orbitals to molecular orbitals
STO	=	Slater-type orbital
GTO	=	Gaussian-type orbital
ECP	=	Effective core potential
PP	=	Pseudopotential
J_u	=	Coulomb operator
K_u	=	Exchange operator
SCF	=	Self-consistent field
RHF	=	Restricted Hartree-Fock

LIST OF ABBREVIATIONS (Continued)

UHF	=	Unrestricted open-shell Hartree-Fock
CI	=	Configuration interaction
MP	=	Many-body perturbation
CC	=	Coupled cluster
KS	=	Kohn-Sham
GGA	=	Generalized gradient approximation
NVE	=	Microcanonical ensemble
NVT	=	Canonical ensemble
μVT	=	Grand canonical ensemble
μ	=	Chemical potential
V	=	Volume
T	=	Temperature
E_{tot}	=	Total interaction energy
E_{MM}	=	Interactions within MM region
E_{QM-MM}	=	Interactions between QM and MM regions
$\left\langle \Psi_{QM} \left| \hat{H} \right| \Psi_{QM} \right\rangle$	=	Interactions within QM region
F_i	=	Forces acting on each particle
$S_m(r)$	=	Smoothing function
F_{QM}	=	QM force
F_{MM}	=	MM force

LIST OF ABBREVIATIONS (Continued)

r_1	=	Distance characterizing the start of QM region
r_0	=	Distance characterizing the end of QM region
PW91	=	Perdew–Wang 1991 correlation function
MP2	=	Second-order Møller-Plesset perturbation theory
aug-cc-pVDZ	=	Additional diffuse basis function and correlation consistent polarized valence double zeta
kcal/mol	=	Kilocalorie per mole
Å	=	Angstrom
°	=	Degree
NBO	=	Natural bond orbital
$\Delta E_{HCOO^- - H_2O}$	=	Pair potential energy between $HCOO^-$ and H_2O
$\Delta E_{CH_3COO^- - H_2O}$	=	Pair potential energy between CH_3COO^- and H_2O
r_{ij}	=	Distances between i-th atom and j-th atom
q	=	Atomic net charges
C_{methyl}	=	Carbon atom of methyl group of CH_3COO^-
$C_{carbonyl}$	=	Carbon atom of carbonyl group of CH_3COO^-
ΔE_{FIT}	=	Fitted pair potential function
ΔE_{HF}	=	Stabilization energies of the HF single point calculations
ΔE_{B3LYP}	=	Stabilization energies of the B3LYP single point calculations
BSSE	=	Basis set superposition error

LIST OF ABBREVIATIONS (Continued)

ΔE_{BSSE}	=	Stabilization energies with basis set superposition error (BSSE) correction
CN	=	Coordination number
t_{sim}	=	Simulation time
N_{ex}	=	Number of exchange events
t^{*}	=	Time for observing the number of exchange water molecules
A	=	Associative exchange mechanism
D	=	Dissociative exchange mechanism
I	=	Interchange mechanism
I_a	=	Associative-like interchange mechanism
I_d	=	Dissociative-like interchange mechanism
K	=	Kelvin
fs	=	Femto second
ps	=	Pico second
$\tau_{H_2O}(O_i)$	=	MRT of water molecules at each of oxygen atoms
$\tau_{H_2O}(H_2O)$	=	MRT of water molecules in the bulk

CHAPTER I

INTRODUCTION

1.1 Literature reviews of RCOO⁻ in aqueous solution

Comprehensive understanding of characteristics of ions solvated in aqueous electrolyte solutions has long been a scientific interest due to their well-known relevance in many chemical and biological processes (Frank, 1956; Williams, 1971; Cowan, 1998). Accordingly, reliable methods for an accurate treatment of such systems have always been in strong demand. In terms of experiments, several techniques such as x-ray diffraction (XRD), neutron diffraction (ND), nuclear magnetic resonance (NMR) spectroscopy (Ohtaki and Radnai, 1993) and x-ray absorption spectroscopy (XAS) have been used to obtain detailed knowledge of ions in aqueous solution. However, most of the experiments often require the resources of a major research laboratory and the results are often prone to large errors, especially for very dilute solution, due to their technical limitations.

In conjunction with experiments, computer simulations, in particular Monte Carlo (MC) and molecular dynamics (MD), have become an alternative way to elucidate such details. During the past decades, most of the simulation works had applied potential functions for describing all kinds of system's interactions. In general, the potential functions employed in the simulations are usually constructed from either experimental data or *ab initio* quantum mechanics calculations, most of which are relied on pairwise additive approximations. To obtain a correct description

of ions solvated in aqueous solution, it has been demonstrated that "quantum effects" are significant and that the inclusion of these effects in the simulations is mandatory (Rode, Schwenk, and Tongraar, 2004).

In practice, the performance of *ab initio* quantum mechanical calculations for a condensed-phase system consisting of a large number of molecules is still beyond the current computational feasibility. With regard to this point, Car-Parrinello molecular dynamics (CPMD) method (Car and Parrinello, 1985) compromise the computational limitation by the usage of a simple density functional and the reduction of the investigated system to a moderate size of 100-200 atoms. However, a severe limitation of the CPMD scheme, in particular the use of simple density functional and of the relatively small system size, for the treatment of electrolyte solutions is well documented (Schwenk, Loeffler, and Rode, 2001; Rode, Schwenk, and Tongraar, 2004; Yoo, Zeng, and Xantheas, 2009).

Consequently, an alternative approach is to apply a so-called combined quantum mechanical and molecular mechanical (QM/MM) method (Warshel and Levitt, 1976; Singh and Kollman, 1986). On the basis of QM/MM technique, the most important part of the system is described as accurately as needed using quantum mechanics (QM), while the rest of the system is handled by molecular mechanics (MM) with an appropriate force field. In recent years, several hybrid QM/MM models have been proposed and successfully applied to various condensed-phase systems (Field, Bash, and Karplus, 1990; Bernardi, Olivucci, and Robb, 1992; Aaqvist and Warshel, 1993; Stanton, Hartspugh, and Merz, 1995; Gao, 1996; Cummins and Gready, 1997). Particularly, the QM/MM MD technique has been applied to various ions in solutions, providing many new insights into the solvation structure and

dynamics of these solvated ions (Gao, 1996; Tongraar, Liedl, and Rode, 1998; Monard and Merz, 1999; Orozco and Luque, 2001; Rode, Schwenk, and Tongraar, 2004).

In the present study, the characteristics of carboxylate ions (RCOO⁻), in particular formate (HCOO⁻) and acetate (CH₃COO⁻), in aqueous solution is of our interest. The role of RCOO⁻ in biology is wide-spread (Andrea, Cavalieri, Goldfine, and Jorgensen, 1980; Beveridge and Heywood, 1993). In many surfactant molecules, it is known that RCOO⁻ constitutes a polar head group, in which its interaction with water is important for the formation of lamella phases, vesicles cell and micelles (Shah, 1998; Ivanov and Klein, 2002; Leung and Rempe, 2004). According to NMR experiments, it has been reported that an average of 5.0-6.5 water molecules is observed for each of carboxylate group (Kuntz, 1971). For HCOO⁻, the simplest species containing RCOO⁻, it has been inferred that the hydration number is below 2.5 per HCOO⁻ oxygen atom for HCOO⁻ concentrations of 15 mole% and lower (Kameda, Fukuhara, Mochiduki, Naganuma, Usuki, and Uemura, 2002; Kameda, Sasaki, Yaegashi, Tsuji, Oomori, Hino, and Usuki, 2004).

In terms of theoretical investigations, MC simulations with "optimized potentials for liquid simulations (OPLS)" empirical force fields have been performed, employing a rigid model for HCOO⁻ and CH₃COO⁻ and a united atom for –CH and –CH₃ groups. Surprisingly, it has been shown that the aqueous CH₃COO⁻ exhibits correlation functions very similar to that of aqueous HCOO⁻ (Alagona, Ghio, and Kollman, 1986; Jorgensen and Gao, 1986). Later, MD simulation with SPC/E water models has been carried out for aqueous solution of CH₃COO⁻, revealing similar results with those obtained by the MC-OPLS technique (Meng and Kollman, 1996).

Next, MC simulations based on a combined quantum mechanical and molecular mechanical approach have been applied to CH_3COO^- in several aqueous solutions (Cubero, Luque, Orozco, and Gao, 2003), providing data similar to those obtained from MC-OPLS and MD-SPC/E simulations (Cubero, Luque, Orozco, and Gao, 2003). Recently, CPMD simulations of aqueous $HCOO^-$ have been carried out, showing that the structural details differed significantly from the previous simulation studies. With regard to the CPMD simulations, however, it should be realized that the system's size under investigation is relatively small, consisting of only 53 water molecules, and the simple BLYP exchange correlation functional is employed. Recently, it has been reported that the simple DFT functional, like BLYP and PBE usually employed in the CPMD scheme are even not capable to describe the solvent water itself. For example, the use of these functionals leads to a "supercooled" state rather than a liquid at room temperature (Yoo, Zeng, and Xantheas, 2009).

As a consequence of the discrepancies found in the previous simulations, more sophisticated simulation technique is apparently required. In this work, it is of particular interest, therefore, to apply the high-level QM/MM technique in order to obtain more reliable description of $HCOO^-$ and CH_3COO^- in aqueous solution.

1.2   Research objectives

1. To test the validity and applicability of the high-level QM/MM MD technique for the treatment of various composite species in aqueous solution.

2. To investigate the structural and dynamical properties of carboxylate ions ($HCOO^-$ and CH_3COO^-) in water using the QM/MM MD technique.

1.3   Scope and limitation of the study

This work focuses on the use of combined *ab initio* QM/MM MD technique for elucidating the structural and dynamical properties of carboxylate ions, in particular $HCOO^-$ and CH_3COO^-, in aqueous solution. By means of the QM/MM MD approach, the most interesting region, the sphere which includes the ion and its surrounding water molecules, is treated at Hatree-Fock (HF) and B3LYP hybrid density function theory (DFT) levels of accuracy using DZV+ basis set, while the rest of the system is described by classical pair potentials. With regard to the computational expense for QM force calculations, the QM size and the basis set employed in this study is considered to be suitable, compromising between the quality of the simulation results and the requirement of CPU time. In this work, the use of basis set augmented with diffuse functions is necessary, in particular for a correct treatment of polarizable systems. The pair potential functions for describing $HCOO^-$-water and CH_3COO^--water interactions were newly constructed. For the description of bulk water, a flexible BJH-CF2 model, which describes intermolecular (Stillinger and Rahman, 1978) and intramolecular (Bopp, Jancso, and Heinzinger, 1983) interactions, is employed. The structural properties of $HCOO^-$ and CH_3COO^- in aqueous solution will be characterized through radial distribution functions (RDFs) and their corresponding integration numbers, together with detailed analysis on angular distribution functions (ADFs) and orientations of water molecules surrounding the ions. The corresponding dynamics details will be analyzed with regards to mean residence times (MRTs) of water molecules as well as to water exchange processes at the ions. The results obtained by the HF/MM and B3LYP/MM simulations will be

discussed and compared to the previous theoretical studies as well as to the available experimental data.

1.4   References

Åqvist, J. and Warshel, A. (1993). Simulation of enzyme reactions using valence bond force fields and other hybrid quantum/classical approaches. **Chemical Reviews**. 93: 2523-2544.

Alagona, G., Ghio, C. and Kollman, P. (1986). Monte Carlo simulation studies of the solvation of ions. 1. Acetate anion and methylammonium cation. **Journal of the American Chemical Society**. 108: 185-191.

Andrea, T. A., Cavalieri, R. R., Goldfine, I. D. and Jorgensen, E. C. (1980). Binding of thyroid hormones and analogues to the human plasma protein prealbumin. **Biochemistry**. 19: 55-63.

Bernardi, F., Olivucci, M. and Robb, M. A. (1992). Simulation of MC-SCF results on covalent organic multi-bond reactions: Molecular mechanics with valence bond (MM-VB). **Journal of the American Chemical Society**. 114: 1606-1616.

Beveridge, A. J. and Heywood, G. C. (1993). A quantum mechanical study of the active site of aspartic proteinases. **Biochemistry**. 32: 3325-3333.

Bopp, P., Jancso, G. and Heinzinger, K. (1983). An improved potential for non-rigid water-molecules in the liquid-phase. **Chemical Physics Letters**. 98: 129-133.

Car, R. and Parrinello, M. (1985). Unified approach for molecular dynamics and density-functional theory. **Physical Review Letters**. 55: 2471-2474.

Cowan, J. A. (1998). Metal activation of enzymes in nucleic acid biochemistry. **Chemical Reviews**. 98 1067-1088.

Cubero, E., Luque, F. J., Orozco, M. and Gao, J. (2003). Perturbation approach to combined QM/MM simulation of solute−solvent interactions in solution. **The Journal of Physical Chemistry B**. 107: 1664-1671.

Cummins, P. L. and Gready, J. E. (1997). Coupled semiempirical molecular orbital and molecular mechanics model (QM/MM) for organic molecules in aqueous solution. **Journal of Computational Chemistry**. 18: 1496-1512.

Field, M. J., Bash, P. A. and Karplus, M. (1990). A combined quantum-mechanical and molecular mechanical potential for molecular-dynamics simulations **Journal of Computational Chemistry**. 11: 700-733.

Frank, H. (1956). **Chemical physics of ionic solutions**. New York: John Wiley & Sons.

Gao, J. (1996). A theoretical investigation of the enol content of acetic acid and the acetate ion in aqueous solution. **Journal of Molecular Structure : THEOCHEM**. 370: 203-208.

Ivanov, I. and Klein, M. L. (2002). Deprotonation of a histidine residue in aqueous solution using constrained *ab initio* molecular dynamics. **Journal of the American Chemical Society**. 124: 13380-13381.

Jorgensen, W. J. and Gao, J. (1986). Monte Carlo simulations of the hydration of ammonium and carboxylate ions. **The Journal of Physical Chemistry**. 90: 2174-2182.

Kameda, Y., Fukuhara, K., Mochiduki, K., Naganuma, H., Usuki, T. and Uemura, O. (2002). Structure of HCOOK hydrated melts. **Journal of Non-Crystalline Solids**. 312-314: 433-437.

Kameda, Y., Sasaki, M., Yaegashi, M., Tsuji, K., Oomori, S., Hino, S. and Usuki, T. (2004). Hydration structure around the carboxyl group studied by neutron diffraction with C-12/C-13 and H/D isotopic substitution methods. **Journal of Solution Chemistry**. 33: 733-745.

Kuntz, I. D. J. (1971). Hydration of macromolecules. III. Hydration of polypeptides. **Journal of the American Chemical Society**. 93: 514-516.

Leung, K. and Rempe, S. B. (2004). *Ab initio* molecular dynamics study of formate ion hydration. **Journal of the American Chemical Society**. 126: 344-351.

Meng, E. C. and Kollman, P. A. (1996). Molecular dynamics studies of the properties of water around simple organic solutes. **The Journal of Physical Chemistry**. 100: 11460-11470.

Monard, G. and Merz, K. M. J. (1999). Combined quantum mechanical/molecular mechanical methodologies applied to biomolecular systems. **Accounts of Chemical Research**. 32: 904-911.

Ohtaki, H. and Radnai, T. (1993). Structure and dynamics of hydrated ions. **Chemical Reviews**. 93: 1157-1204.

Orozco, M. and Luque, F. J. (2001). Theoretical methods for the description of the solvent effect in biomolecular systems. **Chemical Reviews**. 101: 203-203.

Rode, B. M., Schwenk, C. F. and Tongraar, A. (2004). Structure and dynamics of hydrated ions-new insights through quantum mechanical simulations. **Journal of Molecular Liquids**. 110: 105-122.

Schwenk, C. F., Loeffler, H. H. and Rode, B. M. (2001). Molecular dynamics simulations of Ca^{2+} in water: Comparison of a classical simulation including three-body corrections and Born-Oppenheimer *ab initio* and density functional theory quantum mechanical/molecular mechanics simulations. **The Journal of Chemical Physics**. 115: 10808-10813.

Shah, D. O. (1998). **In Micelles, Microemulsions, and Monolayers: Quarter Century Progress at the University of Florida**. Florida: Marcel Dekker.

Singh, U. C. and Kollman, P. A. (1986). A combined *ab initio* quantum mechanical and molecular mechanical method for carrying out simulations on complex molecular systems: Applications to the $CH_3Cl + Cl^-$ exchange reaction and gas phase protonation of polyethers. **Journal of Computational Chemistry**. 7: 718-730.

Stanton, R. V., Hartspugh, D. S. and Merz, K. M., Jr. (1995). An examination of a density functional/molecular mechanical coupled potential. **Journal of Computational Chemistry**. 16: 113-128.

Stillinger, F. H. and Rahman, A. (1978). Revised central force potentials for water. **The Journal of Chemical Physics**. 68: 666-670.

Tongraar, A., Liedl, K. R. and Rode, B. M. (1998). The hydration shell structure of Li^+ investigated by Born–Oppenheimer *ab initio* QM/MM dynamics **Chemical Physics Letters**. 286: 56-64.

Warshel, A. and Levitt, M. (1976). Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme **Journal of Molecular Biology**. 103: 227-249.

Williams, R. J. P. (1971). **Bio-inorganic chemistry**. Washington DC: American Chemical Society.

Yoo, S., Zeng, X. C. and Xantheas, S. S. (2009). On the phase diagram of water with density functional theory potentials: The melting temperature of ice I_h with the Perdew--Burke--Ernzerhof and Becke--Lee--Yang--Parr functionals. **The Journal of Chemical Physics**. 130: 221102-221104.

CHAPTER II

QUANTUM MECHANICS IN CHEMISTRY

Quantum chemistry is based on quantum mechanical principles, defined by the mathematical descriptions of chemistry. Regarding to quantum chemistry, fundamental behavior of matter at molecular scale, in particular the electron behavior, is obtainable. A wavefunction, obtained by solving of *Schrödinger equation* (Schrödinger, 1926), is the tool of quantum chemistry for describing the properties of matter in terms of energies and positions of the nuclei and electrons. The applications of quantum chemistry include solving many chemical problems, especially understanding of chemical bonding, spectral phenomena, molecular reactivities and various other fundamental chemical problems. However, only simple chemical systems that can be determined through the purely quantum chemistry terms. For complicated systems, the simplification of quantum mechanics, such as Hartree-Fock (HF) or density functional theory (DFT), are mostly applied for the convenient investigations.

2.1   Schrödinger equation

The complete description of a wavefunction can be given through the solution of the Schrödinger equation, which describing the atom system. Schrödinger obtained an equation by taking the classical time-independent wavefunction equation and substituting *de Brogile's relation* for λ. Thus, if

$$\nabla^2 \psi = -(2\pi / \lambda)^2 \psi \qquad (2.1)$$

and

$$\lambda = \frac{h}{\sqrt{2m(E-V)}} \qquad (2.2)$$

then

$$[-(h^2 / 8\pi^2 m)\nabla^2 + V(x, y, z)\psi(x, y, z)] = E\psi(x, y, z), \qquad (2.3)$$

representing the *Schrödinger's time-independent* wave equation for a single particle of the mass (*m*) moving in the three-dimensional potential field (*V*). The left-hand side of the equation is called the *Hamiltonian operator* (*H*),

$$H \equiv -(h^2 / 8\pi^2 m)\nabla^2 + V, \qquad (2.4)$$

which is often written as

$$H\psi = E\psi. \qquad (2.5)$$

The Hamiltonian operator takes into account five contributions to the total energy of the system, namely the kinetic energies of the electrons and nuclei, the

attraction of the electrons to the nuclei and the interelectronic and internuclear repulsions that would be of the form

$$H = -\sum_i \frac{\hbar^2}{2m_e}\nabla_i^2 - \sum_k \frac{\hbar^2}{2m_k}\nabla_k^2 - \sum_i\sum_k \frac{e^2 Z_k}{r_{ik}} + \sum_{i<j}\frac{e^2}{r_{ij}} + \sum_{k<l}\frac{e^2 Z_k Z_l}{r_{kl}}, \qquad (2.6)$$

where i and j defined to electrons, k and l run over nuclei, $\hbar$ is *Plank's constant* divided by 2π, m_e is the mass of the electron, m_k is the mass of nucleus, e is the electron charge, Z is an atomic number, r_{ab} is the distance between particles a and b and ∇^2 is the *Laplacian operator*, which can be defined as

$$\nabla^2 \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}. \qquad (2.7)$$

2.2 The variation theory

In general, the energy of the system can be calculated through equation (2.1) by operating the Hamiltonian operator on the wavefunction. The useful method for estimating the lowest energy is based on *variation theory*, which assessing and improving guesses about the forms of wavefunctions in systems. The theory starts with a trial function (Φ), which can be written in terms of a linear combination of the wavefunctions (ψ),

$$\Phi = \sum_i c_i \psi_i, \qquad (2.8)$$

where the individual ψ_i and coefficients c_i are unknown. Then, the normality of Φ

imposes a constraint on the coefficients, deriving from

$$\int \Phi^2 dr = 1 = \int \sum_i c_i \psi_i \sum_j c_j \psi_j dr$$

$$= \sum_{ij} c_i c_j \int \psi_i \psi_j dr$$

$$= \sum_{ij} c_i c_j \delta_{ij}$$

$$= \sum_i c_i^2. \tag{2.9}$$

Then, considering the energy associated with the wavefunction (Φ) as

$$\int \Phi H \Phi dr = \int (\sum_i c_i \psi_i) H (\sum_j c_j \psi_j) dr$$

$$= \sum_{ij} c_i c_j \int \psi_i H \psi_j dr$$

$$= \sum_{ij} c_i c_j E_j \delta_{ij}$$

$$= \sum_i c_i^2 E_i. \tag{2.10}$$

After that, combining the results from equation (2.9) and (2.10), give

$$\int \Phi H \Phi dr - E_0 \int \Phi^2 dr = \sum_i c_i^2 (E_i - E_0). \tag{2.11}$$

The coefficients have been assumed to be real number, thus, c_i^2 and the result of $(E_i - E_0)$ must be greater than or equal to zero. Therefore,

$$\int \Phi H \Phi dr - E_0 \int \Phi^2 dr \geq 0 \qquad (2.12)$$

or

$$\frac{\int \Phi H \Phi dr}{\int \Phi^2 dr} \geq E_0. \qquad (2.13)$$

According to equation (2.13), the quality of wavefunctions for describing the ground state of a system can be defined by their associated energies as the better wavefunction could provide the lower energy. Moreover, the guess of the trial wavefunction can be constructed in any manner, which determined the quality by the integral in equation (2.13).

2.3   Born-Oppenheimer approximation

The difficulty for solving the Schrödinger equation appears in the many-particle molecular systems involving correlated motions of particles. In fact, the nuclei are heavier than electrons, thus, the nuclei are moving slowly than the electrons. According to this property, the approximation has been made by separating the nuclei and electrons motions, called *Born-Oppenheimer approximation*.

Based on Born-Oppenheimer approximation, the electronic energies are computed by fixing nuclear position. Consequently, the nuclear kinetic energy term is independent, and thus, can be neglected. The attractive electron-nuclear potential energy term is eliminated and the repulsive nuclear-nuclear potential energy term can be considered to be constant. Thus, the electronic Schrödinger equation can be defined as

$$(H_{el} + V_N)\Psi_{el}(q_i; q_k) = E_{el}\Psi_{el}(q_i; q_k), \qquad (2.14)$$

where the subscript *el* refer to the Born-Oppenheimer approximation, H_{el} includes only the first, third and fourth terms on the equation (2.6) as

$$H_{el} = -\sum_i \frac{\hbar^2}{2m_e}\nabla_i^2 - \sum_i \sum_k \frac{e^2 Z_k}{r_{ik}} + \sum_{i<j} \frac{e^2}{r_{ij}}, \qquad (2.15)$$

where V_N is the nuclear-nuclear repulsion energy, q_i is the electronic coordinates and q_k is the nuclear coordinates. The eigenvalue of the electronic Schrödinger equation is called the *electronic energy* (E_{el}). Since the term V_N is a constant for a given set of fixed nuclear coordinates, the wavefunction can be solved without the inclusion of V_N. In this respect, the eigenvalue is called the *pure electronic energy*. The term V_N can be added to this eigenvalue in order to obtain the total electronic and nuclear-nuclear repulsion energy.

2.4   Molecular orbital theory

The molecular orbital theory is a method for determining molecular structure. A molecular orbital is a region in which an electron may be found in a molecule. The molecular orbital can be described by the wavefunction of the electron in a molecule, in particular a spatial distribution ($|\psi_i(r)|^2$) of an electron and energy of up to two electrons within it. The complete wavefunction for an electron is consist of a molecular orbital and a spin function (α and β), which can be defined as a *spin orbital* ($\chi(x)$) where x indicates both space and spin coordinates. Therefore, a spatial orbital can be formed into two different spin orbitals as

$$\chi(x) = \begin{cases} \psi(r)\alpha(\omega) \\ \\ \psi(r)\beta(\omega). \end{cases} \tag{2.16}$$

For *N*-electron wavefunction, the Hamiltonian of the simpler system, which contained noninteracting electrons, can be defined as

$$H = \sum_{i=1}^{N} h(i), \tag{2.17}$$

where *h(i)* is the operator that describes the kinetic energy and potential energy of electron *i*. Then, the set of spin orbitals ($\chi_j(x)$) have been added to the operator, which presented in equation (2.18),

$$h(i)\chi_j(x_i) = \varepsilon_j \chi_j(x_i). \qquad (2.18)$$

Therefore, the wavefunction is a simple product of spin orbital wavefunction for each electron as

$$\Psi^{HP}(x_1, x_2, ..., x_N) = \chi_i(x_1)\chi_j(x_2)\cdots\chi_k(x_N). \qquad (2.19)$$

Thus, the equation (2.5) can be written as

$$H\Psi^{HP} = E\Psi^{HP}, \qquad (2.20)$$

where E is the sum of the spin orbital energies of each spin orbital in Ψ^{HP}, as

$$E = \varepsilon_i + \varepsilon_j + \cdots + \varepsilon_k. \qquad (2.21)$$

Accordingly, a N-electron wavefunction is termed a *Hartree product*, where the electron-one has been described by the spin orbital (χ_i), electron-two has been described by the spin orbital (χ_j), etc. However, this wavefunction does not allow the antisymmetry principle.

To satisfy correcting the antisymmetry principle, considering a two-electron case in order to put electron-one in χ_i and electron-two in χ_j as

$$\Psi_{12}^{HP}(x_1, x_2) = \chi_i(x_1)\chi_j(x_2). \tag{2.22}$$

In the opposite way, putting electron-one in χ_j and electron-two in χ_i as

$$\Psi_{21}^{HP}(x_1, x_2) = \chi_i(x_2)\chi_j(x_1). \tag{2.23}$$

After that, taking the appropriate linear combination of these two Hartree products,

$$\Psi(x_1, x_2) = 2^{-1/2}(\chi_i(x_1)\chi_j(x_2) - \chi_i(x_2)\chi_j(x_1)), \tag{2.24}$$

where the factor $2^{-1/2}$ is a normalization factor and the minus sign insures that $\Psi(x_1, x_2)$ is antisymmetric with respect to the interchange of the coordinates of electrons one and two. From equation (2.24), the wavefunction disappears if both electrons occupy the same spin orbital, *i.e.*, following the *Pauli exclusion principle*. Moreover, the antisymmetric wavefunction can be rewritten in terms of a determinant,

$$\Psi(x_1, x_2) = 2^{-1/2}\begin{vmatrix} \chi_i(x_1) & \chi_j(x_1) \\ \chi_i(x_2) & \chi_j(x_2) \end{vmatrix}, \tag{2.25}$$

which is called a *Slater determinant* (Slater, 1929). For an *N*-electron system, the generalization is

$$\Psi(x_1, x_2, ..., x_N) = (N!)^{-1/2} \begin{vmatrix} \chi_i(x_1) & \chi_j(x_1) & \cdots & \chi_k(x_1) \\ \chi_i(x_2) & \chi_j(x_2) & \cdots & \chi_k(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_i(x_N) & \chi_j(x_N) & \cdots & \chi_k(x_N) \end{vmatrix}, \qquad (2.26)$$

where the factor $(N!)^{-1/2}$ is the normalization factor.

2.5   The LCAO approach and basis sets

The molecular orbitals can be built from the atomic orbitals by using the *linear combination of atomic orbitals to molecular orbitals* (LCAO-MO) method. The relation can be written as

$$\psi_i = \sum_{\mu=1}^{N} C_{\mu i} \phi_\mu, \qquad (2.27)$$

where $C_{\mu i}$ are the molecular orbital expansion coefficients, N is the number of atomic basis function and the set of N function ϕ_μ is called *basis set*.

The common types of basis function, as also called *atomic orbital*, used in electronic structure calculations are *Slater-type orbitals* (STOs) (Slater, 1930) and *Gaussian-type orbitals* (GTOs) (Boys, 1950).

For the STOs, they are constructed as

$$\psi(n, l, m_l; r, \theta, \phi) = N r^{n_{eff} - 1} e^{-Z_{eff} \rho / n_{eff}} Y_{lm_l}(\theta, \phi), \qquad (2.28)$$

where *n*, *l*, and m_l are the quantum numbers, *N* is the normalization constant and Y_{lm_l} is a spherical harmonic. The exponential dependence on the distance between the nucleus and electron mirrors the exact orbitals for the hydrogen atom, where Z_{eff} is the effective nuclear charge in which the effective principal quantum number (n_{eff}) is related to the true principal quantum (*n*) by the following mapping as

$$n \rightarrow n_{eff} : 1 \rightarrow 1 \quad 2 \rightarrow 2 \quad 3 \rightarrow 3 \quad 4 \rightarrow 3.7 \quad 5 \rightarrow 4.0 \quad 6 \rightarrow 4.2,$$

and the value of ρ equal to r/a_0, where a_0 is *Bohr radius*.

The STOs are usually applied for atomic and diatomic systems, which high accuracy, as well as in semi-empirical methods, where all three- and four-center integrals are neglected. In density functional methods, exact exchange is not included and the coulomb energy is calculated by fitting the density to a set of auxiliary functions. However, the STOs do not satisfy in two-electron integral problem. The feasible basis function is GTOs, which are function of the form

$$\theta_{ijk}(r_1 - r_c) = (x_1 - x_c)^i (y_1 - y_c)^j (z_1 - z_c)^k e^{-\alpha|r_1 - r_c|^2}, \qquad (2.29)$$

where (x_c, y_c, z_c) are the Cartesian coordinates of the center of the Gaussian function at r_c, (x_1, y_1, z_1) are the Cartesian coordinates of an electron at r_1, *i, j* and *k* are non-negative integers and α is a positive exponent. The advantage of GTOs is that the product of two Gaussians at different centers is equivalent to a single Gaussian function centered at a point between the two centers. Therefore, the two-electron

integral problem on three and four or more different atomic centers can be reduced to integrals over two different centers. However, the GTO gives an inferior representation of the orbitals at the atomic nuclei, which can be considered at 1s-orbital. A 1s-orbital of STO has a *cusp* at the atomic nucleus but a GTO does not, as can be seen in Figure 2.1. In this respect, the larger basis must be used to achieve an accuracy comparable to that obtained from STOs.

$exp(-\alpha r)$ $exp(-\alpha r^2)$

Slater-type 1s orbital   **Gaussian-type 1s orbital**
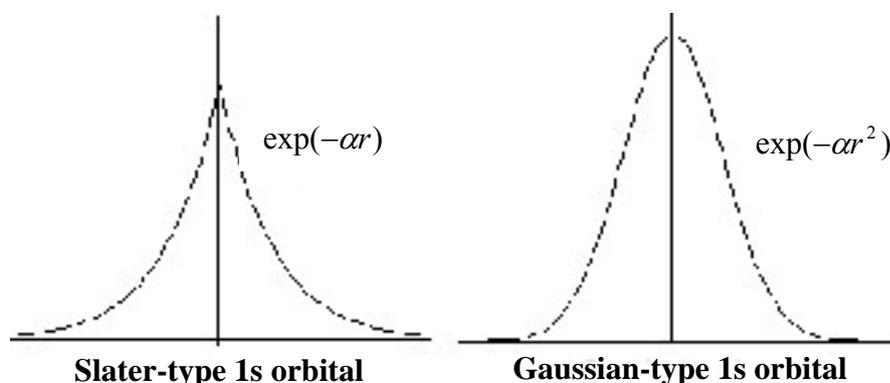
Figure 2.1 The Slater-type and Gaussian-type for 1s orbital.

The most important factor for creating the molecular orbital is the set of parameters when applied to the basis function, called *basis set*. The smallest number of function possible is a *minimum basis set*. The improvement of the basis set can be achieved by replacing two basis functions into each basis function in the minimal basis set, called *double zeta* (DZ). Accordingly, a *triple zeta* (TZ) refers to three basis functions that are used to represent each of the minimal basis sets. The compromise between the DZ and TZ basis sets is called a *split valence* (SV) basis set, in which each valence atomic orbital is represented by two basis functions whereas each core orbital is represented by a single basis function.

In 1969, Pople and coworkers (Hehre, Stewart, and Pople, 1969) designed the basis set by expanding the STO in terms of n primitive Gaussians, called the STO-nG basis set. The primitive Gaussian has been derived for n = 2-6. However, the STO-3G basis set is a widely used minimum basis set, which can be seen in Figure 2.2. The STO-3G basis set takes into account the *cusp* of s-type orbital, which is quite reasonable for representing the orbitals at the atomic nuclei.

STO-3G 1s Basis function
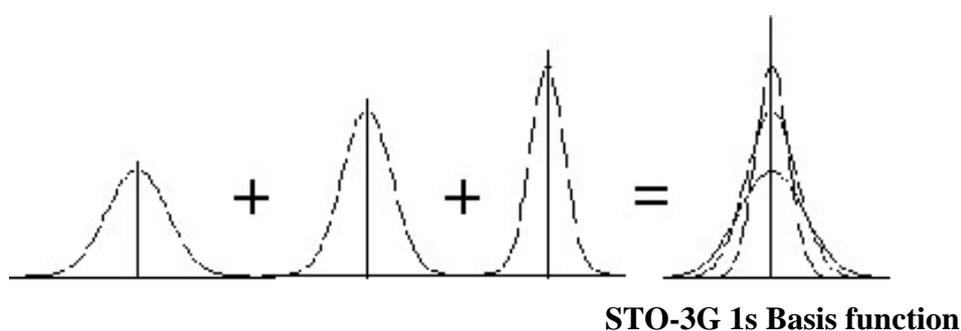
Figure 2.2 The STO-3G basis set representing the desired STO.

In addition, Pople and coworkers applied the split valence for increasing flexibility in the basis set, which can be designed to *k-nlm*G basis set. The first parameter (*k*) indicates the number of primitives used in the contracted core. Two values (*nl*) refer to a split valence, while three values (*nlm*) refer to a triple split valence such as 6-311G that is a triple split valence basis, where the core orbital is a contraction of six primitives and the valence split into three functions, represented by three, one and one primitive GTOs, respectively. The Pople style basis sets can be augmented by diffuse and/or polarization functions. The diffuse function can be denoted by + or ++ before the G, where the first + indicating one set of diffuse s- and p-function on heavy atoms and the second + refers to a diffuse s-function adding to hydrogen. The polarization function can be added after the G, separating for heavy

atom and hydrogen. For example, 6-31+G(d) basis set refers to a split valence with additional diffuse sp-function and a single d-type polarization function on heavy atoms. The largest standard Pople style basis set is 6-311++G(3df,3pd). Alternatively, the polarization function can be replaced with notation *. For example, the 6-311G* basis set is identical to 6-311G(d) and 6-311G** basis set is identical to 6-311G(d,p).

Furthermore, the *contracted Gaussian function* has been applied to *Dunning-Huzinaga* (DZ) basis set (Huzinaga, 1965; Dunning, 1970; Dunning, 1971). The DZ basis set can be made by a contraction such as the (9s5p) primitive GTO to [4s, 2p]. The contraction scheme is 6,1,1,1 for s-functions and 4,1 for the p-functions. In addition, the development of basis set by Dunning and coworkers for recovering the correlation energy of the valence electrons is named as *correlation consistent* (cc) basis sets. A general form can be written as cc-pVnZ, where n = D for double zeta, T for triple zeta, Q for quadruple zeta, 5 for quintuple zeta or 6 for sextuple zeta. The diffuse functions can be augmented into the correlation consistent basis set by adding the prefix aug-, such as aug-cc-pVnZ. The correlation consistent basis sets provide accurate description of anions and weak interactions, in particular for systems involving van der Waals forces and hydrogen bonding.

For the systems having large number of core electrons elements, it is necessary to use a large number of basis functions for describing them. However, the deep core electrons do not important in a chemical sense. This problem can be solved by replacing the core electrons with analytical functions, called an *effective core Potential* (ECP) or *pseudopotential* (PP), which would reasonably accurately, and much more efficiently, representing the combined nuclear-electronic core to the remaining electrons.

2.6   Hartree-Fock method

The important factor in the electronic structure calculations is the electron-electron repulsions, which must be included in any accurate electronic structure treatment. The *Hartree-Fock* (HF) method treats the electron-electron repulsions in an average way. The HF equation for spin orbital (ϕ_a), which assigning electron 1 to spin orbital (ϕ_a), is

$$f_1\phi_a(1) = \varepsilon_a\phi_a(1), \tag{2.30}$$

where ε_a is the spin orbital energy and f_1 is the *Fock operator*. The f_1 can be defined as

$$f_1 = h_1 + \sum_u \{J_u(1) - K_u(1)\}, \tag{2.31}$$

where h_1 is the core Hamiltonian for electron 1, the sum is over all spin orbital $u = a, b, \ldots, z$, J_u is the Coulomb operator and K_u is the exchange operator. The J_u and K_u operator can be defined as

$$J_u(1)\phi_a(1) = j_0\left\{\int \phi_u^*(2)\frac{1}{r_{12}}\phi_u(2)dx_2\right\}\phi_a(1), \tag{2.32}$$

and

$$K_u(1)\phi_a(1) = j_0 \left\{ \int \phi_u^*(2) \frac{1}{r_{12}} \phi_a(2)dx_2 \right\} \phi_u(1), \qquad (2.33)$$

where j_0 is $\dfrac{e^2}{4\pi\varepsilon_0}$.

The Coulomb operator takes into account the coulombic repulsion and the exchange operator represents the modification of this energy that can be ascribed to the effect of spin correlation. In equation (2.31), the sum represents the average potential energy of electron 1 due to the presence of the other n-1 electrons. Since the Fock operator depends on the spin orbitals of all the other n-1 electrons, the HF method must already know the solution beforehand, thus, the iterative style of solution has been carried out and stopping when the solution is self-consistent, as called *self-consistent field* (SCF). The self-consistent is started with a trial set of spin orbitals and used to construct the Fock operator. After that, the HF equation is solved to obtain the new set of spin orbitals, which are used to construct a revised Fock operator, and so on. The calculation is repeated until a convergence criterion is satisfied.

In addition, the HF equations for the spin orbitals in equation (2.31)-(2.33) are converted to the set of spatial eigenvalue equation by integration over the spin functions and using the orthonormality of α and β, thus, the HF equation of spatial wavefunction (ψ_s) is occupied by electron 1 as

$$\left\{ h_1 + \sum_r (2J_r - K_r) \right\} \psi_s(1) = \varepsilon_s \psi_s(1), \qquad (2.34)$$

the Coulumb and exchange operators can be defined as follows

$$J_r \psi_s(1) = \left\{ \int \psi_r^*(2) \left(\frac{e^2}{4\pi\varepsilon_0 r_{12}} \right) \psi_r(2) d\tau_2 \right\} \psi_s(1), \qquad (2.35)$$

and

$$K_r \psi_s(1) = \left\{ \int \psi_r^*(2) \left(\frac{e^2}{4\pi\varepsilon_0 r_{12}} \right) \psi_s(2) d\tau_2 \right\} \psi_r(1). \qquad (2.36)$$

In the HF-SCF calculations on closed- shell states of atoms, each spin orbitals are contained a pair of electrons, which are $\frac{1}{2}n$ spatial orbitals of the form $\psi_a(r_1)$ where n is the number of electrons, and the HF ground-state wavefunction (Φ_0) is

$$\Phi_0 = (n!)^{-1/2} \det \left| \psi_a^\alpha(1)\psi_a^\beta(2)\psi_b^\alpha(3)\ldots\psi_z^\alpha(n) \right|. \qquad (2.37)$$

This wavefunction is called a *restricted Hartree-Fock* (RHF) wavefunction.

For open-shell states of atoms, all electrons except those occupying open-shell orbitals are doubly occupied in spatial orbitals, called the *restricted open-shell formalism* such as the restricted open-shell wavefunction for atomic lithium (Li) would be of the form

$$\Phi_0 = (6)^{-1/2} \det\left|\psi_{1s}^{\alpha}(1)\psi_{1s}^{\beta}(2)\psi_{2s}^{\alpha}(3)\right|. \tag{2.38}$$

However, 1s α-spin orbital has an exchange interaction with 2s α-spin orbital, thus, the variational ground state energy is not accurate. For the *unrestricted open-shell Hartree-Fock* (UHF) formalism, the two 1s-electrons are not constrained to the same spatial wavefunction. For example, the UHF wavefunction for Li would be of the form

$$\Phi_0 = (6)^{-1/2} \det\left|\psi_a^{\alpha}(1)\psi_b^{\beta}(2)\psi_c^{\alpha}(3)\right|. \tag{2.39}$$

With regard to the unconstraint of occupying two electrons in the same spin orbital, the open-shell UHF formalism gives a lower variational energy than the open-shell RHF formalism. However, the total spin angular momentum is not well-defined for a UHF wavefunction.

The HF equation can be solved numerically for atom, but complicated for molecules. For expanding the HF equation, a known set of basis function has been used to expand the spin orbitals. The procedure starts from occupying electron 1 into the spatial function ($\psi_a(1)$), which can be written as

$$f_1\psi_a(1) = \varepsilon_a\psi_a(1). \tag{2.40}$$

The Fock operator in equation (2.40) can be expressed in terms of the spatial wavefunction as

$$f_1 = h_1 + \sum_u \left\{ 2J_u(1) - K_u(1) \right\}. \qquad (2.41)$$

Then, each spatial wavefunction (ψ_i) can be expanded to a linear combination of a set of M basis function θ_j as

$$\psi_i = \sum_{j=1}^{M} c_{ji}\theta_j, \qquad (2.42)$$

where c_{ji} are unknown coefficients. Substituting equation (2.42) into equation (2.40), gives

$$f_1 \sum_{j=1}^{M} c_{ja}\theta_j(1) = \varepsilon_a \sum_{j=1}^{M} c_{ja}\theta_j(1). \qquad (2.43)$$

Then, multiplication of both sides of equation (2.43) by $\theta_i^*(1)$ and integration over r_1 as

$$\sum_{j=1}^{M} c_{ja} \int \theta_i^*(1) f_1 \theta_j(1) dr_1 = \varepsilon_a \sum_{j=1}^{M} c_{ja} \int \theta_i^*(1) f_1 \theta_j(1) dr_1. \qquad (2.44)$$

The *overlap matrix* (S) can be defined as

$$S_{ij} = \int \theta_i^*(1)\theta_j(1) dr_1 \qquad (2.45)$$

and the *Fock matrix* (F) can be of the form

$$F_{ij} = \int \theta_i^*(1) f_1 \theta_j(1) dr_1. \tag{2.46}$$

Then, equation (2.44) becomes

$$\sum_{j=1}^{M} F_{ij} c_{ja} = \varepsilon_a \sum_{j=1}^{M} S_{ij} c_{ja}. \tag{2.47}$$

This expansion is know as the *Roothaan equation*, which can be written as the single matrix equation

$$Fc = Sc\varepsilon, \tag{2.48}$$

where c is an M x M matrix composed of elements c_{ja} and ε is an M x M diagonal matrix of the orbital energies ε_a.

According to the properties of matrix equations, the Roothaan equations have a non-trivial solution until this secular equation is satisfied,

$$\det|F - \varepsilon_a S| = 0. \tag{2.49}$$

However, the matrix elements F_{ji} involve integral over the Coulomb and exchange operators, whose depend on the spatial wavefunctions. Therefore, the self-

consistent field approach must be adopted in order to obtain a new set of coefficient c_{ja} and the procedure is continued until a convergence criterion has been reached (see Figure 2.3)

In general, the SCF calculation produces the different energy value, depending on the basis set. For example, using a minimal basis set yields a total electronic energy E1. The energy E1 can be improved by choosing a new basis set like a double zeta basis, *i.e.*, to compute the lower energy E2. Moreover, the polarization function can be added into the basis set to give the lower energy E3. On the other hand, the expansion of the basis set will decrease the total electronic energy. Nevertheless, since the basis sets used in the calculations are finite, thus, the energy will approach a limiting value. This limiting energy is called a *Hartree-Fock limit*. The molecular orbitals that correspond to this limit are called *Hartree-Fock orbitals* (HF orbitals) and the determinant is called the HF wavefunction.
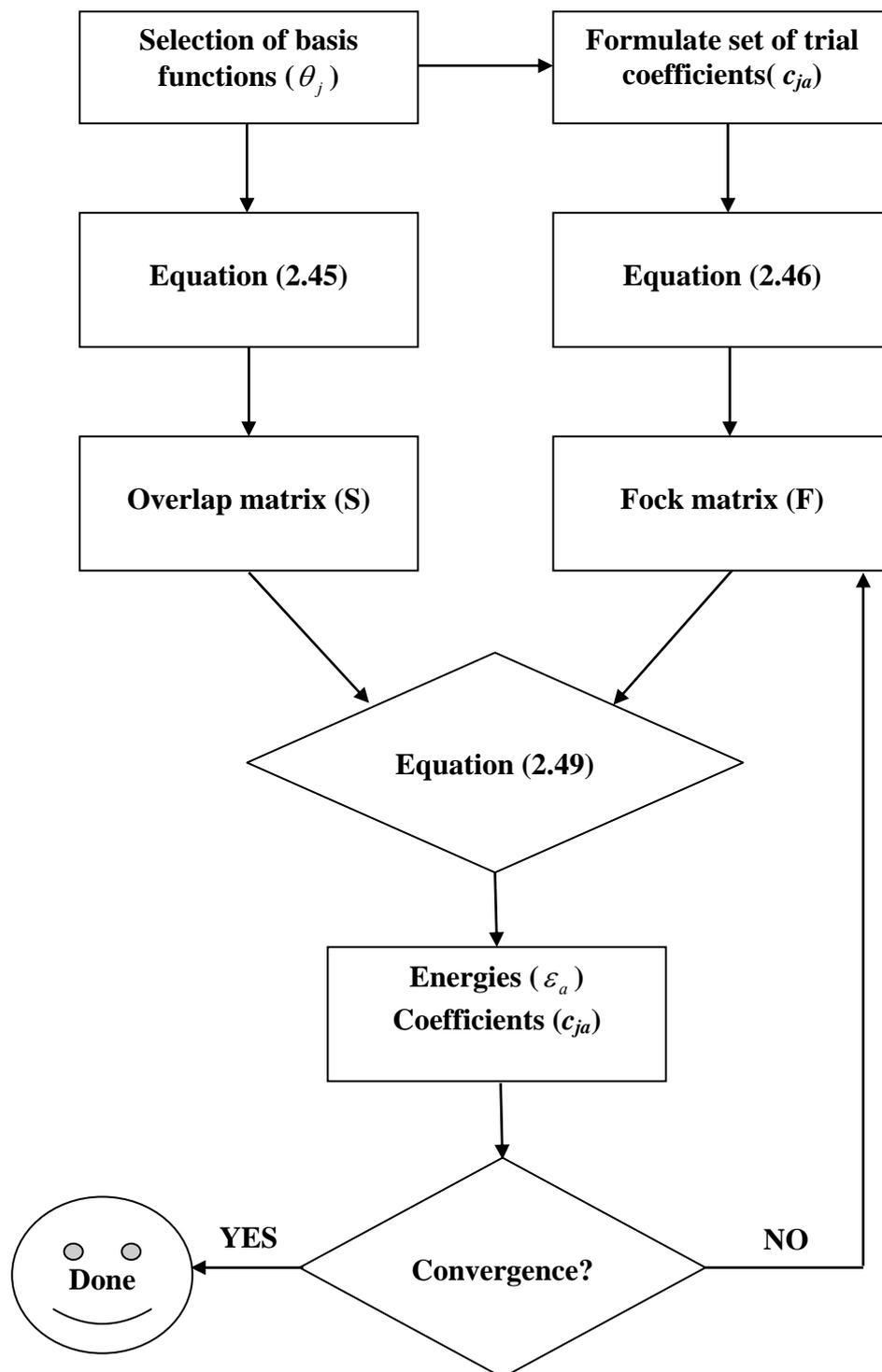
Figure 2.3 Iteration procedure for Hartee-Fock self-consistent field calculation.

2.7   Electron correlation

The Hartree-Fock limit is not the exact ground-state energy of the molecule, *i.e.*, the exact ground-stat energy must include the effects of electron correlation, especially for the open-shell RHF type of wavefunction where all electrons are paired in molecular orbitals. Because of the orthomormality of molecular orbitals, the spatial overlap between orbitals of two such pair-electrons is one, while the overlap between electrons in different molecular orbitals is zero. However, the meaning of this definition relies on the correlation between pairs of electrons in the same spatial molecular orbital as the main of the electron correlation. In addition, the correlations between opposite spins in both intraorbitals and interorbitals are larger than the correlations between electrons in the same spin. According to the Pauli principle, there is no intraobital correlation from electron pairs within the same spin. The opposite spin correlation is called the *Coulomb correlation*, while the same spin correlation is called the *Fermi correlation*. In the HF calculation, the one-determinant wavefinction has been used as trial wavefunction. To improve the HF calculation, the trial wavefunction must contain more than one Slater determinant. Therefore, the electron in the orbital should be considered in terms of electron density. The generic multi-determinant trial wavefunction can be of the form

$$\Psi = a_0 \Phi_{\text{HF}} + \sum_{i=1} a_i \Phi_i, \qquad (2.50)$$

where a_0 is the coefficient, which usually close to one. The electron correlation methods calculate the coefficient in front of the other determinants in difference way

such as *configuration interaction* (CI) (Sherrill and Schaefer, 1999), *many-body perturbation* (MP) (Møller and Plesset, 1934), *coupled cluster* (CC) (Bartlett, 1989) and *density function theory* (DFT).

In general, a basis with *M* member results in *M* spatial wavefunction and 2*M* different spin orbitals. The HF wavefunction (Φ_0) can be formed by ordering the spin orbital energetically and taking the *n* lowest in energy. However, the molecular orbital are occupied by *n* electrons, while the virtual orbitals are occupied by 2*M-n* electrons. The HF determinant should be used as reference determinant of classification of excited determinant, *i.e.*, to decide how many occupied HF molecule orbitals that will be replaced by unoccupied molecular orbitals. The excited Slater determinant can be written as *singles* (S), *double* (D), *triples* (T), *quadruples* (Q), etc. depending on the number of excited electrons (see Figure 2.4).
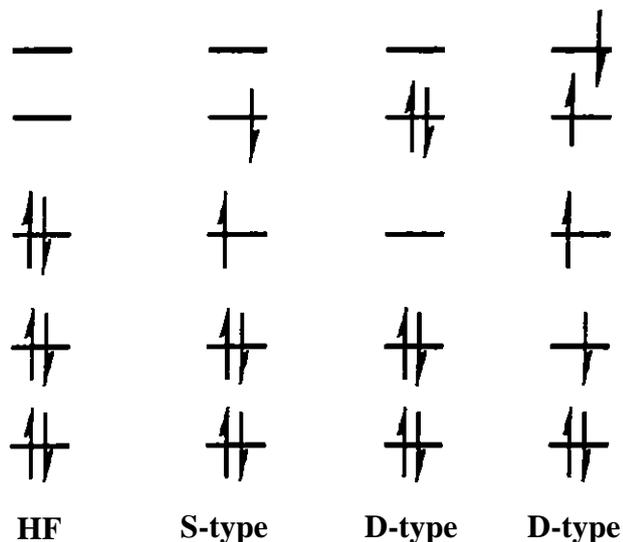
Figure 2.4 Excited Slater determinants generated from HF reference.

In general, the number of determinant depends on the size of basis set. The larger basis set refers to the more virtual molecular orbitals and more constructing excited determinants, *i.e.*, recovering all the electron correlation. Therefore, the exact solution can be obtained from this wavefunction. However, the exact solution is related to the size of basis set and the number of determinants. As can be seen in Figure 2.5, the use of larger basis set size together with larger number of determinant leads to better result.
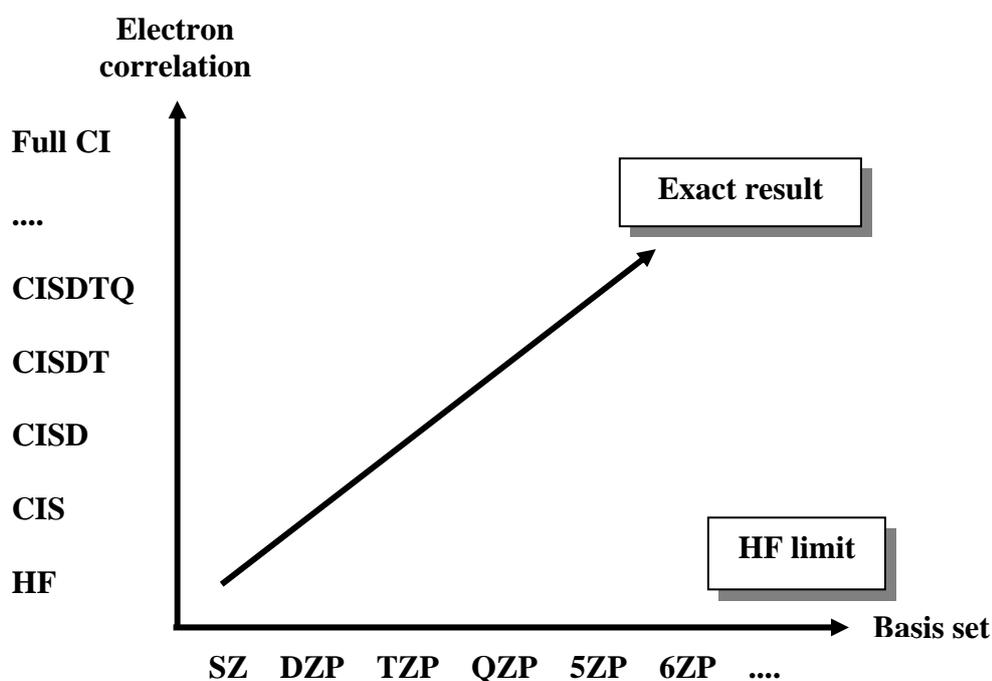
Figure 2.5 Common approach to the exact energy of a system.

The oldest electron correlation method is *configuration interaction* (CI) (Sherrill and Schaefer, 1999). This method has been used to determine the minimum energy of the trial wavefunction, according to the linear combination of determinants

together with the expansion coefficients. The excited Slater determinants, as taken from the HF calculation, should be of the form

$$\Psi_{CI} = a_0 \Phi_{HF} + \sum_S a_S \Phi_S + \sum_D a_D \Phi_D + \sum_T a_T \Phi_T + \dots = \sum_{i=0} a_i \Phi_i, \qquad (2.51)$$

where the subscripts S, D, T, etc., indicate determinants that are Singly, Doubly, Triply, etc. To determine the minimum energy, the Lagrange multiplier has been applied to the total CI wavefunction as

$$L = \langle \Psi_{CI} | H | \Psi_{CI} \rangle - \lambda(\langle \Psi_{CI} | \Psi_{CI} \rangle - 1). \qquad (2.52)$$

The first term is the energy of the CI wavefunction and the second term is the norm of the wavefunction. The determinant in equation (2.51) can be written in terms of equation (2.52) as

$$\langle \Psi_{CI} | H | \Psi_{CI} \rangle = \sum_{i=0} \sum_{j=0} a_i a_j \langle \Phi_i | H | \Phi_j \rangle$$

$$= \sum_{i=0} a_i^2 E_i + \sum_{i=0} \sum_{j \neq i} a_i a_j \langle \Phi_i | H | \Phi_j \rangle \qquad (2.53)$$

and

$$\left\langle \Psi_{\mathrm{CI}} \middle| \Psi_{\mathrm{CI}} \right\rangle = \sum_{i=0}\sum_{j=0} a_i a_j \left\langle \Phi_i \middle| \Phi_j \right\rangle$$

$$= \sum_{i=0} a_i^2 \left\langle \Phi_i \middle| \Phi_i \right\rangle$$

$$= \sum_{i=0} a_i^2 . \tag{2.54}$$

The variational procedure has been applied by setting all the derivations of the Lagrange function with respect to the a_i expansion coefficient equal to zero.

$$\frac{\partial L}{\partial a_i} = 2\sum_j a_j \left\langle \Phi_i \middle| H \middle| \Phi_j \right\rangle - 2\lambda a_i = 0$$

$$a_i (E_i - \lambda) + \sum_{j\neq 0} a_j \left\langle \Phi_i \middle| H \middle| \Phi_j \right\rangle = 0. \tag{2.55}$$

If there is only one determinant in the expansion ($a_0 = 1$, $a_{i\neq 0} = 0$), the Lagrange multiplier λ is the CI energy. In many determinants, the variational problem has been solved by a set of CI secular equation as

$$\begin{pmatrix} H_{00} - E & H_{01} & \cdots & H_{0j} & \cdots \\ H_{10} & H_{11} - E & \cdots & H_{1j} & \cdots \\ \vdots & \vdots & \ddots & \vdots & \cdots \\ H_{j0} & \vdots & \cdots & H_{jj} - E & \cdots \\ \vdots & \vdots & \cdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_j \\ \vdots \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \vdots \end{pmatrix} . \tag{2.56}$$

This can be written as a matrix equation as

$$(H\text{-}E\mathrm{I})a = 0$$

$$Ha = Ea. \tag{2.57}$$

Consequently, The CI energy can be obtained by the lowest eigenvalue of the CI matrix.

The disadvantage of CI is the lack of size-consistency. This problem has been taken into account in the *perturbation theory* (PT). The application of PT to a system composed of many interacting particles is called *many-body perturbation theory* (MBPT). To find the correlation energy of the ground state, the zero-order Hamiltonian from the Fock operator of the HF calculation has been designed using *Møller-Plesset perturbation theory* (MPPT) (Møller and Plesset, 1934).

Based on the MPPT, the zero-order Hamiltonian ($H^{(0)}$) can be calculated by the sum of one-electron Fock operators as defined in equation (2.41), which is written as

$$H_{\mathrm{HF}} = \sum_{i=1}^{n} f_i. \tag{2.58}$$

In the first-order Hamiltonian ($H^{(1)}$), the perturbation is given by

$$H^{(1)} = H - \sum_{i=1}^{n} f_i, \tag{2.59}$$

where H is the electronic Hamiltonian. The HF energy (E_{HF}) is associated with the ground-state HF wavefunction (Φ_0) as defined by

$$E_{\text{HF}} = \langle \Phi_0 | H | \Phi_0 \rangle, \tag{2.60}$$

or

$$E_{\text{HF}} = \langle \Phi_0 | H_{\text{HF}} + H^{(1)} | \Phi_0 \rangle. \tag{2.61}$$

Consequently, the total energy of the first-order perturbation is equal to the sum of the zero-order energy ($E_0^{(0)}$) and the first-order energy correction ($E_0^{(1)}$), and can be written as

$$E_{\text{HF}} = E_0^{(0)} + E_0^{(1)}. \tag{2.62}$$

The next improvement of the perturbation theory is the second-order perturbation theory, which can be defined as

$$E^{(2)} = \sum_{J \neq 0} \frac{\langle \Phi_J | H^{(1)} | \Phi_0 \rangle \langle \Phi_0 | H^{(1)} | \Phi_J \rangle}{E_0^{(0)} - E_J^{(0)}}, \tag{2.63}$$

where Φ_J is a multiply excited determinant of an eigenfunction of H_{HF} with eigenvalue $E_J^{(0)}$. To evaluate the off-diagonal matrix element $\langle \Phi_J | H^{(1)} | \Phi_0 \rangle$, the matrix element $\langle \Phi_J | H_{\text{HF}} | \Phi_0 \rangle$ equal to zero because Φ_0 is an eigenfunction of H_{HF} and spin orbialts are orthogonal. Thus, if $\langle \Phi_J | H | \Phi_0 \rangle = 0$, then $\langle \Phi_J | H^{(1)} | \Phi_0 \rangle = 0$.

From Brillouin's theorem (Lowe and Peterson, 2005), only the doubly excited determinants have non-zero $H^{(1)}$ matrix elements and therefore only double excitation contributes to $E^{(2)}$. These non-vanishing matrix elements can be expressed as

$$E^{(2)} = \frac{1}{4} \sum_{a,b}^{occ} \sum_{p,q}^{vir} \frac{(ab\|pq)(pq\|ab)}{\varepsilon_a + \varepsilon_b - \varepsilon_p - \varepsilon_q},$$

(2.64)

in which

$$(ab\|pq) = j_0 \int \phi_a^*(1)\phi_b^*(2)\frac{1}{r_{12}}\phi_p(1)\phi_q(2)dx_1 dx_2$$

$$- j_0 \int \phi_a^*(1)\phi_b^*(2)\frac{1}{r_{12}}\phi_q(1)\phi_p(2)dx_1 dx_2,$$

(2.65)

where ϕ_a and ϕ_b are occupied spin orbitals and ϕ_p and ϕ_q are virtual spin orbitals. The second-order energy correlation in MPPT is called MP2. In this respect, the MPPT can be extended to third- and fourth-order energy correlations, known as MP3 and MP4.

An alternative approach for the treatment of electron correlation is the *coupled-cluster* (CC) (Bartlett, 1989) method. This method is size-consistent. The CC method introduces the *cluster operator* (C), in which the equation of the exact electronic wavefunction can be written as

$$\Psi = e^C \Phi_0.$$

(2.66)

This exact electronic wavefunction is related to the HF wavefunction (Φ_0) and the exponential operator e^C is defined as

$$e^C = 1 + C + \frac{1}{2!}C^2 + \frac{1}{3!}C^3 + \ldots \qquad (2.67)$$

The cluster operator can be expressed in terms of a linear combination of Slater determinant when the electrons in the occupied spin orbitals are excited to virtual spin orbitals. Therefore, C is the sum of the many electron excitation operator, which can be written as

$$C = C_1 + C_2 + C_3 + \ldots + C_N. \qquad (2.68)$$

The effects of the excitation operators are

$$C_1\Phi_0 = \sum_{a,p} t_a^p \Phi_a^p, \qquad (2.69)$$

$$C_2\Phi_0 = \sum_{a,b,p,q} t_{ab}^{pq} \Phi_{ab}^{pq}, \qquad (2.70)$$

and so on, where t_a^p are *single-excitation amplitudes* and t_{ab}^{pq} are *double-excitation amplitudes*, etc.

There are two approximations in CC application. First, a finite basis set is used in the determinantion of Φ_0. Second, the expression for the cluster operator (C) is truncated to include only specified electron excitation operators. This refers to

coupled cluster singles and doubles (CCSD), where C is approximated by $C_1 + C_2$. In CCD, only the C_2 is employed, and in CCSDT, C is composed of $C_1 + C_2 + C_3$.

2.8   Density functional theory

For the treatment of system containing many atoms and many electrons, the *ab initio* methods are found to be very time-consuming. The *density functional theory* (DFT) is then used as an alternative approach, which takes into account the electron correlation using the concept of electron probability density. The energy of an electronic system is described in terms of the electron probability (ρ). In many electrons system, the total electron density at a particular point r in space can be denoted as $\rho(r)$. The electronic energy (E) is the functional of the electron density, which can be defined as $E(\rho)$.

The DFT method considers the pair electrons in the same spatial one-electron orbitals. Kohn and Sham suggest that the exact ground-state electronic energy (E) of an n-electrons system can be of the form

$$E(\rho) = -\frac{\hbar_2}{2m_e} \sum_{i=1}^{n} \int \psi_i^*(r_1)\nabla_1^2\psi_i(r_1)dr_1 - j_0 \sum_{I=1}^{N} \frac{Z_I}{r_{I1}}\rho(r_1)dr_1$$

$$+ \frac{1}{2}j_0 \int \frac{\rho(r_1)\rho(r_2)}{r_{12}}dr_1dr_2 + E_{XC}[\rho] \qquad (2.71)$$

The first term of equation (2.71) describes the kinetic energy of the electron. The second term represents the electron-nucleus attraction in which the sum is over all N nuclei with index I and atomic number Z_I. The third term refers to the Coulumb

interaction between the total charge distribution at r_1 and r_2 and the last term is the *exchange-correlation energy* of the system. The one-electron spatial orbitals (ψ_i; $i = 1, 2, \ldots, n$) are the *Kohn-Sham* (KS) (Kohn and Sham, 1965) orbital. The exact ground-state electron density can be defined by

$$\rho(r) = \sum_{i=1}^{n} |\psi_i(r)|^2. \tag{2.72}$$

The KS orbital can be described by solving the *Kohn-Sham* (Kohn and Sham, 1965) equation and the one-electron orbital ($\psi_i(r_1)$) can be of the form

$$\left\{ -\frac{\hbar^2}{2m_e}\nabla_1^2 - j_0 \sum_{I=1}^{N} \frac{Z_I}{r_{I1}} + j_0 \int \frac{\rho(r_2)}{r_{12}} dr_2 + V_{XC}(r_1) \right\} \psi_i(r_1) = \varepsilon_i \psi_i(r_1), \tag{2.73}$$

where ε_i are the KS orbital energies and V_{XC} is the exchange-correlation potential, which can be derived from the exchange-correlation energy,

$$V_{XC}[\rho] = \frac{\delta E_{XC}[\rho]}{\delta \rho}. \tag{2.74}$$

Starting with the guess electron density (ρ), a self-consistent fashion is employed for calculating the KS equations. By using an appropriate form of the $E_{XC}[\rho]$, the V_{XC} can be calculated as the function of r. The set of KS equation is solved in order to obtain an initial set of KS orbitals. Then, the set of orbitals is used

to compute an improved density from equation (2.72). These procedures reach convergence when the density and exchange-correlation energy are satisfied.

According to the calculation process, the main error of DFT is the approximation of the $E_{XC}[\rho]$. This function can be separated into an exchange functional and a correlation function. In the *local density approximation* (LDA), the exchange-correlation can be defined as

$$E_{XC} = \int \rho(r)\varepsilon_{XC}[\rho(r)]dr, \tag{2.75}$$

where $\varepsilon_{XC}[\rho(r)]$ is the exchange-correlation energy per electron in a homogeneous electron gas of constant density. To improve the exchange-correlation function, a non-local correction involving the gradient of ρ is added to the exchange-correlation energy. The LDA with gradient-corrections is called the *generalized gradient approximation* (GGA). The exchange-correlation functionals have been developed for use in DFT calculations, such as mPWPW91, B3LYP, MPW1K, PBE1PBE, BLYP, BP91 and PBE. The name of each function refers to the pairing of an exchange function and correlation function. For example, the BLYP function is a combination of the gradient-corrected exchange functional, developed by Becke (Becke, 1986; Becke, 1988), and the gradient-corrected correlation functional developed by Lee, Yang and Parr (Lee, Yang, and Parr, 1988). The B3LYP function makes use of Hartree-Fock corrections in conjunction with density function correlation and exchange. Nowadays, the DFT calculations are widely used for large molecular systems, such as protein.

2.9   References

Bartlett, R. J. (1989). Coupled-cluster approach to molecular structure and spectra: A step toward predictive quantum chemistry. **The Journal of Physical Chemistry**. 93: 1697-1708.

Becke, A. D. (1986). Density function calculations of molecular-bond energies. **The Journal of Chemical Physics**. 84: 4524-4529.

Becke, A. D. (1988). Density-functional exchange-energy approximation with correct asymptotic behavior. **Physical Review A**. 38: 3098-3100.

Boys, S. F. (1950). Electronic wavefunctions. I. A general method of calculation for stationary states of any molecular system. **Proceedings of the Royal Society of London Series A**. 200: 542-554.

Dunning, T. H. J. (1970). Gaussian basis functions for use in molecular calculations. I. Contraction of (9s5p) atomic basis sets for the first-row atoms. **The Journal of Chemical Physics**. 53: 2823-2833.

Dunning, T. H. J. (1971). Gaussian basis set function for use in molecular calculations. III. Contraction of (10s6p) atomic basis sets for the first-row atoms. **The Journal of Chemical Physics**. 55: 716-723.

Hehre, W. J., Stewart, R. F. and Pople, J. A. (1969). Self-consistent molecular-orbital methods. I. Use of gaussian expansions of slater-type atomic orbitals. **The Journal of Chemical Physics**. 51: 2657-2664.

Huzinaga, S. (1965). Gaussian-type functions for polyatomic systems. I. **The Journal of Chemical Physics**. 42: 1293-1302.

Kohn, W. and Sham, L. J. (1965). Self-consistent equations including exchange and correlation effects. **Physical Review**. 140: A1133-A1138.

Lee, C., Yang, W. and Parr, R. G. (1988). Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. **Physical Review B**. 37: 785-789.

Lowe, J. P. and Peterson, K. A. (2005). **Quantum Chemistry**. US: Elsevier Academic Press.

Møller, C. and Plesset, M. S. (1934). Note on an approximation treatment for many-electron systems. **Physical Review**. 46: 618-622.

Schrödinger, E. (1926). Quantisierung als Eigenwertproblem. **Annalen der Physik**. 79: 361-376.

Sherrill, C. D. and Schaefer, H. F. (1999). The configuration interaction method: Advances in highly correlated approaches. **Advances in Quantum Chemistry**. 34: 143-269.

Slater, J. C. (1929). The theory of complex spectra. **Physical Review**. 34: 1293-1322.

Slater, J. C. (1930). Atomic shielding constants. **Physical Review**. 36: 57-64.

CHAPTER III

MOLECULAR DYNAMICS SIMULATIONS

In terms of computation simulations, the molecular dynamics (MD) technique is well-known. This technique is widely used for studying various molecular systems. MD simulation provides the time dependent behavior of a molecular system. The MD simulation starts with reading in the initial configuration, such as coordinates, velocities, accelerations and forces. The initial configuration can be obtained from random configurations or a lattice. One of the essential conditions of the simulation is that there are no explicitly time-dependent or velocity dependent forces that shall act on the system. In practice, the trajectories cannot be directly obtained from Newton's equation. Therefore, the *time integration algorithm* will be used to obtain the knowledge of positions, velocities and accelerations of two successive time steps. The energy of the system can be calculated using molecular mechanics (MM) or quantum mechanics (QM) methods. The force on each atom can be obtained from the derivative of the energy with respect to the change in the atom's position. The particles will be moved by their new force to the new configurations. This process will be repeated until the system reaches equilibrium. Then, the coordinates, velocities, accelerations, forces and so on of all particles will be collected for further structural and dynamical property calculations. In most cases, only positions and velocities are usually stored since most important properties can be obtained from these two quantities. The schematic of molecular dynamics simulation is shown in Figure 3.1.

starting configuration, velocities, accelerations and forces

predict configurations, velocities, accelerations, etc., at a time t+δt using the current values of these quantities

calculate interaction energy E (using either MM or QM method) and force on an atom $\left(F_i \left(-\dfrac{dE}{dr_i} = F_i \right) \right)$ and hence acceleration $\left(a_i = \dfrac{F_i}{m_i} \right)$

correct the predicted configurations, velocities, accelerations, etc., using the new accelerations

time integration algorithm

move particles by force F_i to the new configurations

equilibrium?

no

yes

store coordinates, velocities, accelerations and forces, etc. of all particles

calculate properties of the system

Figure 3.1 The schematic of molecular dynamics simulation.

3.1 Time average and ensemble average

The properties of the system will depend upon the positions and the momenta of N particles that comprise the system. The value of the property A can thus be written as

$$A(\mathrm{p}^N(t), \mathrm{r}^N(t)), \tag{3.1}$$

where $\mathrm{p}^N(t)$ and $\mathrm{r}^N(t)$ represent the N momenta and positions, respectively. The value of property A is the average of the A over the time of the measurement, known as a *time average*. In principle, if the time measurement reach infinity, the value of the property A is then the *true value*,

$$A_{ave} = \lim_{\tau \to \infty} \frac{1}{\tau} \int_{t=0}^{\tau} A(\mathrm{p}^N(t), \mathrm{r}^N(t)) dt. \tag{3.2}$$

In practice, the treatment for system consisting of a large number of atoms or molecules is not feasible, even the determination of an initial configuration of the system. Boltzmann and Gibbs developed statistical mechanics, known as *ensemble*. The ensemble is a single system evolving in time that contains a large number of mental copies of a system, considered all at once, each of which represents a possible state of the real system. The time average is then replaced by an *ensemble average* as

$$\langle A \rangle = \iint d\mathrm{p}^N d\mathrm{r}^N A(\mathrm{p}^N, \mathrm{r}^N) \rho(\mathrm{r}^N, \mathrm{r}^N). \tag{3.3}$$

The angle bracket (< >) indicates an ensemble average or *expectation value*, *i.e.*, the average value of the property *A* over all replications of the ensemble generated by the simulation. Different macroscopic environmental constraints lead to different types of ensembles. In general, the ensemble is employed with constraints, such as constant number of particles (N), volume (V), energy (E), temperature (T), chemical potential (μ), pressure (P) and so on. For example, the *microcanonical ensemble* (NVE) fixes the number of particles (N), the volume (V) and the energy (E) of the system. The equilibrium states of NVE ensemble characterize the entropy. Another example is the *canonical ensemble* (NVT), which fixes the number of particles (N), the volume (V) and the temperature (T). The thermodynamic property derived from the NVT ensemble is Helmholtz free energy. Other ensembles include the *grand canonical ensemble* (μVT), which fixes the chemical potential (μ), the volume of the system (V) and the temperature (T). The pressure×volume (PV) quantity can be obtained from this ensemble.

3.2 Intermolecular potentials

For classical simulations based on molecular mechanical force field, system's interactions are usually described by intermolecular potential. The potential energy function is the total interactions comprising all of the pair, three-body, four-body up to N-body interactions,

$$V_{total} = \sum V(i,j) + \sum V(i,j,k) + \ldots + \sum V(i,j,k,\ldots,N). \qquad (3.4)$$

In practice, the non-additive terms (*i.e.* the three-body up to N-body) are assumed to converge slowly and tend to have alternating signs (Kistenmacher, Popkie, and Clementi, 1974). This leads to *pairwise additive approximation,* in which the total interactions of the system are assumed to be the summation of only pair interactions.

The pair potential functions can be constructed from experimental data but this approach requires more information. The popular way is to construct from *ab initio* calculations. The construction of pair potential contains five steps. The first step is to define the pair interacting model. Then, the conformations and/or configurations of the model are generated in order to perform *ab initio* calculations. After that, the suitable quantum mechanics method and basis set are chosen in the calculations of the corresponding pair interaction energies. In the fourth step, a suitable mathematical formula is selected and used in the fitting procedure. Finally, the resulting fitted potentials are evaluated for checking their reliability, before being employed in the simulation.

3.3   Time-integration algorithms

In general, there are many algorithms for integrating the equations of motion, most of which are based on *finite difference method.* By this method, the integration is broken down into many small stages, each separated in time by a fixed time interval Δt. The total force on each particle in the system at time t is calculated as the vector sum of its interactions with other particles. Then, the positions and velocities at a time t are used to calculate the positions and velocities at a time $t + \Delta t$. Then, the new positions and velocities have been calculated at time $t + 2\Delta t$, and so on.

All algorithms assume that the positions and dynamic properties can be approximated as *Taylor series expansion,*

$$r(t + \Delta t) = r(t) + \Delta t v(t) + \frac{1}{2}\Delta t^2 a(t) + \frac{1}{6}\Delta t^3 b(t) + \frac{1}{24}\Delta t^4 c(t) + \cdots, \quad (3.5)$$

$$v(t + \Delta t) = v(t) + \Delta t a(t) + \frac{1}{2}\Delta t^2 b(t) + \frac{1}{6}\Delta t^3 c(t) + \cdots, \quad (3.6)$$

$$a(t + \Delta t) = a(t) + \Delta t b(t) + \frac{1}{2}\Delta t^2 c(t) + \cdots, \quad (3.7)$$

$$b(t + \Delta t) = b(t) + \Delta t c(t) + \cdots. \quad (3.8)$$

The first derivative of the position (r) with respect to time is the velocity (v), the second derivative is the acceleration (a) and the third derivative is b, and so on. The widely used methods for integrating the equation of motion in molecular dynamic simulation are the *Verlet algorithm* (Verlet, 1967) and *Predictor-corrector algorithm* (Gear, 1971).

3.3.1   Verlet algorithm

The Verlet algorithm uses the positions and accelerations at time t and positions from the previous step, $r(t - \Delta t)$, to calculate the new position at $t + \Delta t$, $r(t + \Delta t)$. The following relationships between these quantities and the velocities at time t would be of the form

$$r(t + \Delta t) = r(t) + \Delta t v(t) + \frac{1}{2}\Delta t^2 a(t) + \cdots, \quad (3.9)$$

$$r(t - \Delta t) = r(t) - \Delta t v(t) + \frac{1}{2} \Delta t^2 a(t) - \cdots. \qquad (3.10)$$

Combining these two equations gives

$$r(t + \Delta t) = 2r(t) - r(t - \Delta t) + \Delta t^2 a(t) + \cdots. \qquad (3.11)$$

In this respect, the velocities of the Verlet algorithm do not explicitly appear in the equations. However, the velocities can be calculated in many ways such as dividing the difference in positions at time $t + \Delta t$ and $t - \Delta t$ by $2\Delta t$ as

$$v(t) = \left[r(t + \Delta t) - r(r - \Delta t) \right] / 2\Delta t. \qquad (3.12)$$

Alternatively, the velocities can be obtained at the half-step $(t + \frac{1}{2}\Delta t)$ as

$$v(t + \frac{1}{2}\Delta t) = \left[r(t + \Delta t) - r(t) \right] / \Delta t. \qquad (3.13)$$

The deficiency of Verlet algorithm is the difficulty in calculating the velocities, since these quantities can not be obtained until the positions are computed at the next step. On the other hand, it is not a self-starting algorithm. For example, at t = 0 there is only one set of positions and it needs to know positions at $t - \Delta t$, which can be obtained by the Taylor series as

$$r(-\Delta t) = r(0) - \Delta t v(0) - \Delta t v(0). \tag{3.14}$$

The Verlet algorithm has been developed to the *leap-frog* algorithm (Hockney, 1970), in which the positions and velocities can be written in the forms of

$$r(t + \Delta t) = r(t) + \Delta t v(t + \frac{1}{2}\Delta t), \tag{3.15}$$

$$v(t + \frac{1}{2}\Delta t) = v(t - \frac{1}{2}\Delta t) + \Delta t a(t). \tag{3.16}$$

The leap-frog algorithm starts with the velocities $v(t + \frac{1}{2}\Delta t)$ that calculated from the velocities at time $t - \frac{1}{2}\Delta t$ and the accelerations at time t. Then, the positions $r(t + \Delta t)$ are computed by the velocities that calculated together with the position at time t, $r(t)$. The velocities at time t can be calculated from

$$v(t) = \frac{1}{2}\left[v(t + \frac{1}{2}\Delta t) + v(t - \frac{1}{2}\Delta t) \right]. \tag{3.17}$$

In this respect, the velocities leap-frog over the positions to give their values at $t + \frac{1}{2}\Delta t$ (hence the name). Then, the positions leap-frog over the velocities to give their new values at $t + \Delta t$, ready for the velocities at $t + \frac{3}{2}\Delta t$ and so on.

With regard to the leap-frog method, however, some deficiencies are still remain, *i.e.*, this algorithm can not calculate the positions and velocities at the same time. On the other hand, the kinetic energy contribution can not be calculated and included into the total energy at the same time (unlike for the calculations of positions, velocities and accelerations). An alternative approach is to use the *velocity Verlet method* (Swope, Anderson, Berens, and Wilson, 1982) in which the relationship between the positions and velocities can be expressed as

$$\mathrm{r}(t + \Delta t) = \mathrm{r}(t) + \Delta t \mathrm{v}(t) + \frac{1}{2}\Delta t^2 \mathrm{a}(t), \tag{3.18}$$

$$\mathrm{v}(t + \Delta t) = \mathrm{v}(t) + \frac{1}{2}\Delta t \left[\mathrm{a}(t) + \mathrm{a}(t + \Delta t)\right]. \tag{3.19}$$

The velocity Verlet method is a three-stage procedure. In the first step, the positions at $t + \Delta t$ are computed by using the velocities and the accelerations at time *t*. Then, the velocities at time $t + \frac{1}{2}\Delta t$ are determined by

$$\mathrm{v}(t + \Delta t) = \mathrm{v}(t) + \frac{1}{2}\Delta t \mathrm{a}(t). \tag{3.20}$$

After that, new forces are computed from the current positions, giving accelerations at time $t + \Delta t$, $\mathrm{a}(t + \Delta t)$. Finally, the velocities at time $t + \Delta t$ are determined by

$$\mathrm{v}(t + \Delta t) = \mathrm{v}(t + \frac{1}{2}\Delta t) + \frac{1}{2}\Delta t \mathrm{a}(t + \Delta t). \tag{3.21}$$

A more accurate integration method for the velocity is the *Beeman's algorithm* (Beeman, 1976), which is related to the Verlet method. The Beeman integration scheme gives better energy conservation because the kinetic energy is calculated directly from the velocities. However, this method is more complicate than the Verlet algorithm, as well as more time-consuming. The Beeman's algorithm can be expressed as

$$r(t + \Delta t) = r(t) + \Delta t v(t) + \frac{2}{3} \Delta t^2 a(t) - \frac{1}{6} \Delta t^2 a(t - \Delta t), \qquad (3.22)$$

$$v(t + \Delta t) = v(t) + \frac{1}{3} \Delta t a(t) + \frac{5}{6} \Delta t a(t) - \frac{1}{6} \Delta t a(t - \Delta t). \qquad (3.23)$$

3.3.2 Predictor-corrector algorithm

The predictor-corrector algorithm has three steps. Let's start with the new positions, velocities, accelerations and higher-order terms that are predicted according to the Taylor expansion (see equations (3.5)-(3.8)). Then, the force is evaluated at the new positions to give accelerations $a(t + \Delta t)$. The accelerations are compared with the accelerations that are predicted from the Taylor series expansion ($a^C(t + \Delta t)$). The difference between the predicted and calculated accelerations is used to correct the positions, velocities, etc., using equations (3.24)-(3.28), as followed:

$$\Delta a(t + \Delta t) = a^C(t + \Delta t) - a(t + \Delta t). \qquad (3.24)$$

$$r^C(t + \Delta t) = r(t + \Delta t) + C_0 \Delta a(t + \Delta t). \qquad (3.25)$$

$$v^C(t + \Delta t) = v(t + \Delta t) + C_1 \Delta a(t + \Delta t). \qquad (3.26)$$

$$a^C(t + \Delta t)/2 = a(t + \Delta t)/2 + C_2 \Delta a(t + \Delta t). \qquad (3.27)$$

$$b^C(t + \Delta t)/6 = b(t + \Delta t)/6 + C_3 \Delta a(t + \Delta t). \qquad (3.28)$$

The superscript C marks these as corrected values. The set of coefficients depend upon the order of the Taylor series expansion.

3.4   Periodic boundary conditions

One of the common problems found in computer simulations is the boundary effects or surface effects. This problem can be solved by using periodic boundary conditions. By this scheme, particles in the box are replicated in all directions to give a periodic array (see Figure 3.2).
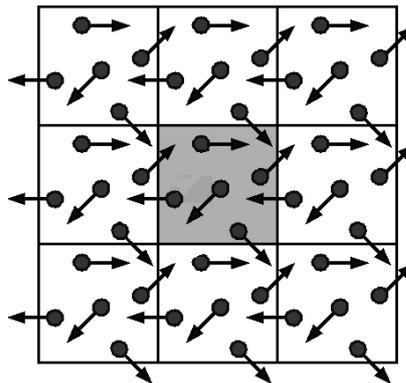
Figure 3.2 Periodic boundary conditions in two dimensions.

The main point of periodic boundary conditions is the coordinates of the particles in the image boxes can be computed by adding or subtracting integral

multiples of the box sides. If a particle leaves the box during the simulation, it is replaced by an image particle that enters from the opposite side in the same time, as illustrated in Figure 3.2. Therefore, the number of particles within the central box remains constant.

3.5   Cut-off and discontinuity at cut-off

In general, the non-bonded interactions are calculated between every pair of atoms in the system. However, the most time-consuming of the simulation is the calculation of the non-bond energies and forces. The simple way of reducing the expense is to use a *cut-off* and to apply the *minimum image convention*. The minimum image convention is a common form of the periodic boundary condition, in which each atom interacts with the most one image of every other atom in the system. The energies and forces are computed with the closest atom or image as shown in Figure 3.3.
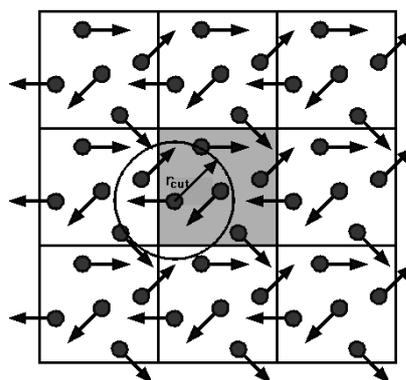
Figure 3.3 The spherical cut-off and the minimum image convention.

For the cut-off, the interactions between all pairs of atoms that are further apart than the cut-off value are set to zero. The cut-off should not be greater than half of the length of their image, which would be calculated for the same atom. The use of cut-off can cause serious errors in some simulations, especially for the simulations involving strong pair interactions, as can be seen in Figure 3.4.
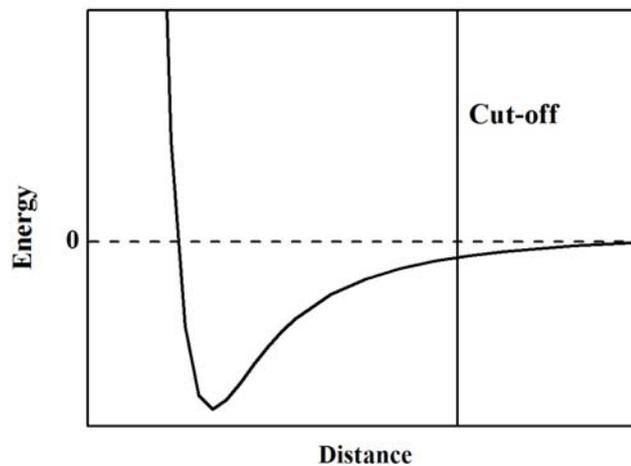
Figure 3.4 A discontinuity at cut-off.

The use of cut-off introduces a discontinuity in both the potential energy and the force after the cut-off value. One way for solving this problem is to use a *shifted potential*, in which a constant term is subtracted from the potential at all values as

$$V'(r) \begin{cases} V(r) - V_c & r \le r_c \\ 0 & r > r_c \end{cases}, \tag{3.29}$$

where r_c is the cut-off distance and V_c is equal to the value of the potential at the cut-off distance. The energy conservation can be improved by the shifted potential but

there is a discontinuity in the force with the shifted potential. At the cut-off distance, the force will have a finite value. Therefore, the suitable shifted potential would be of the form

$$
V'(r)
\begin{cases}
V(r) - V_c - \left(\dfrac{dV(r)}{dr} \right)_{r=r_c} (r - r_c) & r \le r_c \\
0 & r > r_c
\end{cases},
\tag{3.30}
$$

However, the shifted potential is not easy for inhomogeneous systems containing many different types of atom. An alternative way to eliminate discontinuities in the energy and force is a *switching function.* The switched potential ($V^{SF}(r)$) is related to the true potential ($V(r)$) as

$$
V'(r) = V(r)S(r).
\tag{3.31}
$$

Some switching functions are applied to the entire range of the potential up to the cut-off point. The switching function has a value of 1 at $r = 0$ and a value of 0 at $r = r_c$ and the switching function values between two cut-offs have been varied. The potential functions are shown in Figure 3.5.
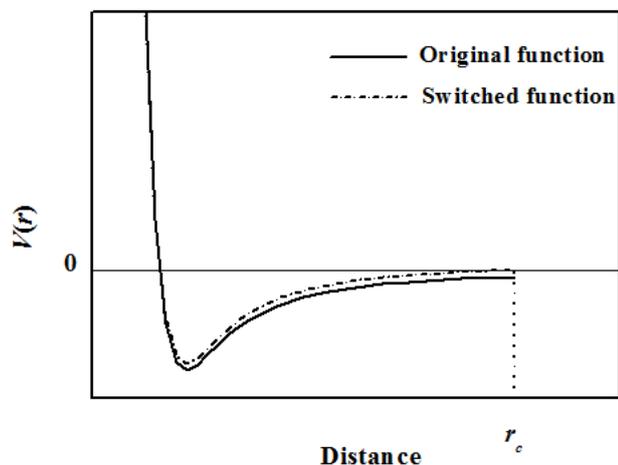
Figure 3.5 A switching function applied on the Leannard-Jones potential.

The switching function multiplies the potential in range between the lower (r_1) an upper cut-off (r_u) distance, which takes the value 1 at the lower cut-off distance and 0 at the upper cut-off distance. The range is usually small such as r_1 might be 9 Å and r_u is 10 Å. The potential that applies the switching function in the range near the cut-off is shown in Figure 3.6.
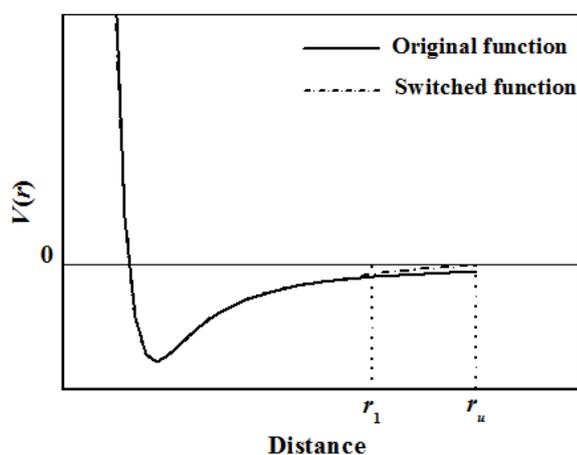
Figure 3.6 A switching function applied over a range near the cut-off and its results on the Lennard-Jones potential.

3.6   Non-bonded neighbour lists

Useful technique in reducing time-consuming in computer simulations is to use the *non-bonded neighbour list*. This technique lists the atoms to be included in the atom's neighbours without having to calculate them for every time. The first non-bonded neighbour list has been proposed by Verlet (Verlet, 1967). The Verlet neighbour list stores all atoms within the cut-off distance (the solid circle in Figure 3.7) and atoms that are slightly further away from the cut-off distance (the dashed shown in Figure 3.7). The neighbour list is updated throughout the simulation. If the update is too often, the procedure is inefficient. On the other hand, if the update is too slow, the energies and forces are calculated incorrectly due to atoms moving within the cut-off region. The suitable update frequency is suggested to be between 10 and 20 steps. The distance used to calculate each atom's neighbours should be larger than the actual cut-off distance in order to ensure the atoms outside the cut-off will not approach closer than the cut-off distance before the neighbour list is updated again.
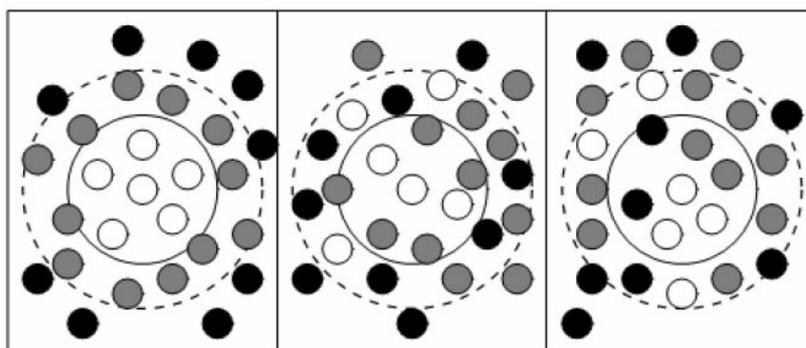
Figure 3.7 The Verlet neighbour list.

3.7   Long-range interactions

For strong interacting systems, the use of small periodic box together with the cut-off can conduct problems in energy and force calculations. One way to treat the long-range forces is to use a larger simulation cell but this results in time-consuming. There are many suitable methods for treating the long-range force. The first method is the *Ewald summation method*, which derived by Ewald in 1921 (Ewald, 1921). This method studies the energetics of ionic crystals, which allow a particle to interact with all the other particles in the simulation box and with all of their images in an infinite array of periodic cells. The charge-charge contribution to the potential energy of the Ewald summation method can be expressed as

$$V = \frac{1}{2} \sum_{|n|=0}^{'} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{q_i q_j}{4\pi\varepsilon_0 |r_{ij} + n|}, \qquad (3.32)$$

where the prime on the first summation indicates that the series does not include the interaction $i = j$ for $n = 0$, q_i and q_j are charges and n is a cubic lattice point. The Ewald summation method includes all the effects of long-range forces in a computer simulation. However this method is quite expensive as the equation (3.32) converges extremely slowly.

Another method for the treatment of long-range interaction is the *reaction field method* (Foulkes and Haydock, 1989). This method constructs the sphere around the molecule with a radius equal to the cut-off distance. The interaction with molecules that are positioned within the sphere is calculated explicitly, whereas the

outside of the sphere is modeled as a homogeneous medium of dielectric constant (ε_s). The electrostatic field due to the surrounding dielectric is given by

$$E_i = \frac{2(\varepsilon_s - 1)}{\varepsilon_s + 1}\left(\frac{1}{r_c^3}\right)\sum_{j;r_{ij}\leq r_c}\mu_j, \tag{3.33}$$

where μ_j are the dipoles of the neighbouring molecules that are positioned within the cut-off distance of the molecules i. The interaction between the molecule i and the reaction field equals $E_i \cdot \mu_i$.

3.8   Combined *ab initio* QM/MM molecular dynamics simulations

According to the QM/MM MD approach, the system is partitioned into two parts, namely QM and MM regions.
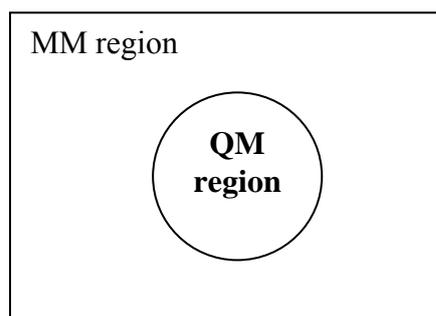
Figure 3.8 System's partition.

The system's interaction energy (E_{tot}) can be calculated from

$$E_{tot} = \left\langle \Psi_{QM} \left| \hat{H} \right| \Psi_{QM} \right\rangle + E_{MM} + E_{QM-MM} , \tag{3.34}$$

where the first term on the right hand side corresponds to the interactions within the QM region derived by means of quantum mechanics and the last two terms represent the interactions within the MM and between the QM and MM regions, respectively, evaluated on the basis of pairwise additive approximations.

During the QM/MM simulations, forces acting on each particle in the system are switched according to which region the solvent molecule was entering or leaving the QM region and is defined as

$$F_i = S_m(r)F_{QM} + (1 - S_m(r))F_{MM} , \tag{3.35}$$

where F_{QM} and F_{MM} are quantum mechanical and molecular mechanical forces, respectively. $S_m(r)$ is a smoothing function described by

$$S_m(r) = 1 \qquad \qquad \text{for } r < r_1,$$

$$S_m(r) = \frac{(r_0^2 - r^2)^2(r_0^2 + 2r^2 - 3r_1^2)}{(r_0^2 - r_1^2)^3} \quad \text{for } r_1 < r \le r_1, \tag{3.36}$$

$$S_m(r) = 0 \qquad \qquad \text{for } r > r_0,$$

where r_1 and r_0 are the distances characterizing the start and the end of the QM region, applied within an interval of 0.2 Å to ensure a continuous change of forces at the transition between QM and MM regions.

3.9   References

Beeman, D. (1976). Some multistep methods for use in molecular dynamics calculations. **Journal of Computational Physics**. 20: 130-139.

Ewald, P. (1921). Due Berechnung optischer und elektrostatischer Gitterpotentiale. **Annalen der Physik**. 64: 253-287.

Foulkes, W. M. and Haydock, R. (1989). Tight-binding models and density-functional theory. **Physical Review B**. 39: 12520-12536.

Gear, C. W. (1971). **Numerical initial value problems in ordinary differential equations**. Englewood Cliffs: NJ: Prentice Hall.

Hockney, R. W. (1970). The potential calculation and some applications. **Methods in Computational Physics**. 9: 136-211.

Kistenmacher, H., Popkie, H. and Clementi, E. (1974). Study of the structure of molecular complexes. VIII. Small clusters of water molecules surrounding Li^+, Na^+, K^+, F^-, and Cl^- ions. **The Journal of Chemical Physics**. 61: 799-815.

Swope, W. C., Anderson, H. C., Berens, P. H. and Wilson, K. R. (1982). A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. **The Journal of Chemical Physics**. 53: 289-298.

Verlet, L. (1967). Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. **Physical Review**. 159: 98-103.

CHAPTER IV

RESEARCH PROCEDURES

4.1 Selection of QM method and geometry optimizations

In general, the simulation with high QM level of accuracy and large basis set can provide more accurate description on the structural and dynamical properties. However, the computational requirement will increase significantly according to the level of QM method and size of basis set. The key rule for selecting QM method, as well as the basis set, is the compromise between the quality of simulation results and the requirement for CPU time. The simple way is to compare the structure and lowest energy of ion-H_2O complex that obtained from geometry optimizations using different QM methods. In this work, three ion-H_2O complexes in gas phase, namely cyclic-, anti- and syn-ion-H_2O conformers, (Markham, Trachtman, Bock, and Bock, 1998) are chosen. These three conformers conduct three different hydrogen bonds (see Figures 4.1 and 4.2). The geometry optimizations of ion-H_2O complexes are carried out at PW91, BLYP, B3LYP, HF, MP2 and CCSD levels of accuracy using DZV+ and aug-cc-pVDZ basis sets. The calculated stabilization energies and some structural parameters of the optimized ion-H_2O complexes are summarized in Tables 4.1 and 4.2.
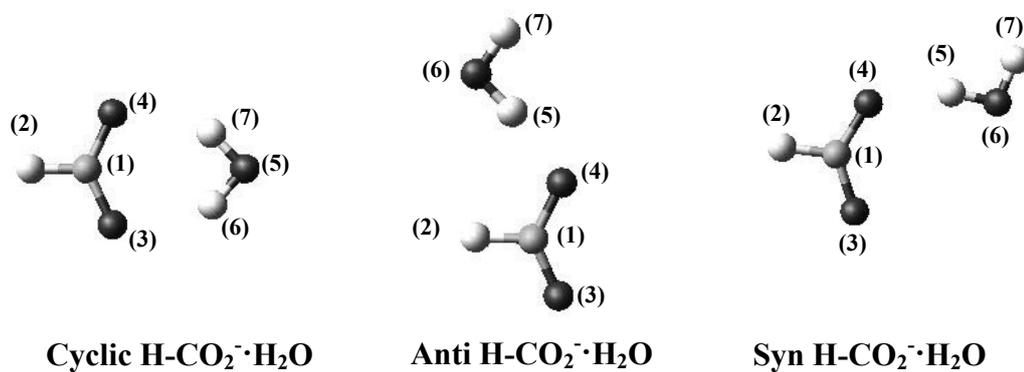
Figure 4.1 Conformations of HCOO⁻-H_2O in gas phase.
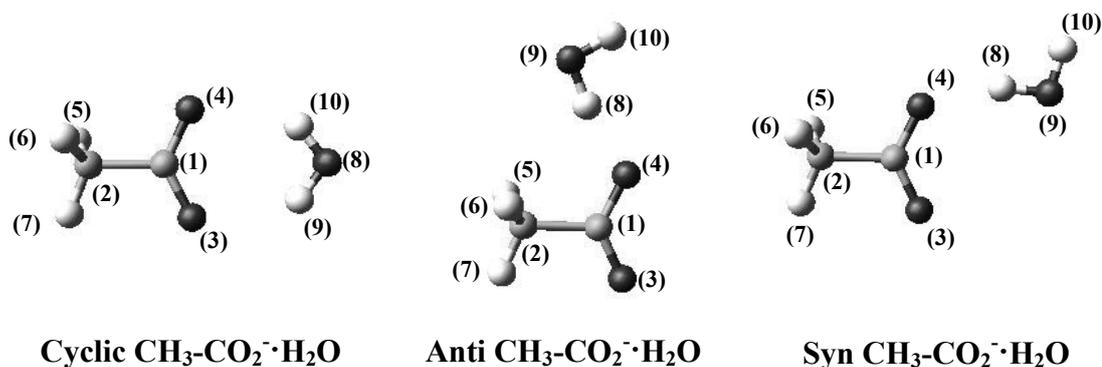
Figure 4.2 Conformations of CH_3COO^--H_2O in gas phase.

According to the data in Tables 4.1 and 4.2, the DFT methods (PW91, BLYP and B3LYP) with DZV+ basis set provide significant difference on the stabilization energies and the hydrogen bond length, as compared to other methods, especially for the resulting anti-HCOO⁻-H_2O, anti-CH_3COO^--H_2O and syn-CH_3COO^--H_2O conformers. However, the results obtained from BLYP and B3LYP methods using larger basis set show fair agreement with those of the correlated data.

In terms of HF calculations, the results show good agreement with the correlated data, in particular when the DZV+ basis set is employed. Overall, the

hydrogen bond length and energy ordering of the three ion-water complexes predicted by the HF and DFT (BLYP and B3LYP) methods are in good accord with the correlated results. On the other hand, this suggests that quantum mechanics calculations by both HF and DFT (BLYP and B3LYP) methods would be reliable enough to be employed in the simulations. The use of HF method has been well demonstrated in previous QM/MM studies, even for the treatment of anion (Kerdcharoen, Liedl, and Rode, 1996; Tongraar, Liedl, and Rode, 1997; Tongraar, Sagarik, and Rode, 2002; Intharathep, Tongraar, and Sagarik, 2005; Xenides, Randolf, and Rode, 2006), assuming that the effects of electron correlation are small enough to be neglected (Tongraar, Liedl, and Rode, 1998; Tongraar and Rode, 2003; Tongraar, Tangkawanwanit, and Rode, 2006).

In addition, the geometry optimizations of $HCOO^-$, CH_3COO^- and water that will be used in the constructions of pair potential functions were performed at HF and B3LYP levels of accuracy using aug-cc-pVDZ basis set (Dunning, 1989; Kendall and Harrison, 1992; Woon and Thom H. Dunning, 1993). The optimized geometry parameters of $HCOO^-$, CH_3COO^- and water are summarized in Tables 4.3-4.5.

Table 4.1 Stabilization energies of the three $HCOO^-$-H_2O and CH_3COO^--H_2O complexes, calculated by PW91, BLYP, B3LYP, HF, MP2 and CCSD methods using DZV+ and aug-cc-pVDZ (data in parentheses) basis sets.

	Stabilization energy (kcal/mol)					
Method	**Cyclic**		**Anti**		**Syn**	
	$HCOO^-$	CH_3COO^-	$HCOO^-$	CH_3COO^-	$HCOO^-$	CH_3COO^-
PW91	-23.17	-24.32	-21.79	-22.40	-20.05	-21.41
	(-19.89)	(-20.30)	(-18.07)	(-18.32)	(-17.06)	(-17.34)
BLYP	-21.37	-21.88	-19.82	-20.13	-18.08	-19.54
	(-17.30)	(-17.65)	(-16.00)	(-16.02)	(-15.99)	(-15.41)
B3LYP	-22.09	-22.65	-20.38	-20.74	-19.36	-20.16
	(-19.09)	(-18.48)	(-16.42)	(-16.54)	(-14.70)	(-15.88)
HF	-20.17	-20.66	-18.27	-18.14	-17.88	-18.30
	(-16.31)	(-16.62)	(-14.13)	(-14.22)	(-13.44)	(-13.93)
MP2	-21.33	-21.78	-18.78	-19.17	-17.36	-18.93
	(-19.22)	(-19.64)	(-16.75)	(-17.34)	(-15.38)	(-16.31)
CCSD	-21.16	-21.64	-18.52	-18.90	-18.14	-18.79
	(-18.80)	(-19.17)	(-16.25)	(-16.71)	(-15.00)	(-15.87)

Table 4.2 Hydrogen bond lengths of the three $HCOO^-$-H_2O and CH_3COO^--H_2O complexes, calculated by PW91, BLYP, B3LYP, HF, MP2 and CCSD methods using DZV+ and aug-cc-pVDZ (data in parentheses) basis sets.

| Method | Hydrogen bond length (Å) | | | | | |
| | Cyclic | | Anti | | Syn | |
	$HCOO^-$	CH_3COO^-	$HCOO^-$	CH_3COO^-	$HCOO^-$	CH_3COO^-
PW91	1.979	1.954	1.543	1.519	1.706	1.586
	(1.953)	(1.924)	(1.604)	(1.589)	(1.539)	(1.651)
BLYP	2.033	2.013	1.585	1.571	1.738	1.636
	(2.021)	(1.994)	(1.652)	(1.640)	(1.658)	(1.700)
B3LYP	2.021	2.001	1.581	1.568	1.704	1.627
	(2.006)	(1.983)	(1.650)	(1.639)	(1.768)	(1.700)
HF	2.137	2.121	1.706	1.700	1.748	1.744
	(2.132)	(2.115)	(1.787)	(1.777)	(1.806)	(1.835)
MP2	2.097	2.080	1.684	1.671	1.839	1.733
	(1.995)	(1.971)	(1.659)	(1.641)	(1.752)	(1.710)
CCSD	2.111	2.097	1.708	1.693	1.766	1.754
	(2.062)	(2.002)	(1.691)	(1.678)	(1.761)	(1.740)

Table 4.3 Optimized geometry parameters of HCOO⁻, calculated by HF and B3LYP

methods using aug-cc-pVDZ basis set.
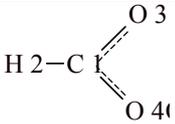
	Method	
O 3 H 2−C 1 O 4(	**HF**	**B3LYP**
$R_{1\text{-}2}$ (Å)	1.1249	1.1417
$R_{1\text{-}3}$ (Å)	1.2652	1.2579
$R_{1\text{-}4}$ (Å)	1.2352	1.2579
$A_{2\text{-}1\text{-}3}$ (°)	114.8641	114.8915
$A_{2\text{-}1\text{-}4}$ (°)	114.8641	114.8915
$A_{3\text{-}1\text{-}4}$ (°)	130.2718	130.2170

Table 4.4 Optimized geometry parameters of CH_3COO^-, calculated by HF and B3LYP methods using aug-cc-pVDZ basis set.
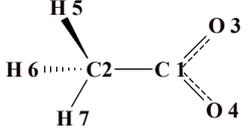
	Method	
	HF	**B3LYP**
$R_{1\text{-}2}$ (Å)	1.5448	1.5597
$R_{1\text{-}3}$ (Å)	1.2381	1.2615
$R_{1\text{-}4}$ (Å)	1.2395	1.2621
$R_{2\text{-}5}$ (Å)	1.0906	1.0998
$R_{2\text{-}6}$ (Å)	1.0933	1.1021
$R_{2\text{-}7}$ (Å)	1.0933	1.10521
$A_{2\text{-}1\text{-}3}$ (°)	116.4148	116.3318
$A_{2\text{-}1\text{-}4}$ (°)	115.0551	115.1272
$A_{3\text{-}1\text{-}4}$ (°)	128.5300	128.5410
$A_{5\text{-}2\text{-}6}$ (°)	109.6132	109.5185
$A_{5\text{-}2\text{-}7}$ (°)	109.6132	109.5185
$A_{5\text{-}2\text{-}1}$ (°)	111.5998	111.5250
$A_{6\text{-}2\text{-}7}$ (°)	107.1271	107.0180
$A_{6\text{-}2\text{-}1}$ (°)	109.3906	109.5768
$A_{7\text{-}2\text{-}1}$ (°)	109.3606	109.5768

The structure diagram shows: H 5, O 3, H 6 C2—C 1, H 7, O 4

Table 4.5 Optimized geometry parameters of H_2O, calculated by PW91, BLYP, B3LYP, HF, MP2 and CCSD methods using DZV+ and aug-cc-pVDZ (data in parentheses) basis sets.
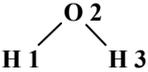
O 2 / H 1 H 3 Water	Method					
	PW91	**BLYP**	**B3LYP**	**HF**	**MP2**	**CCSD**
$R_{1\text{-}2}$ (Å)	0.985	0.989	0.977	0.951	0.979	0.980
	(0.972)	(0.975)	(0.964)	(0.943)	(0.965)	(0.964)
$R_{2\text{-}3}$ (Å)	0.985	0.989	0.977	0.951	0.979	0.980
	(0.972)	(0.975)	(0.964)	(0.943)	(0.965)	(0.964)
$A_{1\text{-}2\text{-}3}$ (°)	109.60	109.39	110.61	112.64	110.93	110.56
	(103.97)	(104.19)	(104.69	(105.92)	(103.83)	(104.09)

4.2   Construction of pair potential functions

In the present study, the pair potential functions for describing ion-water interactions are newly constructed. A flexible BJH-CF2 model, which describes intermolecular (Stillinger and Rahman, 1978) and intramolecular interactions (Bopp, Jancso, and Heinzinger, 1983) was employed for water. This model has been successfully applied for various ions in aqueous solutions (Tongraar, Hannongbua, and Rode, 1997; Remsungnen and Rode, 2003; Rode, Schwenk, and Tongraar, 2004). Of particular useful, this flexible model allows explicit hydrogen movements which ensure a smooth transition of water molecules when they move from the QM region with its full flexibility to the MM region and *vice versa*. The charges on O and H

atoms of water were employed from the CF2 model (Bopp, Jancso, and Heinzinger, 1983) as -0.6598 and 0.3299, respectively.

4.2.1 HCOO⁻-H₂O potential functions

The HCOO⁻-water configurations can be generated by moving a water molecule with respect to the variation of θ (*i.e.*, between 0° and 90°) and Φ (*i.e.*, between 0° and 180°), as depicted in Figure 4.3. With respect to the carbon atom of HCOO⁻, 12 different types of water orientations are obtained (see Figure 4.4).
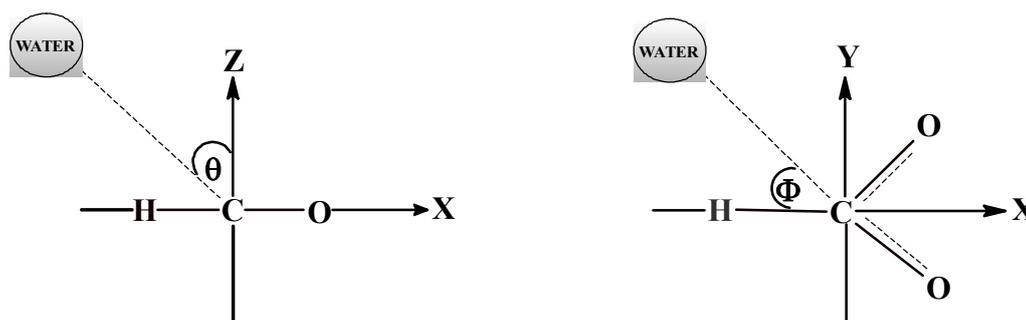
Figure 4.3 Movements of H₂O around HCOO⁻.
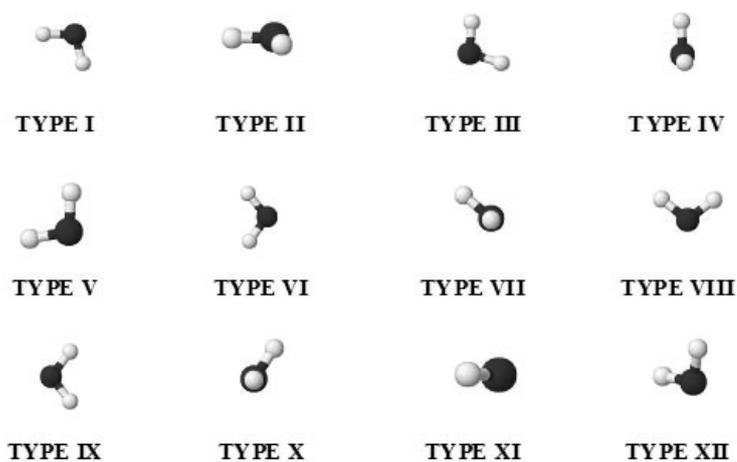
Figure 4.4 Variation of water's orientations.

The 6,015 HF and 5,696 B3LYP interaction energy points for various

HCOO⁻-H₂O configurations, obtained from Gaussian98 (Frisch et al., 2004)

calculations using aug-cc-pVDZ basis set, were fitted to the following analytical

forms

$$\Delta E^{HF}_{HCOO^- - H_2O} = \sum_{i=1}^{4} \sum_{j=1}^{3} \left[\frac{A}{r_{ij}^4} + \frac{B}{r_{ij}^8} + C \exp(-Dr) + \frac{q_i q_j}{r_{ij}} \right] \qquad (4.1)$$

and

$$\Delta E^{B3LYP}_{HCOO^- - H_2O} = \sum_{i=1}^{4} \sum_{j=1}^{3} \left[\frac{A}{r_{ij}^4} + \frac{B}{r_{ij}^5} + C \exp(-Dr) + \frac{q_i q_j}{r_{ij}} \right], \qquad (4.2)$$

respectively, where A, B, C and D are the optimized parameters (see Table 4.6.), r_{ij}

denotes the distances between the i-th atom of HCOO⁻ and the j-th atom of water

molecule and q are atomic net charges.

Table 4.6 Optimized parameters of the analytical pair potential for the interactions of water with HCOO⁻ (Interaction energies in kcal mol⁻¹ and distances in Å).

Pair	*A*	*B*	*C*	*D*
	HF-based method			
	$(kcal\ mol^{-1}\ Å^{-4})$	$(kcal\ mol^{-1}\ Å^{-8})$	$(kcal\ mol^{-1})$	$(Å^{-1})$
C-O$_W$	5.161390×10^5	-3.019987×10^4	-7.559447×10^5	1.955784
O-O$_W$	-2.665243×10^5	-5.276995×10^6	6.778200×10^7	3.556668
H-O$_W$	3.248683×10^5	-4.992158×10^5	-1.720448×10^5	1.318893
C-H$_W$	-2.368418×10^5	4.358412×10^5	7.864110×10^5	2.016779
O-H$_W$	5.356162×10^4	1.264105×10^4	-5.753688×10^5	3.105469
H-H$_W$	-3.889529×10^4	6.586156×10^3	1.236351×10^6	3.485256
	B3LYP-based method			
	$(kcal\ mol^{-1}\ Å^{-4})$	$(kcal\ mol^{-1}\ Å^{-5})$	$(kcal\ mol^{-1})$	$(Å^{-1})$
C-O$_W$	4.136774×10^5	-2.883495×10^5	-1.557409×10^4	0.429679
O-O$_W$	1.336006×10^6	-5.599642×10^6	5.529706×10^7	3.128274
H-O$_W$	9.683479×10^5	-9.298010×10^5	-3.549124×10^5	1.323072
C-H$_W$	2.720650×10^3	5.285593×10^4	2.761879×10^4	1.427650
O-H$_W$	-1.358645×10^5	1.516713×10^5	9.597746×10^4	1.431782
H-H$_W$	-8.338373×10^4	4.962553×10^4	7.038974×10^5	2.987790

The charges on C, H and O atoms of HCOO⁻ were obtained from the Natural Bond Orbital (NBO) of the corresponding HF and B3LYP calculations, as summarized in Table 4.7.

Table 4.7 The NBO charges on C, H and O atoms of HCOO⁻.

Atom	Charge (in atomic unit)	
	HF	**B3LYP**
C	0.86354	0.65834
H	-0.02340	-0.01902
O	-0.92007	-0.81966

4.2.2 CH₃COO⁻-H₂O potential functions

Similar to to construction of HCOO⁻-H₂O pair potentials, the CH₃COO⁻-H₂O configurations are generated by moving a water molecule with respect to the variation of θ (*i.e.*, between 0° and 360°) and Φ (*i.e.*, between 0° and 180°), as depicted in Figure 4.5. This leads to 22 different types of water orientations (see Figure 4.6).
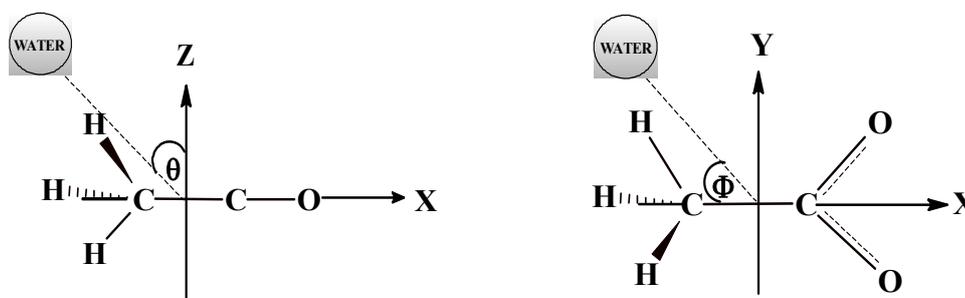
Figure 4.5 Movements of H₂O around CH₃COO⁻.
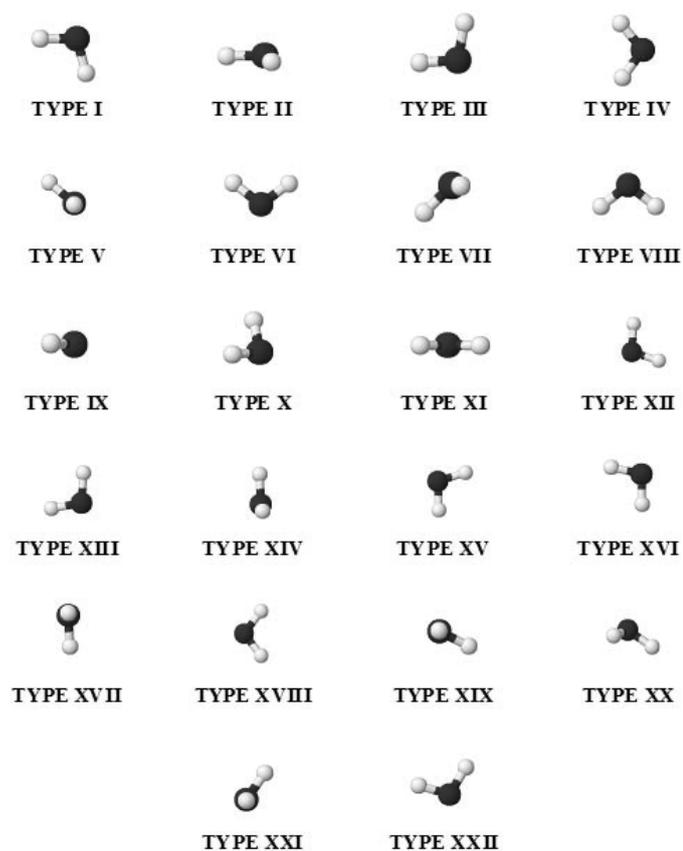
Figure 4.6 Variation of water's orientations.

The total of 24,637 HF and 22,038 B3LYP interaction energy points for various CH_3COO^--water configurations, obtained from Gaussian98 (Frisch et al., 2004) calculations, were fitted with respect to the analytical forms of

$$\Delta E^{HF}_{CH_3COO^- - H_2O} = \sum_{i=1}^{7} \sum_{j=1}^{3} \left[\frac{A}{r_{ij}^4} + \frac{B}{r_{ij}^5} + \frac{C}{r_{ij}^6} + \frac{D}{r_{ij}^{12}} + \frac{q_i q_j}{r_{ij}} \right] \qquad (4.3)$$

and

$$\Delta E^{B3LYP}_{CH_3COO^- - H_2O} = \sum_{i=1}^{7} \sum_{j=1}^{3} \left[\frac{A}{r_{ij}^4} + \frac{B}{r_{ij}^5} + \frac{C}{r_{ij}^6} + \frac{D}{r_{ij}^{12}} + \frac{q_i q_j}{r_{ij}} \right], \qquad (4.4)$$

respectively. All optimized parameters and the corresponding NBO charges are summarized in Tables 4.8 and 4.9.

Table 4.8 Optimized parameters of the analytical pair potential for the interactions of water with CH_3COO^- (Interaction energies in kcal mol^{-1} and distances in Å).

Pair	A kcal mol^{-1} Å^4	B kcal mol^{-1} Å^5	C kcal mol^{-1} Å^6	D kcal mol^{-1} Å^{12}
HF-based method				
C_{methyl}-O_W	6.25575 x 10^6	-2.80636 x 10^7	3.66359 x 10^7	-2.11546 x 10^8
$C_{carbonyl}$-O_W	-3.24728 x 10^6	1.87939 x 10^7	-2.86738 x 10^7	4.74720 x 10^8
O-O_W	6.65129 x 10^5	-6.72760 x 10^6	1.30700 x 10^7	-6.88407 x 10^7
H-O_W	-5.70606 x 10^5	1.68140 x 10^6	-1.08799 x 10^6	-3.39850 x 10^5
C_{methyl}-H_W	-1.07945 x 10^6	3.37053 x 10^6	-2.79560 x 10^6	1.72989 x 10^6
$C_{carbonyl}$-H_W	2.98361 x 10^5	-7.82658 x 10^5	7.53084 x 10^5	-5.07991 x 10^5
O-H_W	6.58235 x 10^4	-2.19426 x 10^5	2.21554 x 10^5	-4.43483 x 10^4
H-H_W	1.87543 x 10^4	-1.39169 x 10^3	-6.36335 x 10^3	5.11901 x 10^2
B3LYP-based method				
C_{methyl}- O_W	6.09669 x 10^6	-2.78939 x 10^7	3.65314 x 10^7	-2.00914 x 10^8
$C_{carbonyl}$- O_W	-4.90291 x 10^6	2.45331 x 10^7	-3.36074 x 10^7	3.84037 x 10^8
O- O_W	7.97158 x 10^5	-7.60038 x 10^6	1.43174 x 10^7	-8.17032 x 10^7
H- O_W	-5.56079 x 10^5	1.77187 x 10^6	-1.27320 x 10^6	-1.87559 x 10^4
C_{methyl}-H_W	-7.51684 x 10^5	2.37607 x 10^6	-1.91117 x 10^6	8.07782 x 10^5
$C_{carbonyl}$-H_W	1.02899 x 10^6	-3.38311 x 10^6	3.26716 x 10^6	-2.67571 x 10^6
O-H_W	6.93278 x 10^4	-2.68761 x 10^5	2.65680 x 10^5	-4.26558 x 10^4
H-H_W	-3.34025 x 10^4	8.37539 x 10^4	-4.57267 x 10^4	1.07909 x 10^3

Table 4.9 The NBO charges on C-methyl, H, C-carbonyl and O atoms of CH_3COO^-.

Atom	Charge (in atomic unit)	
	HF	B3LYP
C-methyl	-0.66444	-0.75055
H	0.17782	0.20281
C-carbonyl	0.97981	0.79449
O	-0.924415	-0.82618

4.2.3 Quality of the pair potential functions

To determine the quality of the fitted pair potentials, the stabilization energies obtained by the fitted functions (ΔE_{FIT}) are compared to the corresponding stabilization energies of the HF (ΔE_{HF}) and B3LYP (ΔE_{B3LYP}) single point calculations. The observed difference in stabilization energies are in acceptable range, as can be seen in Figures 4.5-4.8. In addition, the stabilization energies with basis set superposition error (BSSE) correction (ΔE_{BSSE}) are also given for comparison. It is obvious that the BSSE is small, *i.e.*, all the $HCOO^-$-H_2O and CH_3COO^--H_2O interactions energies are reliable enough can be calculated without the BSSE correction.
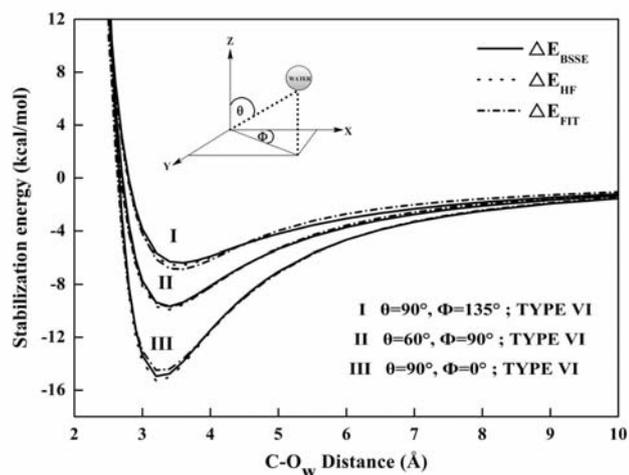
Figure 4.7 Comparison of the stabilization energies of HCOO⁻-H₂O complex, as obtained from the HF calculations without and with basis set superposition error (BSSE) correction, ΔE_{HF} and ΔE_{BSSE}, and from the fitted potential function, ΔE_{FIT}, using the parameters given in Table 4.6 for some values of θ and Φ.
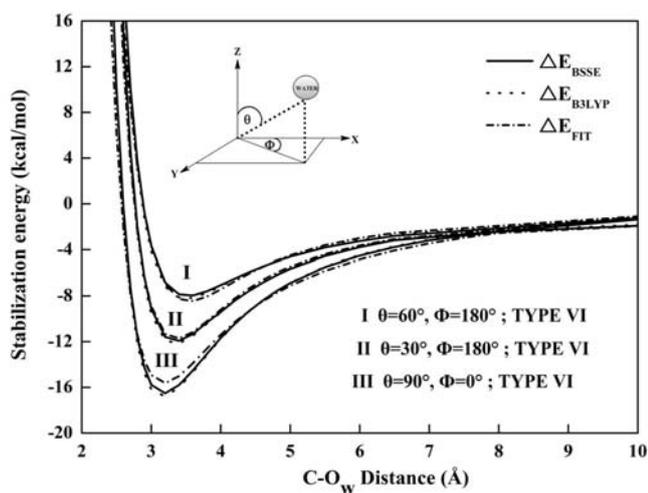
Figure 4.8 Comparison of the stabilization energies of HCOO⁻-H₂O complex, as obtained from the B3LYP calculations without and with basis set superposition error (BSSE) correction, ΔE_{B3LYP} and ΔE_{BSSE}, and from the fitted potential function, ΔE_{FIT}, using the parameters given in Table 4.6 for some values of θ and Φ.
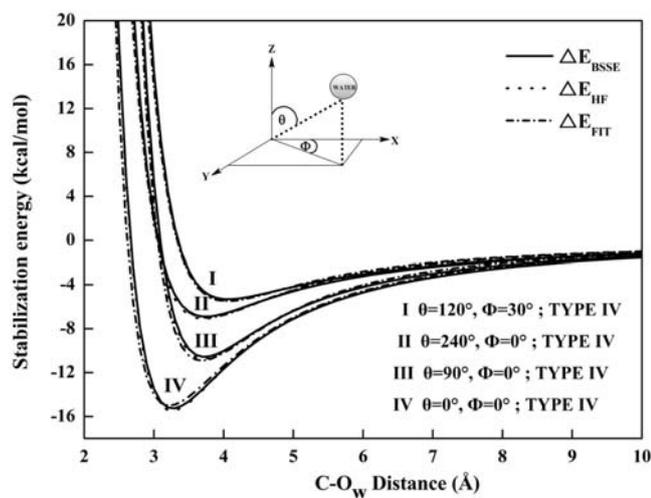
Figure 4.9 Comparison of the stabilization energies of CH_3COO^--H_2O complex, as obtained from the HF calculations without and with basis set superposition error (BSSE) correction, ΔE_{HF} and ΔE_{BSSE}, and from the fitted potential function, ΔE_{FIT}, using the parameters given in Table 4.8 for some values of θ and Φ.
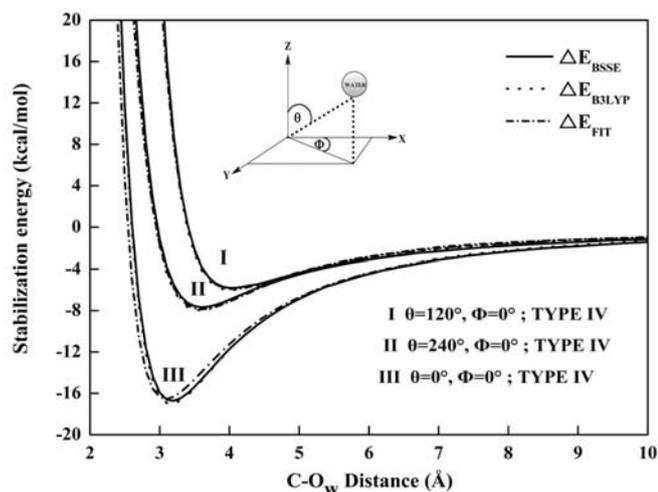
Figure 4.10 Comparison of the stabilization energies of CH_3COO^--H_2O complex, as obtained from the B3LYP calculations without and with basis set superposition error (BSSE) correction, ΔE_{B3LYP} and ΔE_{BSSE}, and from the fitted potential function, ΔE_{FIT}, using the parameters given in Table 4.8 for some values of θ and Φ.

4.3   Selection of QM size

To define the size of QM region, preliminary HF/MM and B3LYP/MM simulations in which only the ion was treated quantum mechanically using HF and B3LYP methods while the rest of the system was described by classical pair potentials, were performed. According to the resulting $C-O_w$ radial distribution functions (RDFs) for aqueous $HCOO^-$ system, as depicted in Figure 4.11a, both HF/MM and B3LYP/MM simulations show first $C-O_w$ minima at around 5.0 Å, where an integrations up to this $C-O_w$ distance yielded about 18-20 water molecules. This implied that a QM size with a radius of 5.0 Å seemed to be desirable for the present study. However, the evaluation of QM forces for all particles within this QM size is still beyond the limit of our computational facility. Therefore, a smaller QM size with diameter of 7.6 Å was chosen, *i.e.*, the values of r_1 and r_0 in equation (3.36) were set to 3.6 and 3.8 Å, which contains about 12-14 nearest-neighbor water molecules.

For the case of CH_3COO^-, the first minima of C_1-O_w RDFs of both HF/MM and B3LYP/MM simulations exhibit between 4.2-4.5 Å, as depicted in Figure 4.11b. The integrations up to this first minima yield about 7-10 water molecules. In this work, the QM size with diameter of 8.4 Å was selected for the system of CH_3COO^-, *i.e.*, r_1 and r_0 in equation (3.36) are 4.0 and 4.2, respectively.
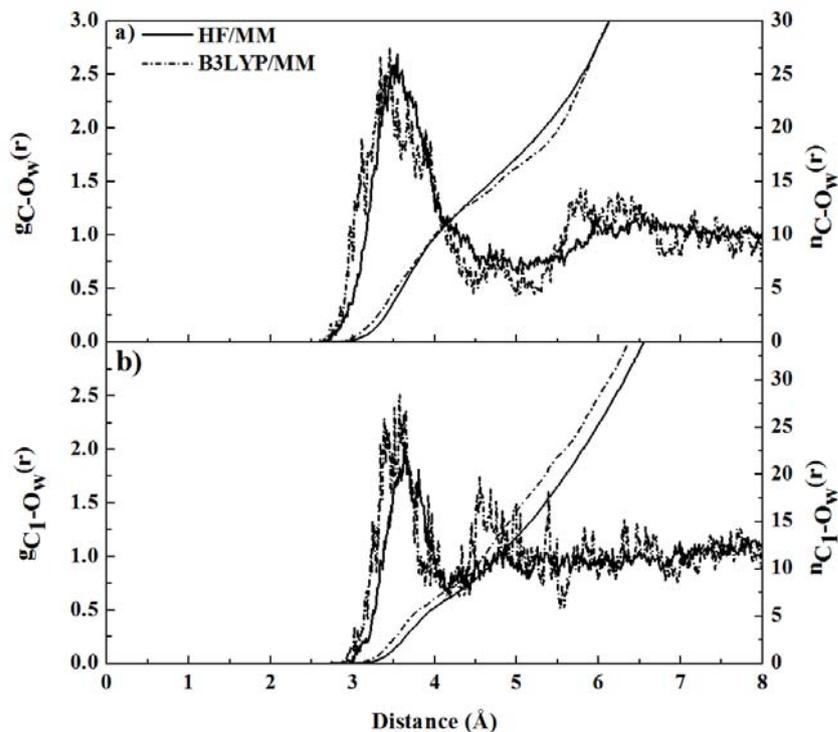
Figure 4.11 a) C-O$_w$ RDFs of aqueous HCOO$^-$ and b) C$_1$-O$_w$ RDFs of aqueous CH$_3$COO$^-$ and their corresponding integration numbers, as obtained from the preliminary HF/MM and B3LYP/MM simulations using DZV+ basis set.

4.4   Determination of structural properties

4.4.1   Correlation functions

Thermodynamic and structural properties of a liquid model depend on the correlation functions, especially the pair correlation function. Correlation functions measure the relationship between two variables. A variety of correlation functions can be defied, as

$$C_{xy} = \frac{1}{M}\sum_{i=1}^{M} x_i y_i, \tag{4.5}$$

which assumed that there are M value of x_i and y_i in the data sets.

Furthermore, correlation functions can be defined as a *cross-correlation function*, in which the quantities of x and y are different. On the other hand, if the x and y are the same quantities then the function is called an *autocorrelation function*.

4.4.2 Distribution functions

The distribution functions can provide detailed information on the structural properties, which measure the average value of a property as a function of an independent variable. The most general one, is the *radial distribution function* (RDF), considering one particle to be fix at r_1, then the probability of observing a second particle in $dr \equiv dr_{12}$ at r_2, about the fixed particles at r_1. Thus,

$$dN = \rho g(r) 4\pi r^2 dr, \tag{4.6}$$

where $g(r)$ is the radial distribution function, ρ is the mean density N/V, dN is the number of particles in the spherical shell of radius r and thickness dr and integration over the volume V as

$$4\rho\pi \int g(r) r^2 dr = N - 1. \tag{4.7}$$

The thermodynamic properties can be calculated using the radial distribution function such as the total energy when pairwise additivity of the forces is assumed, as given by

$$E = \frac{3}{2}Nk_BT + 2\pi N\rho \int_0^\infty r^2 v(r)g(r)dr, \tag{4.8}$$

and the pressure, as defined by

$$PV = Nk_BT - \frac{2\pi N\rho}{3k_BT}\int_0^\infty r^2 r \frac{dv(r)}{dr}g(r)dr. \tag{4.9}$$

4.5 Determination of dynamical properties

4.5.1 Water exchange processes

Detailed information on water exchange processes can be obtained from the calculation of *mean residence times* (MRTs), according to the direct method (Hofer, Tran, Schwenk, and Rode, 2004),

$$MRT(\tau) = \frac{CN \times t_{sim}}{N_{ex}}, \tag{4.10}$$

where CN is the coordination number, t_{sim} is the simulation time and N_{ex} is the number of exchange events. The time for observing the number of exchange water molecules was defined as t^*, being 0.0 ps for studying hydrogen bond lifetime and 0.5 ps for the estimation of ligand exchange processes. In this respect, the number of all transitions through a shell boundary ($N_{ex}^{0.0}$) and the number of changes persisting for 0.5 ps ($N_{ex}^{0.5}$) were calculated.

The mechanisms of water exchange processes can be classified into five ways as proposed by Langford and Gray (Langford and Gray, 1996). Associative exchange (A) prefer the external water molecule joins the complex, before the departing water leaves the complex. In opposite way, if a water molecule leaves the complex, before the external water joins the complex, this process is classified as dissociative exchange (D). In the case where the complex has the incoming and outgoing water molecule at the same time, the mechanism is called interchange mechanism (I), which can be subdivided into two classes depending on the exchange as associative-like (I_a) or dissociative-like (I_d) (see Figure 4.12).
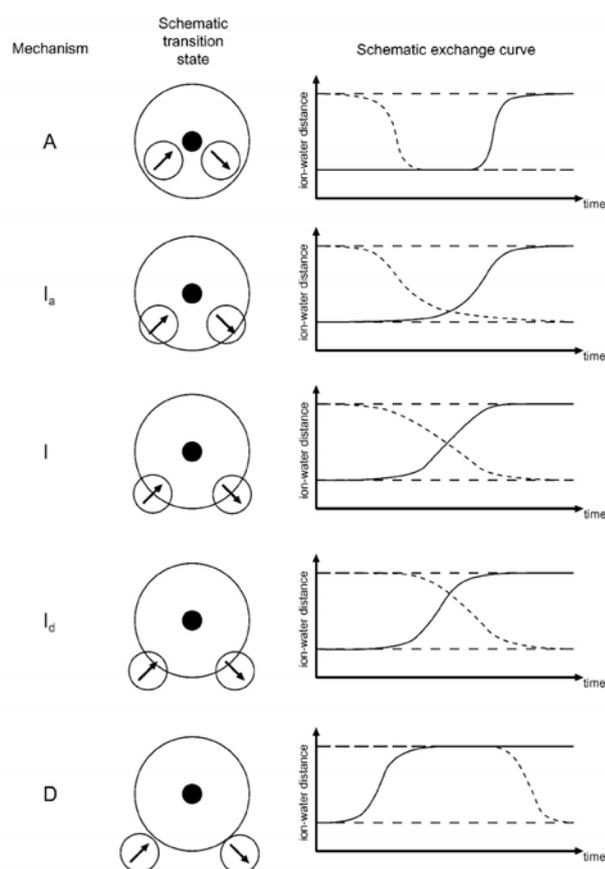
Figure 4.12 Representation of transition state of the five exchange classes (Spångberg, Rey, Hynes, and Hermansson, 2003).

4.6   Simulation details

All QM/MM MD simulations were performed in a canonical ensemble at 298 K with periodic boundary conditions. The system's temperature was kept constant using the Berendsen algorithm (Berendren, Postma, van Gunsteren, and Hermans, 1981). A periodic box, with a box length of 18.17 Å, contains one ion and 199 water molecules, corresponding to the experimental density of pure water. The reaction-field method (Adams, Adams, and Hills, 1979) was employed for the treatment of long-range interactions. The Newtonian equations of motions were treated by a general predictor-corrector algorithm. The time step size was set to 0.2 fs, which allows for the explicit movement of the hydrogen atoms of water molecules. The system was initially equilibrated by performing HF/MM and B3LYP/MM MD simulations for 20,000 time steps. Then, the HF/MM and B3LYP/MM simulations were continued for 200,000 time steps, to collect configurations every 10th step.

4.7   References

Adams, D. J., Adams, E. M. and Hills, G. J. (1979). The computer simulation of polar liquids. **Molecular Physics**. 38: 387-400.

Berendren, H., Postma, J., van Gunsteren, W. and Hermans, J. (1981). **Interaction models for water in relation to proteins hydration. In: Pullman B (ed) Intermolecular forces**. Dordrecht: Reidel.

Bopp, P., Jancso, G. and Heinzinger, K. (1983). An improved potential for non-rigid water-molecules in the liquid-phase. **Chemical Physics Letters**. 98: 129-133.

Dunning, T. H. J. (1989). Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. **The Journal of Chemical Physics**. 90: 1007-1023.

Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Zakrewski, V. G., Montgomery, J. A., Stratmann, R. E., Burant, J. C., Dapprich, S., Millam, J. M., Daniels, A. D., Kudin, K. N., Strain, M. C., Farkas, O., Tomasi, J., Barone, V., Cossi, M., Cammi, R., Mennucci, B., Pomelli, C., Adamo, C., Clifford, S., Ochterski, J., Peterson, G. A., Ayala, P. Y., Cui, Q., Morokuma, K., Malick, D. K., Rabuck, A. D., Raghavachari, K., Foresman, J. B., Cioslowski, J., Ortiz, J. V., Stefanov, B. B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Gomperts, R., Martin, R. L., Fox, D. J., Keith, T., Al-Laham, M. A., Peng, C. Y., Nanayakkara, A., Gonzalez, C., Challacombe, M., Gill, P. M. W., Johnson, B. G., Chen, W., Wong, M. W., Andres, J. L., Head-Gordon, M., Replogle, E. S., Pople, J. A. (1998). **Gaussian 98** package (revision A.7) [Computer Software]. Pittsburgh, PA, USA: Gaussian.

Hofer, T. S., Tran, H. T., Schwenk, C. F. and Rode, B. M. (2004). Characterization of dynamics and reactivities of solvated ions by *ab initio* simulations. **Journal of Computational Chemistry**. 25: 211-217.

Intharathep, P., Tongraar, A. and Sagarik, K. (2005). Structure and dynamics of hydrated NH_4^+: An *ab initio* QM/MM molecular dynamics simulation. **Journal of Computational Chemistry**. 26: 1329-1338.

Kendall, R. A., Jr., T. H. D. and Harrison, R. J. (1992). Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. **The Journal of Chemical Physics**. 96: 6796-6806.

Kerdcharoen, T., Liedl, K. R. and Rode, B. M. (1996). A QM/MM simulation method applied to the solution of Li^+ in liquid ammonia. **Chemical Physics**. 211: 313-323.

Langford, C. H. and Gray, H. B. (1996). **Ligand Substitution Processes**. New York: W.A. Benjamin.

Markham, G. D., Trachtman, M., Bock, C. L. and Bock, C. W. (1998). The binding of water to the carboxylate group in $R-CO_2^-$ (R = H, CH_3, NH_2, OH, and F): An *ab initio* molecular orbital study. **Journal of Molecular Structure : THEOCHEM**. 455: 239-256.

Remsungnen, T. and Rode, B. M. (2003). QM/MM molecular dynamics simulation of the structure of hydrated Fe(II) and Fe(III) ions. **The Journal of Physical Chemistry A**. 107: 2324-2328.

Rode, B. M., Schwenk, C. F. and Tongraar, A. (2004). Structure and dynamics of hydrated ions-new insights through quantum mechanical simulations. **Journal of Molecular Liquids**. 110: 105-122.

Spångberg, D., Rey, R., Hynes, J. T. and Hermansson, K. (2003). Rate and mechanisms for water exchange around Li^+(aq) from MD simulations. **The Journal of Physical Chemistry B**. 107: 4470-4477.

Stillinger, F. H. and Rahman, A. (1978). Revised central force potentials for water. **The Journal of Chemical Physics**. 68: 666-670.

Tongraar, A. and Rode, B. M. (2003). The hydration structures of F$^-$ and Cl$^-$ investigated by *ab initio* QM/MM molecular dynamics simulations. **Physical Chemistry Chemical Physics**. 5: 357-362.

Tongraar, A., Hannongbua, S. and Rode, B. M. (1997). Molecular dynamics simulations of a potassium ion and an iodide ion in liquid ammonia. **Chemical Physics**. 219: 279-290.

Tongraar, A., Liedl, K. R. and Rode, B. M. (1997). Solvation of Ca^{2+} in water studied by Born-Oppenheimer *ab initio* QM/MM dynamics. **The Journal of Physical Chemistry A**. 101: 6299-6309.

Tongraar, A., Liedl, K. R. and Rode, B. M. (1998). The hydration shell structure of Li$^+$ investigated by Born–Oppenheimer *ab initio* QM/MM dynamics **Chemical Physics Letters**. 286: 56-64

Tongraar, A., Sagarik, K. and Rode, B. M. (2002). Preferential solvation of Ca^{2+} in aqueous ammonia solution: Classical and combined *ab initio* quantum mechanical/molecular mechanical molecular dynamics simulations. **Physical Chemistry Chemical Physics**. 4: 628-634.

Tongraar, A., Tangkawanwanit, P. and Rode, B. M. (2006). A combined QM/MM molecular dynamics simulations study of nitrate anion (NO$_3^-$) in aqueous solution. **The Journal of Physical Chemistry A**. 110: 12918-12926.

Woon, D. E. and Thom H. Dunning, J. (1993). Gaussian basis sets for use in correlated molecular calculations. III. The atoms aluminum through argon. **The Journal of Chemical Physics**. 98: 1358-1371.

Xenides, D., Randolf, B. R. and Rode, B. M. (2006). Hydrogen bonding in liquid water: An *ab initio* QM/MM MD simulation study. **Journal of Molecular Liquids**. 123: 61-67.

CHAPTER V

RESULTS AND DISCUSSION

5.1   HCOO⁻ in water

5.1.1   Structural details

5.1.1.1 Radial distribution functions

The hydration structure of HCOO⁻ can be explained through the detailed analysis of the radial distribution functions (RDFs) and their corresponding integration numbers. In this context, the first atom in the RDFs refers to the atom of HCOO⁻, and the latter, with the subscript "w", represents the atom of water molecules. The C-O_w and C-H_w RDFs of HF/MM and B3LYP/MM simulations are plotted in Figure 5.1. The first C-O_w peaks exhibit at 3.50 and 3.42 Å for the HF/MM and B3LYP/MM simulations, respectively, as depicted in Figure 5.1a. The integrations up to their corresponding first minima yield about 16.0 and 17.0 water molecules, respectively. The features of the C-H_w RDFs are depicted in Figure 5.1b, showing consistency with the corresponding C-O_w peaks. The H-O_w and H-H_w RDFs are plotted in Figure 5.3. Both HF/MM and B3LYP/MM simulations supply information that the interactions between HCOO⁻ and its surrounding water molecules are weak.

The characteristics of hydrogen bonds between HCOO⁻ and water can be analyzed through the O-O_w and O-H_w RDFs, as depicted in Figures 5.3a and 5.3b. In the HF/MM simulation, the first O-O_w peak is located at 2.76 Å, where integration

up to the corresponding first $O-O_w$ minimum yields an average coordination number of 3.45. In the B3LYP/MM simulation, the first $O-O_w$ peak is found at the shorter distance of 2.67 Å, with a lower coordination number of 2.90. The second peaks in the $O-O_w$ RDFs are broad and less pronounced. These data correspond to the contributions of both bulk waters and water molecules in the first hydration layer of the other $HCOO^-$ oxygen.

In accordance with the $O-O_w$ RDFs, the HF/MM and B3LYP/MM simulations reveal first $O-H_w$ peaks, an indicative of $HCOO^-$-water hydrogen bonds, with maxima at 1.84 and 1.71 Å, respectively. Integrations up to the corresponding first $O-H_w$ minima give the average values of 3.14 and 2.71, respectively. With respect to both HF/MM and B3LYP/MM simulations, the observed numbers of water oxygen and hydrogen atoms suggest that the first-shell waters are linearly hydrogen bonded to each of the $HCOO^-$ oxygens, *i.e.*, they are acting as hydrogen-bond donors. In comparison to the first peak of pure water O_w-H_w RDFs (cf. Figure 5.4b), *e.g.*, in terms of shape and peak height, it is obvious that the $O\cdots H_w-O_w$ hydrogen bond interactions are relatively stronger. In the HF/MM and B3LYP/MM simulations, the closest $O\cdots H_w$ distances are 1.38 and 1.17 Å, respectively, compared to the shortest $O_w\cdots H_w$ distance of 1.45 Å for bulk water (Tongraar and Rode, 2004). The second peak in the $O-H_w$ RDFs near 3.2 Å can be assigned to the hydrogen atoms of first-shell waters that are not hydrogen-bonded to the $HCOO^-$ oxygens.

According to both HF/MM and B3LYP/MM simulations, each $HCOO^-$ oxygen atom is predicted to form more hydrogen bonds with surrounding water molecules than the experimental value of about 2.5 (Kameda, Mori, Nishiyama, Usuki, and Uemura, 1996; Kameda, Fukuhara, Mochiduki, Naganuma, Usuki, and Uemura,

2002). In fact, as the first minima peaks of O-O$_w$ and O-H$_w$ RDFs are rather broad, the number of first-shell waters is quite sensitive to the defined O-O$_w$ and O-H$_w$ minima. For example, integrations up to O-O$_w$ distance of about 0.2 Å shorter than the corresponding O-O$_w$ minima yield 2.53 and 2.47 water oxygens for the HF/MM and B3LYP/MM simulations, respectively. Thus, the differences between our data and experiment are well within the methodical limits of integration on the one hand and interpretation of experimental data on the other. The recent CPMD study (Leung and Rempe, 2004), although providing a coordination number apparently identical with experiment, delivers deviating results when a different simulation cell size as well as different functionals (*i.e.*, BLYP and PW91) are employed in the simulations. For example, the use of the BLYP functional had predicted a hydration number of 2.45 per HCOO$^-$ oxygen, while the PW91 method gave a large variation in the hydration numbers ranging from 2.12 to 2.66. Possible weaknesses of the density functional methods could be attributed to the incompleteness of the kinetic energy term, the self-interaction error and the more or less empirical parameterization, which did not contain any hydrogen-bonded system.

The distributions of oxygen and hydrogen atoms of first-shell waters, calculated with respect to the first minimum of the O-O$_w$ and O-H$_w$ RDFs, are depicted in Figures 5.5 and 5.6, respectively. In the HF/MM simulation, the most frequent hydration number per HCOO$^-$ oxygen atom is 3, followed by 4, 2 and 5 in decreasing amounts. In the B3LYP/MM simulation, a prevalent value of 3 is also observed, followed by 2 and 4 in smaller amounts. Figures 5.7 and 5.8 show examples of time dependence of the hydration number at each of HCOO$^-$ oxygens within the first 10 ps of the HF/MM and B3LYP/MM simulations, respectively. In both HF/MM

and B3LYP/MM trajectories, it is found that the two HCOO⁻ oxygen atoms simultaneously form asymmetric solvation shells, i.*e.*, each of them hydrogen bonding to different numbers of water molecules. Consequently, this causes numerous possible species of the HCOO⁻-water complexes to co-exist in aqueous solution. As can be seen in Figures 5.7 and 5.8, the total numbers of water molecules in the vicinity of HCOO⁻ oxygens show large fluctuations, ranging from 5 to 10 and from 4 to 8 for the HF/MM and B3LYP/MM simulations, respectively.
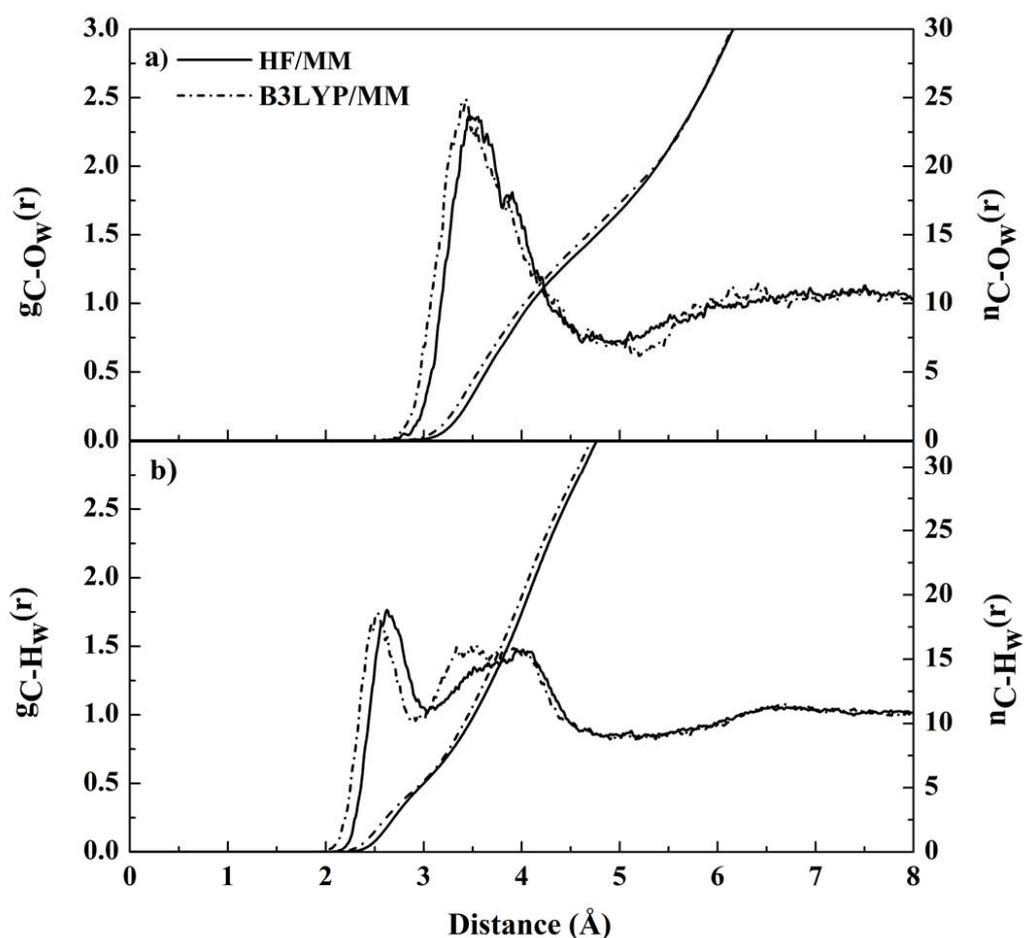
Figure 5.1 a) $C-O_w$ and b) $C-H_w$ RDFs and their corresponding integration numbers, as obtained by HF/MM and B3LYP/MM simulations using DZV+ basis set.
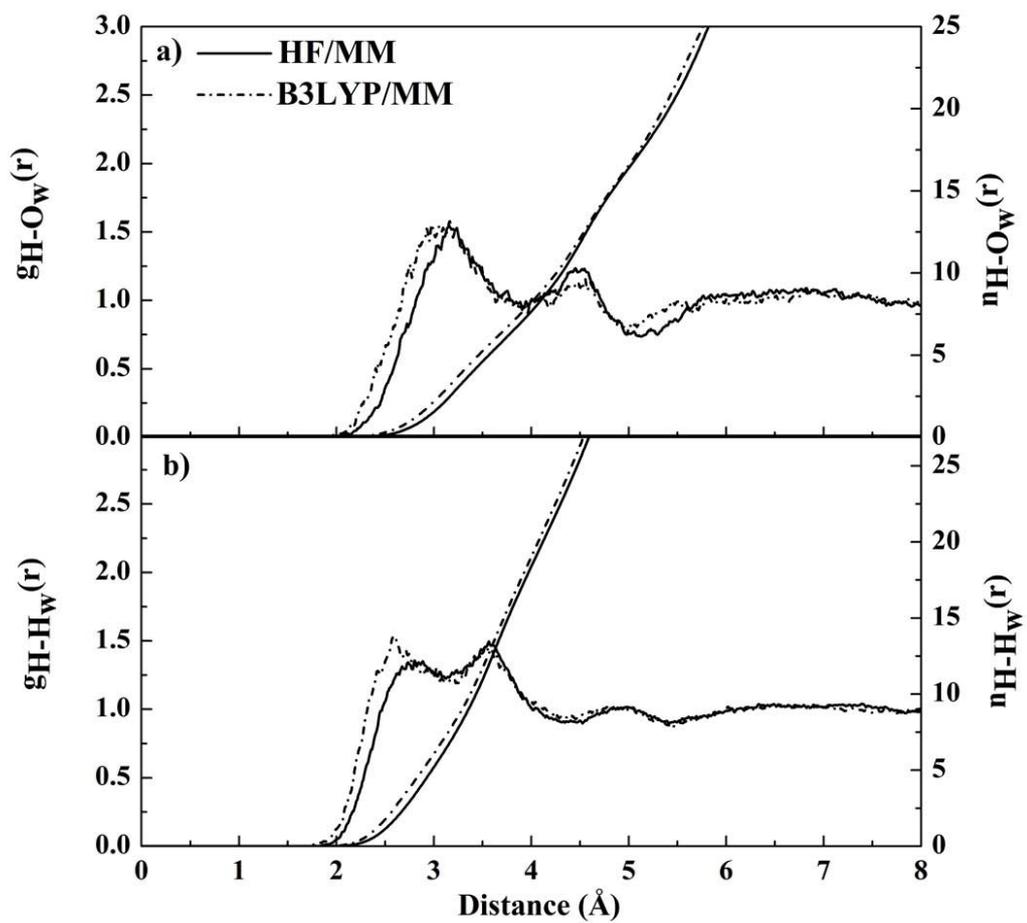
Figure 5.2 a) H-O_w and b) H-H_w RDFs and their corresponding integration numbers, as obtained by HF/MM and B3LYP/MM simulations using DZV+ basis set.
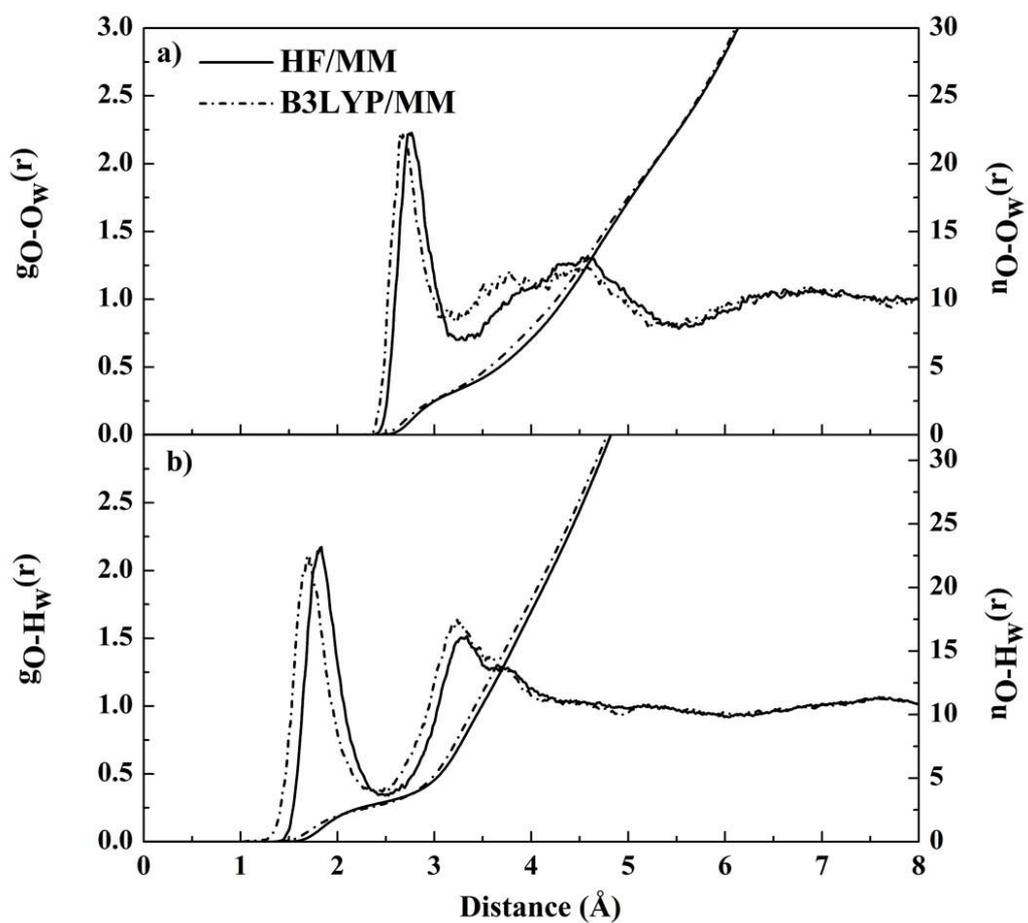
Figure 5.3 a) O-O_w and b) O-H_w RDFs and their corresponding integration numbers, as obtained by HF/MM and B3LYP/MM simulations using DZV+ basis set.
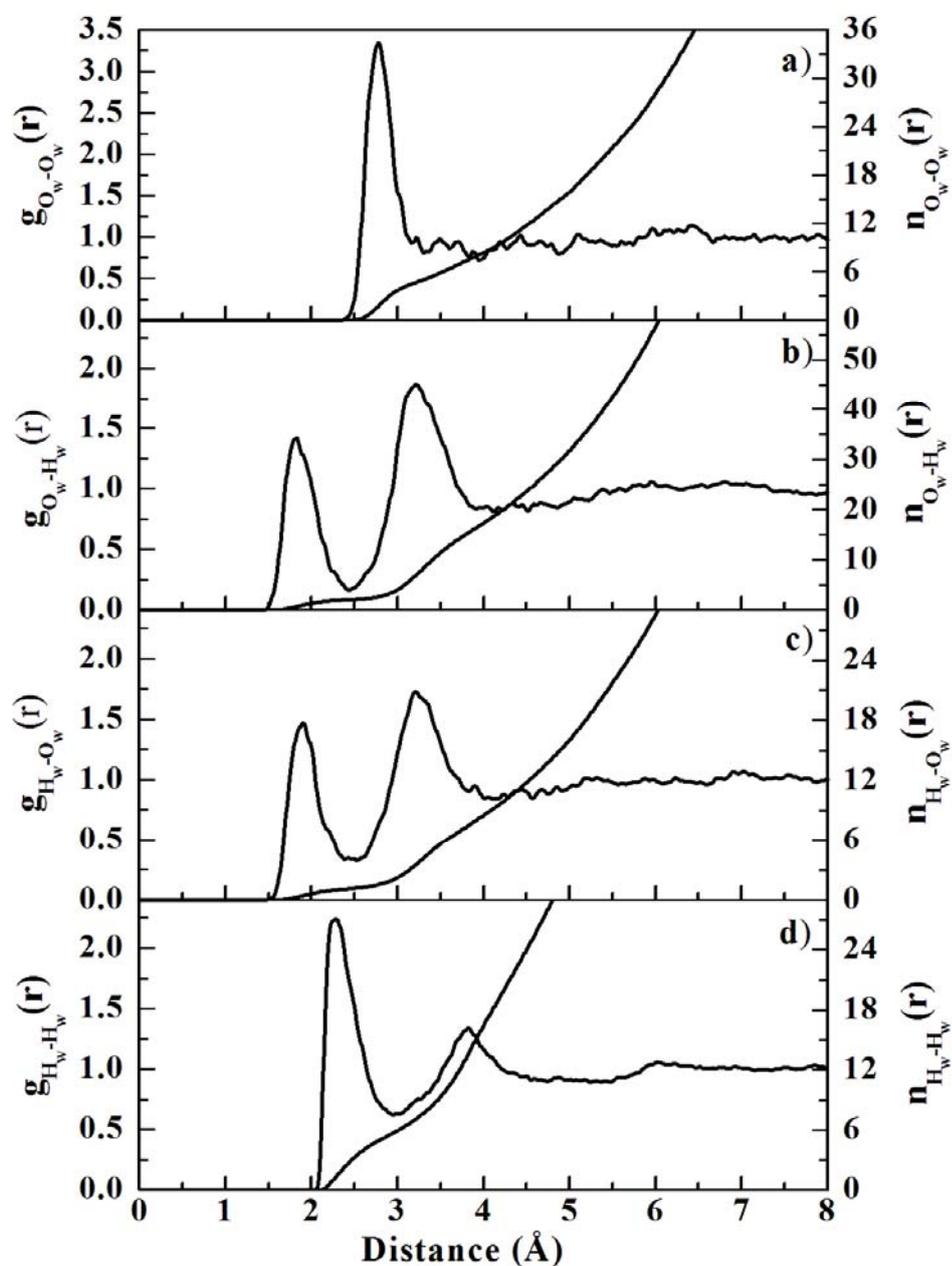
Figure 5.4 a) O_w-O_w, b) O_w-H_w, c) H_w-O_w and d) H_w-H_w RDFs and their corresponding integration numbers. The first atom of each pair refers to the atoms of the water molecule, whose oxygen position was defined as center of the QM region during the QM/MM simulation.
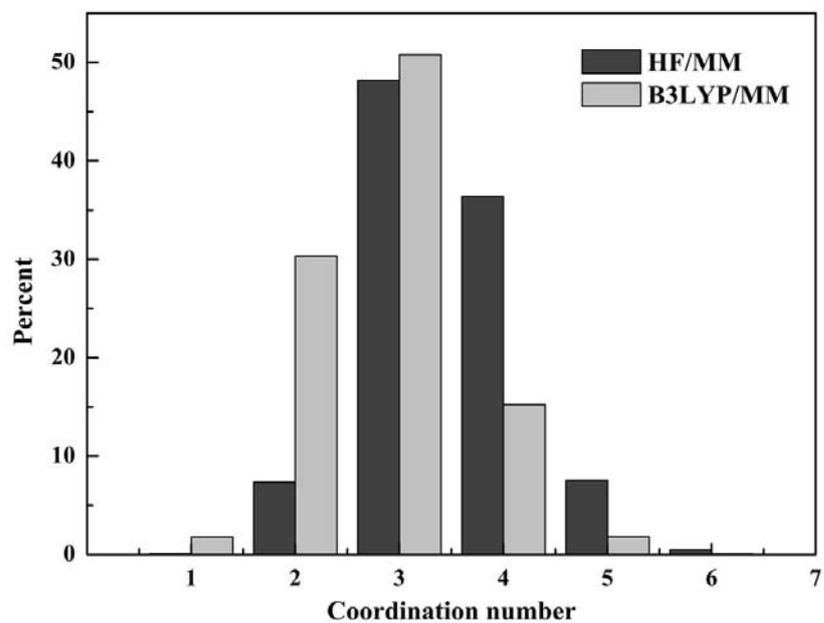
Figure 5.5 Distributions of the number of water's oxygen atoms at each of HCOO$^-$

oxygens, calculated within first minimum of the O-O$_w$ RDFs.
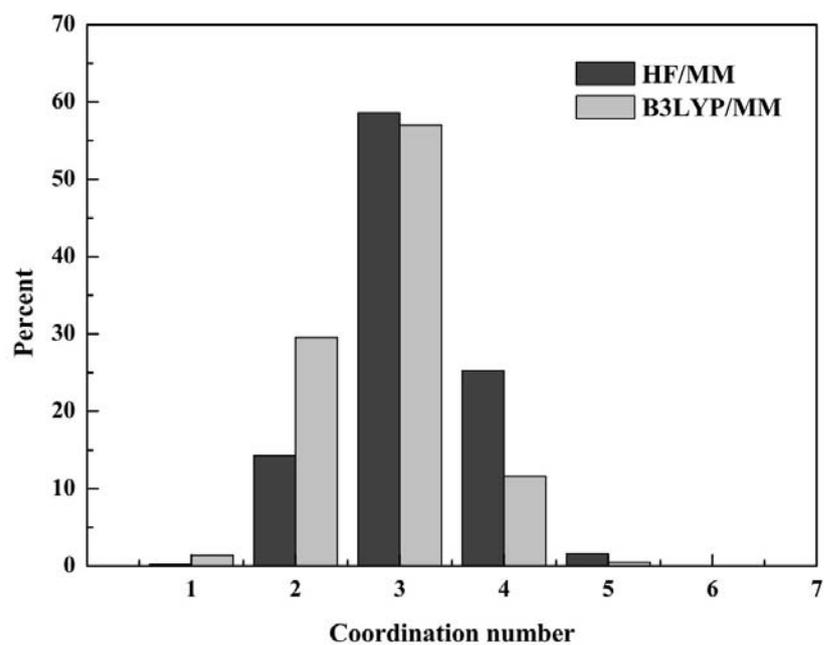
Figure 5.6 Distributions of the number of water's hydrogen atoms at each of HCOO$^-$

oxygens, calculated within first minimum of the O-H$_w$ RDFs.
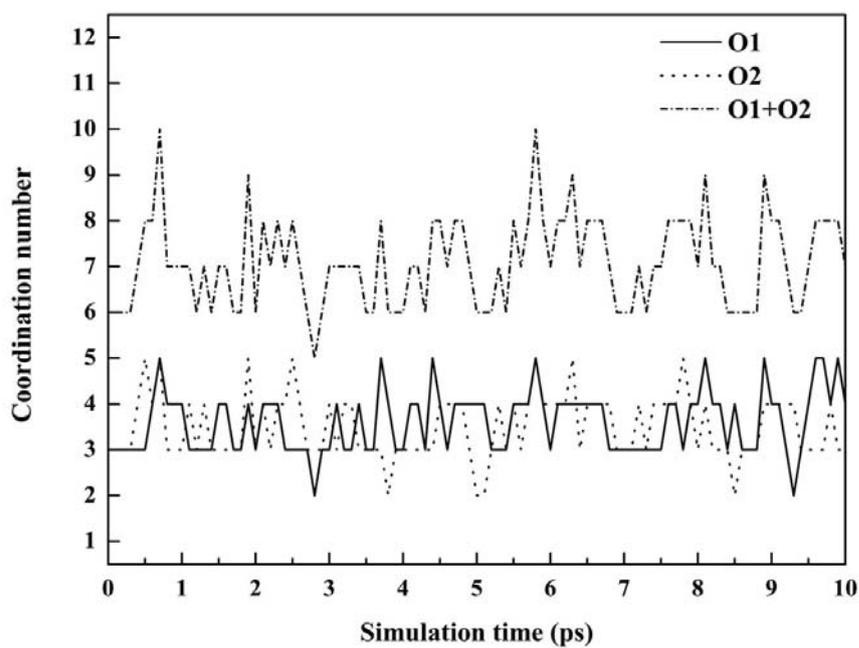
Figure 5.7 Time dependence of the number of first-shell waters at HCOO⁻ oxygen atoms, selecting only for first 10 ps of the HF/MM simulation.
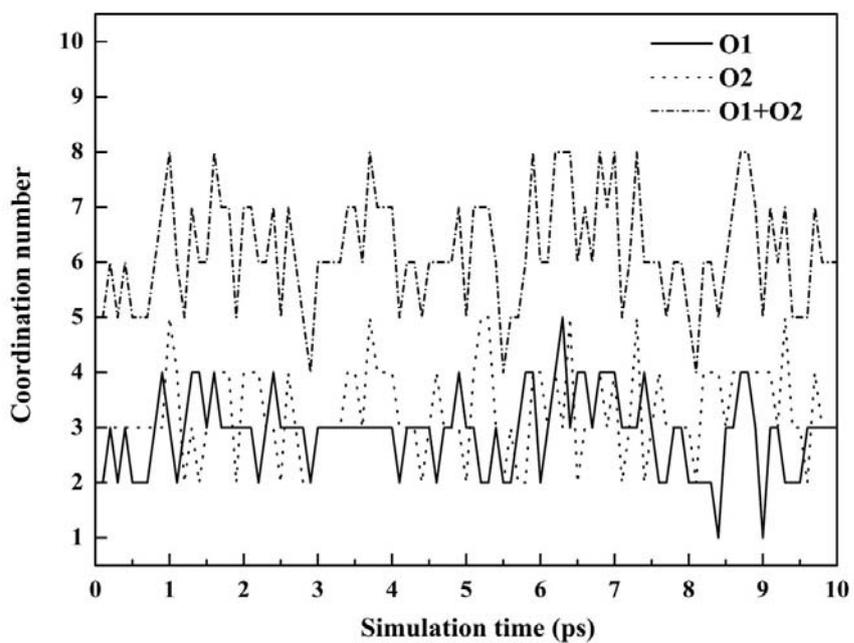
Figure 5.8 Time dependence of the number of first-shell waters at HCOO⁻ oxygen atoms, selecting only for first 10 ps of the B3LYP/MM simulation.

5.1.1.2 Angular distribution functions

A more detailed interpretation of the $HCOO^-$-water hydrogen bonds can be deduced from the probability distributions of the $C-O \cdots H_w$ and $O \cdots O_w-H_w$ angles, calculated within the first minimum of the $O-O_w$ RDFs, as shown in Figures 5.9 and 5.10, respectively. In the case that solvent effects would cause strong charge localization on $HCOO^-$, an asymmetrical charge distribution at the two $HCOO^-$ oxygens, corresponding to a formation of a C-O single and a C=O double bond, could exist in aqueous solution. With regard to this point, one could expect the arrangement of directional $C-O \cdots H_w$ hydrogen bonds that cause the $C-O \cdots H_w$ angle to peak at 109.5º and 120.0º depending on the type of involved oxygen. In both HF/MM and B3LYP/MM simulations, the observed broad $C-O \cdots H_w$ angular distributions (Figure 5.9) clearly suggest the absence of such phenomenon. In addition, with regard to the Mulliken charge analyses of several $HCOO^-$-water complexes, there is no substantial charge concentration at one of the $HCOO^-$ oxygens. Apparently, $HCOO^-$ adopts an electronically delocalized structure in aqueous solution, which may fluctuate due to solvent exchange processes. Figure 5.10 shows the distributions of $O \cdots O_w-H_w$ angle, which clarify the preference for linear $O \cdots H_w-O_w$ arrangements.
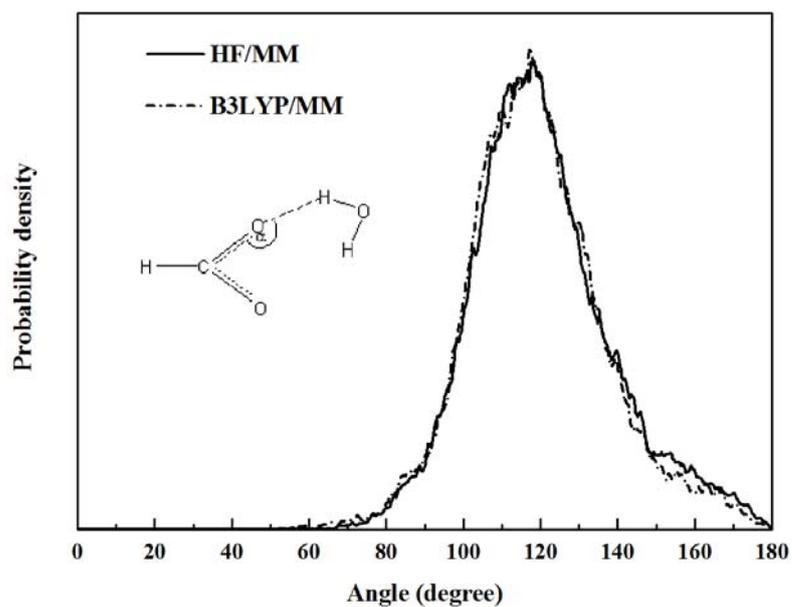
Figure 5.9 Distributions of C-O··H$_w$ angle, calculated within first minimum of the O-O$_w$ RDFs.
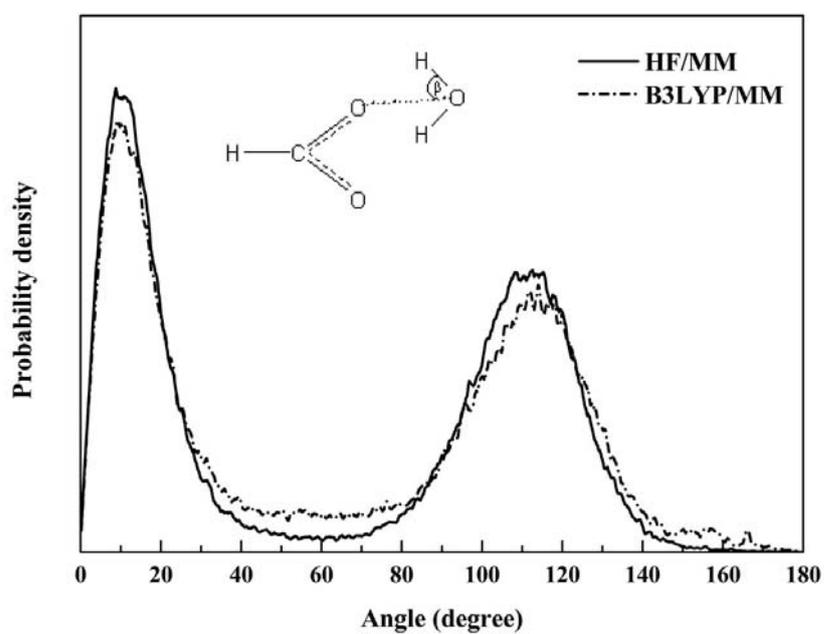
Figure 5.10 Distributions of O··O$_w$-H$_w$ angle, calculated within first minimum of the O-O$_w$ RDFs.

5.1.2 Dynamical Details

5.1.2.1 Intramolecular geometry of HCOO⁻

The geometrical arrangement of HCOO⁻ in aqueous solution is explained in terms of the distributions of C-O and C-H bond lengths as well as of H-C-O and O-C-O angles, as shown in Figures 5.11 and 5.12, respectively. In addition, the distribution of the angle ϕ, as defined by a vector along the C-H bond and a vector pointing outwards between the other two C-O bonds, is also given in Figure 5.13. As compared to the gas phase HCOO⁻ structure, both HF/MM and B3LYP/MM simulations clearly indicate a substantial change in the local structure of HCOO⁻ according to the influence of water environment, in particular a C-O bond lengthening, shortening of the C-H bond and a decrease of the O-C-O angle.
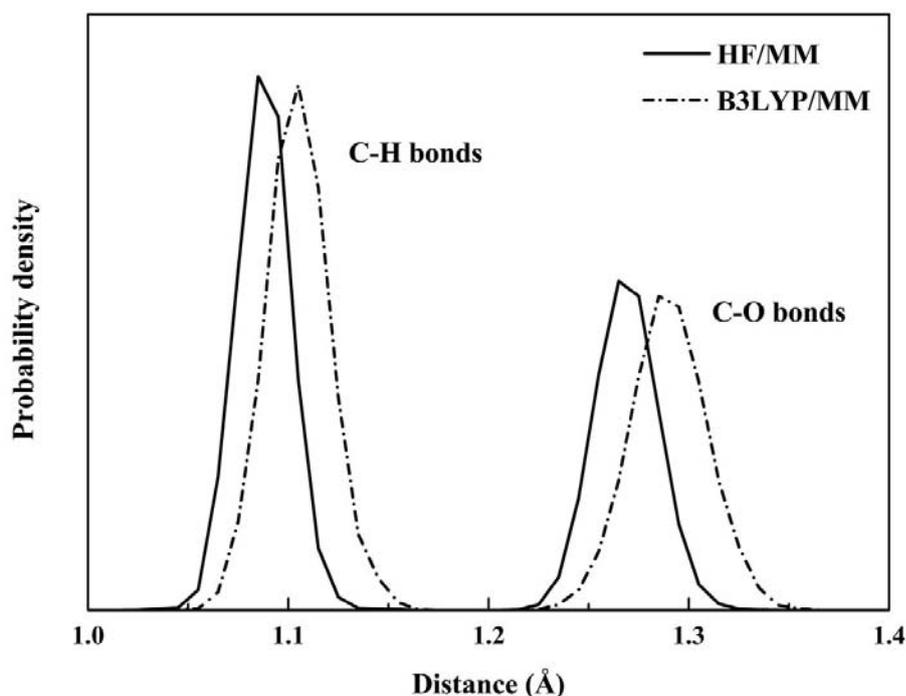
Figure 5.11 Distributions of C-H and C-O bond lengths.
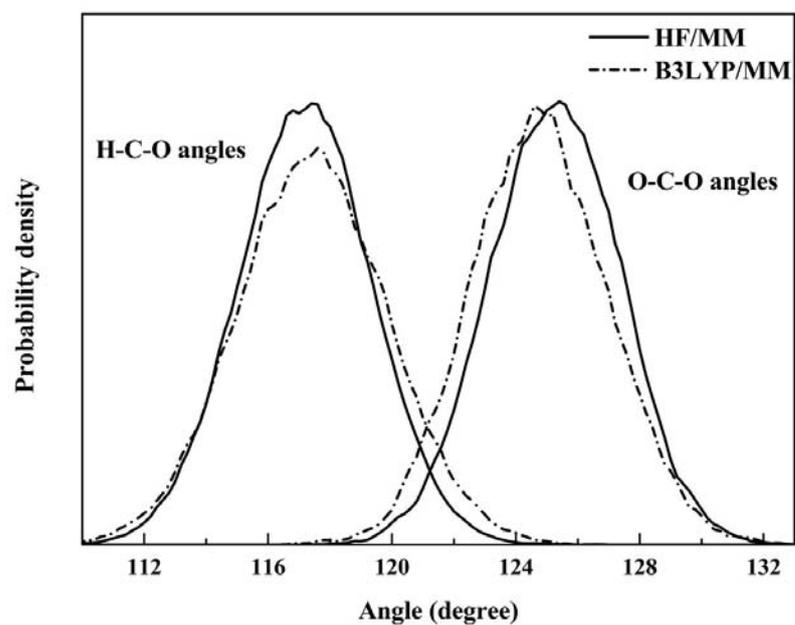
Figure 5.12 Distributions of H-C-O and O-C-O angles.
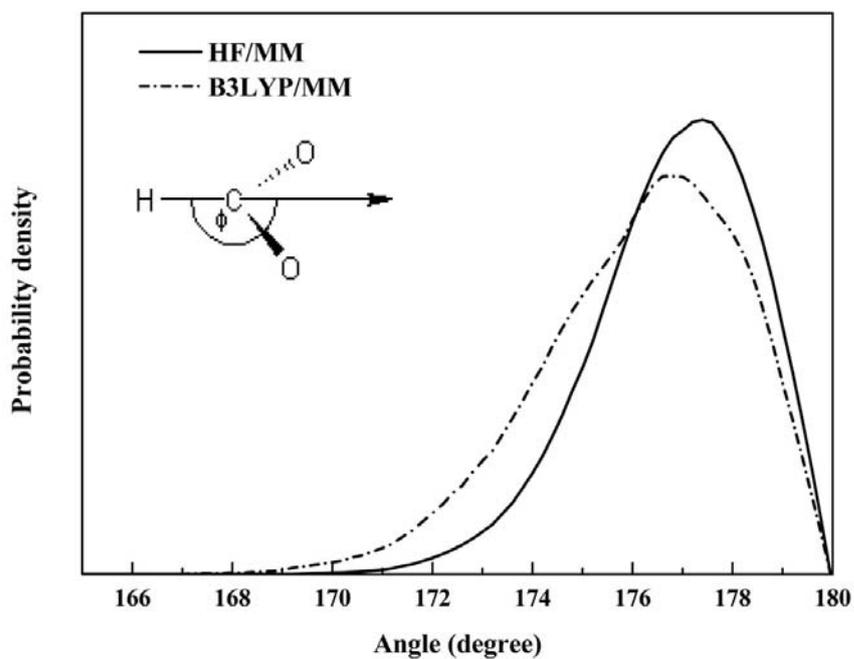
Figure 5.13 Distributions of ϕ, as defined by a vector along the C-H bond and a vector pointing outwards between the two C-O bonds.

5.1.2.2 Water exchange processes at HCOO⁻ oxygens

According to Figure 5.3, the non-zero first minimum of the O-O$_w$ and O-H$_w$ RDFs obtained by both HF/MM and B3LYP/MM simulations clearly suggests an easy exchange of water molecules between the solvation shell and the bulk. The exchange processes of first-shell waters at each of the HCOO⁻ oxygen atoms can be visualized through the plots of the O-O$_w$ distances against the simulation time, as shown in Figures 5.14 and 5.15 for the HF/MM and B3LYP/MM simulations, respectively. During the first 10 ps of the HF/MM and B3LYP/MM trajectories, numerous water molecules can be interchanged between the first shell and the bulk, leading to large fluctuations in the hydration number at each of the HCOO⁻ oxygen atoms (*e.g.*, see inserts in Figures 5.14 and 5.15). Inside the hydration shell, water molecules are either loosely or tightly bound to HCOO⁻ oxygen, *i.e.*, some first-shell waters temporarily form a hydrogen bond with HCOO⁻ oxygen, then leaving or even entering again, while others form longer hydrogen bonds to the respective HCOO⁻ oxygen. On the basis of both HF/MM and B3LYP/MM simulations, the occurrence of bifurcated hydrogen bonds (cf. Cyclic conformer in Figure 4.1) appears rare in aqueous solution. First-shell water molecules also preferentially associate with one HCOO⁻ oxygen atom or the other, rather than adopting simultaneous coordination to both.

The rate of water exchange processes at each HCOO⁻ oxygen atom was evaluated through mean residence times (MRT) of the surrounding water molecules. In this work, the MRT data were calculated using the direct method (Hofer, Tran, Schwenk, and Rode, 2004), as the product of the average number of nearest-neighbor water molecules located within the first minimum of the O-O$_w$ RDFs with the

duration of the simulation, divided by the number of exchange events. With respect to time parameters t^* (*i.e.*, the minimum duration of a ligand's displacement from its original coordination shell to be accounted) of 0.0 and 0.5 ps, the calculated MRT values are summarized in Table 5.1.

In the HF/MM simulation, the calculated MRT values with respect to $t^* = 0.0$ and 0.5 ps are slightly larger than that of pure water (Tongraar and Rode, 2004). These data correspond to the observed stronger hydrogen bonds between HCOO⁻ oxygens and their first-shell water molecules. In the B3LYP/MM simulation, as compared to the B3LYP/MM data for pure water (Xenides, Randolf, and Rode, 2005), a clear opposite order of $\tau_{H_2O}(O_i) < \tau_{H_2O}(H_2O)$ is observed. Here it should be noted that the B3LYP/MM results for pure water have produced too slow exchange rates compared to the experimental values (Lock, Woutersen, and Bakker, 2001). This failure of the B3LYP method to predict the dynamical properties of pure water could be considered as an example for the inadequacy of the DFT methods to correctly describe the characteristics of any aqueous hydrogen-bonded systems. Slow dynamics of aqueous HCOO⁻, in particular the HCOO⁻ rotation, were also observed in the recent CPMD study when the PW91 functional was employed (Leung and Rempe, 2004).
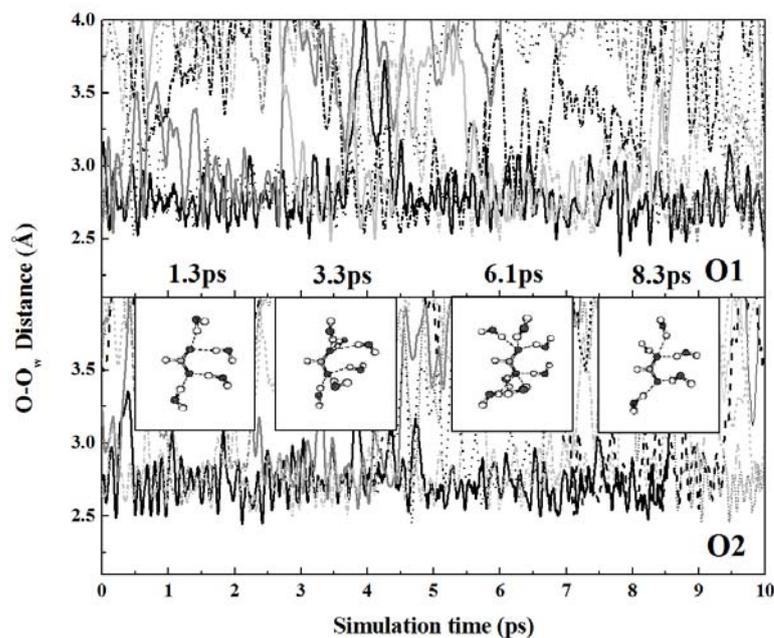
Figure 5.14 Time dependence of O··Oᵥᵥ distances, selecting only for first 10 ps of the HF/MM trajectories.
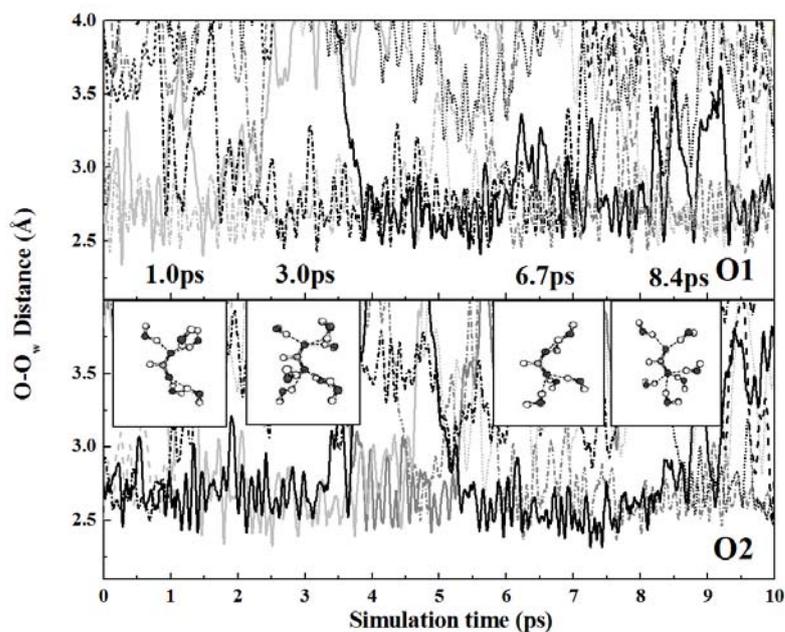
Figure 5.15 Time dependence of O··Oᵥᵥ distances, selecting only for first 10 ps of the B3LYP/MM trajectories.

Table 5.1 Mean residence times of water molecules in the bulk and in the vicinity of

HCOO⁻ oxygens, calculated within first minimum of the O-O$_w$ RDFs.

Atom/Solute	CN	t_{sim}	$t^* = 0.0$ ps		$t^* = 0.5$ ps	
			$N_{ex}^{0.0}$	$\tau_{H_2O}^{0.0}$	$N_{ex}^{0.5}$	$\tau_{H_2O}^{0.5}$
HF/MM MD						
O1	3.45	70.00	970	0.25	132	1.83
O2	3.44	70.00	1042	0.23	112	2.15
Pure H₂O[*]	4.60	12.00	292	0.20	31	1.80
Pure H₂O[**]	4.20	40.00	-	0.33	-	1.51
B3LYP/MM MD						
O1	2.84	50.00	810	0.17	57	2.49
O2	2.97	50.00	799	0.19	63	2.36
Pure H₂O[**]	4.20	30.00	-	1.07	-	7.84

[*] (Tongraar and Rode, 2004)

[**] (Xenides, Randolf, and Rode, 2005)

5.2   CH$_3$COO$^-$ in water

5.2.1   Structural details

5.2.1.1 Radial distribution functions

For the case of CH$_3$COO$^-$, two carbon atoms of CH$_3$COO$^-$ are defined as C$_1$ and C$_2$, which represent the carbon atom of carbonyl and methyl group, respectively. The C$_1$-O$_w$ and C$_1$-H$_w$ RDFs are plotted in Figures 5.16a and 5.16b, respectively. In the HF/MM simulation, the first C$_1$-O$_w$ is located at 3.58 Å, where an integration up to the corresponding first C$_1$-O$_w$ minimum yields about 7.78 water molecules. In the B3LYP/MM simulation, the C$_1$-O$_w$ peak is observed at shorter distance of 3.42 Å, consisting of about 8.43 water molecules within the first shell. In both HF/MM and B3LYP/MM simulations, the C$_1$-H$_w$ RDFs are exhibited at distances correspond to the C$_1$-O$_w$ RDFs, being 2.71 and 2.54 Å. The integrations up to first minimum of the C$_1$-H$_w$ RDFs yield 5.89 and 5.25 water molecules for the HF/MM and B3LYP/MM simulations, respectively. With regard to Figure 5.16, the shape and height of the C$_1$-O$_w$ and C$_1$-H$_w$ RDFs clearly suggest the recognizable hydrogen bonding between the carbonyl group and its surrounding water molecules.

The interactions between methyl group and its surrounding water molecules can be explained in Figures 5.17 and 5.18. As compared to that of carbonyl group, the observed broad and less pronounced C$_2$-O$_w$ and C$_2$-H$_w$ RDFs clearly indicate less influence of methyl group on its surrounding water, *i.e.*, water molecules in this region would prefer to form water-water hydrogen bonds rather than to form hydrogen bond with the methyl group.

The hydrogen bonds between CH$_3$COO$^-$ and water can be visualized through the O-O$_w$ and O-H$_w$ RDFs, together with their corresponding integration

numbers, as depicted in Figure 5.19. In the HF/MM simulation, the first O-O$_w$ peak is located at 2.75 Å, where integration up to the corresponding first O-O$_w$ minimum yields an average coordination number of 3.16. In the B3LYP/MM simulation, the first O-O$_w$ peak is observed at the shorter distance of 2.71 Å, with a lower coordination number of 3.12. In both HF/MM and B3LYP/MM simulations, the second O-O$_w$ peaks are broad and less pronounced. These peaks correspond to the contribution of both bulk waters and water molecules in the hydration layer of the other CH$_3$COO$^-$ oxygen.

The indication of CH$_3$COO$^-$-water hydrogen bonds is apparent, as can be seen from the O-H$_w$ RDFs shown in Figure 5.19b. In accordance with the O-O$_w$ RDFs, the HF/MM and B3LYP/MM simulations reveal first O-H$_w$ peaks with maximum at 1.80 and 1.68 Å, respectively. Integrations up to the corresponding first O-H$_w$ minima give the average coordination numbers of 2.94 and 2.65, respectively. According to both HF/MM and B3LYP/MM results, the observed numbers of water oxygen and hydrogen atoms supply information that the first-shell waters are linearly hydrogen bounded to each of the CH$_3$COO$^-$ oxygens, *i.e.*, they are acting as hydrogen-bond donors. In comparison to the first O-H$_w$ RDF of pure water (cf. Figure 5.4b), *i.e.*, in terms of shape and peak height, it is obvious that the O$^{...}$H$_w$-O$_w$ hydrogen-bond interactions are relatively stronger. In addition, the strength of the CH$_3$COO$^-$-water hydrogen bond is even larger than that of between HCOO$^-$ and waters. This phenomenon could be described due to the electronic effect in CH$_3$COO$^-$, *i.e.*, methyl group acts as electron donating group which release electrons into the CH$_3$COO$^-$ carbonyl group when CH$_3$COO$^-$ oxygen atoms form hydrogen bond with their first-shell water molecules. Consequently, this leads to stronger CH$_3$COO$^-$-water hydrogen

bonds. In both HF/MM and B3LYP/MM simulations, the second O-H_w RDFs around 3.2 Å can be attributed as to the hydrogen atoms of first-shell waters that are not hydrogen-bonded to the CH_3COO^- oxygens.

Similar to the case for $HCOO^-$, the observed higher coordination numbers than the experimental data (of about 2.5) could be described due to the broad O-O_w and O-H_w minima, *i.e.*, the integrations up to slight different O-O_w or O-H_w distances (slightly deviated from the minimum position) can yield significant difference in the numbering of water molecules. For example, the integrations up to shorter O-O_w minima of about 0.2 Å yield the coordination numbers of 2.70 and 2.64 for the HF/MM and B3LYP/MM simulations, respectively.

The distributions of the number of water's oxygen and hydrogen atoms around each of CH_3COO^- oxygens are shown in Figures 5.20 and 5.21, respectively. According to Figure 5.20, the most population of water's oxygen atoms of HF/MM and B3LYP/MM simulations are 3, followed by 4 and 2 in decreasing amounts. For the distributions of hydrogen atoms of neighbouring water molecules (Figure 5.21), the prevalent value of 3 is also observed, followed by 2 and 4 in smaller amounts. More detailed interpretation on the distributions of hydration number at each of CH_3COO^- oxygens can be found in Figures 5.22 and 5.23 for the HF/MM and B3LYP/MM simulations, respectively.

With regard to both HF/MM and B3LYP/MM simulations, it is observed that each of CH_3COO^- oxygens simultaneously hydrogen bond with different number of water molecules, leading to numerous possible species of the CH_3COO^--water complexes. As can be seen in Figure 5.22 and 5.23, the total numbers of water molecules in the vicinity of CH_3COO^- oxygens show large

fluctuations, ranging from 5 to 8 and from 4 to 9 for the HF/MM and B3LYP/MM simulations, respectively.
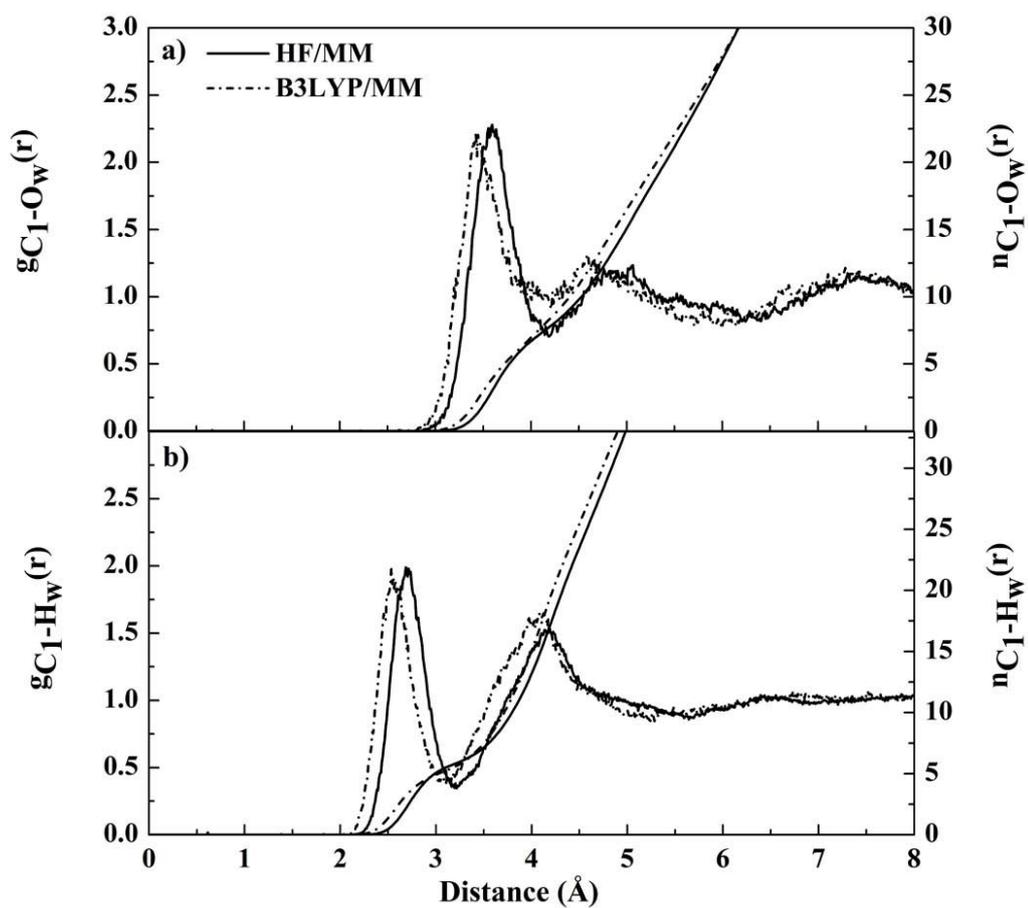
Figure 5.16 a) C_1-O_w and b) C_1-H_w RDFs and their corresponding integration numbers, as obtained by HF/MM and B3LYP/MM simulations using DZV+ basis set.
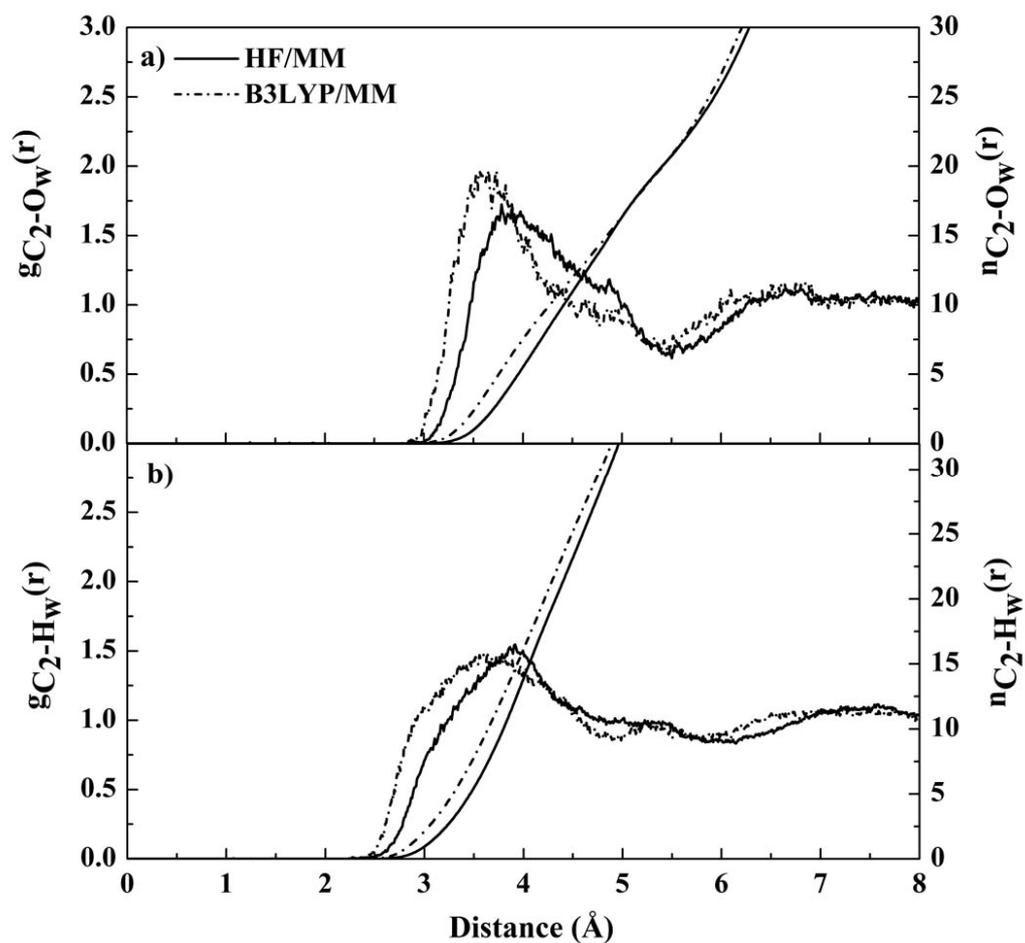
Figure 5.17 a) C_2-O_w and b) C_2-H_w RDFs and their corresponding integration numbers, as obtained by HF/MM and B3LYP/MM simulations using DZV+ basis set.
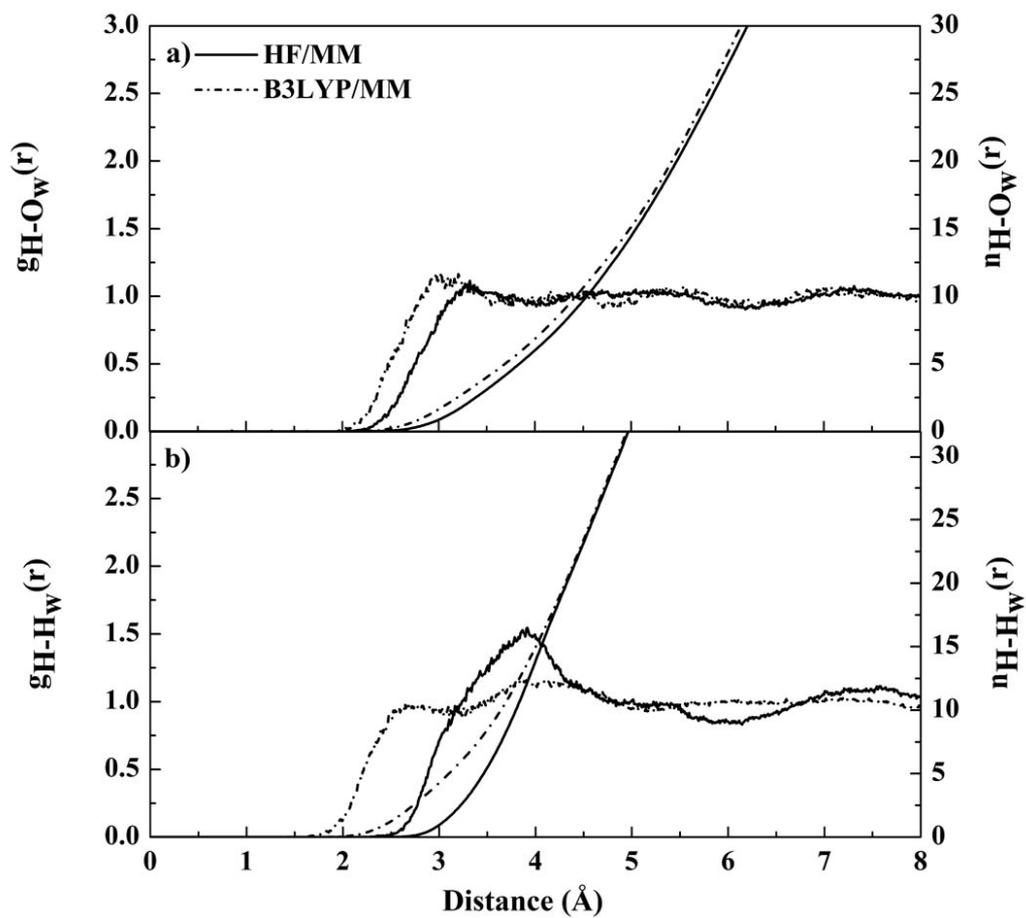
Figure 5.18 a) H-O$_w$ and b) H-H$_w$ RDFs and their corresponding integration numbers, as obtained by HF/MM and B3LYP/MM simulations using DZV+ basis set.
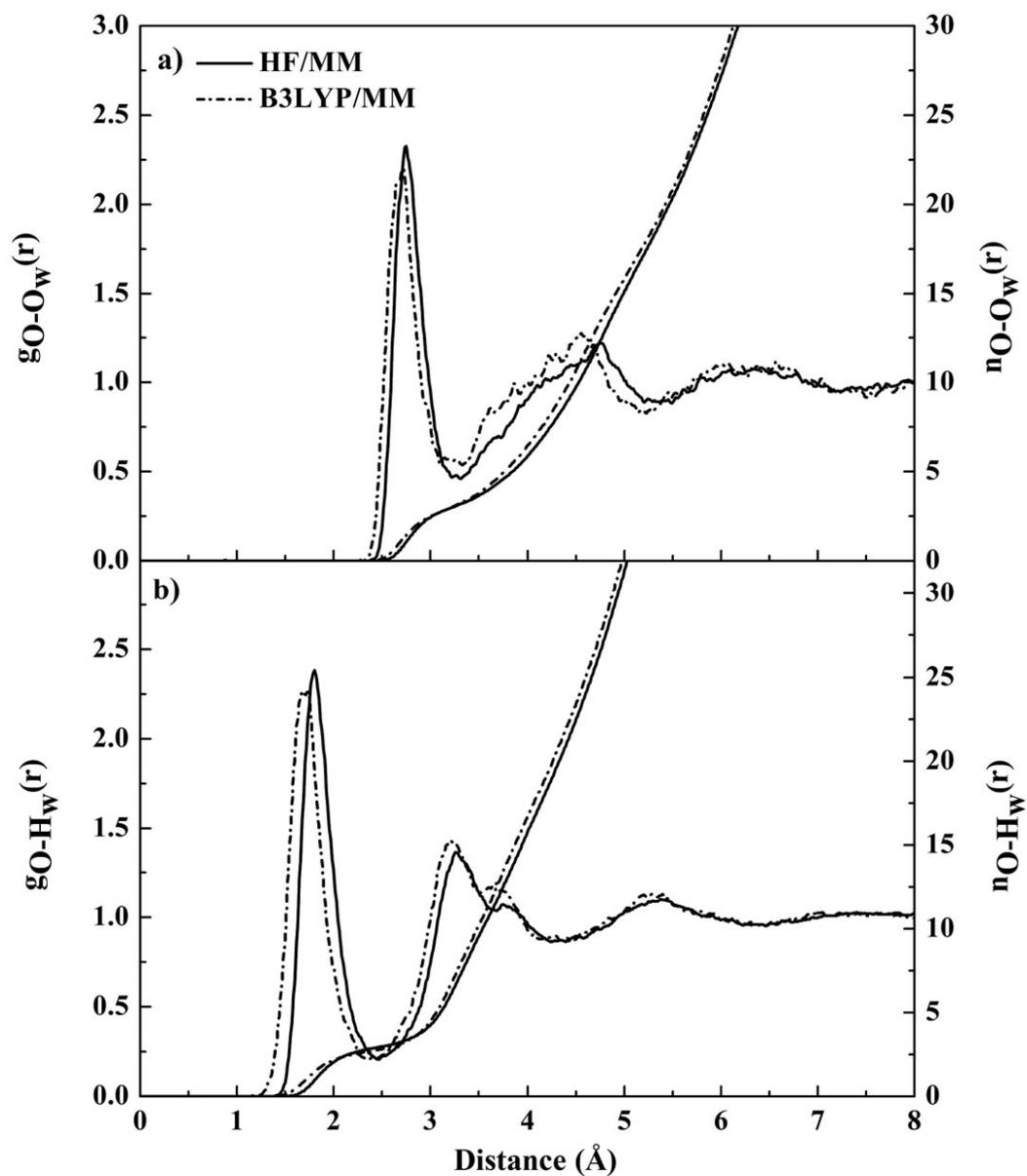
Figure 5.19 a) O-O$_w$ and b) O-H$_w$ and their RDFs corresponding integration numbers, as obtained by HF/MM and B3LYP/MM simulations using DZV+ basis set.
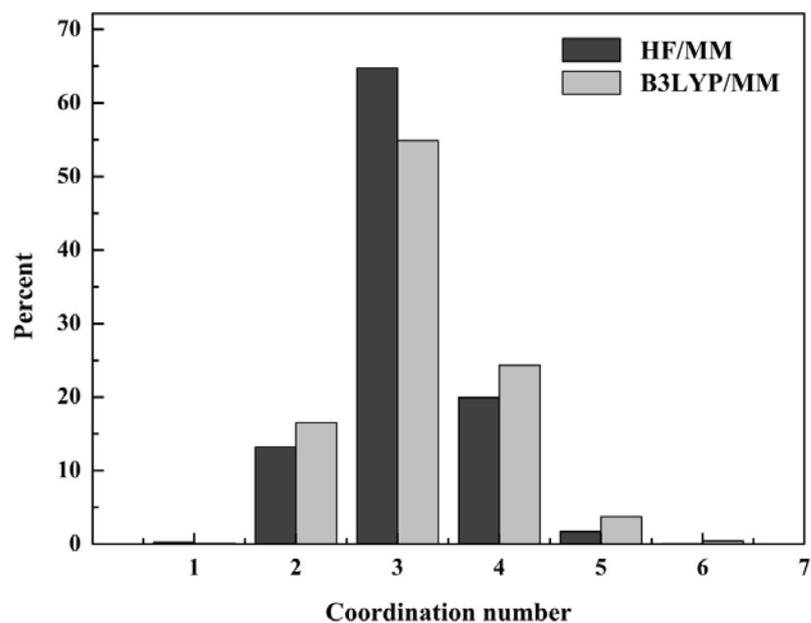
Figure 5.20 Distributions of the number of water's oxygen atoms at each of CH$_3$COO$^-$ oxygens, calculated within first minimum of the O-O$_w$ RDFs.
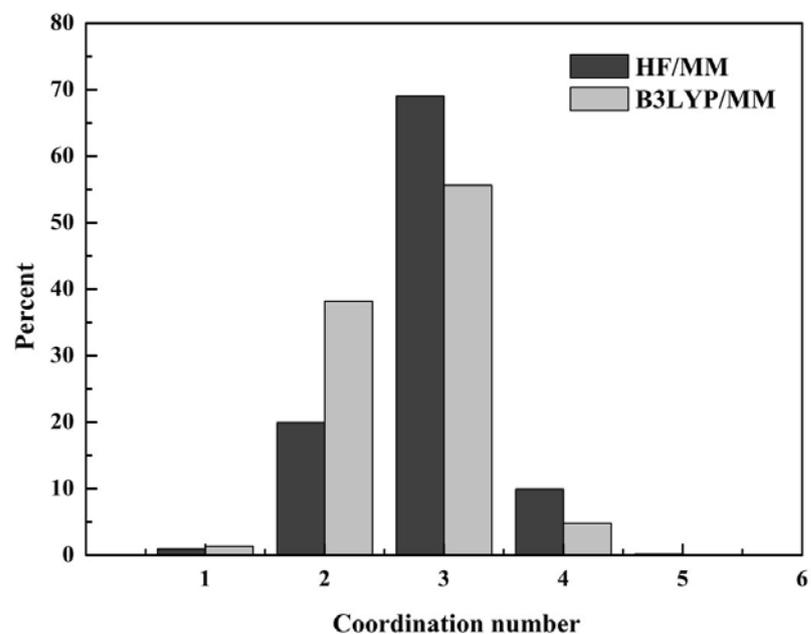
Figure 5.21 Distributions of the number of water's hydrogen atoms at each of CH$_3$COO$^-$ oxygens, calculated within first minimum of the O-O$_w$ RDFs.
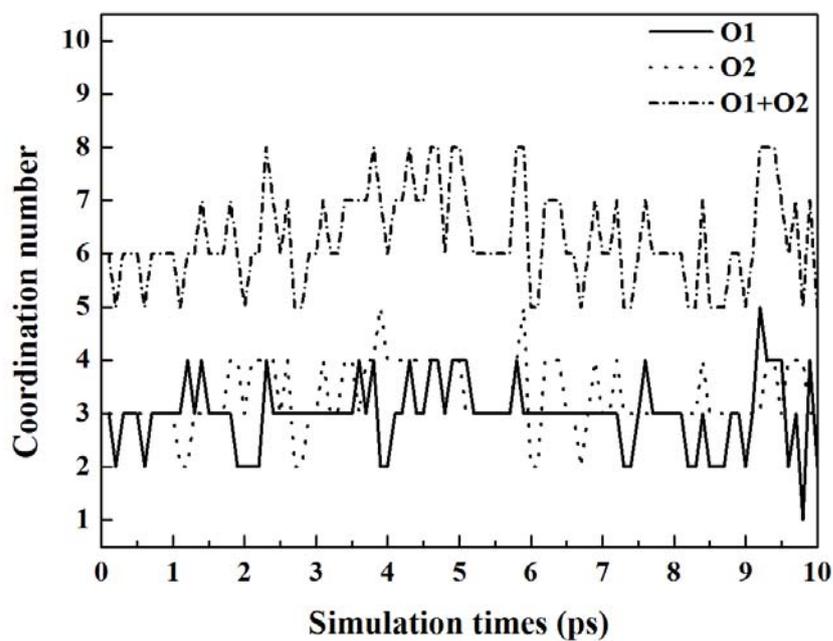
Figure 5.22 Time dependence of the number of first-shell waters at CH_3COO^- oxygen atoms, selecting only for first 10 ps of the HF/MM simulation.
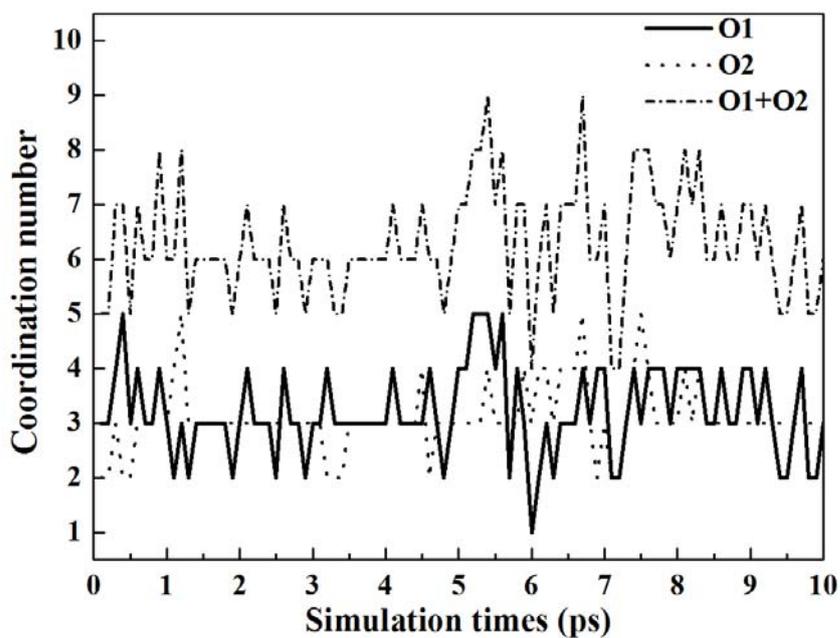
Figure 5.23 Time dependence of the number of first-shell waters at CH_3COO^- oxygen atoms, selecting only for first 10 ps of the B3LYP/MM simulation.

5.2.1.2 Angular distribution functions

More details regarding the hydrogen bonds between CH_3COO^- and water can be visualized through the probability distributions of the $O\cdots O_w$-H_w and C-$O\cdots H_w$ angles, as depicted in Figures 5.24 and 5.25, respectively. In both HF/MM and B3LYP/MM simulations (Figure 5.24), the distributions of $O\cdots O_w$-H_w angle clearly confirm the preference for linear $O\cdots O_w$-H_w hydrogen bonds. Furthermore, the observed broad C-$O\cdots H_w$ angle clearly indicate the absence of C-O single and C=O double bond, *i.e.*, there is no significant splitting peaks at 109.5° and 120° that correspond to the formation of the C-O single and C=O double bonds in aqueous solution. On the other hand, the CH_3COO^- adopts the electrically delocalized structure in aqueous solution, which may more or less fluctuate according to the process of water exchange.
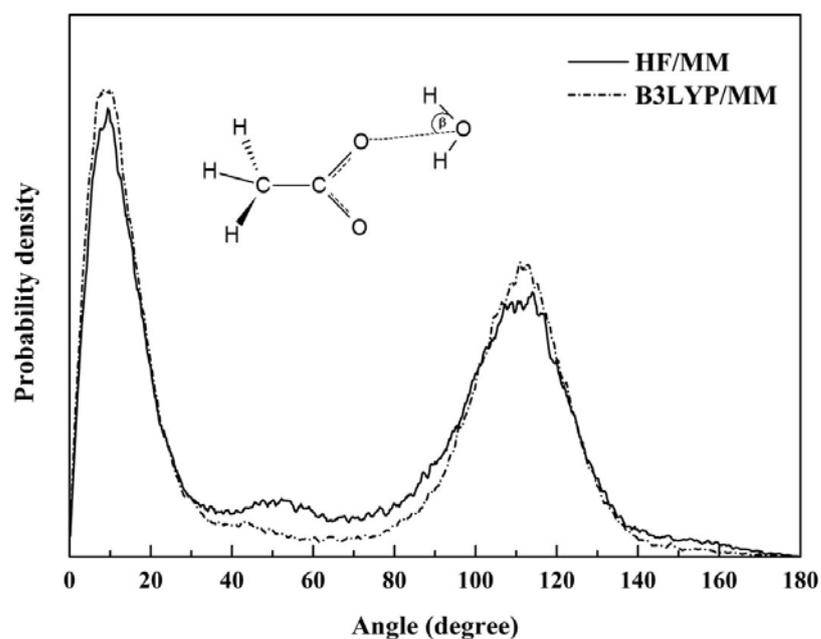
Figure 5.24 Distributions of $O\cdots O_w$-H_w angle, calculated within first minimum of the O-O_w RDFs.
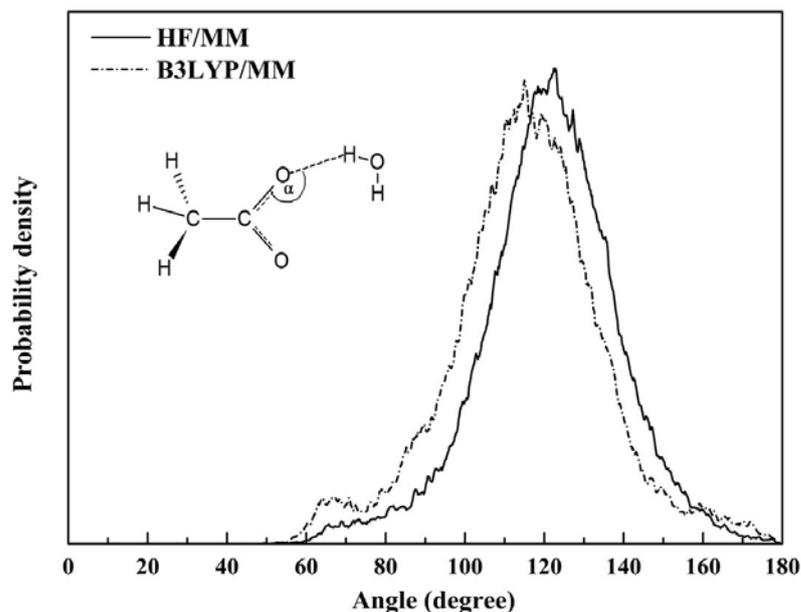
Figure 5.25 Distributions of C-H··Ow angle, calculated within first minimum of the O-Ow RDFs.

5.2.2 Dynamical Details

5.2.2.1 Intramolecular geometry of CH$_3$COO$^-$

The flexibility of CH$_3$COO$^-$ structure in aqueous solution can be seen from the probability distributions of intermolecular bond lengths and bond angles, as shown in Figures 5.26, 5.27 and 5.28. As compared to the gas phase CH$_3$COO$^-$ structures (cf. Table 4.4), both HF/MM and B3LYP/MM simulations clearly confirm the significant influence of surrounding water molecules that cause substantial modification in the local structure of CH$_3$COO$^-$.
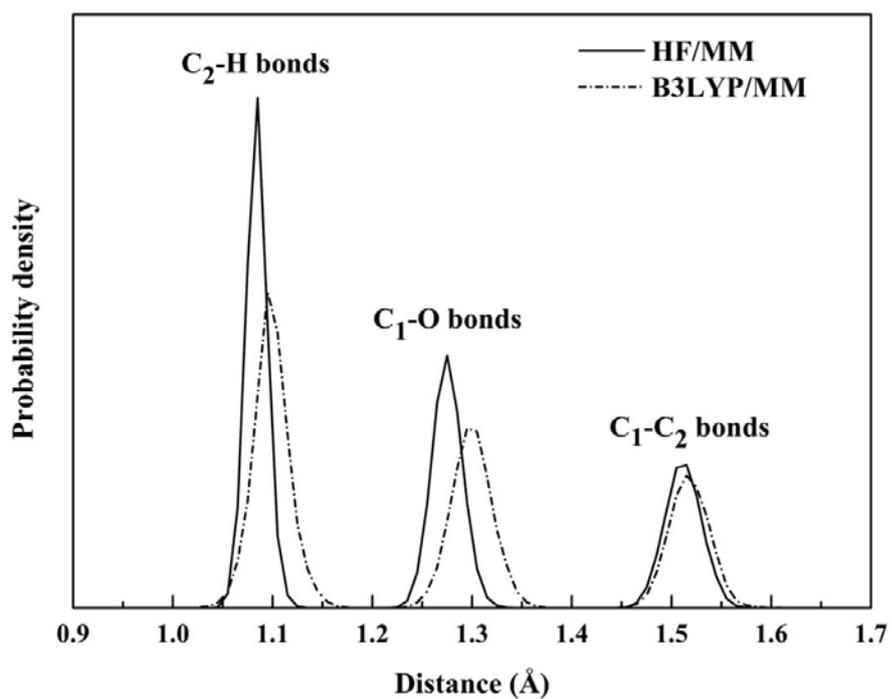
Figure 5.26 Distributions of C_1-O, C_1-C_2 and C_2-H bond lengths of CH_3COO^-.
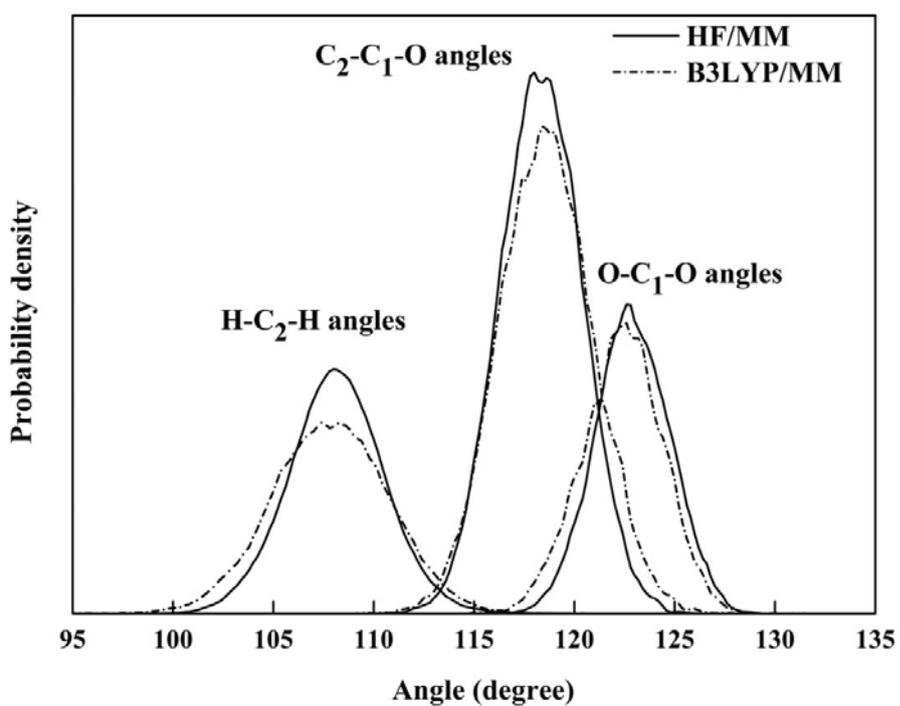
Figure 5.27 Distributions of C_2-C_1-O, O-C_1-O and H-C_2-H angles of CH_3COO^-.
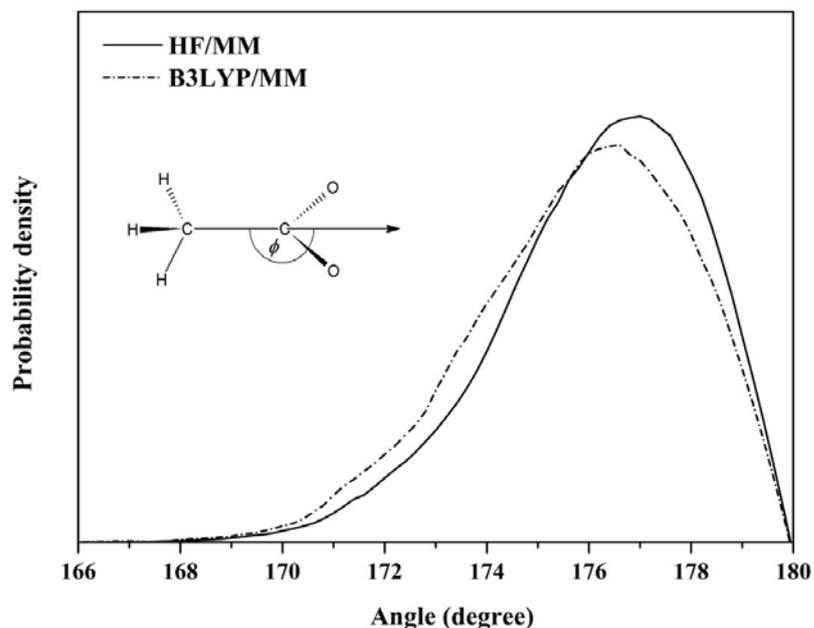
Figure 5.28 Distributions of ϕ, as defined by a vector along the C_1-C_2 bond and a vector pointing outwards between the other two C-O bonds.

5.2.2.2 Water exchange processes at CH$_3$COO$^-$ oxygens

In both HF/MM and B3LYP/MM simulations, the observed large fluctuation in the hydration numbers at each of CH$_3$COO$^-$ oxygen atoms clearly suggests the numerous water exchange processes between the first hydration shell and the bulk. The details of the water exchange processes at each of the CH$_3$COO$^-$ oxygens can be visualized from the plots of O-O$_w$ distances against the simulation time, as shown in Figures 5.29 and 5.30 for the HF/MM and B3LYP/MM simulations, respectively. In both HF/MM and B3LYP/MM simulations, it is obvious that first-shell water molecules are either "loosely" or "tightly" bound to the CH$_3$COO$^-$ oxygens. This behavior is similar to the case of HCOO$^-$. However, the water exchange events at CH$_3$COO$^-$ oxygen atoms are less frequent, as a consequence of stronger CH$_3$COO$^-$-water interactions.

The rates of water exchange processes at each of CH_3COO^- oxygen atoms, as calculated by means of MRT values of the surrounding water molecules, are summarized in Table 5.2. In the HF/MM simulation, the MRT values with respect to $t^* = 0.0$ and 0.5 ps clearly demonstrate the strong hydrogen bonds between CH_3COO^- oxygen atoms and their first-shell water molecules. The results obtained by the B3LYP/MM simulation are also in good accord with the HF/MM results. However, since the B3LYP/MM results for pure water have produced too slow exchange rates, the B3LYP/MM results obtained by the present study should be discussed with caution. The failure of the B3LYP method to predict the dynamics details of pure water could be demonstrated as warning when apply the DFT methods for studying condensed phase systems.
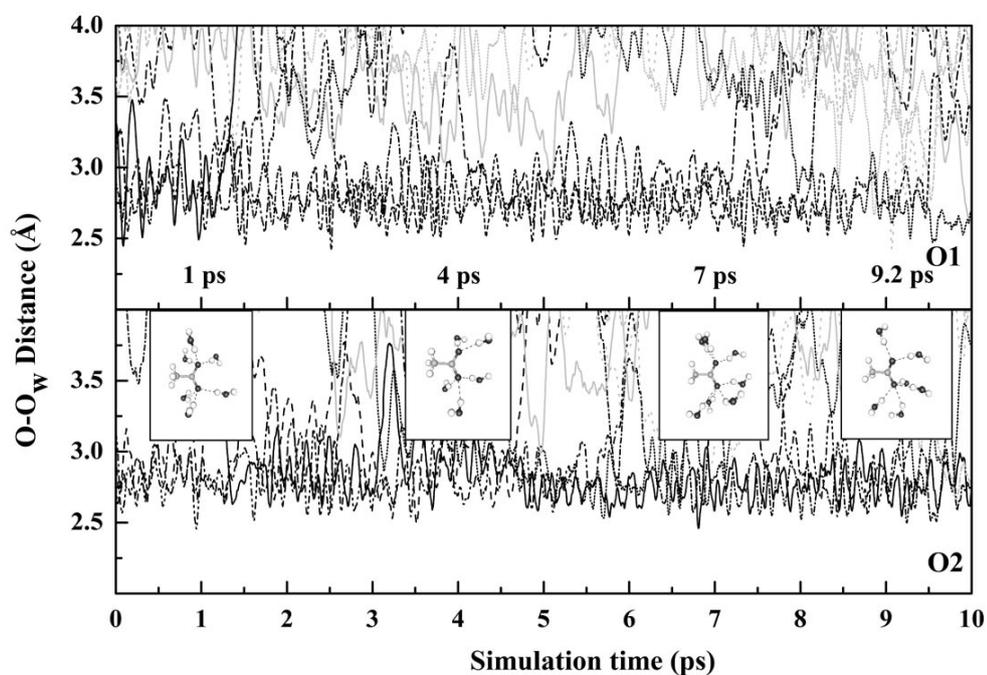
Figure 5.29 Time dependence of $O^{..}O_w$ distances, selecting only for first 10 ps of the HF/MM trajectories.
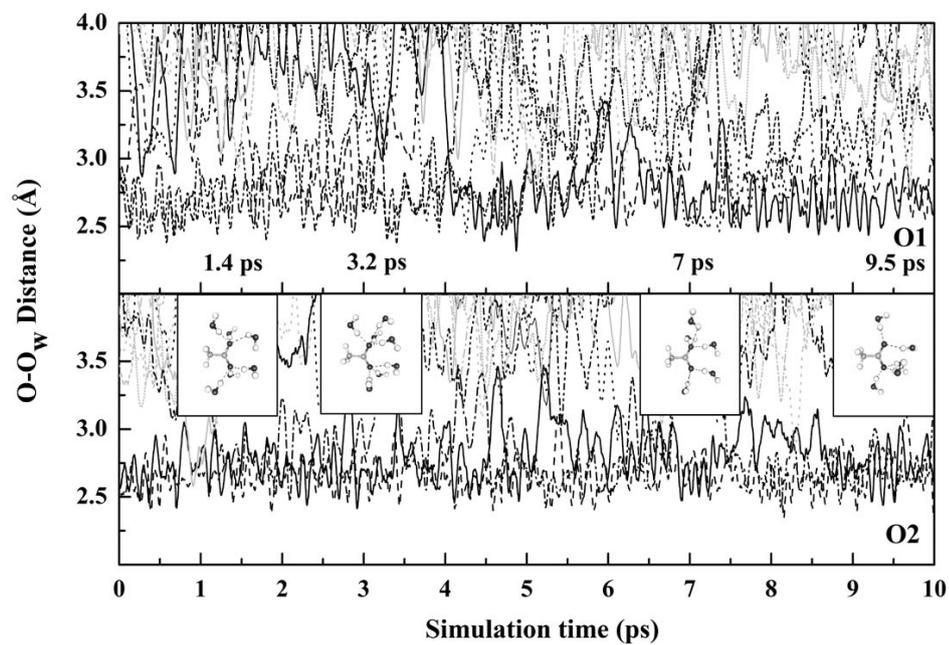
Figure 5.30 Time dependence of O··O$_w$ distances, selecting only for first 10 ps of the B3LYP/MM trajectories.

Table 5.2 Mean residence times of water molecules in the bulk and in the vicinity of CH$_3$COO$^-$ oxygens, calculated within first minimum of the O-O$_w$ RDFs.

Atom/solute	CN	t_{sim}	$t^* = 0.0$ ps		$t^* = 0.5$ ps	
			$N_{ex}^{0.0}$	$\tau_{H_2O}^{0.0}$	$N_{ex}^{0.5}$	$\tau_{H_2O}^{0.5}$
HF/MM MD						
O1	3.10	65.0	642	0.31	79	2.55
O2	3.15	65.0	553	0.37	80	2.56
Pure H$_2$O[*]	4.60	12.0	292	0.20	31	1.80
Pure H$_2$O[**]	4.20	40.0	-	0.33	-	1.51
B3LYP/MM MD						
O1	3.18	55.0	710	0.25	59	2.97
O2	3.09	55.0	641	0.27	53	3.21
Pure H$_2$O[**]	4.20	30.0	-	1.07	-	7.84

[*] (Tongraar and Rode, 2004)

[**] (Xenides, Randolf, and Rode, 2005)

5.3   References

Hofer, T. S., Tran, H. T., Schwenk, C. F. and Rode, B. M. (2004). Characterization of dynamics and reactivities of solvated ions by *ab initio* simulations. **Journal of Computational Chemistry**. 25: 211-217.

Kameda, Y., Fukuhara, K., Mochiduki, K., Naganuma, H., Usuki, T. and Uemura, O. (2002). Structure of HCOOK hydrated melts. **Journal of Non-Crystalline Solids**. 312-314: 433-437.

Kameda, Y., Mori, T., Nishiyama, T., Usuki, T. and Uemura, O. (1996). Structure of concentrated aqueous sodium formate solutions. **Bulletin of the Chemical Society of Japan**. 69: 1495-1504.

Leung, K. and Rempe, S. B. (2004). *Ab initio* molecular dynamics study of formate ion hydration. **Journal of the American Chemical Society**. 126: 344-351.

Lock, A. J., Woutersen, S. and Bakker, H. J. (2001). **Femtochemistry and Femtobiology**. Singapore: World Scientific.

Tongraar, A. and Rode, B. M. (2004). Dynamical properties of water molecules in the hydration shells of Na^+ and K^+: *Ab initio* QM/MM molecular dynamics simulations **Chemical Physics Letters**. 385: 378-383

Xenides, D., Randolf, B. R. and Rode, B. M. (2005). Structure and ultrafast dynamics of liquid water: A quantum mechanics/molecular mechanics molecular dynamics simulations study. **The Journal of Chemical Physics**. 122: 174506-174516.

CHAPTER VI

CONCLUSION

In the present study, the high-level QM/MM MD technique has been successfully applied for studying structural and dynamical properties of carboxylate ions, namely $HCOO^-$ and CH_3COO^-, in aqueous solution. On the basis of both HF/MM and B3LYP/MM simulations, it is observed that the $HCOO^-$-water and CH_3COO^--water hydrogen bonds are relatively strong, especially when compared to that of water-water hydrogen bonds in the bulk. The geometrical arrangements of $HCOO^-$ and CH_3COO^- in aqueous solution are found to be rather flexible and the first-shell water molecules can be either "loosely" or "tightly" bound to each of $HCOO^-$ and CH_3COO^- oxygen atoms, forming an asymmetric solvation structure with varying numbers of hydrogen bonds, with the prevalent value of 3. In this respect, the effects of water environment are demonstrated to play an important role on the $HCOO^-$-water and CH_3COO^--water hydrogen bond formation.

Comparing between the aqueous $HCOO^-$ and CH_3COO^- systems, the CH_3COO^--water hydrogen bonds are found to be stronger than that of between $HCOO^-$ and waters. This corresponds to the electronic effects in CH_3COO^-, *i.e.*, the electron donating methyl group of CH_3COO^- can release electrons to the CH_3COO^- carbonyl group when CH_3COO^- oxygen atoms form hydrogen bonds with their neighbouring water molecules. The observed stronger CH_3COO^--water hydrogen

bonds consequently reflect in the slower water exchange rate at CH_3COO^- oxygens, when compared to that of $HCOO^-$ oxygens.

With respect to the HF/MM and B3LYP/MM results of both $HCOO^-$ and CH_3COO^- systems, the structural details described by means of HF and B3LYP methods are quite consistency. However, despite the lack of electron correlations, which lead to slightly too long and weak hydrogen bonds, the HF method proves more reliable than the B3LYP method, as the latter completely, fails to describe the dynamics of several hydrated ions, *i.e.*, even for the description of bulk water itself. As a consequence of continuous increase in computer capacity and performance, the extension of the QM/MM simulations by increasing the QM size, enlarging the basis set and finally including electron correlations would allow a future improvement of the simulation results.

APPENDICES

APPENDIX A

RESEARCH WORKS AT GAO GROUP,

UNIVERSITY OF MINNESOTA

A.1   A mixed molecular orbital and valence bond (MOVB) method

For the mixed molecular orbital and valence bond (MOVB) method (Mo and Gao, 2000; Song and Gao, 2008; Song, Mo, and Gao, 2009), the block-localized wave function has been used to define individual VB configuration (Mo and Peyerimhoff, 1998; Mo, Zhang, and Gao, 1999; Mo and Gao, 2000; Song and Gao, 2008) or Lewis resonance structure or diabatic states that represented by a single Slater determinant wavefunction. The MOVB wavefunction can be written in terms of a linear combination of these MOVB diabatic states as,

$$\Phi_{MOVB} = \sum_K b_K \Psi_{MOVB}^K, \qquad\qquad A.1$$

where Ψ_{MOVB}^K is the wavefunction for the Kth MOVB diabatic state that described by the block-locailized wavefunction (BLW) method (Mo and Peyerimhoff, 1998; Mo, Zhang, and Gao, 1999; Mo and Gao, 2000; Song and Gao, 2008; Song, Mo, and Gao, 2009). b_K is the coefficient for state K.

The BLW method is widely used for small molecules in gas phase and applied in MOVB method for studying condensed phase systems. By the BLW approach of

the diabatic state K, the system is divided into Ω_K subgroups. The molecular wavefunction for diabatic state K is

$$\Psi_{MOVB}^{K} = \hat{A}\left\{\chi_1^K \chi_2^K \chi_3^K \cdots \chi_{\Omega_K}^K\right\}, \qquad \text{A.2}$$

where $\hat{A}$ is an antisymmetrizing operator and χ_a^K is a product of molecular orbitals in the ath subgroup.

$$\chi_a^K = \varphi_1^K \alpha \varphi_1^K \beta \cdots \varphi_{n_a/2}^K \beta, \qquad \text{A.3}$$

where $\left\{\varphi_i^K, i = 1, \ldots, n_a/2\right\}$ are molecular orbitals, α and β are electronic spin orbitals and n_a is the number of electrons in subgroup a. For the molecular orbitals, this corresponds to the orthonormal constraints as

$$\left\langle \varphi_i^a \middle| \varphi_j^b \right\rangle = \begin{cases} \delta_{ij}, \\ S_{ij}, \end{cases} \quad \text{when} \quad \begin{matrix} a = b \\ a \neq b \end{matrix}, \qquad \text{A.4}$$

where S_{ij} is the overlap integral between two molecular orbitals i and j, $i.e.$, molecular orbitals within each block are orthogonal, whereas orbitals in different subgroups are non-orthogonal. These features can be illustrated by the transformation matrix as

$$C^R = \begin{pmatrix} C_1^K & 0 & \cdots & 0 \\ 0 & C_2^K & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & C_{\Omega_K}^K \end{pmatrix},$$ A.5

where C_a^K is a block matrix of orbital coefficients for orbitals in χ_a^K (Mo and Gao, 2000; Song and Gao, 2008).

Due to the efficiency and useful of MOVB method, it was applied to study the detailed mechanism of histone lysine demethylation. The histone demethylations can be catalyzed by two classes of enzymes, the FAD-dependent amino oxidases (Shi, Lan, Matson, Mulligan, Whetstine, Cole, Casero, and Shi, 2004) and the Fe(II):α-ketoglutarate dioxygenases (Klose, Kallin, and Zhang, 2006; Tsukada, Fang, Erdjument-Bromage, Warren, Borchers, Tempst, and Zhang, 2006).

For the FAD-dependent amino oxidases, the mechanisms can be reported into three major alternative pathways, *i.e.*, electron transfer (Silverman, 1995), nucleophilic addition at the flavin C4a carbon and direct hydride transfer to the flavin N5 nitrogen. However, the hydride transfer mechanism is the most interesting pathway because an iminium ion can be produced, released and hydrolyzed by water (Fitzpatrick, 2004; Scrutton, 2004).

The detailed electronic structures of reactant, transition and product state of hydride transfer mechanism at flavin N5 nitrogen are shown in Figures A.1-A.2. All geometry optimization structures were calculated by B3LYP and MO62X methods using 6-31+G* basis set that performed with GAUSSIAN03 program calculation. The observed geometry optimization energies of each structure are shown in Table A.1.
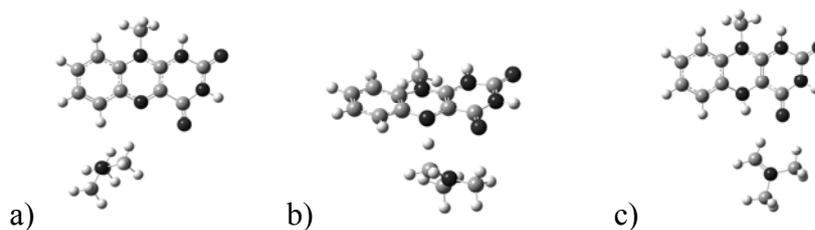
Figure A.1 Optimized structures of a) reactant, b) transition and c) product states of hydride transfer mechanism at flavin N5 nitrogen, as obtained from B3LYP calculations using 6-31+G* basis set.
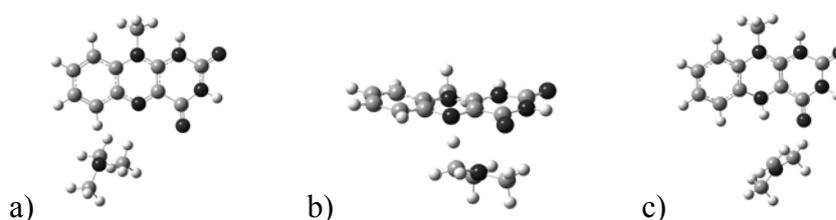
Figure A.2 Optimized structures of a) reactant, b) transition and c) product states of hydride transfer mechanism at flavin N5 nitrogen, as obtained from MO62X calculations using 6-31+G* basis set.

Table A.1 Geometry optimization energies of reactant, transition and product states of hydride transfer mechanism at flavin N5 nitrogen, obtained from B3LYP and MO62X methods using 6-31+G* basis set.

Structure	Optimization energies (kcal/mol)	
	B3LYP	MO62X
Reactant state	-2.0475281	-4.899245628
Transition state	13.02336053	12.4565179
Product state	-6.231869137	-11.02230048

For the Fe(II):α-ketoglutarate dioxygenases, a three-step kinetic mechanism has been proposed by Taud, Krebs *et al*., as can be seen in Figure A.3 that the Fe(IV)=O reactive species and a Fe(II) form to be complex (Fitzpatrick, 2004; Price, Barr, Hoffart, Krebs, and Bollinger, 2005; Sinnecker, Svensen, Barr, Ye, Bollinger, Neese, and Krebs, 2007). At the Fe(IV)=O reactive complex, the iron complex may has high spin (S=2) or low spin (S=1) (Decker, Rohde, Que, and Solomon, 2004).
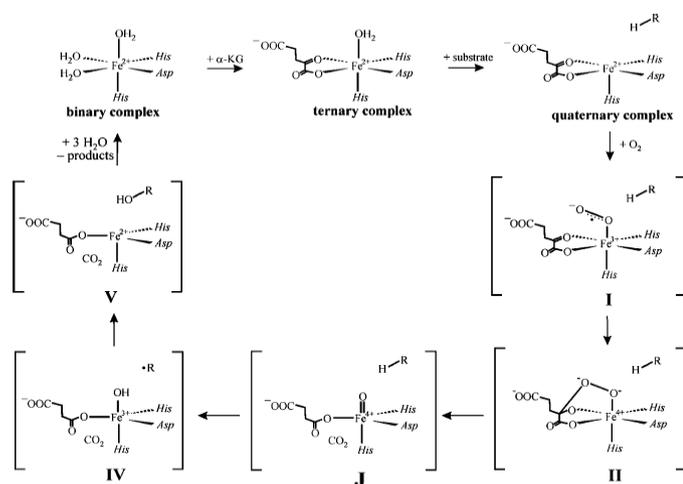
Figure A.3 Proposed mechanism of the Fe(II)- and R-ketoglutarate-dependent dioxygenases (Sinnecker, Svensen, Barr, Ye, Bollinger, Neese, and Krebs, 2007).

To investigate the electronic structure of the iron-oxo reactive species, the S=2 and S=1 states structures were optimized at B3LYP level of accuracy using 6-31+G* basis set. The optimized geometries of low and high spin of Fe(IV)=O reactive complex are shown in Figures A.4 and A.5, respectively. In addition, the optimized energies of low and high spin of Fe(IV)=O reactive complex can be seen in Table A.2.
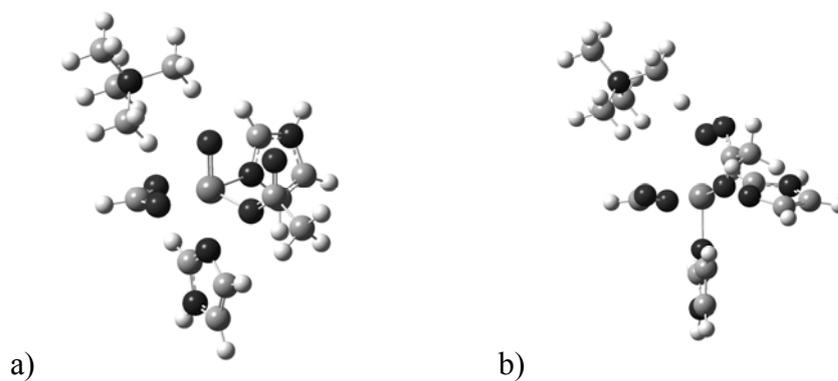
Figure A.4 Optimized low spin structures of a) reactant state and b) transition state of Fe(IV)=O reactive complex, as obtained by B3LYP calculations using 6-31+G* basis set.
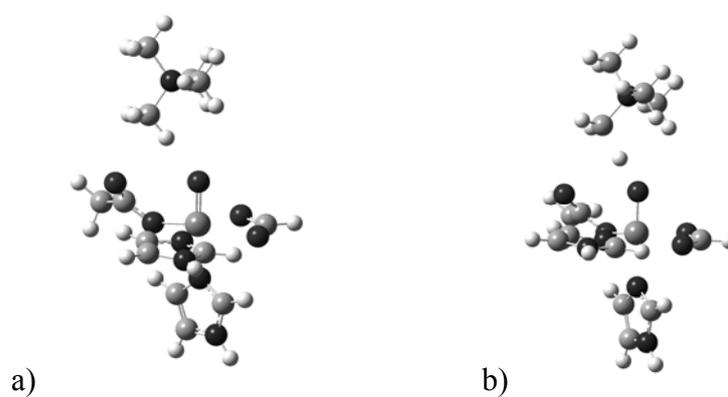
Figure A.5 Optimized high spin structures of a) reactant state and b) transition state of Fe(IV)=O reactive complex, as obtained by B3LYP calculations using 6-31+G* basis set.

Table A.2 Optimized energies of reactant and transition states of low and high spin of Fe(IV)=O reactive complex, obtained from B3LYP method with 6-31+G* basis set.

Structures	Optimization energies (A.U.)	
	Low spin	High Spin
Reactant state	-2423.19953245	-2423.25444134
Transition state	-2423.14134460	-2423.20775968

For further works, the MOVB approach will be applied to investigate the mechanism of histone lysine demethylation, *i.e.*, by using the geometry optimized structures of reactant, transition and product state of hydride transfer mechanism at flavin N5 nitrogen for the FAD-dependent amino oxidases catalysis and the geometry optimized structures of the Fe(II):α-ketoglutarate dioxygenases catalysis.

A.2   References

Decker, A., Rohde, J.-U., Que, L. and Solomon, E. I. (2004). Spectroscopic and quantum chemical characterization of the electronic structure and bonding in a non-Heme FeIVO complex. **Journal of the American Chemical Society**. 126: 5378-5379.

Fitzpatrick, P. F. (2004). Carbanion versus hydride transfer mechanisms in flavoprotein-catalyzed dehydrogenations. **Bioorganic Chemistry**. 32: 125-139.

Klose, R. J., Kallin, E. M. and Zhang, Y. (2006). JmjC-domain-containing proteins and histone demethylation. **Nature Reviews Genetics**. 7: 715-727.

Mo, Y. and Gao, J. (2000). An *ab initio* molecular orbital−valence bond (MOVB) method for simulating chemical reactions in solution. **The Journal of Physical Chemistry A**. 104: 3012-3020.

Mo, Y. and Gao, J. (2000). Energy decomposition analysis of intermolecular interactions using a block-localized wave function approach. **The Journal of Chemical Physics**. 112: 5530-5538.

Mo, Y. and Peyerimhoff, S. D. (1998). Theoretical analysis of electronic elocalization. **The Journal of Chemical Physics**. 109: 1687-1697.

Mo, Y., Zhang, Y. and Gao, J. (1999). A simple electrostatic model for trisilylamine: Theoretical examinations of the n→σ* negative hyperconjugation, p_π→d_π bonding, and stereoelectronic interaction. **Journal of the American Chemical Society**. 21: 5737-5742.

Price, J. C., Barr, E. W., Hoffart, L. M., Krebs, C. and Bollinger, J. M. (2005). Kinetic dissection of the catalytic mechanism of taurine:α-ketoglutarate dioxygenase (TauD) from Escherichia coli. **Biochemistry**. 44: 8138-8147.

Scrutton, N. S. (2004). Chemical aspects of amine oxidation by flavoprotein enzymes. **Natural Product Reports**. 21: 722-730.

Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J. R., Cole, P. A., Casero, R. A. and Shi, Y. (2004). Histone Demethylation Mediated by the Nuclear Amine Oxidase Homolog LSD1. **Cell**. 119: 941-953.

Silverman, R. B. (1995). Radical Ideas about Monoamine Oxidase. **Accounts of Chemical Research**. 28: 335-342.

Sinnecker, S., Svensen, N., Barr, E. W., Ye, S., Bollinger, J. M., Neese, F. and Krebs, C. (2007). Spectroscopic and computational evaluation of the structure of

the high-spin Fe(IV)-oxo intermediates in taurine: alpha-Ketoglutarate dioxygenase from Escherichia coli and its His99Ala ligand variant. **Journal of the American Chemical Society**. 129: 6168-6179.

Song, L. and Gao, J. (2008). On the construction of diabatic and adiabatic potential energy surfaces based on *ab initio* valence bond theory. **The Journal of Physical Chemistry A**. 112: 12925-12935.

Song, L., Mo, Y. and Gao, J. (2009). An effective hamiltonian molecular orbital-valence bond (MOVB) approach for chemical reactions as applied to the nucleophilic substitution reaction of hydrosulfide ion and chloromethane. **Journal of Chemical Theory and Computation**. 5: 174-185.

Tsukada, Y., Fang, J., Erdjument-Bromage, H., Warren, M. E., Borchers, C. H., Tempst, P. and Zhang, Y. (2006). Histone demethylation by a family of JmjC domain-containing proteins. **Nature**. 439: 811-816.

APPENDIX B

LIST OF PRESENTATIONS

1. Apirak Payaka and Anan Tongraar. (March 27-29, 2008). Hydration structure of formate ion (HCOO$^-$) studied by QM/MM molecular dynamics simulations. **12th Annual Symposium on Computational Science and Engineering (ANSCSE12)**. Ubon Rajathanee University, Ubon Rajathanee, Thailand.

2. Apirak Payaka and Anan Tongraar. (March 23-26, 2010). QM/MM dynamics of HCOO$^-$-water hydrogen bonds in aqueous solution. **14th Annual Symposium on Computational Science and Engineering (ANSCSE14)**. Mae Fah Luang University, Chiang Rai, Thailand.

3. Apirak Payaka and Anan Tongraar. (April 1-3, 2010). A combined QM/MM MD study of HCOO$^-$-water hydrogen bonds in aqueous solution. **RGJ-Ph.D. Congress XI**. Jomthien Palm Beach Hotel and Resort, Chonburi, Thailand.

APPENDIX C

PUBLICATION

Combined QM/MM MD Study of HCOO⁻–Water Hydrogen Bonds in Aqueous Solution

Apirak Payaka,[†] Anan Tongraar,*,[†] and Bernd Michael Rode[‡]

School of Chemistry, Institute of Science, Suranaree University of Technology,
Nakhon Ratchasima 30000, Thailand, and Department of Theoretical Chemistry, Institute of General, Inorganic
and Theoretical Chemistry, University of Innsbruck, Innrain 52a, A-6020 Innsbruck, Austria

Characteristics of HCOO⁻–water hydrogen bonds in dilute aqueous solution have been investigated by means of combined HF/MM and B3LYP/MM molecular dynamics simulations, in which the central HCOO⁻ and its surrounding water molecules were treated at HF and B3LYP levels of accuracy, respectively, using DZV+ basis set. Both HF/MM and B3LYP/MM simulations supply information that the hydrogen bonds between HCOO⁻ oxygens and first-shell waters are relatively strong, that is, compared to the water–water hydrogen bonds. Regarding to the HF/MM and B3LYP/MM trajectories, it is observed that first-shell waters are either "loosely" or "tightly" bound to their respective HCOO⁻ oxygen atoms, showing large fluctuations in the hydration number, varying from 2 to 6 (HF/MM) and 1 to 5 (B3LYP/MM), with the prevalent value of 3. Comparing the HF and B3LYP methods for the description of QM treated region, the first one leads to slightly too weak and thus longer hydrogen bonds, while the latter predicts them stronger but with the wrong dynamical data.

1. Introduction

Investigations of the microstructure and dynamics of ions solvated in aqueous electrolyte solutions have long been a scientific interest because of their diverse functions in many chemical and biological processes.[1–3] In experiments, several techniques such as X-ray (XD) and neutron diffraction (ND), as well as nuclear magnetic resonance (NMR), infrared absorption, X-ray absorption fine structure (EXAFS), and X-ray absorption near edge structure (XANES) spectroscopy have been used to obtain detailed knowledge of ions in aqueous solution.[4–6] However, most of these experimental techniques require major resources of laboratory equipment, but even then the results often show large discrepancies, especially for very dilute solution, due to technical limitations.[7,8]

Alternatively, computer simulations by means of Monte Carlo (MC) or molecular dynamics (MD) have become a powerful tool to study such solutions. For more than three decades, a large number of MC and MD simulations have been carried out, providing microscopic details for numerous solvated ions.[9–13] Because most of the earlier works had to rely on classical molecular mechanical force fields for describing all kinds of interactions, the simulation results, in particular, the hydration structure, as well as the dynamics of solvent molecules surrounding the ions, crucially depended on the quality and completeness of the potential functions employed in the simulations.[14–16] To accurately describe the properties of ions in aqueous solution, it has been demonstrated that "quantum effects" are significant and that inclusion of these effects in the simulations is mandatory.[17] Since the performance of ab initio quantum mechanical calculations for a condensed-phase system consisting of a large number of molecules is still beyond the current computational feasibility, the Car–Parrinello molecular dynamics (CPMD) method[18,19] reduced the computational

expense by using a simple density functional and a moderate system size containing about 30–60 solvent molecules. Albeit the CPMD technique has meanwhile been well-established for the study of solvated ions,[20–23] the simplification of the quantum mechanics employed has led to some severe limitations in accuracy of this scheme for the treatment of electrolyte solutions.[17]

With regard to the limits in computing power, another approach is to apply a hybrid quantum mechanics/molecular mechanics (QM/MM) method.[24–27] By the QM/MM technique, the chemically most relevant region, that is, a sphere including the ion and its surrounding solvent molecules, is described as accurately as needed using quantum mechanics (QM), whereas the rest of the system is handled by molecular mechanics (MM) with appropriate force fields. In the course of QM/MM scheme, the complicated many-body contributions as well as the polarization effects, which are hardly accessible through the basic assumptions underlying the classical models, can be reliably included into the specified region. In recent years, a number of QM/MM MD simulations have been carried out for various ions in solutions, providing many new insights into the solvation structure and dynamics of the solvated ions.[28–35]

In the present study, the behavior of formate ion (HCOO⁻), the simplest species containing the carboxylate (COO⁻) functional group, solvated in water was of interest. According to NMR experiments with carboxylic acids, it has been suggested that each carboxylate group is surrounded by 5.0–6.5 water molecules.[36] For HCOO⁻, the results obtained from X-ray and neutron scattering of aqueous NaHCOO[37] and KHCOO[38] solutions as well as from infrared measurements of a series of KHCOO concentrations[38] have inferred that the hydration number per HCOO⁻ oxygen atom is below 2.5. In terms of theoretical investigations, early MC simulations with "optimized potentials for liquid simulations" (OPLS) empirical force fields have been performed, revealing strong pronounced pair correlation functions between each HCOO⁻ oxygen and the oxygen/hydrogen atoms of water molecules, with a hydration

* Corresponding author. E-mail: anan_tongraar@yahoo.com. Fax: 0066-44-224017.
† Suranaree University of Technology.
‡ University of Innsbruck.

Payaka et al.

number of 3.6.[39] Recently, CPMD simulations of aqueous HCOO⁻ have been carried out,[40] showing structural features significantly different from those predicted by the OPLS potentials. According to the CPMD simulation using the BLYP functional, a hydration number of 2.45 per HCOO⁻ oxygen was predicted, which is in good accord with recent experiments.[37,38] However, the use of another exchange correlation functional, namely, PW91, led to hydration numbers ranging from 2.12 to 2.66. With regard to the CPMD method, it should be taken into account that the system's size under investigation was rather small, consisting of only 53 water molecules, and that only the simple GGA density functionals were employed. In this work, it was of particular interest, therefore, to apply the QM/MM technique in order to obtain a description of the HCOO⁻-water coordination in aqueous solution based on noncorrelated ab initio quantum mechanics and at DFT level using a more accurate hybrid functional.

2. Methods

According to the QM/MM MD technique,[17,28–35] the system is partitioned into two parts, namely, QM and MM regions. The total interaction energy of the system is defined as

$$E_{total} = \langle \Psi_{QM} | \hat{H} | \Psi_{QM} \rangle + E_{MM} + E_{QM-MM} \qquad (1)$$

where $\langle \Psi_{QM} | \hat{H} | \Psi_{QM} \rangle$ refers to the interactions within the QM region, while E_{MM} and E_{QM-MM} represent the interactions within the MM and between the QM and MM regions, respectively. The QM region, the most interesting subsystem, which includes HCOO⁻ and its nearest-neighbor water molecules, is treated quantum mechanically, while the rest of the system is described by classical pair potentials. Considering the exchange of water molecules between the QM and MM regions, which can occur frequently during the QM/MM simulations, the forces acting on each particle in the system are switched according to which region the water molecule is entering or leaving and can be defined as

$$F_i = S_m(r)F_{QM} + (1 - S_m(r))F_{MM} \qquad (2)$$

where F_{QM} and F_{MM} are quantum mechanical and molecular mechanical forces, respectively. $S_m(r)$ is a smoothing function[41]

$$
\begin{aligned}
S_m(r) &= 1, \quad \text{for } r \le r_1 \\
S_m(r) &= \frac{(r_0^2 - r^2)^2(r_0^2 + 2r^2 - 3r_1^2)}{(r_0^2 - r_1^2)^3}, \quad \text{for } r_1 < r \le r_0 \\
S_m(r) &= 0, \quad \text{for } r > r_0
\end{aligned}
\qquad (3)
$$

where r_1 and r_0 are distances characterizing the start and the end of the smoothing region, for which an interval of 0.2 Å has been found to be optimal to ensure a continuous change of forces at the boundary between QM and MM regions.

In the QM/MM technique, the quality of the simulation results crucially depends on the selection of QM method, basis set, and QM size. Because the performance of QM/MM MD simulations in conjunction with correlated ab initio methods is still far too time-consuming, the HF and the hybrid density functional B3LYP methods were the only possible alternatives for the present study. To simply check whether the HF and B3LYP methods are adequate for this particular system,
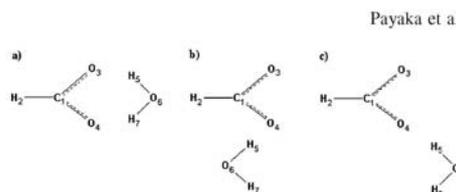
Figure 1. (a) Cyclic, (b) anti, and (c) syn HCOO⁻−H₂O complexes.

geometry optimizations of cyclic, anti, and syn HCOO⁻−H₂O clusters (see Figure 1) were carried out at HF, B3LYP, MP2, and CCSD levels of accuracy using DZV+[42] and aug-cc-pvdz[43–45] basis sets. As can be seen from the optimized parameters in Table 1, the B3LYP hydrogen-bond lengths and energies are close to those of the correlated ab initio methods using the larger basis set, while the HF results show good agreement with the correlated data when the smaller basis set, DZV+, is employed. Overall, the H-bond length and energy ordering of the three clusters predicted by the HF and B3LYP methods are in accord with the correlated results. This suggests that quantum mechanical calculations by both HF and B3LYP methods would be reliable enough to achieve a sufficient level of accuracy in the QM/MM simulations. The quality of the HF method has been well demonstrated in previous QM/MM studies,[17,28–35] even for the treatment of anions, implying that the effects of electron correlation are small enough to be neglected.[30,31] In a recent QM/MM MD simulation of pure water,[46] it has been shown that the HF method with a sufficiently large QM size could provide detailed information of pure water in good agreement with the MP2-based simulation and with experimental data concerning hydrogen-bond structure and lifetime. The B3LYP method, although inferior for most hydrated cations,[17,34] was also employed in the present work because it has been claimed that this method could predict reasonable data for weakly bound H-bonded systems.[35,47]

Because a satisfactory description of anions usually requires diffuse basis functions, the DZV+ basis set[42] was chosen, considered as a suitable compromise between the quality of the simulation results and the requirement of CPU time. To define the size of the QM region, preliminary HF/MM and B3LYP/MM simulations in which only the HCOO⁻ ion was treated quantum mechanically using HF and B3LYP methods, while the rest of the system was described by classical pair potentials, were performed. In the resulting C−O_w radial distribution functions (RDFs) (data not shown), both HF/MM and B3LYP/MM simulations showed first C−O_w minima at around 5.0 Å, and integrations up to this C−O_w distance yielded about 18−20 water molecules. This implied that a QM size with a radius of 5.0 Å seemed to be desirable for the present study. However, the evaluation of QM forces for all particles within this QM size is still beyond the limit of our computational facility. Therefore, a smaller QM size with radius of 4.0 Å was chosen (i.e., the values of r_1 and r_0 in eq 3 were set with respect to the C···O_w distances to 3.8 and 4.0 Å, respectively), which contains the central HCOO⁻ and about 12−14 nearest-neighbor water molecules.

A flexible model, which describes intermolecular[48] and intramolecular[49] interactions, was employed for water. This flexible water model allows explicit hydrogen movements, thus ensuring a smooth transition when water molecules move from the QM region with its full flexibility to the MM region. The pair potential functions for describing HCOO⁻−H₂O interactions were newly constructed. The 6015 HF and 5966 B3LYP

TABLE 1: Stabilization Energies and some Selected Structural Parameters of the Optimized Cyclic, Anti, and Syn HCOO⁻−H₂O complexes[a]

method	HF	B3LYP	MP2	CCSD
		Cyclic Complex		
ΔE (kcal·mol⁻¹)	−20.17 (−16.31)	−22.09 (−18.09)	−21.33 (−19.22)	−21.16 (−18.80)
R_{1-2} (Å)	1.1008 (1.1176)	1.1218 (1.1318)	1.1233 (1.1282)	1.1262 (1.1301)
R_{1-3}, R_{1-4} (Å)	1.2665 (1.2370)	1.2913 (1.2602)	1.3121 (1.2715)	1.3047 (1.2645)
R_{3-5}, R_{4-7} (Å)	2.1372 (2.1326)	2.0211 (2.0060)	2.0977 (1.9955)	2.1115 (2.0624)
R_{5-6}, R_{6-7} (Å)	0.9592 (0.9514)	0.9911 (0.9786)	0.9915 (0.9799)	0.9892 (0.9755)
A_{3-1-4} (deg)	128.12 (129.36)	128.16 (129.27)	128.18 (129.22)	128.18 (129.33)
A_{3-5-6}, A_{4-7-6} (deg)	137.87 (141.38)	140.74 (143.70)	140.25 (144.63)	139.84 (143.87)
A_{5-6-7} (deg)	104.71 (98.67)	101.11 (96.20)	102.01 (95.31)	102.33 (96.07)
		Anti Complex		
ΔE (kcal.mol⁻¹)	−18.278 (−14.13)	−20.38 (−16.42)	−18.78 (−16.75)	−18.52 (−16.25)
R_{1-2} (Å)	1.1034 (1.1185)	1.1236 (1.1315)	1.1258 (1.1286)	1.1289 (1.1306)
R_{1-3} (Å)	1.2544 (1.2274)	1.2756 (1.2482)	1.2981 (1.259)	1.2902 (1.2525)
R_{1-4} (Å)	1.2726 (1.244)	1.3004 (1.2703)	1.3189 (1.2808)	1.3125 (1.2740)
R_{4-5} (Å)	1.7066 (1.7871)	1.5815 (1.6503)	1.6844 (1.6596)	1.7081 (1.6912)
R_{5-6} (Å)	0.9782 (0.9666)	1.0284 (1.0087)	1.0166 (1.0075)	1.0094 (0.9983)
R_{6-7} (Å)	0.9506 (0.9420)	0.9770 (0.9631)	0.9796 (0.9644)	0.9799 (0.9630)
A_{3-1-4} (deg)	128.87 (129.42)	128.25 (128.82)	128.42 (128.91)	128.60 (129.09)
A_{5-6-7} (deg)	109.37 (103.41)	108.19 (102.71)	108.17 (101.95)	107.75 (102.16)
		Syn Complex		
ΔE (kcal.mol⁻¹)	−17.88 (−13.44)	−19.36 (−14.70)	−17.36 (−15.38)	−18.14 (−15.00)
R_{1-2} (Å)	1.1086 (1.1174)	1.1259 (1.1247)	1.1279 (1.1371)	1.1294 (1.1235)
R_{1-3} (Å)	1.2413 (1.2313)	1.2784 (1.2539)	1.3034 (1.2620)	1.2905 (1.2463)
R_{1-4} (Å)	1.2490 (1.2220)	1.3002 (1.2673)	1.3192 (1.2836)	1.3123 (1.2677)
R_{4-5} (Å)	1.7483 (1.8068)	1.7044 (1.7685)	1.8397 (1.7523)	1.7663 (1.7611)
R_{5-6} (Å)	0.9682 (0.9527)	1.0125 (0.9907)	1.0009 (0.9938)	1.0034 (0.9668)
R_{6-7} (Å)	0.9477 (0.9426)	0.9777 (0.9624)	0.9823 (0.9665)	0.9789 (0.9605)
A_{3-1-4} (deg)	129.45 (129.91)	129.27 (129.85)	129.07 (130.15)	129.22 (129.61)
A_{5-6-7} (deg)	108.27 (102.51)	106.41 (102.31)	105.60 (101.84)	106.48 (101.08)

[a] Calculated at HF, B3LYP, MP2, and CCSD methods using DZV+ and aug-cc-pvdz (data in parentheses) basis sets.

interaction energy points for various HCOO⁻−H₂O configurations, obtained from Gaussian98[50] calculations using aug-cc-pvdz basis set,[43−45] were fitted to the analytical forms of

$$\Delta E^{HF}_{HCOO^- - H_2O} = \sum_{i=1}^{4} \sum_{j=1}^{3} \left[\frac{A_{ij}}{r_{ij}^4} + \frac{B_{ij}}{r_{ij}^8} + C_{ij}\exp(-D_{ij}r_{ij}) + \frac{q_i q_j}{r_{ij}} \right] \quad (4)$$

and

$$\Delta E^{B3LYP}_{HCOO^- - H_2O} = \sum_{i=1}^{4} \sum_{j=1}^{3} \left[\frac{A_{ij}}{r_{ij}^4} + \frac{B_{ij}}{r_{ij}^5} + C_{ij}\exp(-D_{ij}r_{ij}) + \frac{q_i q_j}{r_{ij}} \right] \quad (5)$$

where A, B, C, and D are fitting parameters, r_{ij} denotes the distances between the ith atoms of HCOO⁻ and the jth atoms of the water molecule and q are atomic net charges. In the present study, the charges on C, O, and H of HCOO⁻ were obtained from Natural Bond Orbital (NBO) analysis[51−53] of the corresponding HF and B3LYP calculations, as 0.8635, −0.9201, and −0.0234 (HF) and 0.6583, −0.8197 and −0.0190 (B3LYP), respectively. The charges on O and H of water molecule were adopted from the BJH-CF2 water model[46] as −0.6598 and 0.3299, respectively. The optimized parameters for the intermolecular potentials (4) and (5) are listed in Table 2.

All simulations were performed in a canonical ensemble at 298 K with a time step of 0.2 fs. The system's temperature was kept constant using the Berendsen algorithm.[54] The periodic box, with a box length of 18.17 Å, contained one HCOO⁻ and 199 water molecules, corresponding to the experimental density of

pure water. Long-range interactions were treated using the reaction-field procedure.[55] In the present study, the HF/MM and B3LYP/MM simulations were carried out independently with system re-equilibration for 30000 time steps, followed by another 450000 (HF/MM) and 250000 (B3LYP/MM) time steps to collect configurations every tenth step.

3. Results and Discussion

3.1. Structural Details. The characteristics of hydrogen bonds between HCOO⁻ and water can be interpreted through the O−O$_w$ and O−H$_w$ RDFs, together with their corresponding integration numbers, as shown in Figure 2a and b, respectively. In this context, the first atom in the RDFs refers to the atom of HCOO⁻, and the latter, with the subscript "w", represents the atom of water molecules. Because the behavior of hydrogen bonds in pure solvent represents a most important reference, the corresponding atom−atom RDFs for pure water obtained at a similar QM/MM level of accuracy[56] were utilized for comparison, as shown in Figure 3. In the HF/MM simulation, the first O−O$_w$ peak is located at 2.76 Å, where integration up to the corresponding first O−O$_w$ minimum yields an average coordination number of 3.45. In the B3LYP/MM simulation, the first O−O$_w$ peak is found at the shorter distance of 2.67 Å, with a lower coordination number of 2.90. The second peaks in the O−O$_w$ RDFs are broad and less pronounced, and they correspond to the contributions from both bulk waters and water molecules in the first hydration layer of the other HCOO⁻ oxygen.

In accordance with the O−O$_w$ RDFs, the HF/MM and B3LYP/MM simulations reveal first O−H$_w$ peaks, an indicative of HCOO⁻−water hydrogen bonds, with maxima at 1.84 and 1.71 Å, respectively. Integrations up to the corresponding first

TABLE 2: Optimized Parameters of the Analytical Pair Potentials for the Interaction of Water with HCOO⁻[a]

pair	A	B	C	D
		HF-Based Method		
	kcal mol⁻¹ Å⁴	kcal mol⁻¹ Å⁸	kcal mol⁻¹	Å⁻¹
C−O$_w$	5.161390×10^5	-3.019987×10^4	-7.559447×10^5	1.955784
O−O$_w$	-2.665243×10^5	-5.276995×10^6	6.778200×10^7	3.556668
H−O$_w$	3.248683×10^5	-4.992158×10^5	-1.720448×10^5	1.318893
C -H$_w$	-2.368418×10^5	4.358412×10^5	7.864110×10^5	2.016779
O−H$_w$	5.356162×10^4	1.264105×10^4	-5.753688×10^5	3.105469
H−H$_w$	-3.889529×10^4	6.586156×10^3	1.236351×10^6	3.485256
		B3LYP-Based Method		
	kcal mol⁻¹ Å⁴	kcal mol⁻¹ Å⁵	kcal mol⁻¹	Å⁻¹
C−O$_w$	4.136774×10^5	-2.883495×10^5	-1.557409×10^4	0.429679
O−O$_w$	1.336006×10^6	-5.599642×10^6	5.529706×10^7	3.128274
H−O$_w$	9.683479×10^5	-9.298010×10^5	-3.549124×10^5	1.323072
C -H$_w$	2.720650×10^3	5.285593×10^4	2.761879×10^4	1.427650
O−H$_w$	-1.358645×10^5	1.516713×10^5	9.597746×10^5	1.431782
H−H$_w$	-8.338373×10^4	4.962553×10^4	7.038974×10^5	2.987790

[a] Interaction energies in kcal·mol⁻¹ and distances in Å.

Figure 2. (a) O−O$_w$ and (b) O−H$_w$ radial distribution functions and their corresponding integration numbers.

O−H$_w$ minima give the average values of 3.14 and 2.71, respectively. With respect to both HF/MM and B3LYP/MM simulations, the observed numbers of water oxygen and hydrogen atoms suggest that the first-shell waters are linearly hydrogen bonded to each of the HCOO⁻ oxygens, that is, they are acting as hydrogen-bond donors. In comparison to the first peak of pure water O$_w$−H$_w$ RDFs (cf. Figure 3), for example, in terms of shape and peak height, it is obvious that the O···H$_w$−O$_w$ hydrogen bond interactions are relatively stronger. In the HF/MM and B3LYP/MM simulations, the closest O···H$_w$ distances are 1.38 and 1.17 Å, respectively, compared to the shortest O$_w$···H$_w$ distance of 1.45 Å for bulk water. The second peak in the O−H$_w$ RDFs near 3.2 Å can be assigned to the hydrogen atoms of first-shell waters that are not hydrogen-bonded to the HCOO⁻ oxygens.

According to both HF/MM and B3LYP/MM simulations, each HCOO⁻ oxygen atom is predicted to form more hydrogen bonds with surrounding water molecules than the experimental
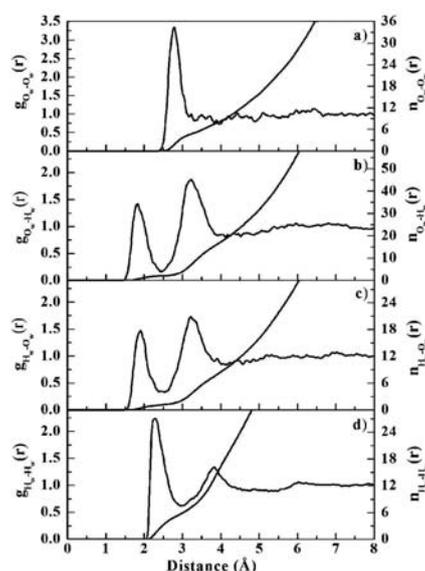
Figure 3. (a) O$_w$−O$_w$, (b) O$_w$−H$_w$, (c) H$_w$−O$_w$, and (d) H$_w$−H$_w$ radial distribution functions and their corresponding integration numbers. The first atom of each pair refers to the atoms of the water molecule, whose oxygen position was defined as the center of the QM region during the QM/MM simulation.

value of about 2.5.[37,38] In fact, as the first minimum peaks of O−O$_w$ and O−H$_w$ RDFs are rather broad, the number of first-shell waters is quite sensitive to the defined O−O$_w$ and O−H$_w$ minima. For example, integrations up to O−O$_w$ distance of about 0.2 Å shorter than the corresponding O−O$_w$ minima yield 2.53 and 2.47 water oxygens for the HF/MM and B3LYP/MM simulations, respectively. Thus, the differences between our data and experiment are well within the methodical limits of integration on the one hand and interpretation of experimental data on the other. The recent CPMD study,[40] although providing
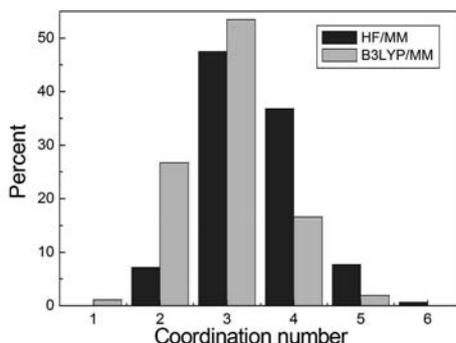
Figure 4. Distributions of the number of water's oxygen atoms at each of HCOO⁻ oxygens, calculated within the first minimum of the $O-O_w$ RDFs.
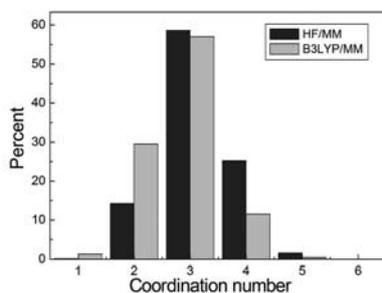
Figure 5. Distributions of the number of water's hydrogen atoms at each of HCOO⁻ oxygens, calculated within the first minimum of the $O-H_w$ RDFs.

Figure 6. Time dependence of the number of first-shell waters at HCOO⁻ oxygen atoms, selecting only for first 10 ps of the HF/MM simulation.

Figure 7. Time dependence of the number of first-shell waters at HCOO⁻ oxygen atoms, selecting only for first 10 ps of the B3LYP/MM simulation.

a coordination number apparently identical with experiment, delivers deviating results when a different simulation cell size as well as different functionals (i.e., BLYP and PW91) are employed in the simulations. For example, the use of the BLYP functional had predicted a hydration number of 2.45 per HCOO⁻ oxygen, while the PW91 method gave a large variation in the hydration numbers ranging from 2.12 to 2.66. Possible weaknesses of the density functional methods could be attributed to the incompleteness of the kinetic energy term, the self-interaction error, and the more or less empirical parametrization, which did not contain any hydrogen-bonded system.

The distributions of oxygen and hydrogen atoms of first-shell waters, calculated with respect to the first minimum of the $O-O_w$ and $O-H_w$ RDFs, are depicted in Figures 4 and 5, respectively. In the HF/MM simulation, the most frequent hydration number per HCOO⁻ oxygen atom is 3, followed by 4, 2, and 5 in decreasing amounts. In the B3LYP/MM simulation, a prevalent value of 3 is also observed, followed by 2 and 4 in smaller amounts. Figures 6 and 7 show examples of time dependence of the hydration number at each of HCOO⁻ oxygens within the first 10 ps of the HF/MM and B3LYP/MM simulations, respectively. In both HF/MM and B3LYP/MM trajectories, it is found that the two HCOO⁻ oxygen atoms simultaneously form asymmetric solvation shells, that is, each of them hydrogen bonding to different numbers of water molecules. Consequently, this causes numerous possible species

of the HCOO⁻−water complexes to coexist in aqueous solution. As can be seen in Figures 6 and 7, the total numbers of water molecules in the vicinity of HCOO⁻ oxygens show large fluctuations, ranging from 5 to 10 and from 4 to 8 for the HF/MM and B3LYP/MM simulations, respectively.

A more detailed interpretation of the HCOO⁻−water hydrogen bonds can be deduced from the probability distributions of the $C-O\cdots H_w$ and $O\cdots O_w-H_w$ angles, calculated within the first minimum of the $O-O_w$ RDFs, as shown in Figures 8 and 9, respectively. In the case that solvent effects would cause strong charge localization on HCOO⁻, an asymmetrical charge distribution at the two HCOO⁻ oxygens, corresponding to a formation of a $C-O$ single and a $C=O$ double bond, could exist in aqueous solution. With regard to this point, one could expect the arrangement of directional $C-O\cdots H_w$ hydrogen bonds that cause the $C-O\cdots H_w$ angle to peak at 109.5° and 120°, depending on the type of involved oxygen. In both HF/MM and B3LYP/MM simulations, the observed broad $C-O\cdots H_w$ angular distributions (Figure 8) clearly suggest the absence of such phenomenon. In addition, with regard to the Mulliken charge analyses of several HCOO⁻−water complexes, there is no substantial charge concentration at one of the HCOO⁻ oxygens. Apparently, HCOO⁻ adopts an electronically delocalized structure in aqueous solution, which may fluctuate due

Payaka et al.
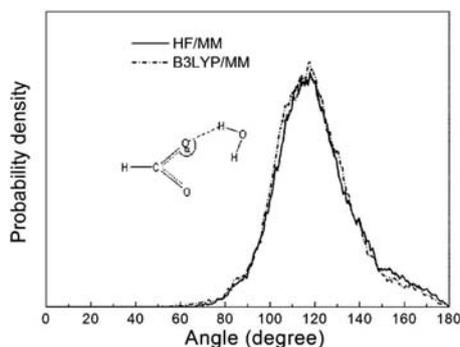
Figure 8. Distributions of C—O···H$_w$ angle, calculated within first minimum of the O—O$_w$ RDFs.
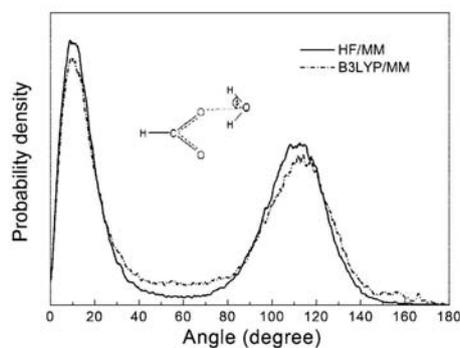
Figure 9. Distributions of O···O$_w$—H$_w$ angle, calculated within first minimum of the O—O$_w$ RDFs.

Figure 10. Distributions of C—H and C—O bond lengths.
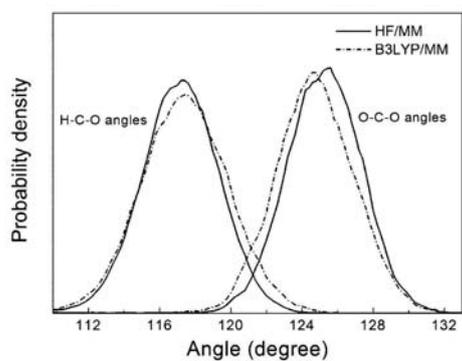
Figure 11. Distributions of H—C—O and O—C—O angles.
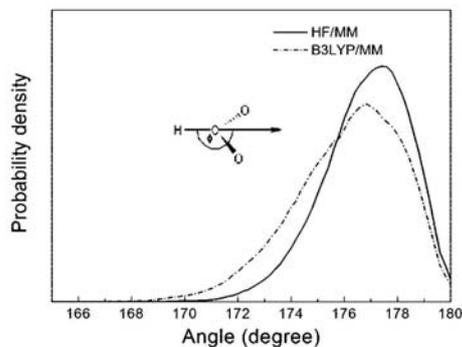
Figure 12. Distributions of ϕ, as defined by a vector along the C—H bond and a vector pointing outward between the two C—O bonds.

to solvent exchange processes. Figure 9 shows the distributions of O···O$_w$—H$_w$ angle, which clarify the preference for linear O···H$_w$—O$_w$ arrangements.

3.2. Dynamical Details. In this section, the dynamical data regarding the intramolecular geometry of HCOO$^-$ and the exchange processes of water molecules at HCOO$^-$ oxygens are reported. The geometrical arrangement of HCOO$^-$ in aqueous solution is explained in terms of the distributions of C—O and C—H bond lengths as well as of H—C—O and O—C—O angles, as shown in Figures 10 and 11, respectively. In addition, the distribution of the angle ϕ, as defined by a vector along the C—H bond and a vector pointing outward between the other two C—O bonds, is also given in Figure 12. As compared to the gas phase HCOO$^-$ structure, both HF/MM and B3LYP/MM simulations clearly indicate a substantial change in the local structure of HCOO$^-$ according to the influence of water environment, in particular, a C—O bond lengthening, shortening of the C—H bond and a decrease of the O—C—O angle.

According to Figure 2, the nonzero first minimum of the O—O$_w$ and O—H$_w$ RDFs obtained by both HF/MM and B3LYP/MM simulations clearly suggests an easy exchange of water molecules between the solvation shell and the bulk. The exchange processes of first-shell waters at each of the HCOO$^-$ oxygen atoms can be visualized through the plots of the O—O$_w$ distances against the simulation time, as shown in Figures 13

and 14 for the HF/MM and B3LYP/MM simulations, respectively. During the first 10 ps of the HF/MM and B3LYP/MM trajectories, numerous water molecules can be interchanged between the first shell and the bulk, leading to large fluctuations in the hydration number at each of the HCOO$^-$ oxygen atoms (e.g., see inserts in Figures 13 and 14). Inside the hydration shell, water molecules are either loosely or tightly bound to

QM/MM MD Study of HCOO⁻−Water Hydrogen Bonds
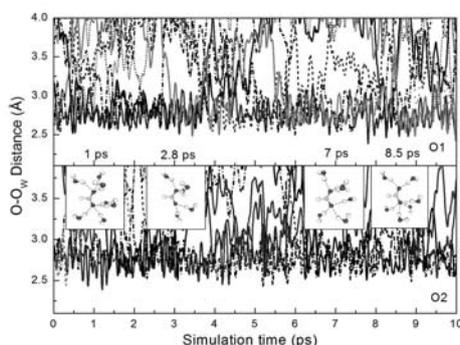
Figure 13. Time dependence of O---O_w distances, selecting only for first 10 ps of the HF/MM trajectories.
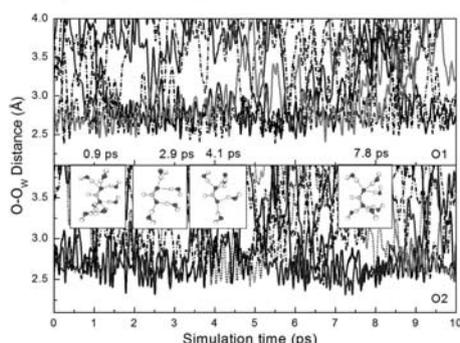
Figure 14. Time dependence of O---O_w distances, selecting only for first 10 ps of the B3LYP/MM trajectories.

HCOO⁻ oxygen, that is, some first-shell waters temporarily form a hydrogen bond with HCOO⁻ oxygen, then leaving or even entering again, while others form longer hydrogen bonds to the respective HCOO⁻ oxygen. On the basis of both HF/MM and B3LYP/MM simulations, the occurrence of bifurcated hydrogen bonds (cf. Figure 1a) appears rare in aqueous solution. First-shell water molecules also preferentially associate with one HCOO⁻ oxygen atom or the other, rather than adopting simultaneous coordination to both.

The rate of water exchange processes at each HCOO⁻ oxygen atom was evaluated through mean residence times (MRT) of the surrounding water molecules. In this work, the MRT data were calculated using the direct method,[57] as the product of the average number of nearest-neighbor water molecules located within the first minimum of the O−O_w RDFs with the duration of the simulation, divided by the number of exchange events. With respect to time parameters t^* (i.e., the minimum duration of a ligand's displacement from its original coordination shell to be accounted) of 0.0 and 0.5 ps, the calculated MRT values are summarized in Table 3. In general, the MRT data obtained using $t^* = 0.0$ ps are used for an estimation of hydrogen bond lifetimes, whereas the data obtained with $t^* = 0.5$ ps are considered as a good estimate for sustainable ligand exchange processes.[57] In the HF/MM simulation, the calculated MRT values with respect to $t^* = 0.0$ and 0.5 ps are slightly larger

TABLE 3: Mean Residence Times of Water Molecules in the Bulk and in the Vicinity of HCOO⁻ Oxygens[a]

atom/solute	CN	t_{sim}	$t^* = 0.0$ ps		$t^* = 0.5$ ps	
			$N_{ex}^{0.0}$	$\tau_{H_2O}^{0.0}$	$N_{ex}^{0.5}$	$\tau_{H_2O}^{0.5}$
HF/MM MD						
O1	3.45	70.0	970	0.25	132	1.83
O2	3.44	70.0	1042	0.23	112	2.15
pure H$_2$O[56]	4.6	12.0	292	0.2	31	1.8
pure H$_2$O[46]	4.2	40.0		0.33		1.51
B3LYP/MM MD						
O1	2.84	50.0	810	0.17	57	2.49
O2	2.97	50.0	799	0.19	63	2.36
pure H$_2$O[46]	4.2	30.0		1.07		7.84

[a] Calculated within the first minimum of the O−O_w RDFs.

than that of pure water.[56] These data correspond to the observed stronger hydrogen bonds between HCOO⁻ oxygens and their first-shell water molecules. In the B3LYP/MM simulation, as compared to the B3LYP/MM data for pure water,[46] a clear opposite order of $\tau_{H_2O}(O_i) < \tau_{H_2O}(H_2O)$ is observed. Here it should be noted that the B3LYP/MM results for pure water have produced too slow exchange rates compared to the experimental values.[58] This failure of the B3LYP method to predict the dynamical properties of pure water could be considered as an example for the inadequacy of the DFT methods to correctly describe the characteristics of any aqueous hydrogen-bonded systems. Slow dynamics of aqueous HCOO⁻, in particular, the HCOO⁻ rotation, were also observed in the recent CPMD study when the PW91 functional was employed.[40]

4. Conclusion

The QM/MM simulations presented here have produced a detailed picture of HCOO⁻-water hydrogen bonds in dilute aqueous solution. Both HF/MM and B3LYP/MM simulations predict relatively strong hydrogen bonds between HCOO⁻ oxygens and first-shell waters. The geometrical arrangement of HCOO⁻ in aqueous solution is found to be rather flexible and first-shell water molecules can be either "loosely" or "tightly" bound to each of HCOO⁻ oxygen atoms, forming an asymmetric solvation structure with a varying number of hydrogen bonds, with the prevalent value of 3. Despite the lack of electron correlations, which leads to slightly too long and weak hydrogen bonds, the ab initio HF method proves more reliable than the B3LYP or other DFT approaches, as the latter completely fails to describe the dynamics of ion hydration. Thus, only an extension of the ab initio QM/MM technique by increasing the QM size, enlarging the basis sets and finally including electron correlations would allow a future improvement of the results.

References and Notes

(1) Frank, H. *Chemical Physics of Ionic Solutions*; John Wiley & Sons: New York, 1956.
(2) Williams, R. J. P. *Bio-inorganic Chemistry*; American Chemical Society: Washington, DC, 1971.
(3) Cowan, J. A. *Chem. Rev.* **1998**, *98*, 1067.
(4) Ohtaki, H.; Radnai, T. *Chem. Rev.* **1993**, *93*, 1157.

Payaka et al.

(5) Dang, L. X.; Schenter, G. K.; Glezakou, V. A.; Fulton, J. L. *J. Phys. Chem. B* **2006**, *110*, 23644.

(6) Angelo, P. D.; Pavel, N. V. *J. Synchrotron Radiat.* **2001**, *8*, 173.

(7) Hewish, N. A.; Neilson, G. W.; Enderby, J. E. *Nature* **1982**, *297*, 138.

(8) Howell, I.; Neilson, G. W. *J. Phys.: Condens. Matter* **1996**, *8*, 4455.

(9) Impey, R. W.; Madden, P. A.; McDonald, I. R. *J. Phys. Chem.* **1983**, *87*, 5071.

(10) Bopp, P. *The Physics and Chemistry of Aqueous Ionic Solutions*; Reidel Publishing Company: Boston, MA, 1987; p 217.

(11) Heinzinger, K. *Pure Appl. Chem.* **1985**, *57*, 1031.

(12) Lee, S. H.; Rasaiah, J. C. *J. Phys. Chem.* **1996**, *100*, 1420.

(13) Obst, S.; Bradaczek, H. *J. Phys. Chem.* **1996**, *100*, 15677.

(14) Ebner, C.; Sansone, R.; Hengrasmee, S.; Probst, M. *Int. J. Quantum Chem.* **1999**, *75*, 805.

(15) Stone, J. A. *The Theory of Intermolecular Forces*; Clarendon Press: Oxford, 1996.

(16) Elrod, M. J.; Saykally, R. J. *Chem. Rev.* **1994**, *94*, 1975.

(17) Rode, B. M.; Schwenk, C. F.; Tongraar, A. *J. Mol. Liq.* **2004**, *110*, 105.

(18) Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985**, *55*, 2471.

(19) Marx, D.; Hutter, J. In *Modern Methods and Algorithms of Quantum Chemistry*; Grotendorst, J., Ed.; NIC: FZ Jülich, 2000.

(20) Chen, B.; Ivanov, I.; Park, J. M.; Parrinello, M.; Klein, M. L. *J. Phys. Chem.* **2000**, *106*, 12006.

(21) Tuckerman, M. E.; Marx, D.; Klein, M. L.; Parrinello, M. *Science* **1997**, *275*, 817.

(22) Brugé, F.; Bernasconi, M.; Parrinello, M. *J. Am. Chem. Soc.* **1999**, *121*, 10883.

(23) Tuckerman, M. E.; Marx, D.; Parrinello, M. *Nature* **2002**, *417*, 925.

(24) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227.

(25) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1986**, *7*, 718.

(26) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700.

(27) Gao, J. *Rev. Comput. Chem.* **1996**, *7*, 119.

(28) Kerdcharoen, T.; Liedl, K. R.; Rode, B. M. *Chem. Phys.* **1996**, *211*, 313.

(29) Tongraar, A.; Liedl, K. R.; Rode, B. M. *J. Phys. Chem. A* **1998**, *102*, 10340.

(30) Tongraar, A.; Rode, B. M. *Phys. Chem. Chem. Phys.* **2003**, *5*, 357.

(31) Tongraar, A.; Rode, B. M. *Chem. Phys. Lett.* **2005**, *403*, 314.

(32) Tongraar, A.; Rode, B. M. *Chem. Phys. Lett.* **2005**, *409*, 304.

(33) Intharathep, P.; Tongraar, A.; Sagarik, K. *J. Comput. Chem.* **2005**, *26*, 1329.

(34) Rode, B. M.; Schwenk, C. F.; Hofer, T. S.; Randolf, B. R. *Coord. Chem. Rev.* **2005**, *249*, 2993.

(35) Tongraar, A.; Tangkawanwanit, P.; Rode, B. M. *J. Phys. Chem. A* **2006**, *110*, 12918.

(36) Kuntz, I. D. *J. Am. Chem. Soc.* **1971**, *93*, 514.

(37) Kameda, Y.; Mori, T.; Nishiyama, T.; Usuki, T.; Uemura, O. *Bull. Chem. Soc. Jpn.* **1996**, *69*, 1495.

(38) Kameda, Y.; Fukuhara, K.; Mochiduki, K.; Naganuma, H.; Usuki, T.; Uemura, O. *J. Non-Cryst. Solids* **2002**, *312–314*, 433.

(39) Jorgensen, W. L.; Gao, J. *J. Phys. Chem.* **1986**, *90*, 2174.

(40) Leung, K.; Rempe, S. B. *J. Am. Chem. Soc.* **2004**, *126*, 344.

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

(42) Dunning, T. H. Jr.; Hay, P. J. *Modern Theoretical Chemistry, III*; Plenum: New York, 1976.

(43) Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007.

(44) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6769.

(45) Woon, D. E.; Dunning, T. H. *J. Chem. Phys.* **1993**, *98*, 1358.

(46) Xenides, D.; Randolf, B. R.; Rode, B. M. *J. Chem. Phys.* **2005**, *122*, 174506.

(47) Tongraar, A.; Kerdcharoen, T.; Hannongbua, S. *J. Phys. Chem. A* **2006**, *110*, 4924.

(48) Stillinger, F. H.; Rahman, A. *J. Chem. Phys.* **1976**, *68*, 666.

(49) Bopp, P.; Jancsó, G.; Heinzinger, K. *Chem. Phys. Lett.* **1983**, *98*, 129.

(50) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrewski, V. G.; Montgomery, J. A.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Peterson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*; Gaussian, Inc.: Pittsburgh, PA, 1998.

(51) Carpenter, J. E.; Weinhold, F. *J. Mol. Struct.: THEOCHEM* **1988**, *169*, 41.

(52) Reed, A. E.; Weinstock, R. B.; Weinhold, F. *J. Chem. Phys.* **1985**, *83*, 735.

(53) Reed, A. E.; Curtiss, L. A.; Weinhold, F. *Chem. Rev.* **1988**, *88*, 899.

(54) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Phys. Chem.* **1984**, *81*, 3684.

(55) Adams, D. J.; Adams, E. H.; Hills, G. J. *Mol. Phys.* **1979**, *38*, 387.

(56) Tongraar, A.; Rode, B. M. *Chem. Phys. Lett.* **2004**, *385*, 378.

(57) Hofer, T. S.; Tran, H. T.; Schwenk, C. F.; Rode, B. M. *J. Comput. Chem.* **2004**, *25*, 211.

(58) Lock, A. J.; Woutersen, S.; Bakker, H. J. *Femtochemistry and Femtobiology*; World Scientific: Singapore, 2001.

CURRICULUM VITAE

Mr. APIRAK PAYAKA

Education Background:

2005	B.Sc. First Class Honor (Chemistry), Maejo University, Chiang Mai, Thailand
2006	Ph.D. Candidate (Chemistry), Suranaree University of Technology, Nakhon Ratchasima, Thailand

Experiences:

2005-2006	Teaching Assistant, School of Chemistry, Institute of Science, Suranaree University of Technology, Nakhon Rachasima, Thailand
2006-2008	Research Assistant, Beam Line8 (BL-8) X-ray absorption spectroscopy Synchrotron Light Research Institute (Public Organization), Nakhon Rachasima, Thailand

Grants and Fellowships:

2006-2010	The Royal Golden Jubilee (RGJ) Ph.D. Scholarship from Thailand Research Found (TRF)
2006-2007	SUT Teaching Assistantship

Awards and Honors:

2002	Professor Dr. Tab Nilanidhi Foundation (The Tab Foundation) for the outstanding first year undergraduate student in chemistry
2005	Professor Dr. Tab Nilanidhi Foundation (The Tab Foundation) for the outstanding undergraduate student in chemistry