5-2-2009

Solvent methods in coupled-cluster theory

Kanchana Sahan Thanthiriwatte

Solvent methods in coupled-cluster theory

Comments

Coupled-Cluster Theory||Solvent Models||Computational Chemistry||SCRF||solvent reaction field method||RISM||Electrophilic Aromatic Nitration of Benzene||reference interaction site model||Quantum Chemistry

SOLVENT METHODS IN COUPLED-CLUSTER THEORY

By

Kanchana Sahan Thanthiriwatte

A Dissertation
Submitted to the Faculty of
Mississippi State University
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
in Chemistry
in the Department of Chemistry

Mississippi State, Mississippi

May 2009

SOLVENT METHODS IN COUPLED-CLUSTER THEORY

By

Kanchana Sahan Thanthiriwatte

Approved:

Steven R. Gwaltney Associate Professor of Physical Chemistry, Department of Chemistry (Major Professor)	Svein Saebø Professor of Physical Chemistry, Department of Chemistry (Committee Member)
Seong-Gon Kim Associate Professor of Physics, Department of Physics & Astronamy (Committee Member)	Edwin A. Lewis Professor of Chemistry & Head of the Department, Department of Chemistry (Committee Member)
Stephen C. Foster Associate Professor of Physical Chemistry & Graduate Coordinator, Department of Chemistry (Committee Member)	Gary L. Myers Dean of the College of Arts & Sciences

Name: Kanchana Sahan Thanthiriwatte

Date of Degree: May 2, 2009

Institution: Mississippi State University

Major Field: Chemistry

Major Professor: Dr. Steven R. Gwaltney

Title of Study: SOLVENT METHODS IN COUPLED-CLUSTER THEORY

Pages in Study: 121

Candidate for Degree of Doctor of Philosophy

This dissertation describes the implementation of the molecular electronic structure calculations with an implicit solvent model using coupled-cluster theory. The theory for and the implementation of the solvent reaction field method (SCRF) and the reference interaction site model (RISM) at the coupled-cluster singles and doubles (CCSD) are presented.

In the SCRF model a solute molecule is placed in a spherical cavity, and the outer solvent is represented by a dielectric continuum, which is characterized by the dielectric constant of the solvent. The reaction field is introduced to the system by using the multipole moment expansion of the electronic structure of the solute molecule and the dielectric constant. The SCRF method has been used to calculate the conformational equilibrium and the rotational barriers of 1,2-dichloroethane in vacuum and in different solvents. The calculated results are compared with experimental values. In addition, the solvent effects

on the energetics of the mechanism of nitration of benzene are reported using the implemented CCSD-SCRF model.

The idea of RISM is to replace the reaction field in continuum models by a microscopic expression in terms of the site-site radial distribution functions between solute and solvent, which can be calculated from the RISM integral equations. The statistical solvent distribution around the solute is determined based on the electronic structure of the solute, while the electronic structure of solute is influenced by the surrounding solvent distribution. Therefore, the wave function and the RISM equations are solved self-consistently with CCSD. Pair correlation functions, partial atomic charges, and solvation free energies of water and *N*-methylacetamide are calculated in liquid water using proposed theory.

Both the CC-SCRF and CC-RISM methods have been implemented in a developmental version of the Q-CHEM 3.2 quantum chemistry package.

Key words: Quantum Chemistry, Computational Chemistry, Coupled-Cluster Theory, Solvent Models, SCRF, solvent reaction field method, Electrophilic Aromatic Nitration of Benzene, RISM, reference interaction site model

DEDICATION

To my parents and my wife.

ACKNOWLEDGMENTS

It gives me immense pleasure to acknowledge many great individuals who have contributed to the success of this dissertation.

First of all, I would like to take this opportunity to thank my advisor Dr. Steven R. Gwaltney for his valuable guidance, support and encouragement through out these years in graduate school. He is an ideal advisor, he provided ample freedom and flexibility to pursue my interests at my own pace. I am grateful for his constructive scientific input that helped me grow as a researcher, and without which this work would not be possible. I want to thank him again for conducting a series of intensive summer classes for me to learn electronic structure methods including coupled-cluster theory and for teaching me to think in many-body methods using diagrams. Amidst a very busy schedule, he always had time to discuss my progress and offered help on problems. Once again, my whole-hearted thanks go to him.

Words fail me to express my appreciation to my wife Chamali, whose dedication, love and persistent confidence in me has taken the load off my shoulder. I am very fortunate to share my life with such an understanding partner who was and is always there for me. Her unwavering support and steadfast belief were crucial for me to realize my dreams. I owe her for unselfishly letting her intelligence, passions, and ambitions collide with mine.

My very special thanks goes to my parents, to whom I owe everything I am today, my beloved mother Rupa and beloved father Indrasena. They are my first teachers, who taught me good and bad, and who sincerely raised me with their caring and gentle love,. Their consistent guidance and unwavering faith and confidence in my abilities and in me is what has shaped me to be the person I am today. Thank you for everything. Many thanks go to my brothers, Kasun and Sharanga for ther love and support. I would also like to thank Piyasena and Asoka for letting me take Chamali's hand in marriage and for accepting me as a member of the family, warmly.

I extend my gratitude to Dr. K. M. Nalin De Silva, my undergraduate research advisor. If he would not have directed me to choose computational chemistry research under him, I could not be a theoretist today. I grateful to my high school chemistry teacher, Mr. D. P. Dewruwan Alwis, who inspired my passion in chemistry. Without his guidance, I would not have choosen chemistry as my major in college.

I must also thank the Department of Chemistry for providing me the financial assistance all these years in graduate school. Especially I express my sincere gratitude to Dr. Stephen Foster, the graduate coordinator of the department for doing a silent but very important service. I would also acknowledge Mr. Steven Holman for his help for making my TA life easier.

I would like to thank the staff of the Mississippi State University Libraries. Their kindness and assistance will always be remembered.

The colleagues I have met while in graduate school have become my closest and dearest friends and counselors, and to all of you I give my love and thanks.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS, ABBREVIATIONS, AND NOMENCLATURE

3D	Three Dimentional
AIMD	Ab Initio Molecular Dynamics
AO	Atomic Orbitals
BO	Born-Oppenheimer (BO) Approximation
CC	Coupled-Cluster
CC-SCRF	Coupled-Cluster Self-Consistent Reaction Field
CHARMM	Chemistry at Harvard using Molecular Mechanics
CI	Configuration Interaction
CISD	Configuration Interaction with all Single and Double Substitution
FCI	Full Configuration Interaction
HF	Hartree-Fock
HF-SCF	Hartree-Fock Self-Consistent Field
MBPT	Many-Body Perturbation Theory
MC	Metropolis Monte Carlo
MCSCF	Multi-Configurational Self-Consistent Field
MD	Molecular Dynamics
MM	Molecular Mechanics

MO	Molecular Orbitals
MP2	Second Order Møller-Plesset Perturbation Theory
NMR	Nuclear Magnetic Resonance
OPDM	One Particle Density Matrix
OZ	Ornstein-Zernike
PES	Potential Energy Surface
QM/MM	Quantum Mechanics/Molecular Mechanics
DC	Dielectric Continuum
RF	Reaction Field
RISM	Reference Interaction Site Model
SCF	Self-Consistent Field
SCRF	Self-Consistent Rreaction Field

CHAPTER 1

INTRODUCTION

An overview of theoretical studies over the last decades shows the extraordinary evolution in quantum mechanical methods and their applications to the study of chemical systems. The extensive research efforts put into molecular electronic structure theory have generated a series of sophisticated methods, that approach to the limit of experimental accuracy in the gas phase. The Nobel Prize awarded to Walter Kohn and John Pople in 1998 [1]is a clear recognition of the impact of quantum chemistry in the chemical science. There is no doubt that quantum chemistry constitutes an extremely powerful tool to understand the structure and reactive properties of molecules in the gas phase.

Even though quantum chemistry for small, isolated molecules has reached the limit of chemical accuracy, the same cannot be said for molecules in a solvent. In condensed phase the molecules of interest are surrounded by and significantly affected by the nature of the solvent. The existence of life depends on complex biochemical reactions taking place within aqueous environments both in vitro and in vivo [1]. The environment plays a key role in the determination of the properties and reactivity of substances in condensed phase [2]. Besides its biological significance, condensed phase chemistry is important to

[1]to Walter Kohn for his development of density functional theory and to John Pople for his development of computational methods in quantum chemistry.

most chemical manufacturing processes, both in the laboratory test tube and in industry. Hence, it is necessary to include the solvent description to molecular electronic structure calculations to get a more detailed understanding in a wide area of chemistry and essentially all of biochemistry.

In recent years there has been growing interest in the interpretation of solvent effects on the properties of molecules [3–7]. The computational calculations of chemical systems in condensed phases is, however, far more difficult than that of isolated molecules. The primary reason that condensed-phase problems are complicated is the intractability of solving the Schrödinger equation for large systems. Although the nuclear degrees of freedom may be rendered separable from the electronic ones by invocation of the Born-Oppenheimer approximation [8], the electronic degrees of freedom remain far too numerous to be handled practically with any quantum mechanical approach and with current computer capability. When considering intermolecular interactions and the size of the whole system, both the solute molecule and the solvent molecules together, the system gets too complicated to handle. Consider, for example, a case of a small solute molecule in water. Defining a cut-off for the electrostatic interaction at 12 Å will amount to considering around 200 explicit water molecules in addition to the solute molecule [9]. If we treat 200 molecules of a solvent explicitly, this adds electronic 6000 degrees of freedom for the water. In addition, the solvent is inherently dynamical by nature, and therefore configurational sampling must be also considered.

The most rigorously correct way of modeling chemistry in solution would be to insert all the solvent molecules explicitly and then run molecular dynamics (MD) [10–14]

or Monte Carlo (MC) [15–17] simulations to give a time-averaged, ensemble average of the property of interest. This can be done using molecular mechanics (MM), where the solvent-solute system is replaced by a classical one in which the electronic energy plus the coulombic interactions of the nuclei, taken together, are modeled by a classical force field. The primary problems with MM explicit solvent calculations are the lack of an ab initio quantum mechanical description of the system and the significant amount of time and substantial amount of computer resources required. Therefore, one can use quantum mechanics/molecular mechanics (QM/MM) [18–22], in which the system is divided into two parts, where the solute is considered as a quantum system, and the solvent is considered as a classical system. Both QM/MM and fully MM methods have computational bottlenecks such as the sampling problem, difficulty to converge, and expense when long-range forces are included [23]. A solution for introducing a quantum mechanical description to MM simulations is ab initio molecular dynamics (AIMD) [24–28]. Traditional "force fields" in MD simulations can be replaced with quantum mechanical techniques in AIMD. In an AIMD calculation, finite-temperature dynamical trajectories are generated using forces obtained directly from electronic structure calculations performed "on the fly" as the simulation proceeds. One can say that AIMD is a solution for the solute-solvent problem. Yes, but AIMD needs much larger amounts of both computational power and time than all other methods described above.

Methods involving an explicit description of the solvent molecules require, analogously with other many-body methods, a sampling of the phase space. Since this is computationally expensive, there is a strong interest in developing methods where the sol-

vent is modeled in a less rigorous fashion. Such complexity, which was described in the above paragraph, has given rise to a wide variety of simplified computational approaches, which can be basically classified into methods based on: (a) a supermolecule description of the solute-solvent system [29–32], which provides limited, but detailed information about specific solute-solvent interactions; (b) methods based on the dielectric continuum (DC) model [3–7, 33–40], where the attention is mainly focused on one component of the system, the solute, whereas the solvent is treated in a very simplified way as a polarizable medium, and (c) statistical mechanical based distribution functions with an integral equation theory [41–45], where the combination is of quantum mechanical treatments of the solute with statistically averaged descriptions of discrete solvent molecules. In this dissertation, we will consider enhancement of a continuum solvent model, which is a self-consistent reaction field (SCRF) solvent model that falls under class (b), and a method that uses statistical mechanical theory-based distribution functions with an integral equation theory, which is the reference interaction site model (RISM) method that falls under class (c).

In the continuum solvent model, the solvent is modeled as a macroscopic dielectric continuum (DC) characterized by a dielectric constant. Thus, any reference to the atomistic nature of the solvent is neglected, and configurational sampling is included implicitly. Therefore, for this model, no explicit simulations have to be conducted, and no explicit sampling is necessary. Because of their speed and simplicity, the continuum models of solvation are very popular in condensed phase computational chemistry [3–5]. As stated above, the solvent may be modeled as a configuration-averaged or time-averaged solvent

environment, where the averaging is Boltzmann weighted at the temperature of interest. The DC approach is thus also sometimes referred to as a "mean-field" approach or "reaction field" approach. The Reaction field DC model provides a simple description for solvation by including the electrostatic effects to the molecule wave function. The classical dielectric continuum is characterized by its dielectric constant, and the solute is embedded in a spherical cavity. The time averaged reaction field (RF) is calculated by the dielectric constant of solvent, the radius of the cavity in which the solute enclosed, and the multipole expansion of the solute charge distribution. The charge distribution of the solute molecule polarizes the solvent, and therefore a new external electric field, the reaction field, perturbs the solute wave function. Then the solute charge distribution from the new wave function again polarizes the solvent. With new RF the solute itself is allowed to be back-polarized by the reaction field. This procedure is called the self-consistent reaction field (SCRF) solvent model.

This model was originally proposed by Born [33], Kirkwood [34, 35], and Onsager [36] and has been established in numerous applications in condensed phases [3–5]. The Onsager-SCRF method was proposed by Tapia et al. [38] and later by Wong et al. [46]. The Onsager model is the simplest version of the DC approach. Solvation is described in terms of a dipole moment, drawn in an iterative way from QM calculations on the molecule. A more general model going beyond the dipole approximation is that developed by Mikkelsen and co-workers [39, 40, 47]. Mikkelsen et al. exploit Kirkwood's idea of describing the interaction between a set of classical charges (described in terms of a multipole expansion) enclosed in a spherical cavity embedded in a DC medium. Mikkelsen

applies the model to QM charge distributions instead of classical point charges, as in the original Kirkwood paper, and the effects of the solvent polarization are described in terms of proper QM operators to be added to the Hamiltonian of the isolated system. It is worth noting that another approach towards SCRF was developed in Nancy, France by Rinaldi and co-workers [37, 48–52]. The main features of this method is an ellipsoidal cavities and ellipsoidal multipole expansions, instead of traditional spherical cavities and multipole expansions. Historically, this method represents the first example of QM continuum solvation methods [37, 48], but over the years it has been continuously developed, and it continues its evolution [51, 52]. The SCRF methods have so far been described in self-consistent field (SCF) theories [37, 38, 46, 47], multiconfigurational SCF (MCSCF) theory [40], second order Møller-Plesset perturbation theory (MP2) [53, 54], and coupled-cluster (CC) theory [55]. Previous work on the CC solvent theory was based on orbital unrelaxed methods in which the orbital response to the reaction field is not included. In this dissertation, we discuss the introduction of solute–solvent interactions into CC methods and introduce a coupled-cluster self-consistent reaction field method (CC-SCRF) including orbital relaxation [56].

Statistical mechanical theory-based distribution functions with integral equation theory for liquids was introduced by Chandler and Andersen in 1971, normally referred to as the the reference interaction site model (RISM) [42]. This theory can be regarded as a natural extension of the Ornstein-Zernike (OZ) equation [57] for simple atomic liquids to a mixture of atoms with chemical bonds represented by intramolecular correlation functions. Introducing this correlation function enables us to take into account the geometry

of molecules. However, it cannot handle electrostatics in its original form, even though the charge distribution in a molecule plays an essential role in determining the chemical specificity of the molecular system. The next important development in the theory was made in 1981 with the extended RISM theory [43–45]. The extended RISM theory takes into account not only the geometry but also the charge distribution of a molecule, which completes the chemical characterization of a species for the statistical mechanics of a molecular liquid [58]. Applications of the theory to a variety of liquids and solutions have demonstrated its capability of describing the chemical specificity of liquids in molecular detail. Such an application includes analysis of the structure and fluid phase behavior of water [59–61], the solvation structure and free energy of ions [62–68], the electronic structure in liquids [69–71], chemical reactions [72–76], dynamics in liquid water [77–80], the structure of a water–alcohol mixture [81], the NMR chemical shift of a water molecule solvated in liquid water, acetone, chloroform, or carbon tetrachloride [82], the partial molar volume of amino acids [83, 84], and the stability and folding of polypeptides [85, 86].

In 1993 Ten-no et al. [69, 70] proposed the original RISM-SCF method. The basic idea of the method is to replace the reaction field in the continuum solvation models with a microscopic expression in terms of the site-site radial distribution functions between solute and solvent, which can be calculated from the RISM theory. In the RISM-SCF theory, the statistical solvent distribution around the solute is determined by the electronic structure of the solute, whereas the electronic structure of the solute is influenced by the surrounding solvent distribution. Therefore, the ab initio MO calculation and the RISM equations must

be solved in a self-consistent manner. The RISM-SCF iteration is continued until both the electronic and solvent structures become self-consistent within given convergence criteria.

The RISM-SCF method has been extended to include analytical gradients for geometry optimizations and to the multiconfigurational self-consistent field (MCSCF) method [71, 87], which can be used for exploring the excited states of a molecule in solution. Sato et al. [88] have also reformulated RISM-SCF/MCSCF for the 3D-RISM formalism to properly include the three-dimensional picture of the solvation structure necessary for complex solutes. The authors found that this reformulation allows one to reliably resolve locations and directions of hydrogen bonding in the hydration shells. At the same time, however, they also found that the results from the original RISM-SCF/MCSCF method are reasonably similar to those of the 3D-RISM/SCF approach after removing of the orientational dependence [58]. The MO part of the method can be extended to the more sophisticated levels beyond Hartree-Fock (HF), such as Configuration Interaction (CI) [89] and Coupled-Cluster (CC) theory [90]. In this dissertation, we report the theory for and the implementation of RISM for solute-solvent systems at the coupled-cluster singles and doubles (CCSD) [91] level.

We chose the nitration of benzene as a classic example of an electrophilic aromatic substitution reaction with which to study solvation effects using CC-SCRF. The nitration of benzene by means of the electrophilic aromatic substitution reaction has attracted great interest for more than 60 years [92]. The first mechanistic proposal of Ingold and Hughes [93, 94] for this reaction assumed the nitronium ion (NO_2^+) as the reactive electrophile that after interaction with the aromatic ring formed a covalent cationic interme-

diate named the σ–complex or Wheland intermediate [95]. Later, Olah and co-workers proposed a modification for the original Ingold-Hughes mechanism, which included the existence of a new intermediate prior to the subsequent formation of the Wheland intermediate [96–98]. This intermediate was considered as a π–complex. In addition to the extensive experimental work on nitration reactions [99, 100], the mechanism of the benzene nitration has been studied theoretically over the past twenty years [101–109]. The computational calculation on solvation effects of benzene nitration reaction has not been given much attention so far. The solvent effects of benzene nitration reaction were taken into account within the proposed self-consistent CC-SCRF method with the coupled-cluster singles and doubles (CCSD) theory in the present study.

CHAPTER 2

COUPLED CLUSTER THEORY

Quantum Chemistry is the application of quantum mechanics to explain the electronic structure and properties of molecules via the non-relativistic time-independent Schrödinger equation

$$\hat{\mathscr{H}}\Psi = \mathscr{E}\Psi, \tag{2.1}$$

where Ψ is the wave function, and $\mathscr{E}$ is the total energy as an eigenvalue of the Hamiltonian operator $\hat{\mathscr{H}}$. The total energy is given by the Hamiltonian operator of a molecule, typically consisting of the kinetic energies of the electrons and nuclei, the attraction of the electrons to the nuclei, and the interelectronic and internuclear repulsions. (In the case of an atom, of course the internuclear repulsions do not exist).

This chapter will focus on finding approximate ground-state solutions to the electronic Schrödinger equation using coupled-cluster theory. First, we will review some basic formalisms in molecular electronic structure theory. A more detailed description can be found in the textbook by Szabo and Ostlund [110].

2.1 The Born-Oppenheimer Approximation

The molecular Hamiltonian operator can be written, while assuming the nuclei and electrons to be point masses and neglecting spin-orbit couplings and other relativistic interactions, as

$$\hat{\mathscr{H}} = -\frac{\hbar^2}{2m_\alpha}\sum_\alpha \nabla_\alpha^2 - \frac{\hbar^2}{2m_e}\sum_i \nabla_i^2 + \sum_\alpha \sum_{\beta>\alpha} \frac{Z_\alpha Z_\beta e'^2}{R_{\alpha\beta}} + \sum_i \sum_{j>i} \frac{e'^2}{r_{ij}} - \sum_\alpha \sum_i \frac{Z_\alpha e'^2}{r_{i\alpha}}. \quad (2.2)$$

In Eq. (2.2) α and β refer to nuclei, i and j refer to electrons, and e'^2 has a value of $e^2/4\pi\epsilon_0$, where e is the charge of an electron and ϵ_0 is the permittivity of vacuum. The first term of Eq. (2.2) is the kinetic energy of the nuclei, and the second term is the kinetic energy of the electrons. The third term is the potential energy of the repulsions between the nuclei, where $R_{\alpha\beta}$ is the distance between nuclei α and β with atomic numbers Z_α and Z_β. The fourth term is the potential energy of the repulsions between the electrons, where r_{ij} is the distance between electrons i and j. The last term is the potential energy of the attractions between the electrons and the nuclei, $r_{i\alpha}$ being the distance between electron i and nucleus α. In atomic units, we can write Eq. (2.2) as

$$\hat{\mathscr{H}} = -\frac{1}{2M_A}\sum_\alpha \nabla_\alpha^2 - \frac{1}{2}\sum_i \nabla_i^2 + \sum_\alpha \sum_{\beta>\alpha} \frac{Z_\alpha Z_\beta}{R_{\alpha\beta}} + \sum_i \sum_{j>i} \frac{1}{r_{ij}} - \sum_\alpha \sum_i \frac{Z_\alpha}{r_{i\alpha}}, \quad (2.3)$$

and can abbreviate Eq. (2.3) as

$$\hat{\mathscr{H}} = \hat{T}_N + \hat{T}_e + \hat{V}(r, R). \quad (2.4)$$

$T_N = -\frac{1}{2M_A}\sum_\alpha \nabla_\alpha^2$ and $T_e = -\frac{1}{2}\sum_i \nabla_i^2$ denote nuclear and electronic kinetic energy operators, respectively, and $\hat{V}(r,R)$ denotes nuclear-nuclear, electron-nuclear, and electron-electron interaction terms

$$\hat{V}(r,R) = \sum_\alpha \sum_{\beta>\alpha} \frac{Z_\alpha Z_\beta}{R_{\alpha\beta}} + \sum_i \sum_{j>i} \frac{1}{r_{ij}} - \sum_\alpha \sum_i \frac{Z_\alpha}{r_{i\alpha}}. \tag{2.5}$$

Since nuclei are much heavier than electrons, they move more slowly. Hence, to a good approximation, one can consider the electrons in a molecule to be moving in the field of fixed nuclei. Within this approximation, the kinetic energy of the nuclei (T_N) can be neglected, and the repulsion between the nuclei $\left(V_{NN} = \sum_\alpha \sum_{\beta>\alpha} \frac{Z_\alpha Z_\beta}{R_{\alpha\beta}}\right)$ can be considered to be constant. Therefore, the Hamiltonian in Eq. (2.4)) can be separated in to an electronic Hamiltonian and a nuclear Hamiltonian. The purely electronic Hamiltonian $\hat{\mathscr{H}}_{elec}$ is

$$\hat{\mathscr{H}}_{elec} = -\frac{1}{2}\sum_i \nabla_i^2 + \sum_i \sum_{j>i} \frac{1}{r_{ij}} - \sum_\alpha \sum_i \frac{Z_\alpha}{r_{i\alpha}}, \tag{2.6}$$

and the nuclear Hamiltonian $\hat{\mathscr{H}}_{nucl}$ is

$$\begin{aligned}
\hat{\mathscr{H}}_{nucl} &= -\frac{1}{2M_A}\sum_\alpha \nabla_\alpha^2 + \sum_\alpha \sum_{\beta>\alpha} \frac{Z_\alpha Z_\beta}{R_{\alpha\beta}} + \langle \psi_{elec}(r,R) | \hat{\mathscr{H}}_{elec} | \psi_{elec}(r,R) \rangle, & (2.7)\\
&= -\frac{1}{2M_A}\sum_\alpha \nabla_\alpha^2 + \sum_\alpha \sum_{\beta>\alpha} \frac{Z_\alpha Z_\beta}{R_{\alpha\beta}} + E_{elec}(R), & (2.8)\\
&= -\frac{1}{2M_A}\sum_\alpha \nabla_\alpha^2 + E_{tot}(R). & (2.9)
\end{aligned}$$

The electronic wave function $\psi_{elec}(r,R)$ describes the motion of the electrons and explicitly depends on the electronic coordinates but depends parametrically on the nuclear coordinates, as does the electronic energy, $E_{elec}(R)$. The total wave function can be written as

$$\psi(r,R) = \psi_{elec}(r,R)\psi_{nucl}(R). \tag{2.10}$$

The total energy $E_{tot}(R)$ includes the potential for nuclear repulsion and the electronic energy. This approximation of separating electronic and nuclear motions is called the Born-Oppenheimer (BO) approximation, which was named after by Max Born and Robert Oppenheimer [8]. The BO approximation is central to quantum chemistry. It is also known as the "clamped nucleus" approximation. A major consequence of the BO is the concept of a potential energy surface (PES). Without the BO approximation, we would lack the concept of a PES. The PES is the surface defined by E_{tot} over all possible nuclear coordinates. We would further lack the concepts of equilibrium and transition state geometries, since these are defined as critical points on the PES; instead we would be reduced to discussing high-probability regions of the nuclear wave functions.

Through out this dissertation we will only consider the electronic problem and simply drop the subscript "elec" from the electronic Hamiltonian $\hat{\mathscr{H}}$.

2.2 Hartree-Fock Theory and the Self-Consistent Field Procedure

Within the BO approximation, the non-relativistic time-independent Hamiltonian operator for an n-electron atom can be written as follows,

$$\hat{\mathscr{H}} = \underbrace{-\frac{1}{2}\sum_{i=1}^{n}\nabla_i^2 - \sum_{i=1}^{n}\frac{Z_A}{r_{iA}}}_{\text{one electron part}} + \underbrace{\sum_{i=1}^{n}\sum_{j>i}^{n}\frac{1}{r_{ij}}}_{\text{two electron part}} \quad . \tag{2.11}$$

Because of the inter-electronic repulsion term $\frac{1}{r_{ij}}$, the Schrödinger equation in Eq. (2.11) is not separable. So we use approximate treatments to obtain the energy.

2.2.1 Hartree Product

Neglecting the electron-electron repulsion, the total Hamiltonian can be written as a sum of the one-electron Hamiltonian operators using only the kinetic energy and the potential energy of electron i

$$\hat{\mathcal{H}} = \sum_{i=1}^{N} \hat{h}(i), \tag{2.12}$$

where $\hat{h}(i)$ is referred to as the core Hamiltonian

$$\hat{h}(i) = -\frac{1}{2}\sum_{i=1}^{n} \nabla_i^2 - \sum_{i=1}^{n} \frac{Z_A}{r_{iA}}. \tag{2.13}$$

The operator $h(i)$ will have a set of eigenfunctions that we can take to be a set of spin-orbitals $\{\varphi_j\}$,

$$\hat{h}(i)\varphi_j(i) = \varepsilon_j \varphi_j(i). \tag{2.14}$$

$\hat{\mathcal{H}}$ is a sum of one-electron Hamiltonians. A wave function which is a product of spin-orbital wave functions for each electron which is often called a 'Hartree-product' wave function,

$$\Psi^{HP} = (\varphi_1(1)\varphi_2(2)\ldots\ldots\varphi_n(n)), \tag{2.15}$$

and is an eigenfunction of $\hat{\mathcal{H}}$ as given in equation 2.12,

$$\hat{\mathcal{H}}\Psi^{HP} = \mathscr{E}\Psi^{HP}, \tag{2.16}$$

$$\mathscr{E} = \varepsilon_1 + \varepsilon_2 + \ldots + \varepsilon_n = \sum_{i=1}^{n} \varepsilon_i. \tag{2.17}$$

2.2.2 Slater Determinants

The molecular wave function for a many-electron system can be written as an antisymmetrized product [111, 112] (Slater determinant) of spin-orbitals,

$$\Psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \varphi_1(\mathbf{x}_1) & \varphi_2(\mathbf{x}_1) & \cdots & \varphi_N(\mathbf{x}_1) \\ \varphi_1(\mathbf{x}_2) & \varphi_2(\mathbf{x}_2) & \cdots & \varphi_N(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_1(\mathbf{x}_N) & \varphi_2(\mathbf{x}_N) & \cdots & \varphi_N(\mathbf{x}_N) \end{vmatrix}, \tag{2.18}$$

$$= |\varphi_i(\mathbf{x})\rangle, \tag{2.19}$$

where $\varphi(\mathbf{x})$ is a spin-orbital, i.e., a product of a spatial orbital and an electron spin eigenfunction, and $\mathbf{x}$ denotes both the spin and spatial coordinates of the electron. This implies that the total electronic wave function must be antisymmetric with respect to interchange of any two electrons' coordinates. The Pauli exclusion principle, which states that two electrons cannot have all quantum numbers equal, is a direct consequence of this antisymmetry requirement.

2.2.3 The Fock operator

In 1928 Hartree proposed an iterative "self-consistent field" (SCF) method [113, 114], and in 1930 Fock proposed extending of Hartree's SCF procedure to use Slater determinantal wave functions [115]. The variational principle states that the best wave function of this functional form is the one which gives the lowest possible energy

$$\mathscr{E}_0 = \langle \Psi_0 | \hat{\mathscr{H}} | \Psi_0 \rangle. \tag{2.20}$$

The variational flexibility in the wave function Ψ_0 is in the choice of spin-orbitals. By minimizing $\mathscr{E}_0$ with respect to the choice of spin-orbitals, one can derive an equation, called the Hartree-Fock (HF) equation, which determines the optimal spin-orbitals. The Hartree-Fock equation is an eigenvalue equation of the form,

$$\hat{f}(i)|\varphi_i\rangle = \varepsilon_i|\varphi_i\rangle, \tag{2.21}$$

where

$$\hat{f}(i) = -\frac{1}{2}\sum_{i=1}^{n}\nabla_i^2 - \sum_{A=1}^{n}\frac{Z_A}{r_{iA}} + v^{HF}(i), \tag{2.22}$$

or

$$\hat{f}(i) = \hat{h}(i) + v^{HF}(i), \tag{2.23}$$

where v^{HF} is the average potential experienced by the ith electron due to the presence of the other electrons, $\hat{h}(i)$ is core Hamiltonian, and $\hat{f}(i)$ is called the Fock operator. The Hartree-Fock potential $v^{HF}(i)$, or equivalently the "field" seen by the ith electron, depends on the spin-orbitals of the other electrons (i.e., the Fock operator depends on its eigenfunctions). Thus, the Hartree-Fock equation is nonlinear and must be solved iteratively. The procedure for solving the Hartree-Fock equation is called the self-consistent field (SCF) method. In this dissertation we will use HF, SCF, or HF-SCF interchangeably for the HF-SCF procedure. The Hartree-Fock potential $v^{HF}(i)$ can be written as follows:

$$v^{HF}(i) = \sum_j \Big(\hat{\mathscr{J}}_j(i) - \hat{\mathscr{K}}_j(i), \tag{2.24}$$

where $\hat{\mathscr{J}}$ and $\hat{\mathscr{K}}$ are the Coulomb and exchange operators, respectively. Now the Hartree-Fock equation Eq. (2.23) can be written as

$$\left[\hat{h}(i) + \sum_{j\neq i}\hat{\mathscr{J}}_j(i) - \sum_{j\neq i}\hat{\mathscr{K}}_j(i)\right]\varphi_i(i) = \varepsilon_i\varphi_i(i), \tag{2.25}$$

where

$$\hat{\mathscr{J}}_j(i)\varphi_i(i) = \left[\int\varphi_j^*(j)\frac{1}{\mathbf{r}_{ij}}\varphi_j(j)\mathrm{d}\tau\right]\varphi_i(i), \tag{2.26}$$

and

$$\hat{\mathscr{K}}_j(i)\varphi_i(i) = \left[\int\varphi_j^*(j)\frac{1}{\mathbf{r}_{ij}}\varphi_i(j)\mathrm{d}\tau\right]\varphi_j(i). \tag{2.27}$$

The basic idea of the HF-SCF method is making an initial guess at the spin-orbitals, calculating the average field (i.e., $v^{HF}(i)$) seen by each electron, and then solving the eigenvalue equation 2.21 for a new set of spin-orbitals. Using these new spin-orbitals, one can obtain new fields and repeat the procedure until self-consistency is reached.

2.2.4 The Variational Treatment

Given any trial function $\tilde{\Phi}$, the expectation value $E[\tilde{\Phi}]$ of the Hamiltonian operator $\hat{\mathscr{H}}$ is a number given by

$$E[\tilde{\Phi}] = \langle\tilde{\Phi}|\hat{\mathscr{H}}|\tilde{\Phi}\rangle. \tag{2.28}$$

The trial wave function can be expressed as liner combination of trial functions

$$|\tilde{\Phi}\rangle = \sum_i^N c_i|\phi_i\rangle. \tag{2.29}$$

We want to minimize the energy subject to the constraint that the trial wave function remains normalized, i.e.,

$$E = \langle\tilde{\Phi}|\hat{\mathscr{H}}|\tilde{\Phi}\rangle = \sum_{ij} c_i^* c_j\langle\phi_i|\hat{\mathscr{H}}|\phi_j\rangle, \tag{2.30}$$

$$\langle\tilde{\Phi}|\tilde{\Phi}\rangle - 1 = \sum_{ij} c_i^* c_j \langle\phi_i|\phi_j\rangle - 1 = 0. \tag{2.31}$$

Using Lagrange's method of undetermined multipliers

$$\begin{aligned} \mathscr{L} &= \langle\tilde{\Phi}|\hat{\mathscr{H}}|\tilde{\Phi}\rangle - E\left(\langle\tilde{\Phi}|\tilde{\Phi}\rangle - 1\right) \\ &= \sum_{ij} c_i^* c_j \langle\phi_i|\hat{\mathscr{H}}|\phi_j\rangle - E \sum_{ij} c_i^* c_j \langle\phi_i|\phi_j\rangle \Big), \end{aligned} \tag{2.32}$$

we minimize with respect to the coefficients c_i

$$\begin{aligned} \delta\mathscr{L} = \delta \sum_{ij} c_i^* c_j \langle\phi_i|\hat{\mathscr{H}}|\phi_j\rangle - E\delta \sum_{ij} c_i^* c_j \langle\phi_i|\phi_j\rangle + \\ \sum_{ij} c_i^* \delta c_j \langle\phi_i|\hat{\mathscr{H}}|\phi_j\rangle - E \sum_{ij} c_i^* \delta c_j \langle\phi_i|\phi_j\rangle = 0. \end{aligned} \tag{2.33}$$

Since E is real, we can rearrange the above equation.

$$\sum_i \delta c_i^* \left[\sum_j H_{ij} c_j - E S_{ij} c_j\right] + \text{complex conjugate} = 0, \tag{2.34}$$

where $H_{ij} = \langle\phi_i|\hat{\mathscr{H}}|\phi_j\rangle$ and $S_{ij} = \langle\phi_i|\phi_j\rangle$. Since δc_i^* is arbitrary, we can write

$$\sum_j H_{ij} c_j = E \sum_j S_{ij} c_j, \tag{2.35}$$

$$\mathbf{Hc} = E\mathbf{Sc}. \tag{2.36}$$

If the Hamiltonian is the Fock operator, we can write

$$\mathbf{Fc} = E\mathbf{Sc}. \tag{2.37}$$

2.2.5 The Roothaan Equations

In 1951 Roothaan proposed representing the HF wave function as a linear combination of a set of previously chosen functions, called basis functions [116]. The Roothaan

expansion procedure allows one to find the HF wave function using matrix algebra and is readily implemented on computers. This is the standard procedure used today to calculate atomic and molecular HF wave functions.

$$\varphi_i = \sum_{i,\mu} c_{\mu i}\chi_\mu, \tag{2.38}$$

where the χ_μ's are some set of basis functions, and where the $c_{i\mu}$'s are expansion coefficients that are found by the SCF iterative procedure.

2.2.6 Closed-Shell Hartree-Fock

A restricted set of spin-orbitals has the form

$$\varphi_i(\mathbf{r}) = \begin{cases} \phi_j(\mathbf{r})\alpha(\omega) \\ \phi_j(\mathbf{r})\beta(\omega) \end{cases}, \tag{2.39}$$

and for a closed shell we can write the wave function as

$$|\Psi_0\rangle = |\varphi_1\varphi_2\ldots\varphi_N\rangle = |\phi_1\bar{\phi}_1\phi_2\bar{\phi}_2\ldots\phi_{N/2}\bar{\phi}_{N/2}\rangle. \tag{2.40}$$

Considering the above statement we can rewrite Eq. (2.21)) as

$$\hat{f}(\mathbf{r}_1)(i)|\varphi_i(\mathbf{r}_1)\rangle = \varepsilon_i|\varphi_i(\mathbf{r}_1)\rangle. \tag{2.41}$$

Introducing a set of K known basis functions $\{\chi_\mu(\mathbf{r}), \mu = 1, 2, \ldots, K\}$ and expanding the unknown molecular orbitals in the linear expansion, to expand the spatial orbitals (ϕ_i) as a linear combinations of a set of one-electron basis functions χ_μ

$$\phi_i = \sum_{\mu=1}^{k} C_{\mu i}\chi_\mu. \tag{2.42}$$

We can write

$$\hat{f}(i)\sum_{\nu} C_{\nu i}\chi_{\nu}(i) = \varepsilon_i \sum_{\nu} C_{\nu i}\chi_{\nu}(i). \tag{2.43}$$

By multiplying by $\chi_{\mu}^{*}(1)$ on the left and integrating, we turn the integro- differential equation into a matrix equation,

$$\sum_{\nu} C_{\nu i} \int \chi_{\mu}^{*}(i)\hat{f}(i)\chi_{\nu}(i)\mathrm{d}\mathbf{r} = \varepsilon_i \sum_{\nu} C_{\nu i} \int \chi_{\mu}^{*}(i)\chi_{\nu}(i)\mathrm{d}\mathbf{r}. \tag{2.44}$$

At this point we need to know some important notation.

- The overlap matrix $\mathbf{S}$

$$\mathbf{S}_{\mu\nu} = \int \chi_{\mu}^{*}\chi_{\nu}\mathrm{d}\mathbf{r}_1 \equiv \langle\chi_{\mu}|\chi_{\nu}\rangle, \tag{2.45}$$

is a $K \times K$ Hermitian matrix. The diagonal elements of $\mathbf{S}$ are unity, and the off diagonal elements are numbers less than one in magnitude.

- The Fock matrix $\mathbf{F}$

$$\mathbf{F}_{\mu\nu} = \int \chi_{\mu}^{*}\hat{f}\chi_{\nu}\mathrm{d}\mathbf{r}_1 \equiv \langle\chi_{\mu}|\hat{f}|\chi_{\nu}\rangle, \tag{2.46}$$

is also a $K \times K$ Hermitian matrix. The Fock matrix is the matrix representation of the Fock operator with the set of basis functions $\{\chi_{\mu}\}$.

Considering the above statements, we can rewrite Eq. (2.43)) as

$$\sum_{\nu} F_{\mu\nu}C_{\nu i} = \varepsilon_i \sum_{\nu} S_{\mu\nu}C_{\nu i}, \quad i = 1, 2, \ldots, K, \tag{2.47}$$

$$\mathbf{FC} = \mathbf{SC}\varepsilon, \tag{2.48}$$

where $\mathbf{C}$ is a $K \times K$ square matrix of the expansion coefficients $C_{\mu i}$,

$$\mathbf{C} = \begin{pmatrix} C_{11} & C_{12} & \cdots & C_{1K} \\ C_{21} & C_{22} & \cdots & C_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ C_{K1} & C_{K2} & \cdots & C_{KK} \end{pmatrix}, \tag{2.49}$$

and ε is a diagonal matrix of the orbital energies ε_i,

$$\varepsilon = \begin{pmatrix} \varepsilon_1 & & & \mathbf{0} \\ & \varepsilon_2 & & \\ & & \ddots & \\ \mathbf{0} & & & \varepsilon_K \end{pmatrix}. \tag{2.50}$$

The formal expression for the charge density of a system is as follows:

$$\begin{aligned} \rho(\mathbf{r}) &= 2\sum_{i}^{N/2} \phi_i^*(\mathbf{r})\phi_i(\mathbf{r}), \\ &= 2\sum_{i}^{N/2}\sum_{\nu} C_{\nu i}^*\chi_\nu^*(\mathbf{r}) \sum_{\mu} C_{\mu i}\chi_\mu(\mathbf{r}), \\ &= \sum_{\mu\nu}\left[2\sum_{i}^{N/2} C_{\mu i}C_{\nu i}^*\right]\chi_\mu(\mathbf{r})\chi_\nu^*(\mathbf{r}), \\ &= P_{\mu\nu}\chi_\mu(\mathbf{r})\chi_\nu^*(\mathbf{r}), \end{aligned} \tag{2.51}$$

where we have defined a density matrix $P_{\mu\nu}$ as,

$$P_{\mu\nu} = 2\sum_{i}^{N/2} C_{\mu i}C_{\nu i}^*. \tag{2.52}$$

With doubly occupied orbitals, and Eqs. (2.25) and (2.46), we can write

$$F_{\mu\nu} = \langle\mu|\hat{h}|\nu\rangle + \sum_{i}^{N/2}\sum_{\lambda\sigma} C_{\lambda i}C^{*}_{\sigma i}\left(2\langle\mu\sigma|\nu\lambda\rangle - \langle\mu\sigma|\lambda\nu\rangle\right), \tag{2.53}$$

$$= H^{core}_{\mu\nu} + \sum_{\lambda\sigma} P_{\lambda\sigma}\left(\langle\mu\sigma|\nu\lambda\rangle - \frac{1}{2}\langle\mu\sigma|\lambda\nu\rangle\right), \tag{2.54}$$

$$= H^{core}_{\mu\nu} + G_{\mu\nu}. \tag{2.55}$$

Since $\mathbf{FC} = \mathbf{SC}\varepsilon$ is non-linear, it must be solved iteratively, through what is termed a self-consistent field (SCF) procedure. The SCF requires an initial guess for the density matrix. Once the density matrix and the Fock matrix have converged, the total Hartree-Fock electronic energy is calculated as

$$E_{HF} = \frac{1}{2}\sum_{\mu\nu} P_{\mu\nu}\left(H^{core}_{\mu\nu} + F_{\mu\nu}\right) \tag{2.56}$$

Repeating the SCF procedure for different nuclear positions gives the total energy as a function of geometry. Such an approach is useful for optimizing molecular geometries, finding transition states, and calculating infrared vibrational spectra.

2.3 Electronic Correlation

A Hartree-Fock SCF procedure takes into account the interactions between electrons only as an average field. Actually, the instantaneous interactions between electrons should be considered. Since the electrons repel each other, they tend to keep out of each other's way. In the helium atom, for example, if one electron is close to the nucleus at a given instant, it is energetically more favorable for the other electron to be far from the nucleus at that instant. One sometimes speaks of a Coulomb hole [110] surrounding each electron in an atom. This is a region in which the probability of finding another electron is very

small. The motions of electrons are correlated with each other, and we speak of "electron correlation".

The HF-SCF method is not a bad approximation; the SCF energy is usually more than 99% of the total (nonrelativistic) electronic energy of an atom or molecule. However, chemists are interested in much smaller energy differences, such as bond dissociation energies, ionization potentials, electron affinities, electronic excitation energies, vibrational energies, etc. For example, if we think that the HF-SCF free energy of formation of a certain molecule is about -1000 kcal/mol, 1% of this total enery is 10 kcal/mol, which is not a small amount in the chemical sense.

The purpose of all many-body methods is to describe electron correlation. Löwdin's definition of electronic correlation [90] is the difference between the exact solution of the non-relativistic time-independent Schrödinger equation and the Hartree-Fock description of the electronic wave function. The correlation energy is

$$\Delta E = E_{exact} - E_{HF}. \tag{2.57}$$

Since the variationally optimized Hartree-Fock energy is an upper bound to the exact energy, the correlation energy must be a negative value. There are three main methods for calculating electron correlation: Configuration Interaction (CI), Many-Body Perturbation Theory (MBPT) and Coupled-Cluster Theory (CC).

2.4 Configuration Interaction

Once we define the correlation energy, one may raise the question of what is the exact energy. So we will now define the exact energy. The exact many-electron wave

function must be antisymmetric like the HF determinant and should consist of all possible higher order excited determinants (i.e., all possible orbital occupancies) with variationally optimized mixing coefficients. Mixing of first- and higher-order contributions from excited configurations produces the configuration interaction (CI) wave function,

$$\Phi_{\text{CI}} = \Phi_0 + \sum_{i,a} C_i^a \Phi_i^a + \sum_{\substack{j>i \\ b>a}} C_{ij}^{ab} \Phi_{ij}^{ab} + \dots, \tag{2.58}$$

which, in addition to the SCF solution, introduces single (S), double(D), triple (T), and higher excitations. Φ_0 is called the reference wavefunction. The single excitation $\Phi_i^a = \mathscr{A}(\varphi_1(1)\dots\varphi_a(i)\dots\varphi_n(n))$ correspond to replacing the occupied SCF orbital φ_i by the unoccupied orbital, φ_a. Φ_{ij}^{ab} represents the double excitation, $\Phi_{ij}^{ab} = \mathscr{A}(\varphi_1(1)\dots\varphi_a(i)\dots\varphi_b(j)\dots\varphi_n(n))$, and so forth through triple Φ_{ijk}^{abc}, quadruple Φ_{ijkl}^{abcd}, ... all the way to n–tuple excitations $\Phi_{ijk\dots n}^{abc\dots z}$ for n electrons. Throughout this dissertation, we adopt the usual convention that indices $i, j, k, \dots$ denote occupied orbitals or holes, indices $a, b, c, \dots$ denote virtual, unoccupied orbitals or particles, and indices $p, q, r, \dots$ denote generic orbitals, which may be occupied, virtual, hole, or particle.

The CI coefficients ($C_i^a, C_{ij}^{ab}, C_{ijk}^{abc}, \dots$) are then optimized to give the lowest variational energy,

$$E_{\text{CI}} = \frac{\langle \Psi_{\text{CI}} | \hat{\mathscr{H}} | \Psi_{\text{CI}} \rangle}{\langle \Psi_{\text{CI}} | \Psi_{\text{CI}} \rangle}. \tag{2.59}$$

Since the CI wave function Φ_{CI} is much more flexible than Φ_{HF}, the CI energy E_{CI} is lower than the SCF energy E_{SCF}.

The "full CI" is defined as the wavefunction that includes all possible excitations through n-fold for n–electrons within a chosen basis. If the basis used to represent the

MOs were complete, then the full CI would be the exact solution to the nonrelativistic Schrödinger equation. FCI is computationally unfeasible beyond systems of just several atoms [117], so we must seek approximations.

In order to turn the FCI equations into practical equations which can be applied to a much wider variety of problems, truncation of the configuration space is necessary, leading to limited CI techniques. The most common treatment is CI with all single and double excitations (CISD) [89]. The CISD method is an iterative technique where the computational dependence of each iteration scales as $\mathscr{O}(N^6)$. The CISD method has thus been used to evaluate a wide range of molecular properties such as geometries, vibrational frequencies, dipole moments, etc [118]. However, all the advantages of the CISD method are offset by its major deficiency, which is the CISD energy is not size-extensive [119, 120]. The energy does not scale linearly with the size of the system, and CISD energy is not additive for infinitely separated systems.

2.5 Many Body Perturbation Theory

In 1934 Møller and Plesset [121] proposed a perturbation theory which treats the electron correlation as a pertubation to the Hartree-Fock solution, and this form of many-body perturbation theory (MBPT) [122] is called Møller-Plesset (MP) perturbation theory. In this approach, the "true" Hamiltonian operator $\hat{\mathscr{H}}$ is expressed as the sum of a "zeroth order" Hamiltonian $\hat{\mathscr{H}}_0$ and a perturbation,

$$\hat{\mathscr{H}} = \hat{\mathscr{H}}_0 + \hat{\mathscr{V}}. \tag{2.60}$$

In order to systematically improve the eigenfunctions and eigenvalues of $\hat{\mathscr{H}_0}$, one can write Eq. (2.60) introducing the parameter λ, which will later be set equal to unity,

$$\hat{\mathscr{H}} = \hat{\mathscr{H}_0} + \lambda\hat{\mathscr{V}}. \tag{2.61}$$

λ can be varied between 0 and 1; when λ is zero, $\hat{\mathscr{H}}$ is the zeroth order Hamiltonian, and when λ is one, then $\hat{\mathscr{H}}$ equals the true Hamiltonian. The exact eigenfunctions Ψ and eigenvalues E of the true Hamiltonian $\hat{\mathscr{H}}$ can be expressed in a Taylor series in powers of λ,

$$|\Phi_i\rangle = |\Psi_i^{(0)}\rangle + \lambda|\Psi_i^{(1)}\rangle + \lambda^2|\Psi_i^{(2)}\rangle + \ldots + \lambda^n|\Psi_i^{(n)}\rangle, \tag{2.62}$$

$$E_i = E_i^{(0)} + \lambda E_i^{(1)} + \lambda^2 E_i^{(2)} + \ldots + \lambda^n E_i^{(n)}, \tag{2.63}$$

where $E^{(n)}$ is nth order energy. Substituting Eq. (2.62), and (2.63) into the Schödinger equation, we have

$$\begin{aligned}\left(\hat{\mathscr{H}_0} + \lambda\hat{\mathscr{V}}\right)\Big(|\Psi_i^{(0)}\rangle + \lambda|\Psi_i^{(1)}\rangle + \lambda^2|\Psi_i^{(2)}\rangle + \ldots\Big)\Big(\\ = \left(E_i^{(0)} + \lambda E_i^{(1)} + \lambda^2 E_i^{(2)} + \ldots\right)\Big(|\Psi_i^{(0)}\rangle + \lambda|\Psi_i^{(1)}\rangle + \lambda^2|\Psi_i^{(2)}\rangle + \ldots\Big)\Big(\end{aligned} \tag{2.64}$$

Considering only equalities of like powers of λ and intermediate normalization, the energies can be obtained from following equations:

$$
\begin{aligned}
E_i^{(0)} &= \langle \Psi_i^{(0)} | \hat{\mathscr{H}_0} | \Psi_i^{(0)} \rangle, && (2.65)\\
E_i^{(1)} &= \langle \Psi_i^{(0)} | \hat{\mathscr{V}} | \Psi_i^{(0)} \rangle, && (2.66)\\
E_i^{(2)} &= \langle \Psi_i^{(0)} | \hat{\mathscr{V}} | \Psi_i^{(1)} \rangle, && (2.67)\\
E_i^{(3)} &= \langle \Psi_i^{(0)} | \hat{\mathscr{V}} | \Psi_i^{(2)} \rangle, && (2.68)\\
\vdots &= \qquad \vdots \\
E_i^{(n)} &= \langle \Psi_i^{(0)} | \hat{\mathscr{V}} | \Psi_i^{(n-1)} \rangle. && (2.69)
\end{aligned}
$$

From Eq. (2.69) it would appear that the $(n-1)$th-order wave function is required for calculating the nth-order energy. However, Löwdin [123] has shown that the $2n$ and $2n+1$th energy can be obtain by only the nth order correction,

$$
\begin{aligned}
E_i^{(2n)} &= \langle \Psi_i^{(n)} | \hat{\mathscr{V}} | \Psi_i^{(n-1)} \rangle - \sum_{k=0}^{n} \sum_{l=0}^{n} \Big(E_i^{(2n-k-1)} \langle \Psi_i^{(k)} | \Psi_i^{(l)} \rangle, && (2.70)\\
E_i^{(2n+1)} &= \langle \Psi_i^{(n)} | \hat{\mathscr{V}} | \Psi_i^{(n)} \rangle - \sum_{k=0}^{n} \sum_{l=0}^{n} \Big(E_i^{(2n+1-k-1)} \langle \Psi_i^{(k)} | \Psi_i^{(l)} \rangle. && (2.71)
\end{aligned}
$$

Now using Eq. (2.70), and (2.71), the nth order energy correction can be solved recursively.

In the MP approach, the unperturbed Hamiltonian $\hat{\mathscr{H}_0}$ is the sum of the one-electron Fock operators for the N electrons, as in Eqs. (2.23) and (2.24),

$$
\begin{aligned}
\hat{\mathscr{H}}^{(0)} &= \sum_{i}^{N} \Big(\hat{h}(i) + v_i^{HF}(j) \Big) \Big(\\
&= \sum_{i}^{N} \hat{h}(i) + \sum_{j}^{N} \Big[\hat{\mathscr{J}}_j(i) - \hat{\mathscr{K}}_j(i) \Big] \Big) \Big(. && (2.72)
\end{aligned}
$$

The Hartree-Fock wave function $\Psi_0^{(0)}$, is an eigenfunction of $\hat{\mathscr{H}_0}$, and the corresponding 0th order MBPT Energy $E_i^{(0)}$ is

$$\begin{aligned} E_0^{(0)} &= \langle \Psi_0^{(0)} | \hat{\mathscr{H}^{(0)}} | \Psi_0^{(0)} \rangle, \\ &= \sum_i^N \left(\langle i|\hat{h}|i\rangle + \langle i|\hat{v_i}^{HF}|i\rangle \right), \\ &= \sum_i \langle i|h|i\rangle + \sum_{j>i} \langle ij||ij\rangle \Bigg), \\ &= \sum_i \epsilon_i, \end{aligned} \tag{2.73}$$

where $\langle ij||kl\rangle$ is the antisymmetrized two-electron integral in Dirac notation, and

$$\langle ij||ij\rangle = \langle ij|kl\rangle - \langle ij|lk\rangle, \tag{2.74}$$

$$= \iint \varphi_i^*(\mathbf{r_1})\varphi_j^*(\mathbf{r_2})\frac{1}{\mathbf{r_{12}}})(1-\hat{\mathscr{P}})\varphi_k(\mathbf{r_1})\varphi_l(\mathbf{r_2})\mathrm{d}\mathbf{r_1}\mathrm{d}\mathbf{r_2}. \tag{2.75}$$

Here, $\mathscr{P}$ is an operator which interchanges the coordinates of electron one and two. The two-electron integrals $\langle ij|kl\rangle$ over spin-orbitals can be written as,

$$\langle ij|kl\rangle = \langle \varphi_i\varphi_j|\varphi_k\varphi_l\rangle, \tag{2.76}$$

$$= \iint \varphi_i^*(\mathbf{r_1})\varphi_j^*(\mathbf{r_2})\frac{1}{\mathbf{r_{12}}})\varphi_k(\mathbf{r_1})\varphi_l(\mathbf{r_2})\mathrm{d}\mathbf{r_1}\mathrm{d}\mathbf{r_2}. \tag{2.77}$$

The perturbation $\hat{\mathscr{V}}$ is the difference between the exact electron-electron interaction and the sum of the Hartree-Fock Coulomb and exchange potentials,

$$
\begin{aligned}
\hat{\mathscr{V}} &= \hat{\mathscr{H}} - \hat{\mathscr{H}}^{(0)}, \\
&= \sum_{i}^{N} \hat{h}(i) + \frac{1}{2}\sum_{ij}^{N}\frac{1}{r_{ij}} - \sum_{i}^{N}\left(\hat{h}(i) + v_i^{HF}(j)\right), \\
&= \frac{1}{2}\sum_{ij}^{N}\frac{1}{r_{ij}} - \sum_{i}^{N} v_i^{HF}(j), \\
&= \frac{1}{2}\sum_{ij}^{N}\frac{1}{r_{ij}} - \sum_{i}^{N}\sum_{j}^{N}\left(\hat{\mathscr{J}}_j(i) - \hat{\mathscr{K}}_j(i)\right)
\end{aligned}
\tag{2.78}
$$

The 1st order MBPT Energy $E_0^{(1)}$ can be expressed as

$$
\begin{aligned}
E_0^{(1)} &= \langle\Phi_0^{(0)}|\hat{\mathscr{V}}|\Phi_0^{(0)}\rangle, \\
&= \frac{1}{2}\sum_{ij}\langle\Phi_0^{(0)}|\frac{1}{r_{ij}}|\Phi_0^{(0)}\rangle - \sum_{i}\langle\Phi_0^{(0)}|v_i^{HF}(j)|\Phi_0^{(0)}\rangle, \\
&= \frac{1}{2}\sum_{j>i}\langle ij||ij\rangle - \sum_{j>i}\langle ij||ij\rangle, \\
&= -\frac{1}{2}\sum_{j>i}\langle ij||ij\rangle.
\end{aligned}
\tag{2.79}
$$

It is worth noting that the Hartree-Fock energy is the sum of the zeroth- and first-order MBPT energies,

$$
E^{HF} = E_0^{(0)} + E_0^{(1)}, \tag{2.80}
$$

$$
= \sum_{i}\varepsilon_i - \frac{1}{2}\sum_{j>i}\langle ij||ij\rangle. \tag{2.81}
$$

Since the first order MBPT energy does not go beyond the Hartree-Fock level, in order to obtain an improvement over the Hartree-Fock energy, at least the second-order correction is needed. This level of theory is referred to as MBPT(2) and involves the integrals

$\langle \Psi_i^{(0)} | \hat{\mathscr{V}} | \Psi_i^{(1)} \rangle$. The sum of 0th order through 2nd order energies defines the MBPT(2) energy.

$$E^{\text{MP2}} = E_0^{(0)} + E_0^{(1)} + E_0^{(2)}, \tag{2.82}$$

$$= E^{HF} + E_0^{(2)}. \tag{2.83}$$

The second-order energy equation is solved from the first-order wave function $|\Phi_0^{(1)}\rangle$. The second-order energy is expressed as

$$E_0^{(2)} = \sum_{\substack{i \\ j>i}} \sum_{\substack{a \\ b>a}} \frac{|\langle ij||ab\rangle|^2}{\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b}, \tag{2.84}$$

where indices i, j, a and b have their usual meaning. By expanding the first-order wave-function in terms of all possible excited determinants, it can be shown that the MP2 energy equation includes only doubly excited determinants [119].

It could be worthwhile to note here that the second order energy for a closed-shell system after transforming the AO integrals to the MO basis can be given by,

$$E^{(2)} = \sum_{\substack{i \\ j>i}} \sum_{\substack{a \\ b>a}} \frac{\langle ij|ab\rangle \left(2\langle ij|ab\rangle - \langle ij|ba\rangle\right)}{\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b}, \tag{2.85}$$

$$= \sum_{\substack{i \\ j>i}} \sum_{\substack{a \\ b>a}} \langle ij|ab\rangle T_2^{(1)}, \tag{2.86}$$

where indices i, j, a, and b have their usual meanings, and $T_2^{(1)}$ is the first order approximation to the doubly excited determinants which we will discuss in Section 2.6.

MP2 calculations can be done reasonably rapidly because Eq. (2.84) can be efficiently evaluated. The scaling behavior of the MP2 method is roughly $\mathscr{O}(N^5)$, where N is the number of basis functions. Analytic gradients [124] and second derivatives [125, 126] are

available for this level of theory, so it can conveniently be used to explore PESs. Higher-order energy terms are computed in a similar manner to the procedure described above with the relatively expensive cost of $\mathscr{O}(N^{(n+3)})$, where n is MPn. The even-ordered series are more commonly used since they will always include new correlation effects: MP2 includes two-electron correlation, MP4 one-, three and four-electron correlation, MP6 five- and six-electron correlation, etc. Although the even members of the series exaggerates the electron correlation, the odd members of the series over-correct the even members, thereby reducing their effects including: stronger separation of electrons, a more diffuse electron density, and a deshielding of the nuclei. Therefore, perturbation theory will generally predict longer bond lengths than those of the true geometry [127].

2.6 The Coupled-Cluster Wave Function

The single-reference formulation of CC theory [90, 128–133] has been proven to be an efficient tool for describing electron correlation effects in nondegenerate systems for molecules in a vacuum, particularly near equilibrium and for vibrational frequencies. In this section we will discuss the basics of CC theory, which has been applied in the present study. The essential idea in CC theory is the ground state wave function $|\Psi_0\rangle$ can be given by the exponential ansatz [131]

$$|\Psi_0\rangle = e^{\hat{T}}|0\rangle, \tag{2.87}$$

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \ldots + \hat{T}_n, \tag{2.88}$$

$$\hat{T}_n = \left(\frac{1}{n!}\right)^2 \sum_{ij\ldots ab\ldots}^{n} t_{ij\ldots}^{ab\ldots} \{a^\dagger b^\dagger \ldots ji \quad , \tag{2.89}$$

where $|0\rangle$ is an independent particle reference state which is normally the ground state Hartree-Fock determinant.

$$\hat{H}|\Psi\rangle = E|\Psi\rangle, \tag{2.90}$$

and can be written as

$$\begin{aligned}
\hat{H}|\Psi_0\rangle &= E|\Psi_0\rangle, \\
\hat{H}e^{\hat{T}}|0\rangle &= Ee^{\hat{T}}|0\rangle, \\
e^{-\hat{T}}\hat{H}e^{\hat{T}}|0\rangle &= Ee^{-\hat{T}}e^{\hat{T}}|0\rangle \\
e^{-\hat{T}}\hat{H}e^{\hat{T}}|0\rangle &= E|0\rangle.
\end{aligned} \tag{2.91}$$

The energy and amplitude expressions can be obtained from Eq. (2.91) left-multiplying by the reference and an excited state determinant, respectively, and integrating over all space

$$\langle 0|e^{-\hat{T}}\hat{H}e^{\hat{T}}|0\rangle = E, \tag{2.92}$$

$$\langle \Phi_{ij\ldots}^{ab\ldots}|e^{-\hat{T}}\hat{H}e^{\hat{T}}|0\rangle = 0, \tag{2.93}$$

where intermediate normalization, $\langle 0|\Psi_0\rangle = 1$, is assumed and $|\Phi_{ij\ldots}^{ab\ldots}\rangle$ represents an excited state determinant. Using the Baker–Campbell–Hausdorff relationship [134–140]

$$e^{-\hat{T}}\hat{H}_N e^{\hat{T}} = \left(\hat{H}_N e^{\hat{T}}\right)_C, \tag{2.94}$$

where $\left(\hat{H}_N e^{\hat{T}}\right)_C$ indicates that only the connected diagrams are included, and $\hat{H}_N$ is the normal product Hamiltonian:

$$\begin{aligned}
\hat{H}_N &= \hat{H} - \langle 0|H|0\rangle, && (2.95)\\
&= \sum_{pq} \hat{f}_{pq} \left\{p^\dagger q\right\} + \frac{1}{4}\sum_{pqrs}\langle pq||rs\rangle \left\{p^\dagger q^\dagger sr\right\}, && (2.96)\\
&= \hat{\mathscr{F}}_N + \hat{\mathscr{W}}_N. && (2.97)
\end{aligned}$$

The operators $\hat{\mathscr{F}}_N$, and $\hat{\mathscr{W}}_N$ are referred to as the one- and two-particle parts of the normal product Hamiltonian. The matrix $\hat{f}_{pq}$ will be referred to as the spin-orbital Fock matrix and is defined by

$$\hat{f}_{pq} = \langle p|h|q\rangle + \sum_m \langle pm||qm\rangle. \tag{2.98}$$

Equations (2.92) and (2.93) can be rewritten as

$$\langle 0| \left(\hat{H}_N e^{\hat{T}}\right)_C |0\rangle = \Delta E, \tag{2.99}$$

$$\langle \Phi^{ab\ldots}_{ij\ldots}| \left(\hat{H}_N e^{\hat{T}}\right)_C |0\rangle = 0, \tag{2.100}$$

where

$$\Delta E = E - \langle 0|\hat{H}|0\rangle. \tag{2.101}$$

The energy and T amplitudes can be obtained by solving Eqs. (2.99) and (2.100), respectively.

In order to satisfy the generalized Hellmann–Feynman theorem and to calculate molecular properties, CC equations can be reformulated while introducing an antisymmetrized, perturbation-independent, deexcitation operator Λ

$$\Lambda = \Lambda_1 + \Lambda_2 + \ldots + \Lambda_n, \tag{2.102}$$

$$\Lambda_n = \left(\frac{1}{n!}\right)^2 \sum_{ij\ldots ab\ldots}^{n} \lambda^{ij\ldots}_{ab\ldots} \left\{i^\dagger j^\dagger \ldots ba \quad . \right. \tag{2.103}$$

Then the energy and the Λ equations can be written as

$$\langle 0| (1+\Lambda) \left(\hat{H}_N e^{\hat{T}}\right)_C |0\rangle = \Delta E, \tag{2.104}$$

$$\langle 0| (1+\Lambda) \left(\hat{H}_N e^{\hat{T}}\right)_C - \Delta E|\Phi^{ab\ldots}_{ij\ldots}\rangle = 0. \tag{2.105}$$

Because of the decoupling of T and Λ in CC theory, only T is required to obtain the energy, but both the T and Λ are required to determine properties.

CC energy gradients are well documented elsewhere [56, 130, 141]. The first derivative of the CC correlation energy is given by [141]

$$\frac{\partial \Delta E}{\partial \chi} = \langle 0| (1+\Lambda) \left(\hat{H}_N^\chi e^{\hat{T}} \right)_C |0\rangle, \tag{2.106}$$

$$= \sum_{pq} D_{pq} \frac{\partial f_{pq}}{\partial \chi} + \sum_{pqrs} \Gamma (pq, rs) \frac{\partial \langle pq||rs\rangle}{\partial \chi}, \tag{2.107}$$

where $\partial \hat{H}_N / \partial \chi = \hat{H}_N^\chi$, which is a closed form expression for the generalized Hellmann–Feynman theorem. The one- and two-particle density matrices, D_{pq} and $\Gamma (pq, rs)$ are functions of the amplitudes of the excitation and deexcitation operators, T and Λ. Gradients are treated as special cases of a "response" or "relaxed" property, wherein the basis functions are dependent upon the nuclear displacement perturbation or external perturbation. Since the molecular orbitals are allowed to "relax" in the presence of the perturbation, the response density necessarily contains all orbital relaxation effects, as opposed to a density constructed by taking the expectation value over a given unperturbed wave function. This concept is called the relaxed density approach [56, 141–144]. Within the relaxed density approach, energy gradients are given by a contraction between integral derivatives with several perturbation-independent quantities, namely, the relaxed density matrix, the effective two-particle density matrix, and additional intermediates which account for the dependence of the chosen basis set on the perturbation. The first derivative of the CC

correlation energy with respect to an external perturbation χ can be given by the general formula [141]

$$\frac{\partial \Delta E}{\partial \chi} = \sum_{pq} D_{pq}^{rlx} \frac{\partial f_{pq}}{\partial \chi} + \sum_{pq} I_{pq} \frac{\partial S_{pq}}{\partial \chi} + \sum_{pqrs} \Gamma\left(pq, rs\right) \frac{\partial \langle pq||rs\rangle}{\partial \chi}, \tag{2.108}$$

where D_{pq}^{rlx} is the one particle relaxed density matrix and $\Gamma\left(pq, rs\right)$ is the two particle density matrix, which are functions of T and Λ. In Eq. (2.108) I_{pq} is a one particle intermediate matrix, which guarantees the fulfillment of the orthonormality condition by the perturbed orbitals, and $\partial S_{pq}/\partial \chi$ is the derivative of the AO overlap integrals rotated into the MO basis.

In this dissertation we use the coupled-cluster method including all single and double substitutions (CCSD), where the cluster operator in Eq. (2.88) becomes $\hat{T} = \hat{T}_1 + \hat{T}_2$ and the CCSD wave function is

$$|\Psi_{\text{CCSD}}\rangle = \exp\big(\hat{T}_1 + \hat{T}_2\big)|0\rangle. \tag{2.109}$$

The CCSD energy and amplitude equation can be written as

$$\begin{aligned}
E_{\text{CCSD}} &= \langle 0|\hat{H}|\Psi_{\text{CCSD}}\rangle \\
&= \left\langle 0 \;\; \hat{H} \left(1 + \hat{T}_1 + \frac{1}{2}\hat{T}_1^2 + \hat{T}_2\right) 0 \right\rangle_{\text{C}} && (2.110) \\
0 &= \left\langle \Phi_i^a \;\; \hat{H} \left(1 + \hat{T}_1 + \frac{1}{2}\hat{T}_1^2 + \hat{T}_2 + \hat{T}_1\hat{T}_2 + \frac{1}{3!}\hat{T}_1^3\right) 0 \right\rangle_{\text{C}} && (2.111) \\
0 &= \left\langle \Phi_{ij}^{ab} \;\; \hat{H} \left(1 + \hat{T}_1 + \frac{1}{2}\hat{T}_1^2 + \hat{T}_2 + \hat{T}_1\hat{T}_2 + \frac{1}{3!}\hat{T}_1^3 \right.\right. \\
&\qquad \left.\left. + \frac{1}{2}\hat{T}_2^2 + \frac{1}{2}\hat{T}_1^2\hat{T}_2 + \frac{1}{4!}\hat{T}_1^4\right) 0 \right\rangle_{\text{C}} && (2.112)
\end{aligned}$$

The first order change in the energy due to a perturbation may be expressed as the expectation value of the perturbation operator $\hat{A}$. The total Hamiltonian, $\hat{H}$, is

$$\hat{H} = \hat{H}_{el} + \alpha\hat{A}, \tag{2.113}$$

and the expectation value can be written as first order derivative of energy, when $|\Psi\rangle$ is an exact wave function or when the generalized Hellmann–Feynman theorem is satisfied,

$$\frac{\partial E}{\partial a}_{\alpha=0} = \langle\Psi|\hat{A}|\Psi\rangle, \tag{2.114}$$

Here $|\Psi\rangle$ is the normalized, unperturbed wave function where $\alpha = 0$. First-order response properties in the CC model are shown to be simple dot products between the relaxed density matrix and the one-electron perturbation matrix of interest [141]. First-order one-electron properties are conveniently expressed as

$$\langle\hat{A}\rangle = Tr(AD), \tag{2.115}$$

or

$$\langle\hat{A}\rangle = \sum_{pq} a_{pq}D_{pq}, \tag{2.116}$$

where a_{pq} are the integrals over the operator associated with the property A. Hence, molecular properties can be calculated by using the relaxed density matrix and the perturbation operator.

CHAPTER 3

IMPLICIT SOLVENT MODELS

When structural and dynamical information about the solvent molecules themselves is not of primary interest, the solute-solvent system may be made simpler by modeling the surrounding solvent as an infinite (usually isotropic) medium characterized by the same dielectric constant as the bulk solvent, i.e., a dielectric continuum. In most applications the continuum may be thought of as a configuration-averaged or time-averaged solvent environment, where the averaging is Boltzmann weighted at the temperature of interest. The dielectric continuum approach is thus also sometimes referred to as a "mean-field" approach or "reaction field" approach. The model includes polarization of the dielectric continuum by the solute's electric field; that polarization and the energetics of the solute-continuum interaction are calculated by classical electrostatic formulae [3, 5].

3.1 Reaction-Field Solvent Models

Implicit continuum reaction field solvent models basically have two advantages. The first is a reduction in the system's number of degrees of freedom. The second advantage of continuum reaction field solvent models is that they provide a very accurate way to treat the strong, long-range electrostatic forces that dominate many solvation phenomena. There is

a third advantage in practice. The reaction field models are much more efficient than both QM/MM and fully MM methods.

Implicit continuum reaction field solvent models can be classified into six categories, namely, (i) apparent surface charge (ASC) methods, (ii) multipole expansion (MPE) methods, (iii) the generalized Born approximation (GBA), (iv) image charge (IMC) methods, (v) finite element methods (FEM), and (vi) finite difference methods (FDM). For good and detailed reviews, see Ref. [3–7].

The MPE methods use an expansion of the electrostatic potential in terms of spherical or elliptical harmonic functions. This method, when properly handled, is fast and precise enough for many chemical applications involving solutes of well-behaved shape [3, 5]. The theory underlying the MPE method will be discussed in more detail in Chapter 4 as "The SCRF Method". However, it was soon recognized that the quantitative applicability of the SCRF method was limited by the specific interactions of the solute with solvent molecules in the region close to the solute, where the solvent has properties different from the bulk, where the average solvent structure is caused by the hydrogen bonding or the steric effects. This region is called the cybotactic region [3]. These effects are dominated basically by the first solvation shell, and we call them first-solvation-shell effects. A variety of approaches to "correct" for first-solvation-shell effects were proposed. The statistical mechanical theories based on distribution functions and integral equations can provide a rigorous framework for incorporating such effects into a description of solvation.

One important class of integral equation theories is based on the reference interaction site model (RISM) proposed by Chandler and Andersen in 1972. The theory is a natural

extension of the Ornstein-Zernike equation of simple liquids. The idea of the RISM is to replace the reaction field in continuum models by a microscopic expression in terms of the site-site radial distribution functions between solute and solvent, which can be calculated from the RISM integral equations. The statistical solvent distribution around the solute is determined based on the electronic structure of the solute, while the electronic structure of the solute is influenced by the surrounding solvent distribution. The theory underlying the RISM method will be discussed in Chapter 6.

3.2 The Free Energy of Solvation

From a microscopic point of view, solvation involves the formation of a set of interactions between a solute and a solvent, as well as a change in the interactions of the solvent molecules in the vicinity of the solute. Conceptually, the free energy of solvation (ΔG_{sol}) can be determined through the addition of several contributions. Generally, the solvation process is partitioned into three different steps: cavitation, dispersion, and electrostatics. [3–6],

$$\Delta G_{sol} = \Delta G_{elec} + \Delta G_{cav} + \Delta G_{dis}. \tag{3.1}$$

(ΔG_{elec}) is the free energy of electrostatics and will be discussed in Section 3.2.1 in detail. The free energy of cavitation (ΔG_{cav}) is the energy needed to form a cavity large enough to accommodate the solute within the solvent. Since this is accomplished by breaking down the cohesive forces between solvent molecules, ΔG_{cav} is unfavorable to solvation and is positive in sign. The free energy of dispersion (ΔG_{dis}) is the energy released by forming the attractive force (van der Waals attractions) between the solute and solvent

molecules. ΔG_{dis} contributes favorably to solvation and is negative in sign. These two terms, cavitation and dispersion, are often referred to as nonelectrostatic contributions. In well-behaved solvents and solutes, the ΔG_{dis} and ΔG_{cav} normally cancel each other [6] and sometimes can be ignored.

3.2.1 Electrostatics

The most important term of ΔG_{sol} in Eq. (3.1) is the free energy of electrostatics (ΔG_{elec}). This term measures the work spent in building up the charge distribution of the solute in solution. ΔG_{elec} includes two components: (i) the work necessary to create the solute's gas-phase charge distribution in solution and (ii) the work required to polarize the solute charge distribution by the solvent. It is worth noting that the electrostatic contribution includes not only the gain of the electrostatic interaction energy between the solute and solvent molecules, but also the work needed to generate the solvent reaction field induced by the solute charge distribution.

The contribution from electrostatics ΔG_{elec} can be computed using the wave function. We consider a vacuum hole (a cavity) in the dielectric continuum with an arbitrary shape. The solute molecule represented by the electron density distribution ρ_M is immersed and located in this cavity. The basic equation for a dielectric continuum model is the Poisson-Boltzmann equation in electrostatics. The electrostatic field in the cavity $\mathbf{V}_{in}$ and outside $\mathbf{V}_{out}$ can be obtained by solving the following equations,

$$-\nabla^2 \mathbf{V}_{in}(\mathbf{r}) = 4\pi \boldsymbol{\rho}_{\sigma}, \tag{3.2}$$

$$-\epsilon \nabla^2 \mathbf{V}_{out}(\mathbf{r}) = 0, \tag{3.3}$$

with the boundary conditions:

$$\mathbf{V}^*_{in} = \mathbf{V}^*_{out}, \tag{3.4}$$

$$\frac{\partial \mathbf{V}^*_{in}}{\partial \mathbf{n}} = \epsilon \frac{\partial \mathbf{V}^*_{out}}{\partial \mathbf{n}}, \tag{3.5}$$

where $\mathbf{n}$ is a vector normal to the surface of the cavity, and $\mathbf{V}^*_{in}$ and $\mathbf{V}^*_{out}$ are the values of the electrostatic potential at neighboring points just inside and outside the cavity. ϵ is the dielectric constant of the medium.

The electronic potential in the cavity can be divided into the direct contribution from the solute molecule $\mathbf{V}_\rho(\mathbf{r})$ and the effect from the continuum medium $\mathbf{V}_\sigma(\mathbf{r})$, which is called the reaction field.

$$\mathbf{V}_{in}(\mathbf{r}) = \mathbf{V}_\rho(\mathbf{r}) + \mathbf{V}_\sigma(\mathbf{r}). \tag{3.6}$$

The electrostatic interaction energy W_{MS} between the solute molecule and the continuum is given by

$$W_{MS} = \iint_{cavity} \rho_M \mathbf{V}_\sigma(\mathbf{r}) \mathrm{d}\boldsymbol{\tau}. \tag{3.7}$$

Within the framework of linear free-energy response theory [145], the electrostatic free-energy contribution to solvation is one-half of the solute-solvent electrostatic interaction;

$$\Delta G_{elec} = -\frac{1}{2} W_{MS}. \tag{3.8}$$

For practical purposes several approaches have been proposed for solving Eq. (3.7). All of them are based on modification of the isolated electronic Hamiltonian.

$$\hat{\mathscr{H}}_0 \Psi_0 = \mathscr{E}_0 \Psi_0 \quad \textit{in vacuo}, \tag{3.9}$$

$$\left[\hat{\mathscr{H}}_0 + \hat{\mathscr{V}}\right] \Psi = \mathscr{E} \Psi \quad \textit{in solution}, \tag{3.10}$$

where $\hat{\mathscr{H}_0}$ is the usual Hamiltonian for an isolated molecule, $\hat{\mathscr{V}}$ is the operator representing a perturbation from solvent and Ψ_0 and Ψ are the solute wave functions in *vacuo* and in solution, respectively. The interaction between the solute and the continuum medium can be calculated by

$$\langle \Psi | \hat{\mathscr{V}} | \Psi \rangle = W_{MS}. \tag{3.11}$$

For the case of a self-consistent field method, the effects of the reaction field can be incorporated as an additional term in the Fock matrix;

$$F_{\mu\nu} = F^0_{\mu\nu} + \langle \phi_\mu | \hat{\mathscr{V}} | \phi_\nu \rangle, \tag{3.12}$$

where $F^0_{\mu\nu}$ is the Fock matrix of the isolated molecule.

3.2.2 Cavitation

The free energy of cavitation accounts for the work necessary to generate a cavity large enough to accommodate the solute in the bulk solvent. In other words, it is the work done by forming a cavity of appropriate "shape" and "volume" inside the liquid, in the absence of solute–solvent interactions. The idea of formation of a cavity in the bulk of the liquid and then the insertion of the molecule in to the cavity, dates back to 1937 [146]. Several approaches were suggested to calculate ΔG_{cav} using surface tension [146], surface tension with microscopic corrections [147], isothermal compressibility [148, 149], geometrical parameters of the solvent molecules [150], and Pierotti formulae [151–156] derived from a theory based on a discrete model of fluids.

The cavitation free energy is determined by following Pierotti's [153–155] scaled particle theory adapted to molecular-shaped cavities by using the procedure proposed by

Claverie [156]. The cavitation free energy of atom i, $\Delta G_{C-P,i}$ is determined by weighting the contribution of the isolated atom, $\Delta G_{P,i}$, by the ratio between the solvent–exposed surface of such an atom, S_i, and the total surface of the molecule, S_T,

$$\Delta G_{cav} = \sum_{i=1}^{N} \Delta G_{C-P,i} = \sum_{i=1}^{N} \left(\frac{S_i}{S_T} \right) \Delta G_{P,i}. \tag{3.13}$$

In the Pierotti's model, the $\Delta G_{P,i}$ is expanded in a series of powers of R_{MS}, which is the radius of a sphere that encloses the solute and excludes the centers of the surrounding solvent molecules [153–155].

$$\Delta G_{P,i} = K_0 + K_1 R_{MS} + K_2 R_{MS}^2 + K_3 R_{MS}^3, \tag{3.14}$$

where

$$R_{MS} = R_M + R_S, \tag{3.15}$$

$$K_0 = RT \left[-\ln(1-y) + \frac{9}{2} \left(\frac{y}{1-y} \right)^2 \right] - \frac{4\pi R_S^3 P}{3}, \tag{3.16}$$

$$K_1 = -\frac{3RT}{R_S} \left[\frac{y}{1-y} + 3 \left(\frac{y}{1-y} \right)^2 \right] + 4\pi R_S^2 P, \tag{3.17}$$

$$K_2 = \frac{3RT}{R_S^2} \left[\frac{y}{1-y} + \frac{3}{2} \left(\frac{y}{1-y} \right)^2 \right] - 4\pi R_S P, \tag{3.18}$$

$$K_3 = \frac{4\pi P}{3}, \tag{3.19}$$

$y = 4\pi R_S^3 n_S/3$, R_M and R_S are the radii of solute and solvent,respectivly, n_S is the number density, P is the pressure, and T is the temperature. R_X, where X is solute (M) or solvent (S), can be obtained by

$$R_S = \left(\frac{S_X}{4\pi} \right)^{\frac{1}{2}}, \tag{3.20}$$

$$R_S = \left(\frac{3V_X}{4\pi} \right)^{\frac{1}{3}}, \tag{3.21}$$

where S_X is the molecular surface area and V_X is the molecular volume.

3.2.3 Dispersion–Repulsion

The van der Waals contribution to the free energy of solvation is computed using a linear relationship that depends on the solvent-exposed surface of the atoms in the solute. Basically, this type of contribution arises from dispersion-repulsion interactions between solute and solvent molecules. Theoretical treatments for recovering dispersion–repulsion free energy of solvation of varying complexity have been proposed, ranging from quantum electrodynamics to simple phenomenological formulations [3–7]. Different solvent models [3, 5] uses various methods to interpret the free energy of solvation arising from dispersion-repulsion [3, 5, 157–159]. The SCRF solvent model completely neglects the van der Waals contribution. The interaction potential function in the RISM model generally expresses the dispersion–repulsion free energy of solvation as the sum of a short-range Lennard–Jones (L–J) energy term and a Coulombic interaction energy. The dispersion potential can be expressed as

$$U_{dis} = \sum_{all\ pairs} 4\epsilon_{\alpha\beta} \left(\frac{\sigma_{\alpha\beta}}{r_{\alpha\beta}} \right)^6, \tag{3.22}$$

where σ, and ϵ are the L–J collisional diameter and attractive well depth parameters, respectively and q_α and q_beta are the charges on atom sites α and β. In all cases the standard combination rules, $\epsilon_{\alpha\beta} = (\epsilon_\alpha \epsilon_\beta)^{1/2}$ and $\sigma_{\alpha\beta} = (\sigma_\alpha + \sigma_\beta)/2$ are employed to estimate the

L–J parameters for a pair of different atoms. A similar expression for the repulsion term can also be written as

$$U_{rep} = \sum_{all\ pairs} \left(4\epsilon_{\alpha\beta} \left(\frac{\sigma_{\alpha\beta}}{r_{\alpha\beta}} \right)^{12} \right). \tag{3.23}$$

The $U_{dis-rep}(r)$ is the solute–solvent interaction potential or free energy of solvation of dispersion–repulsion,

$$U_{dis-rep} = 4\epsilon_{\alpha\beta} \left[\left(\frac{\sigma_{\alpha\beta}}{r} \right)^{12} - \left(\frac{\sigma_{\alpha\beta}}{r} \right)^{6} \right]. \tag{3.24}$$

The dispersion–repulsion interaction energy is built-in to the the RISM distribution functions, total correlation function $h(r)$, and direct correlation function $c(r)$ and will be discussed in Eq. (6.6) and (6.7) in Chapter 6.

CHAPTER 4

SELF-CONSISTENT REACTION FIELD METHOD

Traditionally, continuum models for solvents were focused on dielectric models of electrostatic effects. This model of the solvent assumes a linear response of the solvent to a perturbing electric field. This is the simplest and oldest approach of all. The model was developed and exploited by Born [33], Onsager [36], and Kirkwood [34, 35] 60 to 80 years ago. The solvent reaction field (SCRF) method is one of the approaches to modeling the condensed phase in which the solute is treated in a quantum region, the solvent is a bulk continuum, and the free energy of solvation is calculated from the reaction field. The reaction field represents the interaction between the solute and the solvent. In this model a solute molecule is embedded in a spherical cavity in a dielectric continuum medium.

The expression of the multipole expansion for the potential of a charge distribution placed into a sphere of radius a_0, immersed into a continuum with dielectric constant ϵ, and can be found in numerous textbooks [145, 160, 161]. The electrostatic potential $\Phi(\mathbf{r})$

can be written in terms of Legendre polynomials (or spherical harmonics). The general formula for $\Phi_{in}(\mathbf{r})$ can be simplified in the following way: inside the cavity [145],

$$\mathbf{V}_{in}(\mathbf{r}) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l}\Bigg(\frac{1}{r^{l+1}}M_l^m Y_l^m(\theta,\phi) \; + \sum_{l=0}^{\infty}\sum_{m=-l}^{l}\Bigg(-\frac{(l+1)(\epsilon-1)}{l+(l+1)\epsilon}\frac{r^{l+1}}{a_0^{2l+1}}M_l^m Y_l^m(\theta,\phi), \quad (4.1)$$

where $Y_l^m(\theta,\phi)$ are the spherical harmonic functions with total angular momentum l and projection m along the z axis, and M_l^m is a multipole moment of the solute charge distribution. The second term of Eq. (4.1) represents the solvent reaction field acting on the portion of space occupied by the solute.

The (electrostatic) free energy of solvation, E_{sol}, can be expressed by [40],

$$E_{sol} = -\frac{1}{2}\sum_{l=0}^{\infty}\sum_{m=-l}^{l} M_l^m \cdot \mathscr{R}_l^m, \quad (4.2)$$

where $\mathscr{R}_l^m$ is the solvent reaction field

$$\mathscr{R}_l^m = g_l M_l^m, \quad (4.3)$$

and g is the solute$-$solvent coupling constant, which depends on the shape and the size of the cavity, the dielectric constant of the medium, and the order of the multipole moment expansion. For a spherical cavity, the solute$-$solvent coupling constant can be expressed as [40],

$$g_l = \frac{(l+1)(\epsilon-1)}{l+(l+1)\epsilon}\frac{1}{a_0^{2l+1}}, \quad (4.4)$$

where, a_0 is the radius of the spherical cavity, ϵ is the dielectric constant of the solvent, and l is the order of the multipole moment. The multipole moment of the solute charge distribution M_l^m can be written as

$$M_l^m = P_{\mu\nu}\langle\chi_\mu|\hat{M}_l^m|\chi_\nu\rangle, \tag{4.5}$$

where $P_{\mu\nu}$ is the density matrix as defined in Eq. (2.56), and χ_μ and χ_ν are basis functions as defined in Eq. (2.42). For further details refer to Chapter 2.

The Fock matrix can be expressed in the same way as in the conventional Hartree-Fock equations, with the addition of one additional term in order to introduce the solvent contribution. The total Fock matrix can be written as

$$F_{\mu\nu} = F_{\mu\nu}^0 + F_{\mu\nu}^{sol}, \tag{4.6}$$

where

$$F_{\mu\nu}^{sol} = \sum_{l=0}^{\infty}\sum_{m=-l}^{l}\mathscr{R}_l^m\langle\chi_\mu|\hat{M}_l^m|\chi_\nu\rangle, \tag{4.7}$$

and M_l^m is the spherical multipole operator. In the SCRF method Eq. (4.6) can be solved until both density and reaction field converges. An SCRF can be utilized with any electron correlation method when the multipole moments are given from the electron density.

4.1 Implementation of the CC-SCRF Method

The self-consistent coupled-cluster solvent reaction field (CC-SCRF) code has been implemented in a developmental version of the Q-CHEM 3.2 [162] quantum chemistry program package at the CCSD level. The calculation of the reaction field, $\mathscr{R}$, has been given major consideration. The basic single point HF and CC procedure has been adopted with some

modification to calculate the reaction field. On the first iteration a Hartree-Fock (HF) self-consistent reaction field calculation is performed to produce MO coefficients which include the effect of the reaction field response. Then these MO coefficients are utilized to solve for both the T and Λ in the CCSD equations self-consistently in the presence of the reaction field. Once the CCSD amplitudes (T and Λ) are calculated, the CCSD relaxed density, D_{pq}^{rlx}, is obtained. A new solvent reaction field $\mathscr{R}^{CC}$ is calculated using the relaxed one particle density matrix (OPDM). Then new SCF calculation is performed followed by a CCSD calculation in the presence of the new $\mathscr{R}^{CC}$. The HF calculations before the CCSD calculations guarantee that the MO coefficients include the new reaction field response. This procedure is repeated using the CC solvent response until the solvent reaction field converges. Our CC-SCRF procedure consists of the following steps, which calculate $\mathscr{R}$ and E_{sol} self consistently.

1. Obtain $\mathscr{R}_l^m$ from M_l^{mHF}.
2. Compute the HF-SCF equations with $\mathscr{R}_l^m$.
3. Solve the CCSD T and Λ equations with $\mathscr{R}_l^m$.
4. Obtain the CCSD relaxed density, D_{pq}^{rlx}.
5. Compute M_l^{mCC} from the relaxed density D_{pq}^{rlx}.
6. Construct a new $\mathscr{R}_l^{mCC}$ from M_l^{mCC} and return to step 2 if not converged.

4.2 Solvent Effects on trans/gauche Conformational Equilibria

Conformational equilibria of molecules play a major role in biochemistry, polymer science, and supramolecular chemistry [2, 163–165]. The energetics and thermodynamics of such molecular conformers depend not only on the internal potential energy surface,

but also on the nature of the environment, which is the nature of the solvent. A number of previous experimental [166–173] and theoretical [46, 55, 174–186] studies have been carried out on the conformational equilibria of small polyatomic molecules in both the gas phase and condensed phase. The conformational energy differences between rotamers are usually less than 3 kcal/mol, and the solvation energy of polar solutes in polar solvents are also of same order [184]. The magnitude of the polarity can significantly affect the conformational equilibrium. It is well known that the most stable conformer of some molecules in the gas phase is not the most stable conformer in the condensed phase, and the equilibria between rotamers are changing considerably from nonpolar to polar solvents [2, 163].

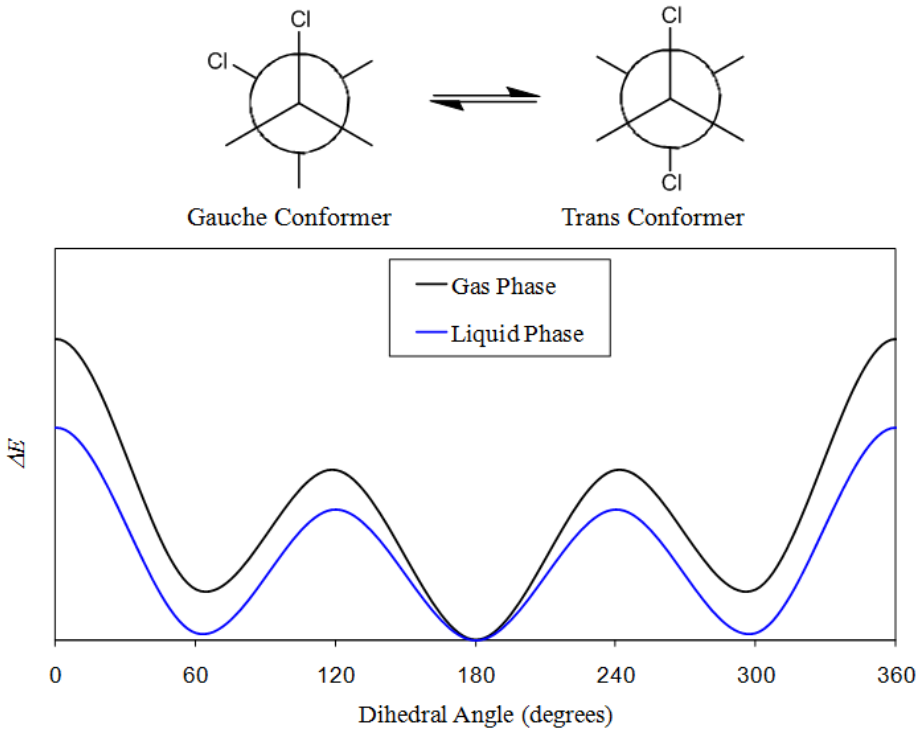


Figure 4.1

Relative rotormer energies of 1,2-dichoroethane

We have considered the solvent effects on trans and gauche conformers of 1,2-dichloroethane (DCE), which shows a strong conformational energy change upon solvation. The gauche and trans structures of DCE and a graphical interpretation of relative conformer energies are shown in Figure 4.1. The graph shows the relative conformer energy vs. the dihedral angle between two chlorine atoms of DCE. The trans conformer is the most stable rotamer in the gas phase because it has the least steric hindrance. The experimental studies shows that the ΔE between trans and gauche conformers in gas phase is 1.20 kcal/mol and in pure liquid is 0.31 kcal/mol [46, 163, 171, 181]. Tanaka et al. [177] carried out a theoretical calculation on the trans/gauche equilibrium of DCE using CI methods. According to their studies the conformational energy difference between trans and gauche conformers, ΔE, is 1.48 kcal/mol. Dixon et al. [174] also studied gas phase conformational equilibria using semiempirical, HF, DFT, and MP2 methods.

Theoretical calculations of solvent effects on the conformational equilibria of DCE have been studied using the SCRF method by Wong et al. [46] with SCF and MP2 methods, and Christiansen et al. [55] with CCSD. We compare our investigation of solvent effects on the conformational equilibrium of DCE on with the previous work done by Wong et al. [46] and Christiansen et al. [55].

Equations (4.2 – 4.7) show that the solvation free energy depends on the dielectric constant (ε) of the solvent, the size of the cavity (a_0), the order of the multipole moment expansion (l), the size the basis set, and the correlation for a particular geometry of the solute. First, we will discuss the question of the choice of cavity radius. The choice of the appropriate cavity size for SCRF calculations has been raised many times [3, 4, 46,

55, 187, 188]. Since many algorithms are available for various solvent methods [3, 4], we restricted our discussion only for describing a spherical cavity in our SCRF calculations. Often, the cavity radius is calculated using

$$a_0^3 = \frac{3V_m}{4\pi N_a}, \tag{4.8}$$

where, V_m is the solute molar volume and can be obtained from experiment (molecular weight/density) or from refractivity [189], and N_a is Avogadro's number. It is also common to add 0.5 Å to the value of a_0 to account for the van der Waals radii [46]. In addition, a_0 can be obtained from the distant between the center of mass and the most distance atom plus the van der Waals radius of that atom. Alternatively, one can obtain $2a_0$ (the diameter of the cavity) by having the greatest value among all the internuclear distances of any two atoms plus the van der Waals radii of the two relevant atoms in the solute molecule. Apart from these empirical cavity radii methods, there are other methods also available for determining the cavity size in the SCRF calculations. In one method the volume of a molecule is defined by the volume occupied by the 0.001 a.u. electron density envelope and is scaled by 1.33 times in order to fit experimental values [190]. In another method the cavity size can be determined by polarizability and the classical Maxwell's field [191], where the cavity size is a function of the dielectric constant, the multipole order, and the basis set used.

For the cavity radius for our investigation, we have chosen the second empirical approach to determine the the cavity size. In our calculations geometries of gauche and trans DCE conformations are optimized in vacuum at the HF level with the 6-31G* basis set [192, 193] unless otherwise specified. Then, CC-SCRF energies were obtained at the

CCSD level with the specified cavity radius and the dielectric constant of medium. Table 4.1 shows total energies and gauche–trans energy differences for several values of the cavity radius in water ($\epsilon = 78.5$) at HF, MP2, and CCSD with the 6-31+G* basis set [194] and with $l = 1$. We note that our CC-SCRF code can be employed through arbitrary order of multipole moments expansion, l, but we restricted ourselves to the $l = 1$ case, which is the Onsager type SCRF [36], for consistency with Wong et al. [46]. The effect of the order of multipole moments expansion will be discussd later.

Table 4.1

Total energies and conformer energy differences of DCE in water with various cavity radii

Cavity radius	Gauche Conformation Energy[a] (hartrees)			ΔE (Gauche−trans) (kcal/mol)		
a_0 (Å)	HF[b]	MP2[b]	CCSD[c]	HF[b]	MP2[b]	CCSD[c]
2.50	−997.04228	−997.59662	−997.64032	−5.91	−5.20	−4.68
2.75	−997.03774	−997.59271	−997.63670	−3.06	−2.75	−2.40
3.00	−997.03531	−997.59066	−997.63435	−1.54	−1.46	−0.93
3.25	−997.03385	−997.58948	−997.63365	−0.62	−0.72	−0.49
3.50	−997.03288	−997.58862	−997.63290	−0.01	−0.18	−0.02
3.75	−997.03221	−997.58784	−997.63238	0.41	0.31	0.30
4.00	−997.03172	−997.58762	−997.63201	0.71	0.44	0.54
4.25	−997.03136	−997.58752	−997.63173	0.94	0.51	0.71
4.50	−997.03109	−997.58725	−997.63152	1.11	0.68	0.85

[a] The HF, MP2, and CCSD energies of the trans rotamer are −997.03286, −997.58833, and −997.63287 hartrees, respectively
[b]Ref. [46]
[c]This work

The conformational energies in Table 4.1 are calculated at 6-31+G* with HF/6-31G* geometries. The energies are given in Hartrees, and the energy differences between gauche

and trans rotamers are in kcal/mol. We kept the core orbitals frozen [195] in the CCSD calculations. The calculated energy is strongly dependent on the size of the cavity. In Table 4.1 the gauche–trans conformational energy differences increase from -4.68 kcal/mol to $+0.85$ kcal/mol in the CCSD calculations as the cavity radius increases. The gauche conformer energies decrease with decreasing cavity radius. This can be readily understood from the relationship between a_0 and g in Eq. (4.4). In practice, as stated above, the gauche–trans conformational energy difference of DCE is reduced to near zero when it is in a condensed phase which has very high dielectric constant [163]. Therefore, a value of a_0 somewhat larger than 3.5 Å may be the ideal value for the cavity radius. Our best estimate for a_0 using the second approach for obtaining the cavity radius (using the center of mass) is 3.54 Å. The third approach (using the cavity diameter) gives us 3.48 Å as cavity radius. But Wong et al. [46] used a much greater a_0 value of 3.65 Å to obtain the conformational energy difference.

We shall next consider the effect of the dielectric medium on the gauche–trans conformational structures of DCE. The conformational equilibria of DCE is found to strongly depend on the dielectric property of the solvent [163]. Table 4.2 shows the the gauche–trans conformational energy difference of DCE in SCRF calculations with HF, MP2, and CCSD using the 6-31+G* basis set and $l = 1$ with different solvents. Here, the conformer energy differences are given in kcal/mol, and the size of the cavity radius is 3.65 Å. We compare our CCSD/6-31+G* CC-SCRF calculations with the HF and MP2 results by Wong et al. [46]. We have used the appropriate dielectric constant to mimic the proper solvent. The CCSD-SCRF calculations are shown to be much closer to the experimental

Table 4.2

Conformer energy differences of DCE in the gas phase and in various solvents

		ΔE (gauche − trans) (kcal/mol)			
Solvent	ε	HF[a]	MP2[a]	CCSD[b]	Exp.[c]
vacuum	1.0	1.96	1.64	1.50	1.20
cyclohexane	2.0	1.32	0.88	1.01	0.91
carbon disulfide	2.6	1.13	0.69	0.86	0.83
diethyl ether	4.3	0.82	0.54	0.62	0.69
ethyl acetate	6.0	0.67	0.41	0.50	0.42
pure liquid	10.1	0.50	0.26	0.37	0.31
acetone	20.7	0.36	0.14	0.27	0.18
acetonitrile	35.9	0.30	0.09	0.22	0.15

[a]Ref. [46] [b]This work [c]Ref. [163]

values than HF and MP2. The root mean square deviation (Δ_{rms}) of the calculated values from the experimental results in the HF calculation is 0.25 kcal/mol, and in of MP2 it is 0.09 kcal/mol [46]. Within our implementation of CCSD-SCRF calculations gave a 0.07 kcal/mol Δ_{rms} value and show more consistency with experimental results.

As we stated earlier, not only ε and a_0, but also the order of the multipole moment expansion (l), the size the basis set, and correlation affects the solvent contribution. The trans rotamer has C_{2h} symmetry, and thus it has no permanent dipole moment. Therefore, the Onsager model ($l = 1$) gives no solvent effect on the trans rotamer, while the gauche rotamer, which has C_2 symmetry, has a significant solvent effect. The results are shown in Table 4.3. We have investigated the variation of the solvent effect with the order of the multipole moment expansion, the size of the basis set, and correlation. We studied the gauche–trans conformational energy difference of DCE in the gas phase and

Table 4.3

Total energies (hartrees) and conformer energy differences (kcal/mol) in the gas phase and in acetonitrile

				Vacuum			ε = 35.9		
Model	Basis	Geometry(vac)	L_{max}	Trans	Gauche	ΔE	Trans	Gauche	ΔE
				Basis set and Correlation effects					
SCF	6-31G*	HF/6-31G*	15	−997.03094	−997.02789	1.91	−997.03429	−997.03236	1.21
CCSD	6-31G*	HF/6-31G*	15	−997.59690	−997.59446	1.53	−997.59967	−997.59800	1.05
CCSD	cc-pVTZ	HF/6-31G*	15	−997.90886	−997.90651	1.47	−997.91124	−997.90968	0.98
				L_{max} dependency					
CCSD	6-31G*	HF/6-31G*	1	−997.59690	−997.59446	1.53	−997.59690	−997.59643	0.30
CCSD	6-31G*	HF/6-31G*	2				−997.59830	−997.59736	0.59
CCSD	6-31G*	HF/6-31G*	5				−997.59863	−997.59786	0.48
CCSD	6-31G*	HF/6-31G*	8				−997.59950	−997.59799	0.95
CCSD	6-31G*	HF/6-31G*	10				−997.59963	−997.59800	1.02
CCSD	6-31G*	HF/6-31G*	12				−997.59966	−997.59800	1.05
CCSD	6-31G*	HF/6-31G*	15				−997.59967	−997.59800	1.05
CCSD	6-31G*	HF/6-31G*	20				−997.59968	−997.59800	1.05
				Previous calculations					
HF[a]	6-31+G*	HF/6-31G*	1	−997.03286	−997.02973	1.96		−997.03238	0.30
MP2[a]	6-31+G*	HF/6-31G*	1	−997.58833	−997.58592	1.51		−997.58819	0.09
QCISD[a]	6-31+G*	HF/6-31G*	1	−997.63378	−997.63639	1.64		−997.63361	0.11
CCSD[b]	6-31G*	HF/6-31G*	1	−997.58820	−997.58575	1.54		−997.58782	0.24
MP2[c]	6-31G*	HF/6-31G*	1	−997.58824	−997.55577	1.55		−997.55730	0.59

[a]Ref. [46] [b]Ref. [55] [c]Ref. [54]

in as acetonitrile ($\epsilon = 35.9$) as a solvent. We used vacuum optimized HF/6-31G* geometries of trans and gauche conformers as in the study of Wong et al. [46]. The electron correlation recovered by CCSD with the 6-31G* basis set and frozen core orbitals is -0.38 kcal/mol in vacuum and -0.16 kcal/mol in acetonitrile. We have furthermore carried out a CCSD(T) single point energy calculations with the 6-31G* basis set and with frozen core orbitals to investigate the effects of triple excitations on the vacuum conformational energy difference. The CCSD(T) results show that the triple excitation contribution to the conformational energy difference in vacuum is about -0.01 kcal/mol. We note that a beter comparison would be with Mikkelsens' [55] unrelaxed approach based CC-SCRF and our relaxed density based CC-SCRF , but unfortunately they incorrectly used spherical functions instead of Cartesian functions with the 6-31G* basis set for both their geometry optimization and single point energy calculations on DCE.

We observed that the calculated solvation free energy depends significantly on the selection of the order of the multipole moment expansion. The Onsager SCRF method [46], where $l = 1$, considerably underestimates the solvent contribution for the conformational energy difference. Since the trans rotamer has no dipole moment, the Onsager SCRF calculation gives no solvent contribution, but when higher order multipole moments are considered, a significant portion of the solvation free energy is recovered. We performed a series of calculations starting from $l = 1$ to $l = 20$ for both the gauche and trans rotamers. For this calculations we have used the gas-phase HF/6-31G* optimized geometries with $a_0 = 3.65$ Å in acetonitrile ($\epsilon = 35.9$). Table 4.3 clearly shows that the gauche–trans conformational energy difference is converged at $l = 12$ to within 0.01 kcal/mol. The Onsager

SCRF method underestimates the gauche–trans conformational energy difference of DCE by 0.75 kcal/mol at convergence. The results of previous studies at [46, 54, 55] are also included in Table 4.3 for comparison.

CHAPTER 5

SOLVENT EFFECTS ON THE NITRATION OF BENZENE

5.1 The Reaction Mechanism

The mechanism of electrophilic aromatic nitration is one of the outstanding important reactions in organic chemistry [92, 99, 100, 196, 197]. Evidence of an electron transfer reaction mechanism has also been reported apart from the electrophilic substitution mechanism [104, 198]. Since the early studies by Ingold and co-workers [197], it has been generally accepted that the electrophile is the NO_2^+ ion and that the reaction involves an intermediate, called the σ–complex or the Wheland intermediate [95]. The original Ingold-Hughes mechanism of benzene nitration [93, 94] was

$$HNO_3 + HA \longrightarrow H_2NO_3^+ + A^-, \tag{5.1}$$

$$H_2NO_3^+ \longrightarrow NO_2^+ + H_2O, \tag{5.2}$$

$$ArH + NO_2^+ \longrightarrow ArHNO_2^+, \tag{5.3}$$

$$ArHNO_2^+ + A^- \longrightarrow ArNO_2 + HA, \tag{5.4}$$

where Ar denotes a phenyl group. Equations (5.1) and (5.2) describe the formation of the electrophile NO_2^+. Basically, the formation of the NO_2^+ is the first step of the Ingold-Hughes mechanism. The electrophile NO_2^+ then attacks to the benzene ring and forms an intermediate $ArHNO_2^+$, which is the Wheland intermediate. The product of Eq. (5.3)

is the Wheland intermediate. The formation of this arenium ion is identified as the rate-determining step [197]. The final step of the Ingold-Hughes mechanism, the deprotonation of the Wheland intermediate, is usually fast and has no effect on the observed kinetics [199].

In the early sixties, Olah and co-workers [96, 97] proposed a modification to the original Ingold-Hughes mechanism, which introduced the existence of a new intermediate prior to the subsequent formation of the Wheland intermediate. Equations (5.5), (5.6), and (5.7) describe the Olah's modified mechanism.

$$\mathrm{ArH} + \mathrm{NO_2^+} \quad \longrightarrow \quad \mathrm{ArH{\cdot}NO_2^+}(\pi), \tag{5.5}$$

$$\mathrm{ArH{\cdot}NO_2^+}(\pi) \quad \longrightarrow \quad \mathrm{ArHNO_2^+}(\sigma), \tag{5.6}$$

$$\mathrm{ArHNO_2^+}(\sigma) + \mathrm{A^-} \quad \longrightarrow \quad \mathrm{ArNO_2} + \mathrm{HA}, \tag{5.7}$$

The new intermediate was called the π-complex, $\mathrm{ArH{\cdot}NO_2^+}(\pi)$. The next step is transformation of the π-complex to the σ-complex. After the transformation the σ-complex undergo the deprotonation step, which yields the substituted product.

The mechanism of aromatic nitration continues to be the subject of active research and debate [92–100, 197, 198]. Cacace and Attinã [200–203] studied a series of gas-phase reactions of alkyl nitrates with aromatics. These studies showed a fundamental similarity between gas-phase and solvent-phase experiments and opened new perspectives for theoretical investigation of mechanistic details of aromatic nitration.

Early theoretical studies were aimed at clarifying the nature of the initial complexes and did not give much attention to reaction pathways. In 1973 Hehre et al. [204] in-

vestigated the mechanism for electrophilic nitration theoretically with HF theory and the STO-3G basis set [205, 206]. Politzer et al. [101] in 1985 carried out calculations of the properties of some of the intermediates in the reactions of benzene and toluene with NO_2^+. They limited their studies to HF with the STO-6G [205, 206] and 5-31G [192, 207] basis sets, based on available computer capability. Gleghorn et al. [102] also studied five benzene/NO_2^+ complexes with HF/4-31G [205], and MP2/4-31G with HF/4-31G geometries. All earlier studies investigated benzene/NO_2^+ intermediates only. Szabõ et al. [103] in 1992 obtained two energy profiles for the nitration of benzene with two possible protonated methyl nitrate isomers at the MP4DQ/3-21G//HF/3-21G [208].

In 2003 three independent studies of the nitration mechanism of benzine were published. Chronologically, they were the work of Gwaltney et al. [104], Esteves et al. [105], and Chen et al. [106]. Gwaltney and co-workers studied the charge transfer mechanism of electrophilic aromatic nitration by means of Marcus-Hush theory with MP2/6-31G** [193, 207] and CCSD(T)/6-31G**. Esteves and co-workers carried out a theoretical investigation of the mechanism of electrophilic aromatic nitration with B3LYP/6-311++G** [209] and MP2(fc)/6-31+G*//B3LYP/6-311++G**// [210]. They identified 37 possible benzene/NO_2^+ stationary points. Chen et al. investigated five stationary points of the benzene nitration mechanism with B3LYP/6-311++G**. In 2004 Yin [108] carried out detailed CCSD(T) [211] calculations with the 6-31G** basis set. These calculations approached the limit of computer power. Yin further studied the effect of the orientation of the approaching NO_2^+ electrophile and substituent effects on electrophilic aromatic ni-

tration. de Queiroz et al. [107] also studied substituent effects on electrophilic aromatic nitration in 2006.

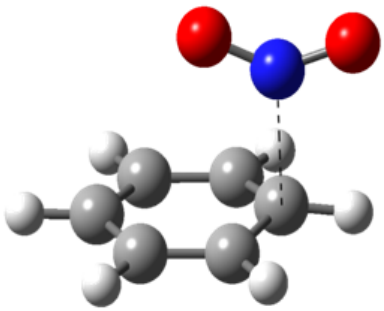
5.2 Solvent Effects

Solvent effects plays a major role in mechanisms of organic reactions [2]. Gwaltney et al. [104] included the Onsager SCRF corrections within B3LYP/6-31G** and Chen et al. [106] also included the same corrections with B3LYP/6-311G** to their reaction energy profile. But according to Table 4.3, Onsager SCRF, where $l = 1$, highly underestimates the solvation free energy. Solvent effects on the mechanism of electrophilic aromatic nitration must be subjected to examination with high-level first-principles molecular electronic structure calculations.

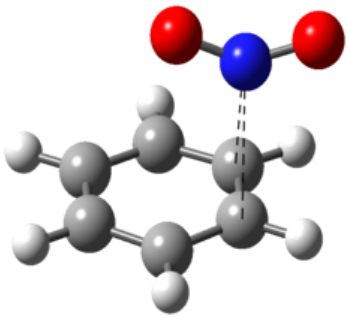
To be consistent with Esteves et al. [105], we adopted the same numbering system for the individual intermediates and transition states on the reaction pathway to benzene nitration. In order to study solvent effects on the benzene nitration reaction, first we have performed gas-phase geometry optimizations on all pre-identified transition states and intermediates, then we have performed a series of CC-SCRF calculations on the converged gas-phase geometries. All gas-phase geometry optimizations were performed with the ACES II program [212] at CCSD(T)/6-31G**, and all CC-SCRF solvent calculations were performed with a development version of Q-CHEM 3.2 [162] with CCSD/6-31G**, $l = 15$, and with the given solvents. The cavity radii ware calculated as discussed in Section 4.2. Initial CCSD(T)/6-31G** optimized structures are obtained from Yin et al. [108, 109]. We have studied 14 stationary points, including a new transition state

which leads to the product phenol. We also included zero point vibrational energy (ZPVE) corrections for the energies of all the stationary points. B3LYP/6-31G** normal mode frequencies were used to obtain the ZPVE. Intrinsic reaction coordinate (IRC) [213] calculations were also performed for those structures to confirm the integrity of the stationary points. We have studied 11 intermediates and transition states, which are directly on the reaction pathway of benzene nitration, and reoptimized them to make sure that the stationary points are connected correctly.

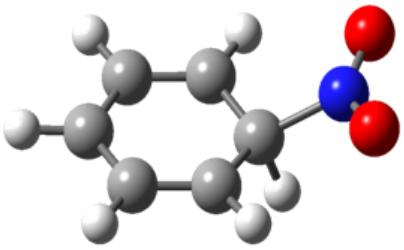
Figures 5.1 and 5.2 show the structures of all 11 stationary points optimized at CCSD(T)/6-31G**. In this mechanism the reactive electrophile NO_2^+ is responsible for the attack on the aromatic benzene ring and the formation of the non-specifically bound intermediate, which we call the π–complex (**4**) and which has C_s symmetry. The NO_2^+ group of the π–complex can migrate to all the carbon atoms of the benzene ring through transition state **7**, which also has C_s symmetry. Then the π–complex becomes the σ–complex or the Wheland intermediate (**8**) via transition state **10**. Intermediate **23** is formed by the migration of the proton on the nitrated carbon atom of **8** to the neighboring carbon atom through transition state **28**. Then intermediate **26** is formed by the migration of the proton from the neighboring carbon atom, which has two protons, to the oxygen atom of the NO_2 group. Intermediate **27** is formed by releasing the conformational stress of **26** via transition state **38**, and intermediate **27** leads to the final product nitrobenzene. Besides these stationary points, a more energetic reaction path through transition state **35** also exists between intermediates **4** and **26**. This is less likely to take place in both vacuum and solvent.



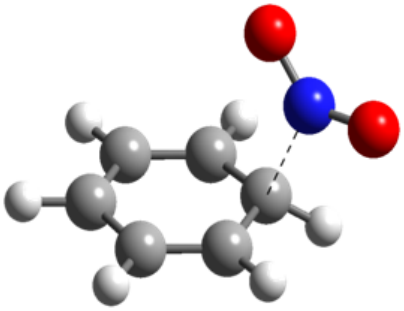
4
C_s, π complex



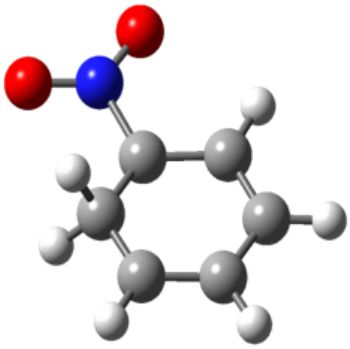
7
C_s, TS



8
C_s, σ complex



10
C_1, TS



23
C_1

Figure 5.1

Stationary structures optimized at CCSD(T)/6-31G** I

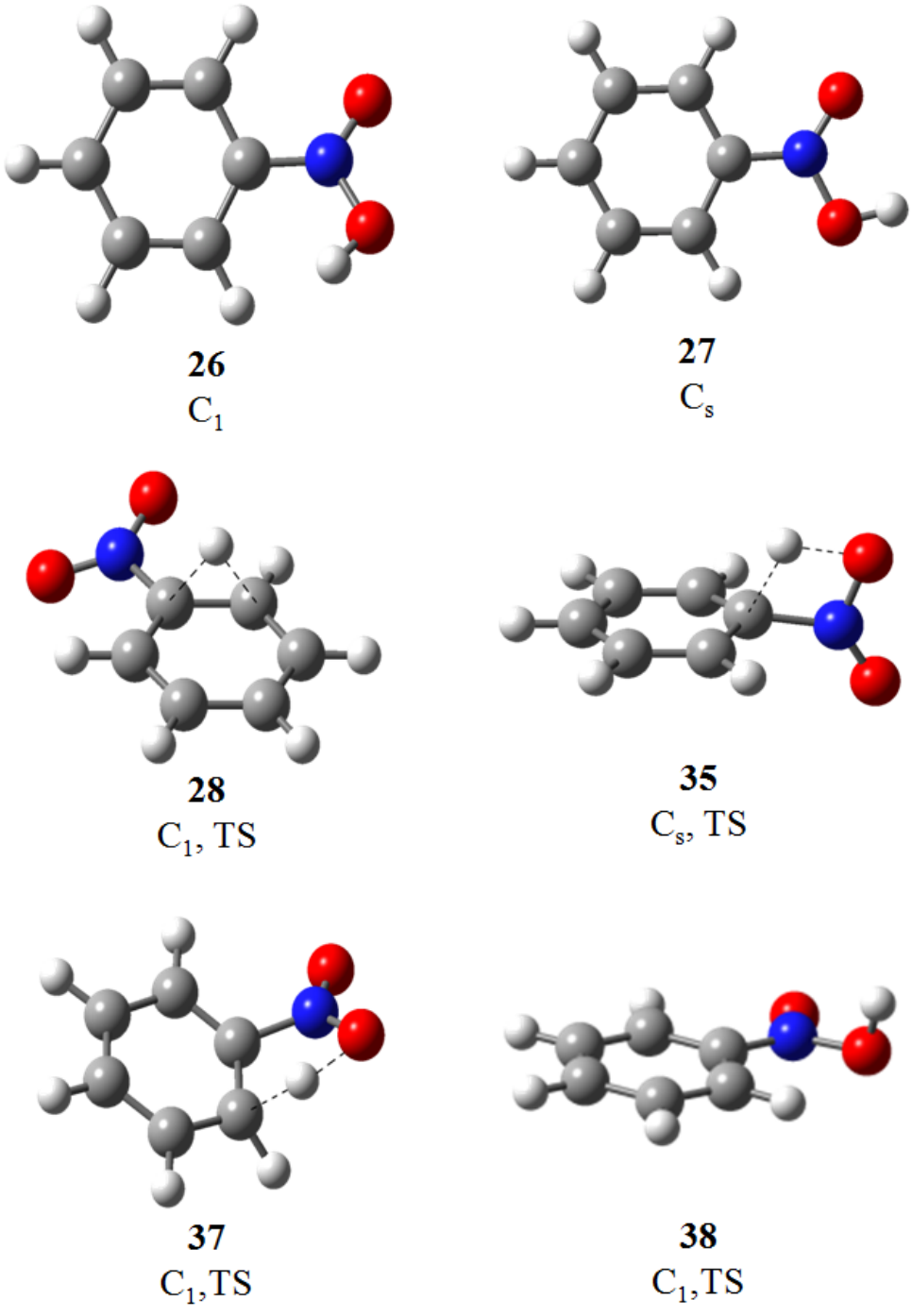


Figure 5.2

Stationary structures optimized at CCSD(T)/6-31G** II

The dielectric constant of a solvent (ϵ) plays a major role in solvent calculations, where the magnitude of the reaction field depends on the dielectric constant and the radius of the cavity. We have carefully chosen our solvent systems, which cover a wide rage of dielectric constants. Methylene chloride (CH_2Cl_2), acetonitrile (CH_3CN), and water (H_2O) were chosen as solvents, for which the dielectric constants are 8.93, 36.64, and 78.54, respectively. All calculations were carried out for 33 stationary points (11×3). Table 5.1 shows the relative energies of stationary points in vacuum, methylene chloride, acetonitrile, and water. Figure 5.3 shows the calculated potential energy profile of the benzine nitration reaction. The black solid line shows the vacuum reaction energy profile, and the blue dashed line shows the reaction energy profile in liquid water.

The CC-SCRF calculations predict that the intermediates and transition states in the condensed phases are considerably more stable than in the gas-phase. This observation can be understand through the phenomenon of solvent relaxation. The stationary points **8**, **28**, **23**, and **37** show more deviation from the gas-phase. This is expected because these have higher dipole moments of 10.45, 9.51, 11.17, and 8.31 Debye (D), respectively, in methylene chloride. When the polarity of the solvent increases, these tend to increases. This is not unusual because solvent polarization upon solvation in polar solvents causes the solute to be more polar. The vacuum and condensed phase dipole moments of all 11 stationary points are shown in Table 5.2.

Among all 11 stationary points, transition state **10** plays a critical roll in determining the crucial step in the mechanism of aromatic nitration. It is the major difference between the mechanisms proposed by Ingold-Hughes and the mechanism proposed by Olah. If

Table 5.1

Relative Energies in the Gas-phase and in Various Solvents

Species	Relative Energies in kcal/mol			
	CCSD(T)/6-31G**	CCSD/6-31G**//CCST(T)/6-31G**		
	Vacuum[a] ($\epsilon = 1$)	CH_2Cl_2 ($\epsilon = 8.93$)	CH_3CN ($\epsilon = 35.65$)	Water ($\epsilon = 78.54$)
C_6H_6/NO_2^+ [b]	0	0	0	0
4	−27.82	−29.76	−30.11	−30.18
7	−26.80	−28.35	−28.63	−28.69
8	−28.95	−38.51	−40.42	−40.77
10	−25.59	−31.19	−32.24	−32.44
23	−34.09	−43.54	−45.36	−45.71
26	−53.37	−56.88	−57.56	−57.69
27	−60.88	−65.22	−65.96	−66.10
28	−18.39	−25.91	−27.38	−27.66
35	−7.29	−11.92	−12.85	−13.03
37	−18.93	−23.68	−24.58	−24.75
38	−50.90	−55.37	−56.24	−56.41

[a]Ref. [108, 109] [b] Separated species in the gas phase

Table 5.2

Vacuum and condensed phase dipole moments in Debye

Species	Vacuum[a] ($\epsilon = 1$)	CH_2Cl_2 ($\epsilon = 8.93$)	CH_3CN ($\epsilon = 35.65$)	Water ($\epsilon = 78.54$)
4	4.35	5.45	5.67	5.71
7	2.34	3.12	3.28	3.31
8	8.45	10.45	10.83	10.90
10	7.51	8.86	9.11	9.15
23	9.21	11.17	11.53	11.59
26	2.73	3.59	3.78	3.82
27	1.22	1.25	1.25	1.26
28	7.79	9.51	9.84	9.90
35	5.85	7.40	7.69	7.75
37	6.35	7.78	8.04	8.09
38	2.62	3.27	3.55	3.63

[a]Ref. [108, 109]

transition state **10** exists, Olah's two intermediate mechanism takes place. If not, Ingold-Hughes's one intermediate mechanism takes place, because **10** creates the energy barrier between the π–complex and the σ–complex.

In the gas phase, the activation energy of intermediate **4** to the intermediate **8** is 2.23 kcal/mol, which is not a considerable amount. The relative energy of transition state **10**, the energy barrier, decreases by a significant amount in the condensed phases due to solvent stabilization. The relative energy differences between **4** and **10** in methylene chloride, acetonitrile, and water are -1.43 kcal/mol, -2.12 kcal/mol, and -2.26 kcal/mol, respectively. This means that the energy of **10** is lower than that of **4**. Thus, no energy barrier exists between **4** and **8**. Therefore, the condensed phase calculations predict that the reactants, which are benzene and nitronium ion, directly yield the σ–complex **8**. Hence, the mechanism of benzene nitration depends on the polarity of solvent used, and polar solvents enhance the Ingold-Hughes mechanism.

The dipole moment of intermediate **4** is 4.35, 5.46, 5.67, and 5.71 D in vacuum, methylene chloride, acetonitrile, and water, respectively. The total increment is 1.36 D from vacuum to liquid water. If we consider transition state **10** and intermediate **8**, it is not the same. Their dipole moments increase by 1.65 D and 2.45 D respectively from vacuum to liquid water. The free energy of solvation of intermediate **8** is greatly effected by its high dipole moment and all the higher order multipole moments contributions in the solvent phase . Therefore, it is expected that the free energy of solvation in liquid water of the intermediate **8** is as high as -11.82 kcal/mol. The dipole moment of transition state **10** is not increased as much, but including all the higher order multipole moments contributions,

the free energy of solvation of **10**, which is -6.85 kcal/mol, overcomes the gas-phase energy barrier, which is 2.23 kcal/mol, between **4** and **10**. The free energy of solvation of **4** is -2.36 kcal/mol. Due to the solvation free every of **10** being greater than the energy barrier between **4** and **10**, the **10** is no longer a stationary point. Hence, these results imply that the Ingold-Hughes's mechanism of nitration of benzene takes place in polar solvent media.

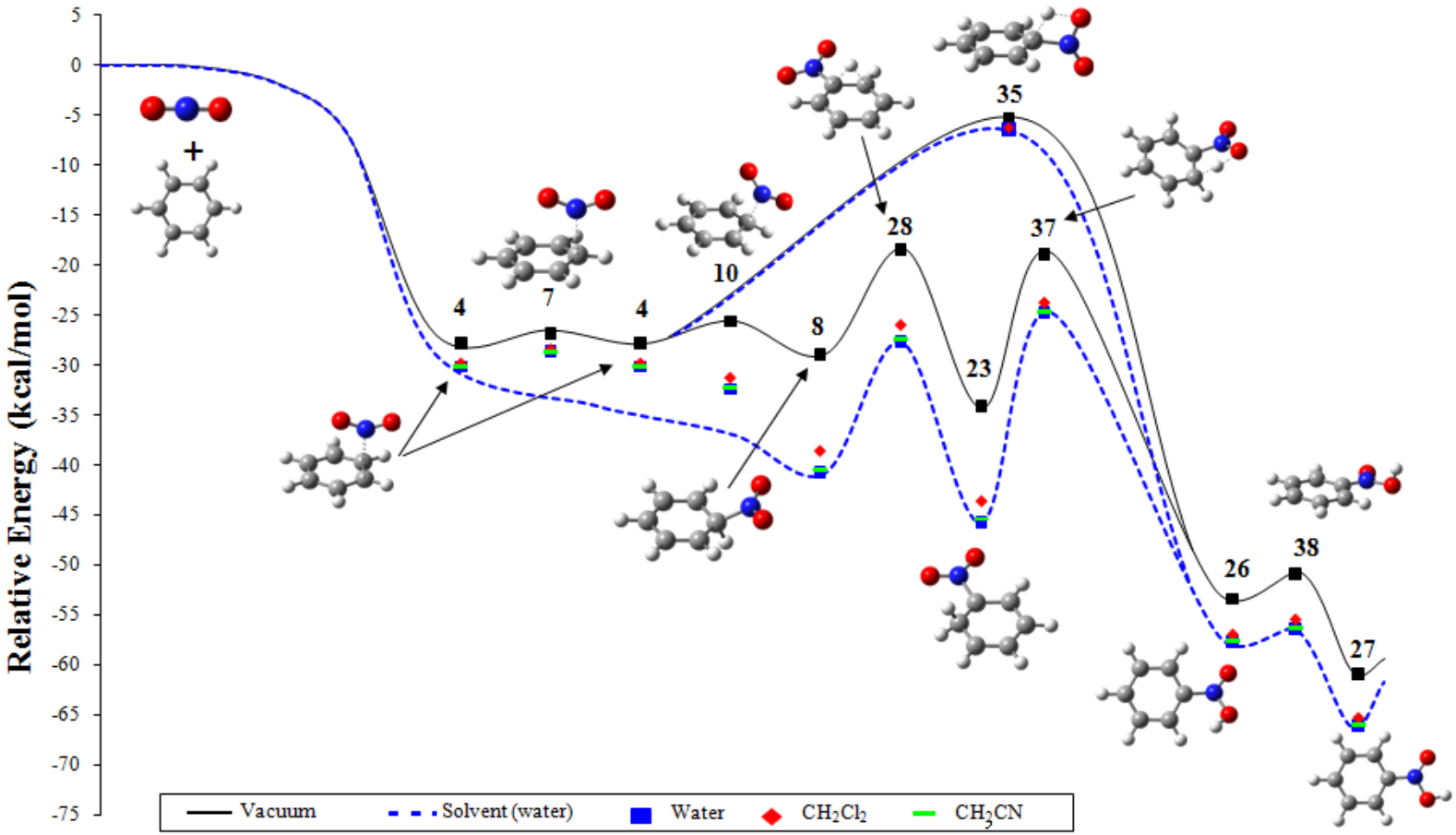


Figure 5.3

Potential Energy Profile with condense phase

CHAPTER 6

REFERENCE INTERACTION SITE MODEL

6.1 Site-Site Ornstein-Zernike (RISM) Theoy

In 1972 Chandler and Andersen [41, 42] first proposed an approximate statistical mechanical model they called reference interaction site model. The RISM equation is a natural extension of the Ornstein-Zernike (OZ) integral equation theory for liquids. In its original forms, the theory takes into account one of the two important chemical aspects of molecules, the geometry, in terms of intramolecular distribution functions. The other chemical aspect of molecules, electrostatics, was introduced as well in 1981 by F. Hirata et al. [43–45] in the extended RISM formulation (XRISM), which considers the charge distribution as well as the molecular geometry contribution. (In the rest of this dissertation RISM always refers to XRISM).

The solvation structure around a molecule is commonly described by a pair correlation function (PCF) or radial distribution function (RDF) $g(r)$. This function represents the probability of finding a specific atom at a distance r from the atom being studied. Figure 6.1 shows the PCF of oxygen-oxygen and hydrogen-oxygen in liquid water.

The first peak at around 1.8 Å in the the O–H PCF shows the probability of finding an H atom from another molecule around an O atom (or vice versa). This sharp peak arises

due to strong hydrogen bonding. The second peak in the O–H PCF ($\sim$3.5 Å) and the first peak in the O–O PCF ($\sim$ 3.0 Å) are consistent with the configuration of two water molecules linked by a hydrogen bond as shown in Figure 6.1.

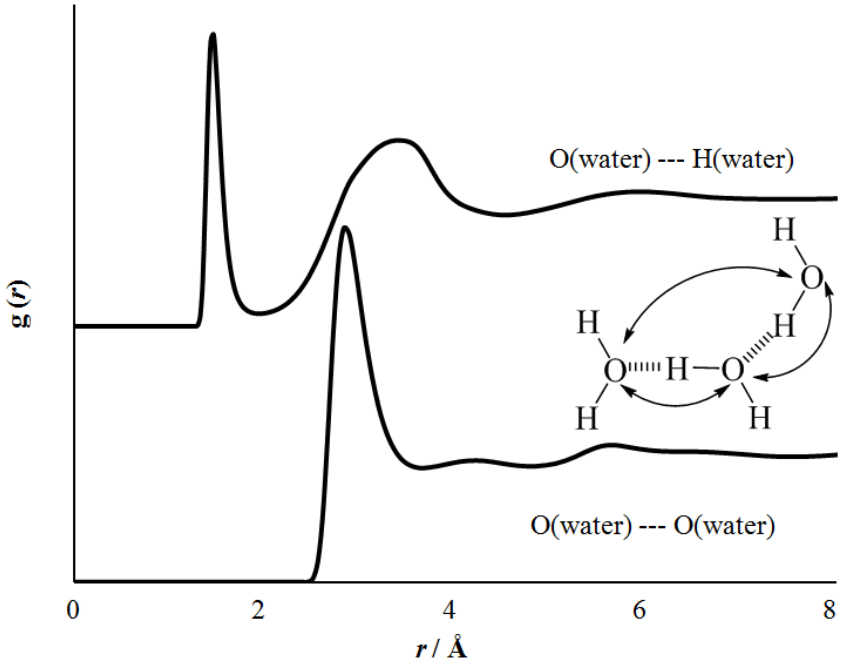


Figure 6.1

Pair correlation function in liquid water obtained from HF/ 6-31G* with CHARMM TIP3P parameters

The Ornstein-Zernike (OZ) equations [41, 42] state that the pair correlation function can be written as [57, 58]

$$h(\mathbf{r},\mathbf{r}') = c(\mathbf{r},\mathbf{r}') + \int\!\!\int c(\mathbf{r},\mathbf{r}'')\rho(\mathbf{r}'')h(\mathbf{r}'',\mathbf{r}')\mathrm{d}r'', \tag{6.1}$$

where ρ is the density of the liquid, and $\mathbf{r}$ represents the positions of the particles. The functions h and c are called the "total" and "direct" correlation functions, respectively. h

is essentially equivalent to the PCF, and both functions are usually regarded as functions of the distance between particles r under isotropic and uniform conditions. Mathematically it is shown that $h(r) = g(r) - 1$.

A general expression of the RISM equation for a system consisting of several molecular species (a dense liquid) can be written as

$$\boldsymbol{\rho}\mathbf{h}\boldsymbol{\rho} = \boldsymbol{\omega} * \mathbf{c} * \boldsymbol{\omega} + \boldsymbol{\omega} * \mathbf{c} * \boldsymbol{\rho}\mathbf{h}\boldsymbol{\rho} \tag{6.2}$$

where an asterisk ($*$) denotes the spatial convolution integral of matrix product in Fourier space. The convolution of the two functions, denoted $g * h$, is defined by $g * h = \int_{-\infty}^{\infty} g(\tau)h(t-\tau)\mathrm{d}\tau$, where $g(\tau)$ and $h(\tau)$ are in Fourier space. $\boldsymbol{\rho}$ denotes a diagonal matrix consisting of the bulk liquid density. $\mathbf{h}$, $\mathbf{c}$, and $\boldsymbol{\omega}$ are the matrices of the site-site total, direct, and intramolecular correlation functions, respectively. In the case of rigid molecules, ω is a delta function and expressed by

$$\omega_{\alpha\beta}(r) = \rho\delta_{\alpha\beta}\delta(r) + (1-\delta_{\alpha\beta})\frac{1}{4\pi L_{\alpha\beta}^2}\delta(r - L_{\alpha\beta}), \tag{6.3}$$

where, $\delta_{\alpha\beta}$ is the Kronecker delta, α and β are atoms or sites, and $L_{\alpha\beta}$ is the distance between atom site pair α and β of same molecule (the chemical bond). $\mathbf{h}$ and $\mathbf{c}$ are described in Eqs. (6.4–6.6).

Since all atomic sites that belong to the same molecular species have the same density, Eq. (6.2) can be simplified considerably if one molecular species and thus all atomic sites associated with it is considered to have zero density (this is called the limit of infinite dilution). This assumption is made for the case of solute–solvent and solute–solute calculations, where the solute species is assumed to have zero density. Using the conven-

tional notation often employed for a sol‘u’te–sol‘v’ent system, one can rewrite the above equation as a set of equations regarding solvent–solvent (vv), and solute–solvent(uv) correlation functions. Whithin the limit of infinite dilution, a new set of formulae can be obtained from dividing Eq. (6.2) by $\boldsymbol{\rho}^2$, when the solvent density is not zero,

$$\mathbf{h}^{vv} = \mathbf{w}^{v} * \mathbf{c}^{vv} * \mathbf{w}^{v} + \mathbf{w}^{v} * \mathbf{c}^{vv} * \boldsymbol{\rho}^{v}\mathbf{h}^{vv}, \tag{6.4}$$

and

$$\mathbf{h}^{uv} = \mathbf{w}^{u} * \mathbf{c}^{uv} * \mathbf{w}^{v} + \mathbf{w}^{u} * \mathbf{c}^{uv} * \boldsymbol{\rho}^{v}\mathbf{h}^{vv}. \tag{6.5}$$

Equations (6.4), and (6.5) contain two unknown functions: the total correlation function $h(r)\,(= g(r) - 1)$, and the direct correlation function $c(r)$. In order to solve them, a second equation is used, which is a closure equation, that completes the system of equations. Normally, the hypernetted chain (HNC) closure [214] is used as the closure equation. The HNC closure can be written as

$$c_{\alpha\beta}(r) \;\; = \;\; \exp\left[-\beta u_{\alpha\beta}(r) + h_{\alpha\beta}(r) - c_{\alpha\beta}(r)\right] - h_{\alpha\beta}(r) - c_{\alpha\beta}(r) - 1, \tag{6.6}$$

where $\beta = 1/k_B T$ and $u_{\alpha\beta}(r)$ is the solute–solvent interaction potential. The interaction potential is the Lennard-Jones (L-J) potential plus the Coulombic interaction:

$$u_{\alpha\beta}(r) = 4\epsilon_{\alpha\beta}\left[\left(\frac{\sigma_{\alpha\beta}}{r}\right)^{12} - \left(\frac{\sigma_{\alpha\beta}}{r}\right)^{6}\right] + \frac{q_\alpha q_\beta}{r}, \tag{6.7}$$

where σ and ϵ are the Lennard-Jones collisional diameter and attractive well depth parameters, respectively, and q_α and q_β are the charges on atom sites α and β. In all cases the standard combination rules

$$\epsilon_{\alpha\beta} = \sqrt{\epsilon_\alpha \epsilon_\beta} \tag{6.8}$$

and

$$\sigma_{\alpha\beta} = \frac{(\sigma_\alpha + \sigma_\beta)}{2} \tag{6.9}$$

are employed to estimate the L–J parameters for a pair of different atoms.

Now, we can solve Eqs. (6.4) and (6.6) self-consistently until both $\mathbf{h}^{vv}$ and $\mathbf{c}^{vv}$ converge. One we have converged $\mathbf{h}^{vv}$, we can now solve Eqs. (6.5) and (6.6) self-consistently until both $\mathbf{h}^{uv}$ and $\mathbf{c}^{uv}$ converge. To solve these equations, the density, temperature, and the Lennard-Jones parameters are needed. When this information is available, the two correlation functions $h(r)$ and $c(r)$ can be determined by solving those simultaneously.

6.2 RISM-SCF Theory

The basic idea of RISM-SCF is to calculate the reaction field from the solvent molecules by using microscopic information of solvation, such as the PCFs between solute and solvent, which are computed by the RISM theory. All thermodynamic functions can be calculated from pair correlation functions. The excess chemical potential or the solvent free energy of a molecule can be express using the total correlation function $h(r)$ and the direct correlation function $c(r)$ as originally defined by Singer and Chandler [214]

$$\Delta\mu = -\frac{\rho}{\beta}\sum_{\alpha\lambda}\int\left[\left(c_{uv}(r) - \frac{1}{2}h_{uv}^2(r) + \frac{1}{2}h_{uv}(r)c_{uv}(r)\right]\mathrm{d}r. \tag{6.10}$$

The Fock operator of a molecule in solution can be expressed as

$$F_{\mu\nu} = F_{\mu\nu}^0 + F_{\mu\nu}^{sol}, \tag{6.11}$$

where,

$$F_{\mu\nu}^{sol} = \sum_{\lambda\in\text{solute}}\left(V_\lambda b_\lambda. \tag{6.12}$$

The first term on the right hand side of the Eq. (6.11) is the operator for an isolated molecule, and the second term represents the effects from the solvent perturbation. V_λ represents the electrostatic potential of the reaction field at solute atom site λ produced by solvent molecules, and b_λ is a population operator measuring the charge density on for the solute atom λ. The gross population of partial charges on site λ is expressed by summing over all the contributions due to the electrons and nuclei,

$$q_\lambda = q_\lambda^N - \sum_i \langle \phi_\mu | b_\lambda | \phi_\nu \rangle, \tag{6.13}$$

where, q^N is the charge of the nucleus, and ϕ_μ, ϕ_μ have their usual meanings. The least squares fitting treatment for partial atomic charge analysis used here is described in the Appendix.

The electrostatic potential term V_λ takes the contribution from the site-site radial distribution functions between solute and solvent:

$$V_{\lambda \in \text{solute}} = \mathbf{V}_{\lambda \in \text{solute}} = \rho \sum_{\alpha \in \text{solvent}} \int \frac{q_\alpha}{r} g_{\lambda\alpha}(r) \mathrm{d}\mathbf{r}, \tag{6.14}$$

where q_α is the partial charge on solvent site α, ρ is the bulk density of the solvent, and $g_{\lambda\alpha}(r) = h_{\lambda\alpha}(r) + 1$ is the RDF.

The statistical solvent distribution around the solute is determined by the electronic structure via the partial charges of the solute, while the electronic structure of the solute is influenced by the solvent distribution. Therefore, the SCF equation and the RISM equations should be solved in a self-consistent manner. In order to distinguish the RISM reaction field from the usual reaction field approach, $\mathbf{V}_{\lambda \in \text{solute}}$ given in Eq. (6.14) can be assumed to be the microscopic mean field. The microscopic mean field includes molecular-

level information as well as bulk properties of the solvent. Therefore, the RISM model can be much more informative compared to the usual SCRF model based on the macroscopic dielectric continuum.

6.3 Implementation of CC-RISM Method

In order to solve the RISM equations, Lennard–Jones (L-J)parameters for both solute and solvent in Eq. (6.7) and partial atomic charges are need. The standard CHARMM TIP3P [215, 216] set of parameters are utilized as L–J parameters in our calculations. TIP3P [217], SPC [218], SPC/E [219], and OPLS [220] parameters could also be used as solvent parameters. The procedure for calculating partial atomic charges is described in the Appendix. The $h(r)$ and and $c(r)$ functions are generated using Benoit Roux's [221, 222] RISM code.

After one SCF cycle, the required partial atomic charges are obtained from the density matrix. Then the RISM equations, Eq. (6.4) with Eq. (6.6) and Eq. (6.5) with Eq. (6.6), are solved self-consistently until converged. The perturbation due to the solvent is calculated using Eqs. (6.12) and (6.14). The one-electron Hamiltonian is built using Eq. (6.11). These SCF cycles are iterated until both the partial atomic charges and the electron density are converged. The MO coefficients obtained from the SCF calculations are utilized to solve for both T and Λ in the CCSD equations self consistently. Once the CCSD amplitudes (T and Λ) are calculated, the CCSD relaxed density D_{pq}^{rlx} is obtained. A new microscopic reaction field $V^{CC}_{\lambda\in\text{solute}}$ is calculated using the relaxed OPDM. $V^{CC}_{\lambda\in\text{solute}}$ is obtained by solving the RISM equations with the CC partial atomic charges q^{CC}. Then

a new HF calculation is performed followed by a CCSD calculation in the presence of the new $V^{CC}_{\lambda\in\text{solute}}$. The HF calculations before CCSD calculations guarantee that the MO coefficients include the new response to the microscopic reaction field response. This procedure is repeated using the CC solvent response until the CC partial atomic charges converge. Once the CC partial atomic charges are converged, the free energy of solvation is calculated using the converged $h(r)$ and $c(r)$, with Eq. (6.10). Our CC-RISM procedure consists of the following steps.

1. Obtain the partial atomic charges.
2. Compute the HF-SCF with RISM equations.
3. Solve the CCSD T and Λ equations.
4. Obtain the CCSD relaxed density, D^{rlx}_{pq}.
5. Compute q^{CC} from the relaxed density D^{rlx}_{pq}.
6. Recalculate the RISM solution with q^{CC}
7. Rerun HF with fixed $V^{CC}_{\lambda\in\text{solute}}$ and then return to step 3 if q^{CC} is not converged.
8. Obtain the free energy of solvation solving the RISM equations.

6.4 Partial Atomic Charge Treatment

Before giving a more detailed description the RDFs, we will discuss about the partial atomic charges. One of the concerns which is common to this type of hybrid approach is an error associated with determination of the charge distribution. The partial atomic charges are the only quantum information that we insert to solve the RISM equations. So we need to pay extra attention to the calculations of the partial atomic charges. The partial atomic charges in the our RISM procedure are determined in such a way that the electrostatic potential (ESP) produced by the partial atomic charges at grid points around

the solute molecule best agree with that calculated from the wave functions in a least squares fitting (LSF) sense. The Appendix describes the procedure for obtaining the partial atomic charges. Typically, in order to calculates the partial atomic charges, we need to calculate the ESP over a large number of grid points in the space surrounding the molecule. The choice of grid points can alter the result for the partial atomic charges for a given set of molecular orbitals. We have studied the dependence of the partial atomic charges on the number of selected grid points. Two parameters should be considered to select a set of grids. They are the number of grid points and the distance of the grid point from the molecule.

Lebedev grids [223–228] are specially constructed grids for quadrature on the surface of a sphere based on the octahedral point group. We selected the Lebedev grids because they reproduce the symmetry at the spherical harmonics in higher order. All grid points are constructed with an atomic nucleus as their origin. We selected the first layer of Lebedev grid points to be on the van der Waals surface of each atom. Then each additional layer or shell of grids is added 0.5 times the van der Walls radius further out from each atom. The grid points contained within overlapping spheres were discarded.

Table 6.1 shows the dependency of the partial atomic charges on $1 - 10$ grid shells with 2030 and 5294 grid points per atom per shell. Table 6.1 clearly shows that the partial atomic charges are converged to 0.0001 au when the total number of grid shells is over five. The other important fact in Table 6.1 is the lack of considerable difference between partial atomic charges that are obtained from 2030 and 5294 grid points per shell.

Table 6.1

Variation of partial atomic charges over various number of grid shells with 2030 and 5294 grid points per atom per grid shell for H_2O with CCSD/aug-cc-pVTZ in water

Number of Shells	Number of Grid Points			
	2030		5294	
	O	H	O	H
1	-0.8562	0.4281	-0.8562	0.4281
2	-0.8563	0.4281	-0.8562	0.4281
3	-0.8559	0.4279	-0.8560	0.4280
4	-0.8556	0.4278	-0.8556	0.4278
5	-0.8554	0.4277	-0.8554	0.4277
6	-0.8554	0.4277	-0.8554	0.4277
7	-0.8554	0.4277	-0.8554	0.4277
8	-0.8554	0.4277	-0.8554	0.4277
9	-0.8554	0.4277	-0.8554	0.4277
10	-0.8554	0.4277	-0.8554	0.4277

Table 6.2 shows the variation of partial atomic charges over various numbers of grid points per grid shell on one grid shell and five grid shells. The partial atomic charges are converged to 0.0001 au after 2030 grid points per shell. There is a significant difference in partial atomic charges between those obtained with one shell and five shells. Both the number of grid points per atom and the number of shells have considerable effects on the partial atomic charges because we reproduce the electrostatic potential around the solute molecule with the least squares fitting procedure to get the best estimation for the partial atomic charges.

The geometry of water was optimized with CCSD(T) and the augmented correlation consistent polarized valence triple zeta (aug-cc-pVTZ) basis set [229, 230] using ACES II. Then, single point calculations are performed with RISM at CCSD with aug-cc-pVTZ

Table 6.2

Variation of partial atomic charges over various number of grid points per shell on one grid shell and five grid shells for H_2O with CCSD/aug-cc-pVTZ in liquid water

Number of Grid Points	Number of Shells			
	1		5	
	O	H	O	H
110	-0.8577	0.4289	-0.8568	0.4284
350	-0.8570	0.4285	-0.8566	0.4283
590	-0.8564	0.4282	-0.8560	0.4280
974	-0.8564	0.4282	-0.8560	0.4280
1202	-0.8562	0.4281	-0.8558	0.4279
1454	-0.8562	0.4281	-0.8556	0.4278
2030	-0.8562	0.4281	-0.8554	0.4277
3470	-0.8562	0.4281	-0.8554	0.4277
4334	-0.8562	0.4281	-0.8554	0.4277
5294	-0.8562	0.4281	-0.8554	0.4277

using the Q-CHEM program, in which we implemented the CCSD-RISM method. From Table 6.1 and 6.2, it appears safe to use 2030 grid points per atom per shell with five grid shells as optimal conditions. Using these conditions, the vacuum partial atomic charges at CCSD/aug-cc-pVTZ for O and H are -0.7370 au and 0.3685 au, respectively. We obtained the partial atomic charges in liquid water using CCSD/aug-cc-pVTZ for O and H as -0.8554 au and 0.4277 au, respectively with the above stated set of grid points. Increasing the magnitude of the partial atomic charges in condensed phase is expected because polar molecules get more polarized in polar solvents.

If we consider the dipole moments in both vacuum and condensed phase we can obtain the same pattern in them too. The experimental dipole moment of an isolated water monomer has a conclusive value of 1.854 D [231, 232], and the theoretical values also

reach the experimental value [233–241]. But the precise condensed phase dipole moment value has been the subject of some debate for over five decades. It is expected that the dipole moment of liquid water considerably increases in the condensed phase as compared to the gas phase. This increase is very important to understand the large dielectric constant of water. Although the experimental measurement of the dipole moment in liquid water is not possible, indirect results can be obtained from X-ray and synchrotron radiation experiments [242], which yield a value of 2.9 ± 0.6 D with quite large error bars. In the crystalline phase the dipole moment of water in the most common phase of ice (Ih) using a multipole expansion gave a the result of 2.6 D [243]. Batista et al. [244] have more recently revised the multipole moments using both experimental (up to quadrupole) and theoretical (up to hexadecapole) moments and obtained a considerably larger value of 3.09 D.

Theoretical estimates [245–262] predict that condensed phase dipole moment increases between 15% and 60%. The average water dipole moment in liquid water has been calculated from ab initio Car–Parrinello molecular dynamics (CPMD) [24] simulations by Laasonen et al. [254]. They obtained an average water dipole moment of 2.66 D while integrating the electron density to the spherical cut-off around each atom in the water molecule. Another CPMD simulation obtained a value of 2.47 D [249] using the flux of the electron density gradient to identify the electron distribution of individual molecules. Parrinello et al. [248] find that the average dipole moment of a water molecule in the liquid should be about 3.0 D, a value much larger than those obtained from other calculations. Jansen et al. [253] reported a value of 2.62 D using the SCRF approach with CISD. In

2002 Poulsen et al. [258] performed MCSCF calculations on molecular dynamics structures.. They obtained a dipole moment of 2.71 D. In order to reproduce the experimental static dielectric constant ($\varepsilon = 78.3$), the water dipole moment is estimated to be about $\sim$2.6 D [262]. An average dipole moment exceeding 2.6 D leads to a significant overestimation of ε [256]. The large uncertainty in the reported values of the dipole moment of liquid water reflects the fact that our basic understanding of the water molecule in condensed phases is still fairly limited. A liquid is statistical by nature, and its structure does not correspond to a minimum-energy configuration. The proper description of the liquid state needs a statistical procedure. Hence, a statistical mechanics-based quantum chemical molecular electronic structure method may give promising results. The best estimation for dipole moments in vacuum and liquid water based on our calculations using CCSD-RISM with the aug-cc-pVTZ basis set is 1.8693 D and 2.3769 D, respectively. A value of 1.8466 D is obtained in vacuum using CCSD(T) with aug-cc-pVTZ basis set.

6.5 Radial Distribution Functions of Water

The structural investigation of water has a strong historical precedence, tracing roots at least as far back as Röntgen's early papers [263–265] on the structure of water and the explanation of its density maximum. In principle, an accurate characterization of the molecular structure of liquid water [266–277] can be found from solution scattering experiments (X-ray and neutron diffraction), MM (both MC and MD), and AIMD simulations. Since the integral equations of RISM theory are adopted from the statistical mechanics and the

liquid state theory, the structure of water, i.e., PCFs, and other thermodynamic properties can be naturally calculated using RISM theory.

In this section we present an application for our CC-RISM method in this section. We have chosen to study the structure of the H_2O molecule as a solute to demonstrate the capability of the CC-RISM method by describing the mutual effect between the electronic and solvation structure. Studying the structure of water in condensed phase is a classic example of all experimental, molecular simulations, and ab initio calculations [266–277].

Table 6.3

CHARMM TIP3P Geometrical and potential parameters for water

r(OH)/Å	0.9572	
$\angle$(HOH)/(°)	104.52	
	O	H
ϵ(kcal/mol)	0.1591	0.0498
σ(Å)	2.8509	1.4254
q(au)	-0.834	$+0.417$

It is very important to understand that there are two types of water in a CC-RISM electronic structure calculation, namely the solute water and the solvent water. The solute water molecule is the water molecule on which we perform all electronic structure calculations and the solvent water is the classic bulk. The geometry and the L–J parameters for solvent water are taken from the CHARMM TIP3P [215, 216] set. The solvent parameters are shown in Table 6.3. All the van der Waals interactions between the solute and solvent molecules were calculated by means of the standard combining rule in Eqs. (6.8),

and (6.9). All the calculations were carried out at 298 K and at the density of bulk water 0.03334 molecule/Å^3.

We have carried out a CCSD(T) geometry optimization without dropping core orbitals using Dunning's cc-pVTZ basis set for the solute water molecule. Then a CCSD-RISM calculation is performed using the cc-pVTZ basis set for the optimized solute molecule with CHARMM TIP3P waters as solvent. The radial distribution functions, partial atomic charges, dipole moments, and solvation free energies are calculated.

Figures 6.2 and 6.3 show the radial distribution functions of water with SCF and CCSD, respectively. The graphs are drawn by the distance from the specified solute atom in Å vs. the probability of finding the specified solvent atom. The blue OU—OV curve denotes the PCF of solute oxygen and solvent oxygen, the red OU—HV curve denotes the PCF of solute hydrogen and solvent oxygen, and the green HU—HV curve denotes the PCF of solute hydrogen and solvent hydrogen. Although both graphs are qualitatively alike, in depth they are different. First we will explain the qualitative picture. One of the important observation is the first distinct peak of the red OU—HV curve. It is a direct proof of the hydrogen bond between a pair of water molecules. The second peak of the same PCF arises from the other hydrogen atom of the same solvent water molecules which are H-bonded to the solute water molecule. Another important observation is the position of the second peak in the blue OU—OV curve, which is caused by the tetrahedral icelike network [275–277] of water. In both Figures 6.2 and 6.3, no water molecules could be recognized. We are no longer able to count the explicit number of molecules surrounding the molecule being studied. Alternatively, the PCFs allow us to understand how many sur-

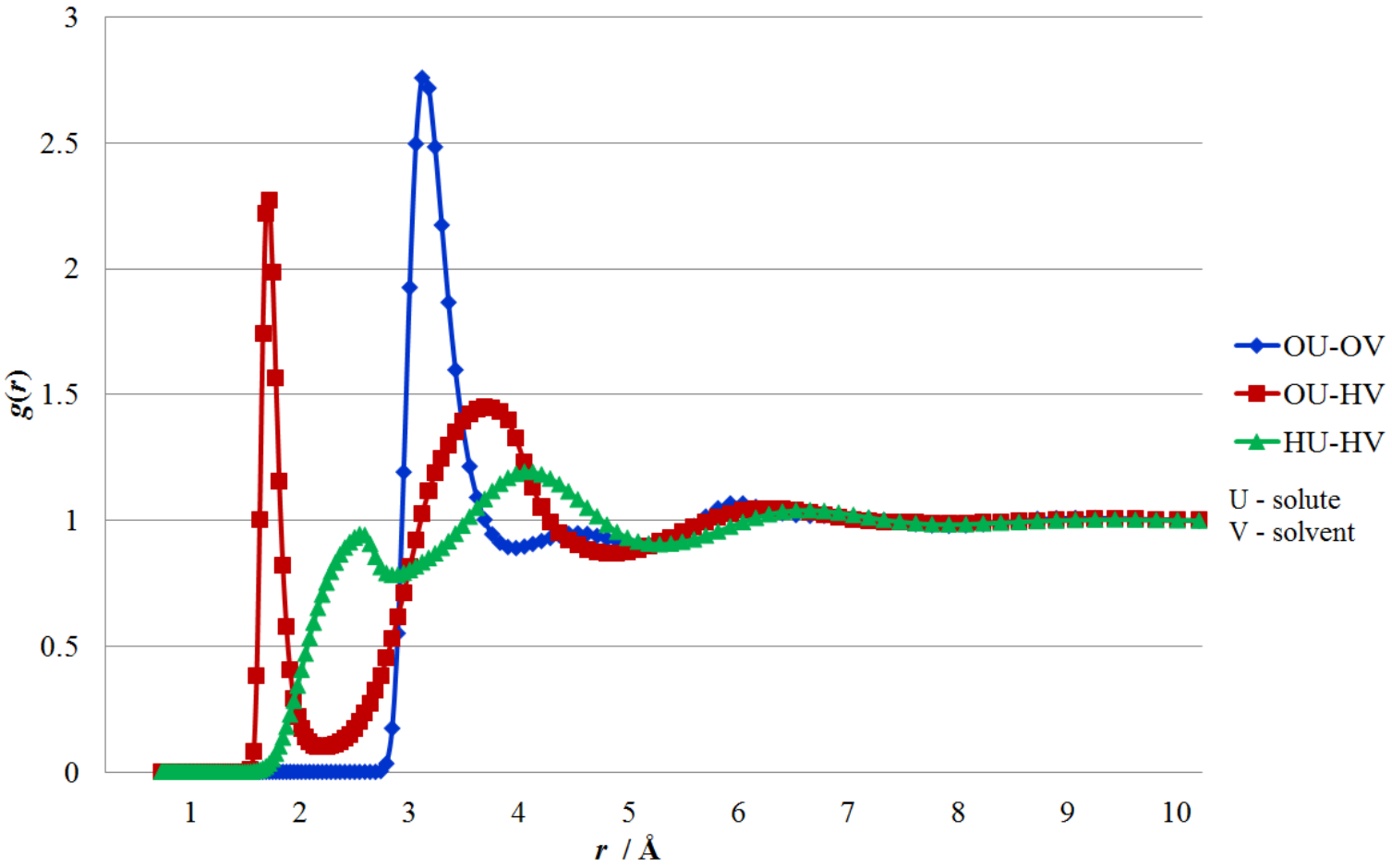


Figure 6.2

PCFs of H_2O with HF/cc-pVTZ in water

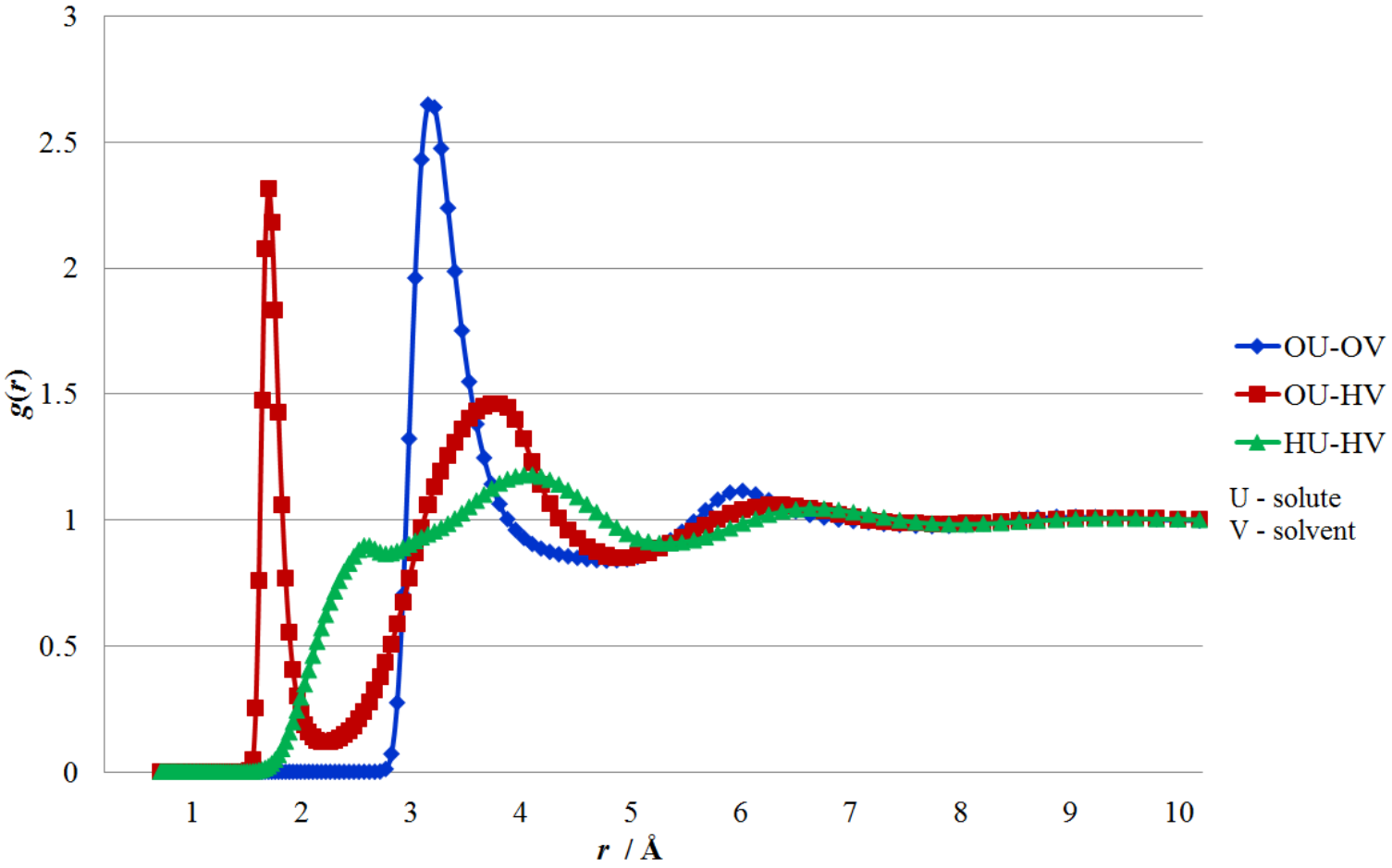


Figure 6.3

PCFs of H_2O with CCSD/cc-pVTZ in water

rounding molecules are found in the statistical sense through the probability distribution function. In the real picture molecules in a condensed phase are in continuous motion. Identifying solvent molecules or determining an integer number of solvent molecules is of no use.

The PCFs in both Figures 6.2, and 6.3 carry not only the statistical nature of the solvent structure, but also information based on the quantum chemical calculations. These are obtained from solving SCF and CCSD equations with RISM self-consistently as discussed in Section 6.3. Unlike MM simulations, QM methods with RISM open many doors to obtain trustworthy electronic molecular properties such as electronic spectrum, properties of excited states, vibrational frequencies, equilibrium geometries, polarizabilities, hyperpolarizabilities, solvation free energies etc.

At a glance both Figures 6.2, and 6.3 look alike. In order to analyze the differences of SCF and CCSD PCFs, we subtracted the CCSD PCFs from the SCF PCFs, and the results are shown in Figure 6.4. Each curve in Figure 6.4 first becomes negative and then becomes positive. This means that the peaks in the SCF PCFs appear a little earlier than the CCSD PCFs. For an example, the distance from a solute oxygen to the first peak of the OU—HV PCF in Figure 6.2 is shorter than the distance from a solute oxygen to the first peak of the OU—HV PCF in Figure 6.3. Hence, it is clear that the SCF-RISM solvation shells are a bit tighter that the CCSD-RISM solvation shells. This conclusion is not hard to understand because CCSD calculations includes electronic correlation whereas SCF does not. Each electron knows where the others are in the CCSD calculations, while in SCF the electrons

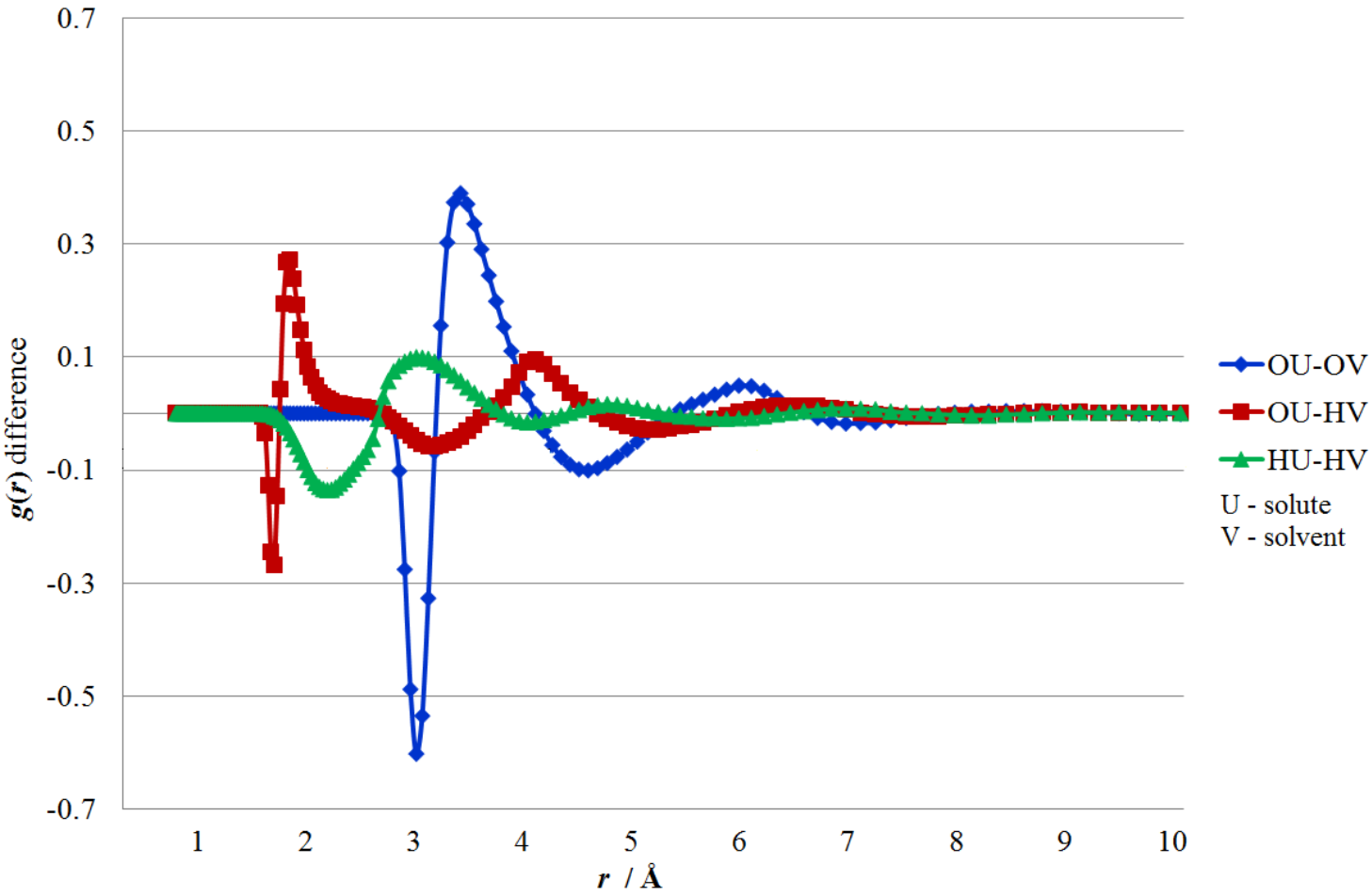


Figure 6.4

Difference between PCFs of water with CCSD and HF with cc-pVTZ

feel the average field of the other electrons. Since the electrons are correlated, they tend to repel or keep away each other, so the CCSD solvation shells are larger than those of SCF.

6.6 Radial Distribution Functions of *N*-methylacetamide

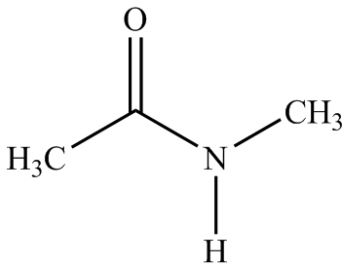
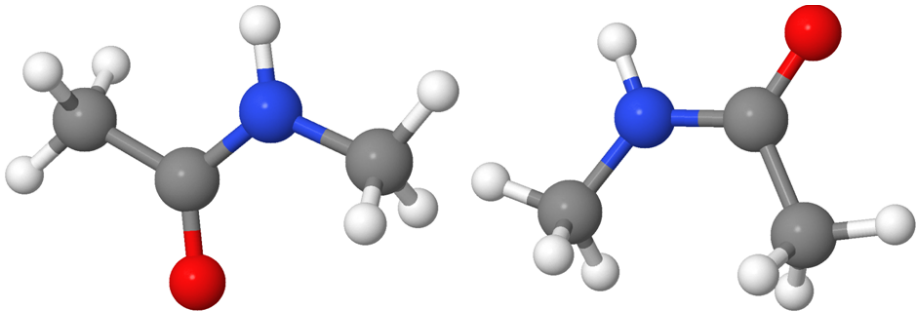


Figure 6.5

N-methylacetamide.

N-methylacetamide (NMA) is a very interesting compound and often serves as a model of the peptide bond. The interaction between NMA and water provides a convenient prototype for the solvation of the peptides in aqueous solutions. The structural information, non-bonded interactions, and solvent nature have been studied experimentally [278–282] as well as theoretically [283–293] over past few decades.

Dixon et al. [283] investigated the strengths of the hydrogen bonds between a water molecule and both *cis*- and *trans*-NMA using MP2 with correlation consistent basis sets. Han et al. [284] studied the conformations, hydrogen bonding effects, and stabilities of different methyl group orientations of isolated *cis*- and *trans*-NMA (the structures of *cis* and *trans* isomers are shown in Figure 6.6) and of NMA$(H_2O)_n$, $(n = 1 - 3)$ complexes,



trans-NMA *cis*-NMA

Figure 6.6

The conformers of isolated NMA.

using all kinds of density functional theory. Gao et al. [289–291] have performed MM simulations as well as QM/MM studies of the solvent nature and the non-bonding interactions of NMA. Vaccum excited state properties and electronic spectra of NMA were studied by Hirst et al. [294] using multireference configuration interaction (MRCI) in a 6-31+G** basis set and by Fülscher and co-wokers [295] using complete active space self-consistent field with second-order perturbation theory (CASSCF/CASPT2). An MCSCF-SCRF/CASPT2 study of the effect of solvation on the electronic spectra of NMA was also reported by Besley et al. [296].

We have performed CCSD-RISM calculations in liquid water to analyze the water–NMA interactions and solvation free energy of both *cis* and *trans* conformers of NMA. All geometrical optimizations have been performed with CCSD(T) using the 6-31G** basis set without frozen core orbitals. CCSD-RISM calculations have been performed using the 6-31G** basis set and also including all orbitals. The CHARMM TIP3P geometry and L–J parameters are used for the solvent water, and CHARMM L–J parameters are used

for NMA. We have given special consideration to the two major types of intermolecular hydrogen bonds which can occur in NMA in liquid water, namely C=O···H−O−H and N−H···OH_2. The former h-bond forms between the amide oxygen in NMA and a water hydrogen, and the latter H-bond forms between amide hydrogen in NMA molecule and a water oxygen. Both are key to the stabilities of the secondary and tertiary structures of a protein in the aqueous phase [289–291].

Figures 6.7, 6.9, and 6.11 show the pair correlation functions of *cis*-NMA, and Figures 6.8, 6.10, and 6.12 show the pair correlation functions of *trans*-NMA in liquid water. Because of the number of interactions, we have divided the PCF of *cis*-NMA as the PCFs for the carbonyl group (C=O), the PCFs for the amine group (N−H), and PCFs for the methyl groups (CH_3) with liquid water. We studied the CH_3 group as a united atom with the C atom and H atoms grouped together because it does not show any special interaction with liquid water. All graphs are plotted as the probability distribution vs. the distance between atom pairs in Å.

The first prominent peaks of the OU—HV PCF and the OU—OV PCF in Figure 6.7 clearly show the evidence of C=O···H−O−H hydrogen bonding. The first peak of the red the OU—HV PCF is at ~ 1.9 Å, and the distance between the red OU—HV PCF peak and the blue OU—OV PCF peak is about 1 Å. The subsequent peaks of both OU—HV and OU—OV PCFs shows the proper packing in the solvent structure. The C atom of the C=O group normally does not show any nonbonded interactions with liquid water. So the green CU—OV and the purple CU—HV PCFs do not have prominent peaks in the nonbonding region. The first peak of the green CU—OV PCF may have occurred due to the solvent

structure around the molecule. In particular, the solvent hydrogen and the carbonyl carbon do not show any interaction, but the peaks arise from solvent packing near the solute.

The major difference between Figures 6.7 and 6.8, which are the PCFs of the carbonyl (C=O) group with liquid water of *cis*- and *trans*-NMA, respectively, is that the *cis* -NMA has higher probability of finding a solvent molecule within the range of nonbonding interactions than *trans*-NMA. This is acceptable because both the O atom of the carbonyl group and the H atom of the (N−H) group are on the same side of the peptide bond, and all non-bonded interactions accumulate to that region. The other important point is that the peaks of the PCFs of the *trans*-NMA isomer with liquid water are shown to be a little closer to the NMA molecule than those of the *cis* -NMA isomer. Figures 6.9 and 6.10, which are the PCFs of the amine (N−H) group with liquid water of *cis*- and *trans*-NMA, respectively, are almost the same, but the *trans*-NMA isomer PCFS are bit tighter. The sharp peak of the HU—OV PCF shows the hydrogen bond formation in both *cis* and *trans* NMA isomers between N−H and H_2O. Figures 6.11 and 6.12, which are the PCFs of the methyl (CH_3) group with liquid water of *cis*- and *trans*-NMA, respectively, also appear almost the same and a bit tighter than usual. The most important observation is that the methyl groups do not show any preferable interactions with either the H or O atoms of H_2O. This is a general chemical observation, which is that methyl groups are usually hydrophobic.

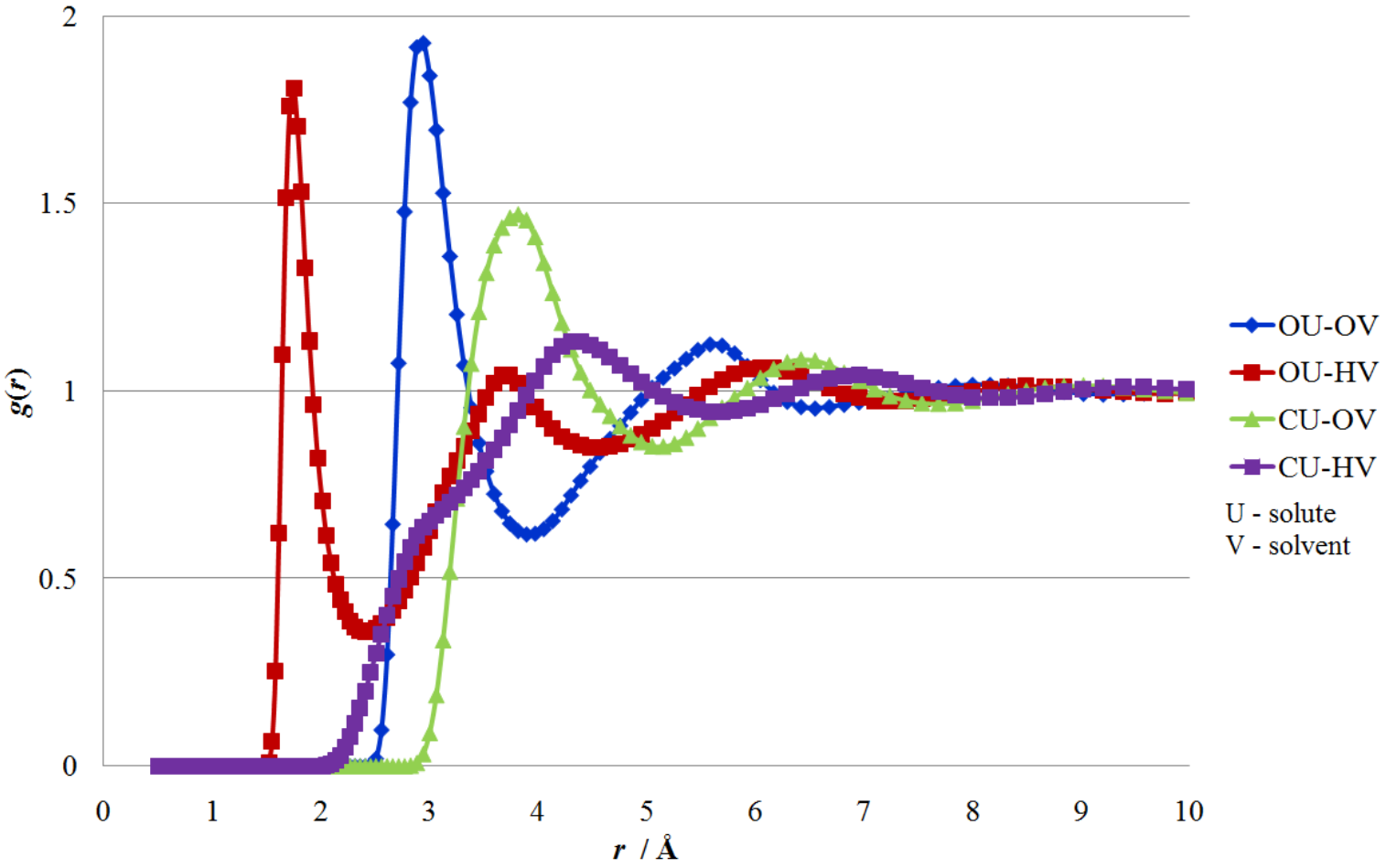


Figure 6.7

PCFs of the carbonyl group of *cis*-NMA with CCSD/6-31G** in water

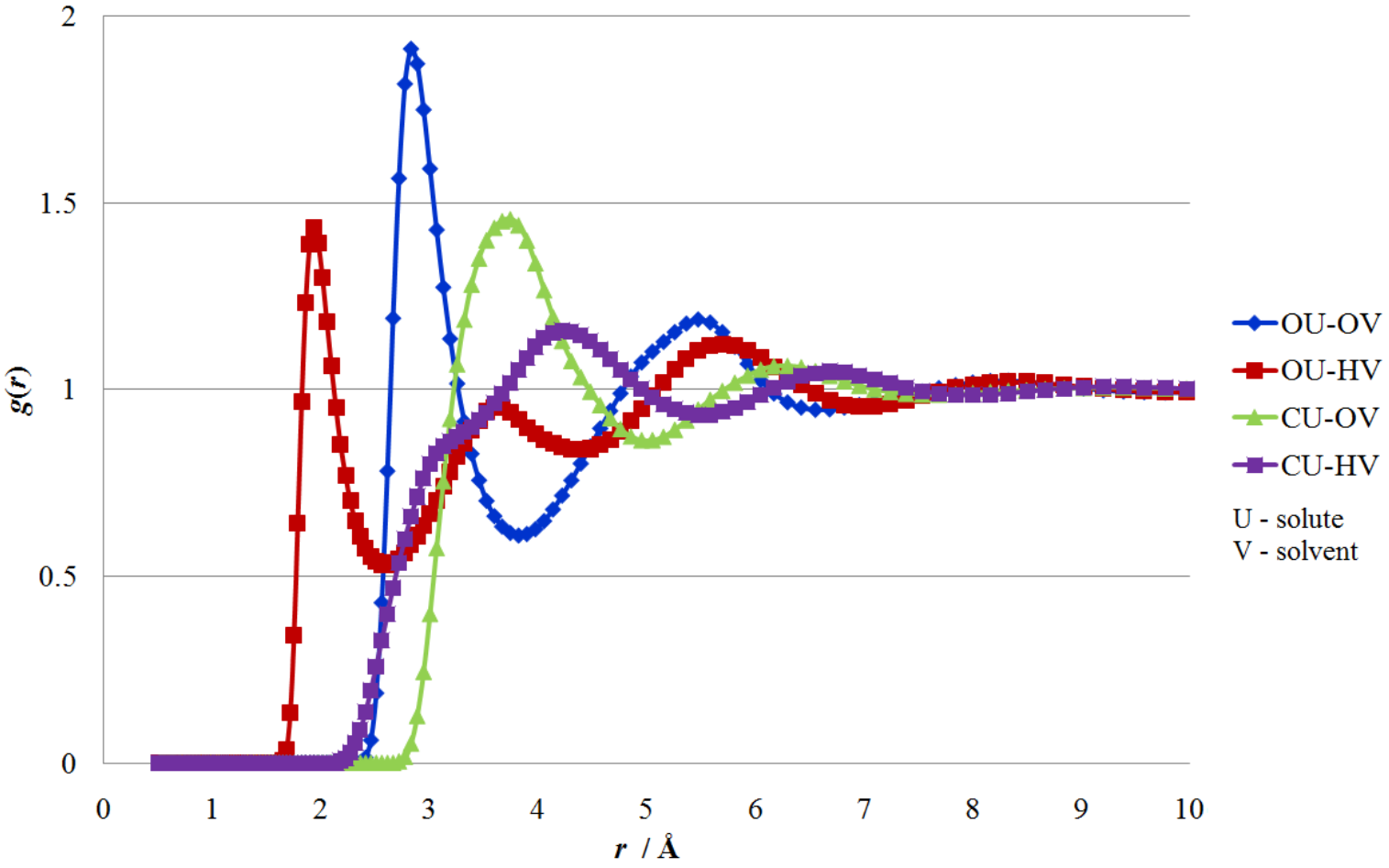


Figure 6.8

PCFs of the carbonyl group of *trans*-NMA with CCSD/6-31G** in water

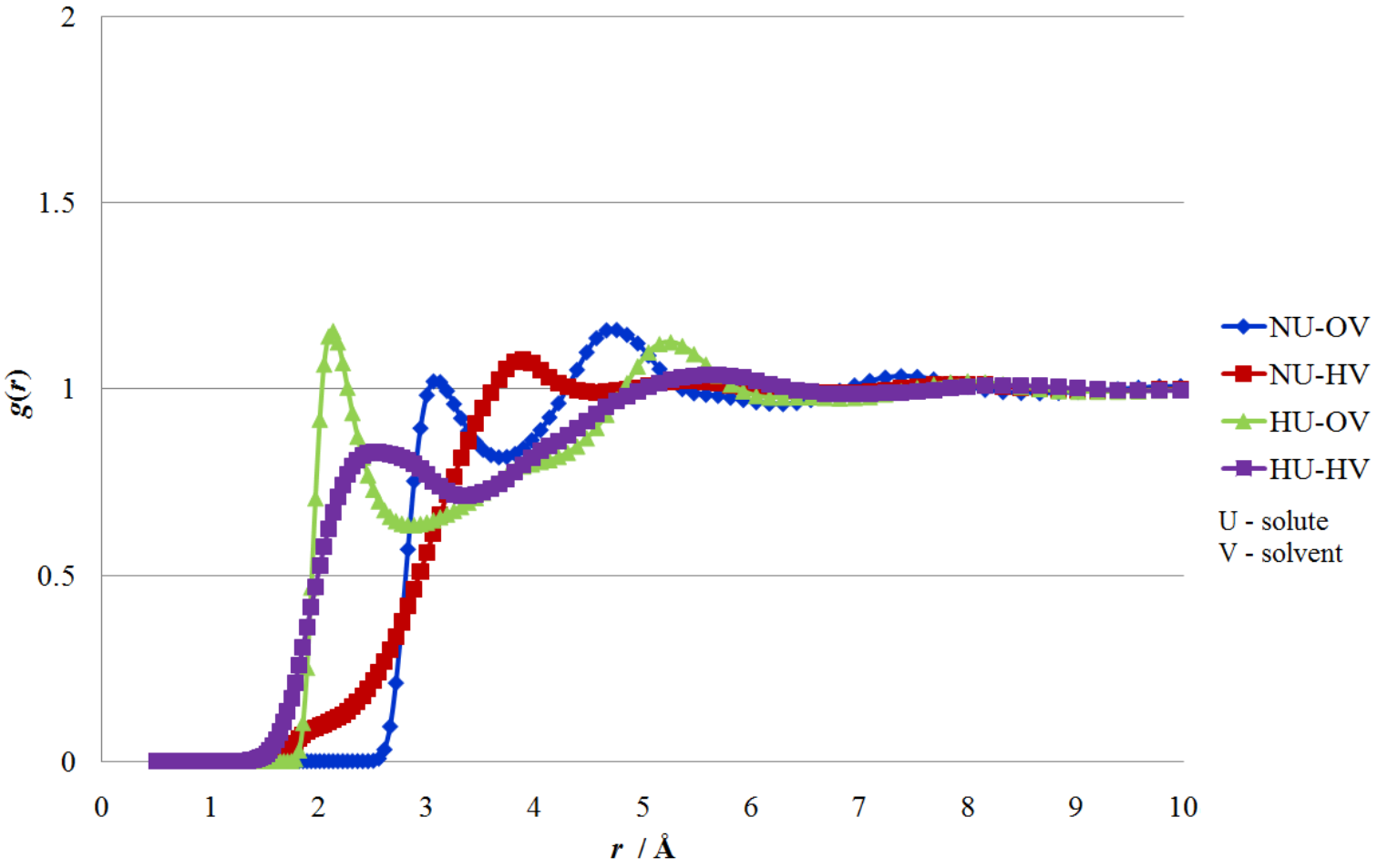


Figure 6.9

PCFs of the amide group of *cis*-NMA with CCSD/6-31G** in water

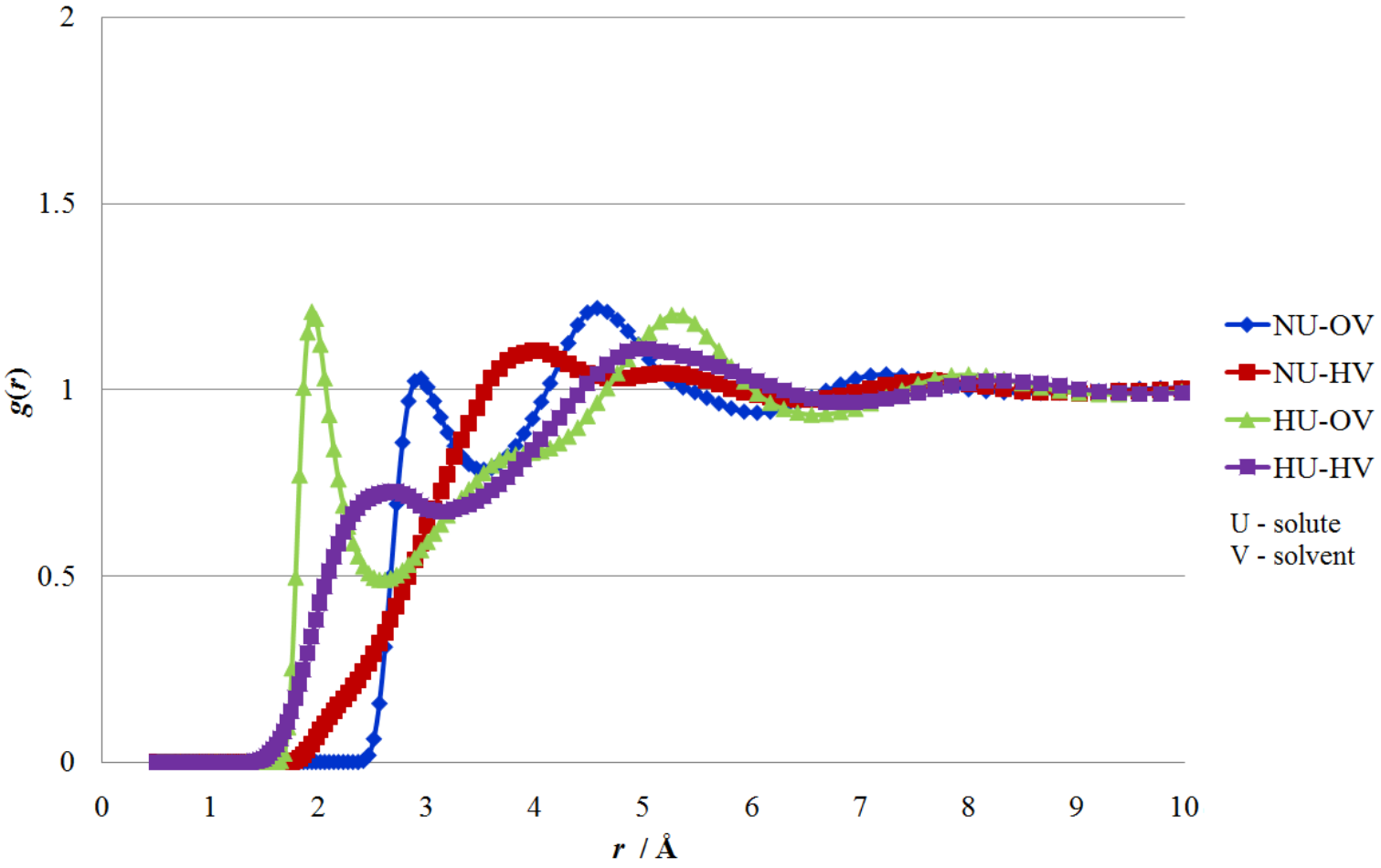


Figure 6.10

PCFs of the amide group of *trans*-NMA with CCSD/6-31G** in water

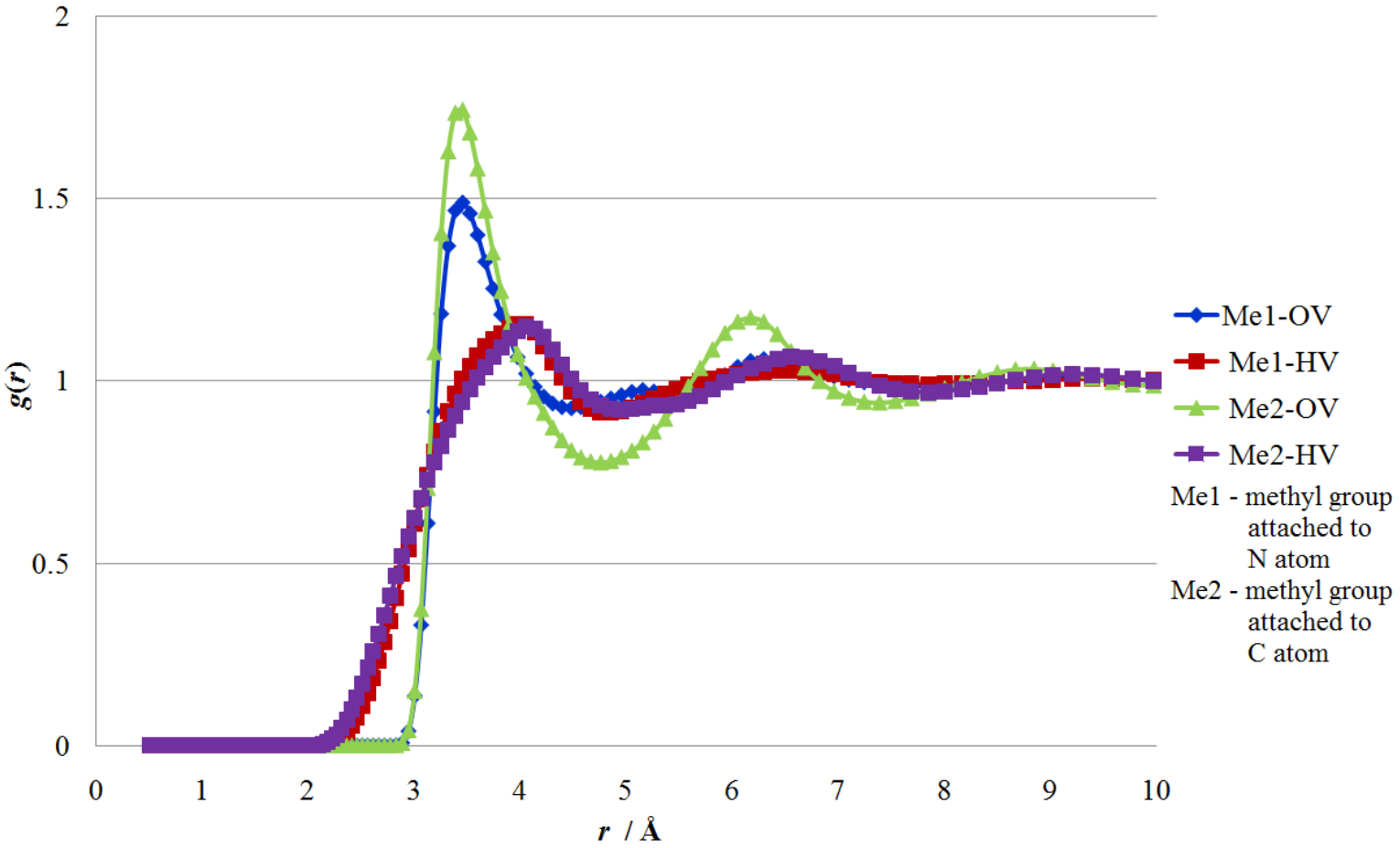


Figure 6.11

PCFs of the methyl groups of *cis*-NMA with CCSD/6-31G** in water

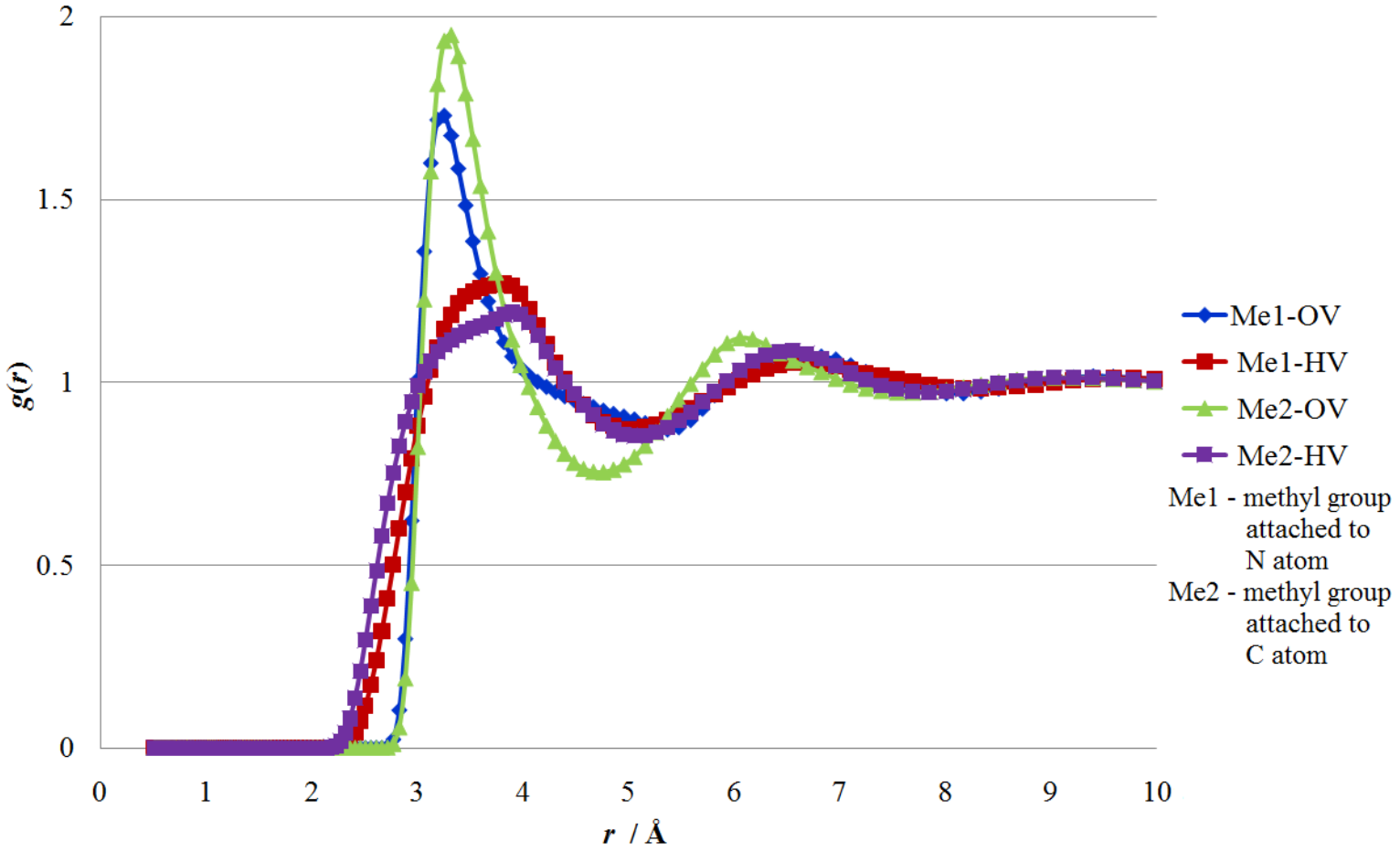


Figure 6.12

PCFs of the methyl groups of *trans*-NMA with CCSD/6-31G** in water

CHAPTER 7

CONCLUSIONS

Molecular electronic structure theory determines the structure of a molecule. However, changes in the electronic energy associated with a chemical process are comparable, in many cases, with those due to solvation in solution. This subtle balance between changes in electronic energy and in the solvation free energy sometimes causes drastic changes in the stability of a chemical species, as we have seen in several examples throughout this dissertation. This dissertation presents the theory for and the implementation of the solvent reaction field method (SCRF) and the reference interaction site model (RISM) for coupled-cluster singles and doubles (CCSD) calculations. It is worth noting here that, once the theory is implemented, solvent calculations of an enormous number of examples can be performed. A minimum number of examples has been chosen to prove the accuracy of the proposed theory and to demonstrate the capability of the theory.

The self-consistent solvent reaction field method at the CCSD level provides a simple quantitative description of solvation. Calculations on the conformational energy differences of 1,2-dichloroethane show that the equilibrium between the *cis* and *trans* rotamers is strongly affected by the nature of the solvent. The Onsager SCRF method, where $l = 1$, considerably underestimates the solvent contribution for the conformational energy difference. Hence, a higher order multipole moments expansion (at least $l = 12$) should be

considered for better accuracy. The CCSD-SCRF calculations of 1,2-dichloroethane are consistent with the experimental data. Therefore, introduction of electronic correlation to electronic structure calculations is necessary for accurate solvation free energies.

Gas phase calculations for the potential energy surface of the nitration of benzene support the two intermediate model for the nitration reaction. Increasing the dielectric strength of the solvent favors the σ-intermediate over the π-intermediate. At very high dielectric strength, the π-intermediate structure may no longer be a minimum. Hence, solvent effects change the mechanism for the nitration of benzene.

RISM-CC provides a promising combination of accurate molecular electronic structure theory with a physically justified model of a liquid solvent. Unlike the SCRF method, RISM provides a description of the local solvent structure by means of a microscopic reaction field, which is built on the radial distribution function of the solvent around the solute. One important solvent–solute interaction is hydrogen bonding. CC-RISM calculations provide better explanations for hydrogen bonding formation in the local solvent structure around the solute. RISM calculations show that the SCF solvation shells are smaller than solvation shells from CCSD. The introduction of electronic correlation to solvent calculation gives accurate radial distribution functions, partial atomic chargers, and solvation free energies.

REFERENCES

[1] Warshel, A. *Computer Modeling of Chemical Reactions in Enzymes and Solutions*; John Wiley & Sons Ltd: New York, 1997.

[2] Reichardt, C. *Solvents and Solvent Effects in Organic Chemistry*, 3rd ed.; WILEY-VCH Verlag GmbH & Co. KGaA.: Weinheim, Germany, 2003.

[3] Tomasi, J.; Persico, M. *Chem. Rev.* **1994**, *94*, 2027–2094.

[4] Tomasi, J. *Theor. Chem. Acc.* **2004**, *112*, 184–203.

[5] Tomasi, J.; Mennucci, B.; Cammi, R. *Chem. Rev.* **2005**, *105*, 2999–3094.

[6] Cramer, C. J.; Truhlar, D. G. *Chem. Rev.* **1999**, *99*, 2161–2200.

[7] Orozco, M.; Luque, F. J. *Chem. Rev.* **2000**, *100*, 4187–4226.

[8] Born, M.; Oppenheimer, J. R. *Annalen der Physik* **1927**, *84*, 457–484.

[9] Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Clarendon Press: Oxford, 2000.

[10] Alder, B. J.; Wainwright, T. E. *J. Chem. Phys.* **1957**, *27*, 1208–1209.

[11] Alder, B. J.; Wainwright, T. E. *J. Chem. Phys.* **1959**, *31*, 459–466.

[12] Alder, B. J.; Wainwright, T. E. *J. Chem. Phys.* **1960**, *33*, 1439–1451.

[13] Rahman, A. *Phys. Rev.* **1964**, *136*, A405–A411.

[14] Rahman, A.; Stillinger, F. H. *J. Chem. Phys.* **1971**, *55*, 3336–3359.

[15] Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. *J. Chem. Phys.* **1953**, *21*, 1087–1092.

[16] Wood, W. W.; Parker, F. R. *J. Phys. Chem.* **1957**, *27*, 720–733.

[17] Barker, J. A.; Watts, R. O. *Chem. Phys. Lett.* **1969**, *3*, 144–145.

[18] Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1986**, *7*, 718–730.

[19] Bash, P. A.; Field, M. J.; Karplus, M. *J. Am. Chem. Soc.* **1987**, *109*, 8092–8094.

[20] Åqvist, J.; Warshel, A. *Chem. Rev.* **1993**, *93*, 2523–2544.

[21] Gao, J. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; VCH: New York, 1995; Vol. 7, Chapter 7, pp 119–185.

[22] Eichinger, M.; Tavan, P.; Hutter, J.; Parrinello, M. *J. Chem. Phys.* **1999**, *110*, 10452–10467.

[23] Cramer, C. J.; Truhlar, D. G. In *Solvent Effects and Chemical Reactivity*; Tapia, O., Bertrán, J., Eds.; Kluwer Academic Publishers: Dordrecht, 1996; Chapter 1, pp 1–80.

[24] Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985**, *55*, 2471–2474.

[25] Remler, D. K.; Madden, P. A. *Mol. Phys.* **1990**, *70*, 921–966.

[26] Marx, D.; Hutter, J. In *Modern Methods and Algorithms of Quantum Chemistry*; Grotendorst, J., Ed., 2nd ed.; John von Neumann Institute für Computing: Jülich, 2000; Vol. 1, pp 301–449.

[27] Tuckerman, M. E. *J. Phys.: Condens. Matter* **2002**, *14*, R1297R1355.

[28] Iftimie, R.; Minary, P.; Tuckerman, M. E. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6654–6659.

[29] Alagona, G.; Pullman, A.; Scrocco, E.; Tomasi, J. *Int. J. Pept. Protein Res.* **1973**, *5*, 251–259.

[30] Pullman, A.; Pullman, B. *Q. Rev. Biophys.* **1975**, *7*, 505–560.

[31] Romesberg, F. E.; Collum, D. B. *J. Am. Chem. Soc.* **1994**, *116*, 9187–9197.

[32] Woon, D. E.; Dunning, T. H. J. *J. Am. Chem. Soc.* **1995**, *117*, 1090–1097.

[33] Born, M. *Z. Phys.* **1920**, *1*, 45.

[34] Kirkwood, J. G. *J. Chem. Phys.* **1934**, *2*, 767–781.

[35] Kirkwood, J. G. *J. Chem. Phys.* **1939**, *7*, 911–919.

[36] Onsager, L. *J. Am. Chem. Soc.* **1936**, *58*, 1486–1493.

[37] Rinaldi, D.; Rivail, J. L. *Theoret. Chim. Acta* **1973**, *32*, 57–70.

[38] Tapia, O.; Goscinski, O. *Mol. Phys.* **1975**, *29*, 1653.

[39] Mikkelsen, K. V.; Dalgaard, E.; Swanstroem, P. *J. Phys. Chem.* **1987**, *91*, 3081–3092.

[40] Mikkelsen, K. V.; Ågren, H.; Jensen, H. J. A.; Helgaker, T. *J. Chem. Phys.* **1988**, *89*, 3086–3095.

[41] Andersen, H. C.; Chandler, D. *J. Chem. Phys.* **1972**, *57*, 1918–1929.

[42] Chandler, D.; Andersen, H. C. *J. Chem. Phys.* **1972**, *57*, 1930–1937.

[43] Hirata, F.; Rossky, P. J. *Chem. Phys. Lett.* **1981**, *83*, 329–334.

[44] Hirata, F.; Pettitt, B. M.; Rossky, P. J. *J. Chem. Phys.* **1982**, *77*, 509–520.

[45] Hirata, F.; Rossky, P. J.; Pettitt, B. M. *J. Chem. Phys.* **1983**, *78*, 4133–4144.

[46] Wong, M. W.; Frisch, M. J.; Wiberg, K. B. *J. Am. Chem. Soc.* **1991**, *113*, 4776–4782.

[47] Ågren, H.; Llanos, C. M.; Mikkelsen, K. V. *Chem. Phys.* **1987**, *115*, 43–55.

[48] Rivail, J.-L.; Rinaldi, D. *Chem. Phys.* **1976**, *18*, 233–242.

[49] Rinaldi, D.; Ruiz-Lopez, M. F.; Rivail, J. L. *J. Chem. Phys.* **1983**, *78*, 834–838.

[50] Rinaldi, D.; Ruiz Lopez, M. F.; Rivail, J. L. *J. Chim. Phys. (Paris)* **1984**, *81*, 295–301.

[51] Rinaldi, D.; Bouchy, A.; Rivail, J.-L.; Dillet, V. *J. Chem. Phys.* **2004**, *120*, 2343–2350.

[52] Rinaldi, D.; Bouchy, A.; Rivail, J.-L. *Theor. Chem. Acc.* **2006**, *116*, 664–669.

[53] Ángyán, J. G. *Chem. Phys. Lett.* **1995**, *241*, 51–56.

[54] Nielsen, C. B.; Mikkelsen, K. V.; Sauer, S. P. A. *J. Chem. Phys.* **2001**, *114*, 7753–7760.

[55] Christiansen, O.; Mikkelsen, K. V. *J. Chem. Phys.* **1999**, *110*, 1365.

[56] Gauss, J.; Stanton, J. F.; Bartlett, R. J. *J. Chem. Phys.* **1991**, *95*, 2623–2638.

[57] Hansen, J.-P.; McDonald, I. R. *Theory of Simple Liquids*; Kluwer, Dordrecht, 2003.

[58] Hirata, F. *Molecular Theory of Solvation*; Kluwer Academic Publishers, Dortrecht, 2003.

[59] Pettitt, B. M.; Rossky, P. J. *J. Phys. Chem.* **1982**, *77*, 1451–1457.

[60] Sato, H.; Hirata, F. *J. Chem. Phys.* **1999**, *111*, 8545–8555.

[61] Kovalenko, A.; Hirata, F. *Chem. Phys. Lett.* **2001**, *349*, 496–502.

[62] Hirata, F.; Redfern, P.; Levy, R. M. *Int. J. Quant. Chem.* **1988**, *34*, 179–190.

[63] Roux, B.; Yu, H. A.; Karplus, M. *J. Phys. Chem.* **1990**, *94*, 4683–4688.

[64] Hirata, F.; Levy, R. M. *Chem. Phys. Lett.* **1987**, *136*, 267–273.

[65] Hirata, F.; Levy, R. M. *J. Phys. Chem.* **1989**, *93*, 479–484.

[66] Chong, S.-H.; Hirata, F. *J. Phys. Chem. B* **1997**, *101*, 3209–3220.

[67] Chiles, R. A.; Rossky, P. J. *J. Am. Chem. Soc.* **1984**, *106*, 6867–6868.

[68] Pettitt, B. M.; Rossky, P. J. *J. Phys. Chem.* **1986**, *84*, 5836–5844.

[69] Ten-no, S.; Hirata, F.; Kato, S. *Chem. Phys. Lett.* **1993**, *214*, 391–396.

[70] Ten-no, S.; Hirata, F.; Kato, S. *J. Chem. Phys.* **1994**, *100*, 7443–7453.

[71] Sato, H.; Hirata, F.; Kato, S. *J. Chem. Phys.* **1996**, *105*, 1546–1551.

[72] Naka, K.; Sato, H.; Morita, A.; Hirata, F.; Kato, S. *Theor. Chem. Acc.* **1999**, *102*, 165–169.

[73] Sato, H.; Hirata, F. *J. Am. Chem. Soc.* **1999**, *121*, 3460–3467.

[74] Harano, Y.; Sato, H.; Hirata, F. *J. Am. Chem. Soc.* **2000**, *122*, 2289–2293.

[75] Ishida, T.; Hirata, F.; Sato, H.; Kato, S. *J. Phys. Chem. B* **1998**, *102*, 2045–2050.

[76] Chong, S.-H.; Miura, S.-i.; Basu, G.; Hirata, F. *J. Phys. Chem.* **1995**, *99*, 10526–9.

[77] Hirata, F. *J. Chem. Phys.* **1992**, *96*, 4619–24.

[78] Chong, S.-H.; Hirata, F. *J. Chem. Phys.* **1999**, *111*, 3083–3094.

[79] Chong, S.-H.; Hirata, F. *J. Chem. Phys.* **1999**, *111*, 3095–3104.

[80] Chong, S.-H.; Hirata, F. *J. Chem. Phys.* **1999**, *111*, 3654–3667.

[81] Yoshida, K.; Yamaguchi, T.; Kovalenko, A.; Hirata, F. *J. Phys. Chem. B* **2002**, *106*, 5042–5049.

[82] Yamazaki, T.; Sato, H.; Hirata, F. *J. Chem. Phys.* **2001**, *115*, 8949–8957.

[83] Imai, T.; Kinoshita, M.; Hirata, F. *J. Chem. Phys.* **2000**, *112*, 9469–9478.

[84] Harano, Y.; Imai, T.; Kovalenko, A.; Kinoshita, M.; Hirata, F. *J. Chem. Phys.* **2001**, *114*, 9506–9511.

[85] Kinoshita, M.; Okamoto, Y.; Hirata, F. *J. Am. Chem. Soc.* **1998**, *120*, 1855–1863.

[86] Kinoshita, M.; Okamoto, Y.; Hirata, F. *J. Am. Chem. Soc.* **2000**, *122*, 2773–2779.

[87] Hirata, F. *Bull. Chem. Soc. Jpn.* **1998**, *71*, 1483–1499.

[88] Sato, H.; Kovalenko, A.; Hirata, F. *J. Chem. Phys.* **2000**, *112*, 9463–9468.

[89] Shavitt, I. In *Methods of Electronic Structure Theory*; Schaefer, H. F., Ed.; Plenum Press: New York, 1977; Vol. 3.

[90] Bartlett, R. J.; Stanton, J. F. Applications of Post-Hartree-Fock Methods: A Tutorial. In *Reviews of Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; VCH Publishers: New York, 1994; Vol. 5, Chapter 2, pp 65–169, and references therein.

[91] III, G. D. P.; Bartlett, R. J. *J. Chem. Phys.* **1982**, *76*, 1910–1918.

[92] Taylor, R. *Electropilic Aromatic Substitution*; Wiley: New York, 1995.

[93] Hughes, E. D.; Ingold, C. K.; Reed, R. I. *Nature* **1946**, *158*, 448.

[94] Ingold, C. K.; Hughes, E. D. *J. Chem. Soc.* **1950**, 2400.

[95] Wheland, G. W. *J. Am. Chem. Soc.* **1942**, *64*, 900–908.

[96] Olah, G. A.; Kuhn, S.; Flood, S. H. *J. Am. Chem. Soc.* **1962**, *83*, 451.

[97] Olah, G. A. *Acc. Chem. Rev.* **1971**, *4*, 240.

[98] Coombes, R. D.; Moodie, R. B.; Schofied, K. *J. Chem. Soc. B* **1968**, 800.

[99] Schofield, K. *Aromatic Nitration*; Cambridge University Press: Cambridge, 1980.

[100] Olah, G. A.; Malhotra, R.; Narang, S. C. *Nitration Methods and Mechanisms*; VCH: New York, 1989.

[101] Politzer, P.; Jayasuriya, K.; Sjoberg, P.; Laurence, P. R. *J. Am. Chem. Soc.* **1985**, *107*, 1174.

[102] Gleghorn, J. T.; Torossian, G. *J. Chem. Soc., Perkin Trans. II* **1987**, 1303.

[103] Szabo, K. J.; Hornfeldt, A.-B.; Gronowitz, S. J. *J. Am. Chem. Soc.* **1992**, *114*, 6827.

[104] Gwaltney, S. R.; Rosokha, S. V.; Head-Gordon, M.; Kochi, J. K. *J. Am. Chem. Soc.* **2003**, *125*, 3273–3283.

[105] Esteves, P.; de M.Carneiro, J.; Cardoso, S.; Barbosa, A.; Laali, K.; Rasul, G.; Prakash, G.; Olah, G. *J. Am. Chem. Soc.* **2003**, *125*, 4836–4849.

[106] Chen, L.; Xiao, H.; Xiao, J.; Gong, X. *J. Phys. Chem. A* **2003**, *107*, 11440–11444.

[107] de Queiroz, J. F.; de M. Carneiro, J. W.; Sabino, A. A.; Sparrapan, R.; Eberlin, M. N.; Esteves, P. M. *J. Org. Chem.* **2006**, *71*, 6192.

[108] Yin, F. M.Sc. thesis, Department of Chemistry, Mississippi State University: Mississippi State, Mississippi, 2004.

[109] Yin, F.; Thanthiriwatte, K. S.; Gwaltney, S. R. *J. Chem. Theory Comput.* Submitted.

[110] Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*; Dover Publications, Inc.: New York, 1996.

[111] Slater, J. C. *Phys. Rev.* **1928**, *32*, 339–348.

[112] Slater, J. C. *Phys. Rev.* **1951**, *81*, 385–390.

[113] Hartree, D. R. *Proc. Cambridge Phil. Soc.* **1927**, *24*, 89–110.

[114] Hartree, D. R. *Proc. Cambridge Phil. Soc.* **1927**, *24*, 112–132.

[115] Fock, V. *Zeitschrift fuer Physik* **1930**, *62*, 795–805.

[116] Roothaan, C. C. J. *Rev. Mod. Phys.* **1951**, *23*, 69–89.

[117] Jr, C. W. B.; Langhoff, S. R.; Taylor, P. R. Accurate Quantum Chemical Calculations. In *Advances in Chemical Physics*; Prigogine, I., Rice, S. A., Eds.; 2007; pp 103–161.

[118] Schaefer, H. F.; Thomas, J. R.; Yamaguchi, Y.; DeLeeuw, B. J.; Vacek, G. In *Modern Electronic Structure Theory*; Yarkony, D. R., Ed.; World Scientific: Singapore, 1995.

[119] Crawford, T. D.; Schaefer, H. F. An Introduction to Coupled Cluster Theory for Computational Chemists. In *Reviews of Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; VCH Publishers: New York, 1999; Vol. 14, Chapter 2, pp 33–136.

[120] Pople, J. A. In *Energy, Structure, and Reactivity*; Smith, D. W., McRae, W. B., Eds.; Theoretical models for chemistry; John Wiley: New York, 1973.

[121] Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618–622.

[122] Bartlett, R. J. *Annu. Rev. Phys. Chem.* **1981**, *32*, 359, and references therein.

[123] Löwdin, P.-O. *J. Math. Phys.* **1965**, *6*, 1341–1353.

[124] Gauss, J.; Stanton, J. F.; Bartlett, R. J. *J. Chem. Phys.* **1992**, *97*, 7825–7828.

[125] Handy, N. C.; Amos, R. D.; Gaw, J. F.; Rice, J. E.; Simandiras, E. D.; Lee, T. J.; Harrison, R. J.; Laidig, W. D.; Fitzgerald, G. B.; Bartlett, R. J. In *Geometrical Derivatives of Energy Surfaces and Molecular Properties*; Jørgensen, P., Simons, J., Eds.; Reidel Publishing Company: Dortrecht, 1986.

[126] Harrison, R. J.; Fitzgerald, G. B.; Laidig, W. D.; Barteltt, R. J. *Chem. Phys. Lett.* **1986**, *124*, 291–294.

[127] Cramer, C. J. *Essentials of Computational Chemistry, Theories and Models*, 2nd ed.; John Wiley & Sons Ltd: England, 2003.

[128] Coester, F. *Nucl. Phys* **1958**, *1*, 421.

[129] Coester, F.; Kömmel, H. *Nucl. Phys* **1960**, *17*, 477.

[130] Bartlett, R. J. In *Geometrical Derivatives of Energy Surfaces and Molecular Properties*; Jørgensen, P., Simons, J., Eds.; Reidel: Dordrecht, 1986; p 35 ff.

[131] Bartlett, R. J. Coupled-Cluster Theory: An Overview of Recent Developments. In *Modern Electronic Structure Theory, Part II*; Yarkony, D. R., Ed.; World Scientific Publishing Co.: Singapore, 1995; Chapter 16, pp 1047–1131, and references therein.

[132] Bartlett, R. J. *J. Phys. Chem.* **1989**, *93*, 1697–1708.

[133] Bartlett, R. J.; Musiał, M. *Rev. Mod. Phys.* **2007**, *79*, 291–352, and references therein.

[134] Wilcox, R. M. *J. Math. Phys.* **1967**, *8*, 962–982.

[135] Campbell, J. E. *Proc. London Math. Soc.* **1897**, *28*, 381–390.

[136] Campbell, J. E. *Proc. London Math. Soc.* **1898**, *29*, 14–32.

[137] Baker, H. F. *Proc. London Math. Soc.* **1902**, *34*, 347–360.

[138] Baker, H. F. *Proc. London Math. Soc.* **1903**, *35*, 333–374.

[139] Baker, H. F. *Proc. London Math. Soc. (Ser2)* **1905**, *3*, 24–47.

[140] Hausdorff, F. *Ber. Verh Sächs Akad. Wiss. Leipzig.* **1906**, *58*, 19–48.

[141] Salter, E. A.; Trucks, G. W.; Bartlett, R. J. *J. Chem. Phys.* **1989**, *90*, 1752–1766.

[142] Salter, E. A.; Trucks, G. W.; Fitzgerald, G.; Bartlett, R. J. *Chem. Phys. Lett.* **1987**, *141*, 61–70.

[143] Trucks, G. W.; Salter, E. A.; Noga, J.; Bartlett, R. J. *Chem. Phys. Lett.* **1988**, *150*, 37–44.

[144] Trucks, G. W.; Salter, E. A.; Sosa, C.; Bartlett, R. J. *Chem. Phys. Lett.* **1988**, *147*, 359–366.

[145] Böttcher, C. J. F. *Theory of Electric Polarization*; Elsevier, Amsterdam, 1973; Vol. 1.

[146] Uhlig, H. H. *J. Phys. Chem.* **1937**, *41*, 1215–1226.

[147] Sinanoglu, . *Chem. Phys. Lett.* **1967**, *1*, 283.

[148] Rao, R. V. G.; Subba Rao, V. V. *Z. Phys. Chem.* **357**, *221*, 1967.

[149] Sanchez, I. C. *J. Chem. Phys.* **1983**, *79*, 405.

[150] Gogonea, E., V.; Osawa *J. Mol. Struct. (Theochem)* **1994**, *311*, 305.

[151] Reiss, H.; Frisch, H. L.; Lebowitz, J. L. *J. Chem. Phys.* **1959**, *31*, 369–380.

[152] Reiss, H.; Frisch, H. L.; Helfand, E.; Lebowitz, J. L. *J. Chem. Phys.* **1960**, *32*, 119–124.

[153] Pierotti, R. A. *Chem. Rev.* **1976**, *76*, 717–726.

[154] Pierotti, R. A. *J. Phys. Chem.* **1965**, *69*, 281–288.

[155] Pierotti, R. A. *J. Phys. Chem.* **1963**, *67*, 1840–1845.

[156] Claverie, P. Diatomics to Biomolecules. In *Intermolecular Interactions*; Pullman, B., Ed.; Wiley: Chichester, 1978.

[157] Curutchet, C.; Bidon-Chanal, A.; Soteras, I.; Orozco, M.; Luque, F. J. *J. Phys. Chem. B* **2005**, *109*, 3565–3574.

[158] Carles Curutchet, M. O. F. J. L. *J.Comput. Chem.* **2001**, *22*, 1180–1193.

[159] Cramer, C. J.; Truhlar, D. G. *Science* **1992**, *256*, 213–217.

[160] Böttcher, C. J. F. *Theory of Electric Polarization*; Elsevier, Amsterdam, 1978; Vol. 2.

[161] Tapia, O.; Bertrán, J. *Solvent Effects and Chemical Reactivity*; Kluwer Academic Publishers: Dordrecht, 1996.

[162] Shao, Y.; Molnar, L. F.; Jung, Y.; Kussmann, J.; Ochsenfeld, C.; Brown, S. T.; Gilbert, A. T. B.; Slipchenko, L. V.; Levchenko, S. V.; O'Neill, D. P.; DiStasio, J., Robert A.; Lochan, R. C.; Wang, T.; Beran, G. J. O.; Besley, N. A.; Herbert, J. M.; Lin, C. Y.; Van Voorhis, T.; Chien, S. H.; Sodt, A.; Steele, R. P.; Rassolov, V. A.; Maslen, P. E.; Korambath, P. P.; Adamson, R. D.; Austin, B.; Baker, J.; Byrd, E. F. C.; Dachsel, H.; Doerksen, R. J.; Dreuw, A.; Dunietz, B. D.; Dutoi, A. D.; Furlani, T. R.; Gwaltney, S. R.; Heyden, A.; Hirata, S.; Hsu, C.-P.; Kedziora, G.; Khalliulin, R. Z.; Klunzinger, P.; Lee, A. M.; Lee, M. S.; Liang, W.; Lotan, I.; Nair, N.; Peters, B.; Proynov, E. I.; Pieniazek, P. A.; Rhee, Y. M.; Ritchie, J.; Rosta, E.; Sherrill, C. D.; Simmonett, A. C.; Subotnik, J. E.; Woodcock, I., H. Lee; Zhang, W.; Bell, A. T.; Chakraborty, A. K.; Chipman, D. M.; Keil, F. J.; Warshel, A.; Hehre, W. J.; Schaefer III, H. F.; Kong, J.; Krylov, A. I.; Gill, P. M. W.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2006**, *8*, 3172–3191.

[163] Abraham, R. J.; Bretschneider, E. In *Internal Rotational Barriers in Molecules*; Orville-Thomas, W. J., Ed.; Wiley-Interscience: New York, 1974; Chapter 13.

[164] E. Bright Wilson, J. The Problem of Barriers to Internal Rotation in Molecules. In *Advances in Chemical Physics*; Prigogine, I., Ed.; 2007; pp 367–393.

[165] Lowe, J. P. *Science* **1973**, *179*, 527–532.

[166] Lewis, G. J.; Whalley, E. *J. Chem. Phys.* **1978**, *68*, 1119–1127.

[167] Taniguchi, Y.; Takaya, H.; Wong, P. T. T.; Whalley, E. *J. Chem. Phys.* **1981**, *75*, 4815–4822.

[168] Takaya, H.; Taniguchi, Y.; Wong, P. T. T.; Whalley, E. *J. Chem. Phys.* **1981**, *75*, 4823–4828.

[169] Ikawa, S.; Whalley, E. *J. Chem. Phys.* **1984**, *81*, 1620–1625.

[170] Katō, M.; Taniguchi, Y. *J. Chem. Phys.* **1990**, *93*, 4345–4350.

[171] Stolov, A. A.; Remizov, A. B. *Spectrochimi. Acta, Part A* **1995**, *51*, 1919–1932.

[172] Herrebout, W. A.; van der Veken, B. J. *J. Phys. Chem.* **1996**, *100*, 9671–9677.

[173] Kato, M.; Abe, I. *J. Chem. Phys.* **1999**, *110*, 11982.

[174] Dixon, D. A.; Matsuzawa, N.; Walker, S. C. *J. Phys. Chem.* **1992**, *96*, 1074010746.

[175] Lawrence, R. P.; Hsu, C. S.; David, C. *J. Chem. Phys.* **1978**, *68*, 4202–4212.

[176] Jorgensen, W. L. *J. Phys. Chem.* **1983**, *87*, 5304–5314.

[177] Tanaka, C.; Tanaka, J.; Hirao, K. *J. Mol. Struct. THEOCHEM* **1986**, *146*, 309–327.

[178] Millot, C.; Rivail, J. L. *Mol. Phys.* **1992**, *77*, 157 – 175.

[179] Depaepe, J. M.; Ryckaert, J. P. *Chem. Phys. Lett.* **1995**, *245*, 653 – 659.

[180] Wiberg, K. B.; Murcko, M. A. *J. Phys. Chem.* **1987**, *91*, 36163620.

[181] Wiberg, K. B.; Keith, T. A.; Frisch, M. J.; Murcko, M. *J. Phys. Chem.* **1995**, *99*, 9072–9079.

[182] Scarsi, M.; Apostolakis, J.; Caflisch, A. *J. Phys. Chem. B* **1998**, *102*, 3637–3641.

[183] Vilaseca, E. *Mol. Phy.* **1999**, *97*, 667 – 676.

[184] Cappelli, C.; Corni, S.; Tomasi, J. *J. Phys. Chem. A* **2001**, *105*, 10807–10815.

[185] Melendez-Pagan, Y.; Taylor, B.; Ben-Amotz, D. *J. Phys. Chem. B* **2001**, *105*, 520–526.

[186] Jedlovszky, P.; Árpád Vincze,; Horvai, G. *J. Chem. Phys.* **2002**, *117*, 2271–2280.

[187] Mierts, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* **1981**, *55*, 117–129.

[188] Tomasi, J.; Bonaccorsi, R.; Cammi, R.; del Valle, F. J. O. *J. Mol. Struc.* **1991**, *234*, 401–424.

[189] Katritzky, A. R.; Zerner, M. C.; Karelson, M. M. *J. Am. Chem. Soc.* **1986**, *108*, 7213–7214.

[190] Wong, M. W.; Wiberg, K. B.; Frisch, M. J. *J. Comput. Chem.* **1995**, *16*, 385.

[191] Luo, Y.; Agren, H.; Mikkelsen, K. V. *Chem. Phys. Lett.* **1997**, *275*, 145.

[192] Ditchfield, R.; Hehre, W. J.; Pople, J. A. *J. Chem. Phys.* **1971**, *54*, 724–728.

[193] Hariharan, P. C.; Pople, J. A. *Mol. Phys.* **1974**, *27*, 209 – 214.

[194] Hariharan, P. C.; Pople, J. A. *Theoret. Chim. Acta* **1973**, *28*, 213 – 222.

[195] Baeck, K. K.; Watts, J. D.; Bartlett, R. J. *J. Chem. Phys.* **1997**, *107*, 3853–3863.

[196] Carey, F. A.; Sundberg, R. J. *Advanced Organic Chemistry, Part A: Structure and Mechanisms*, 5th ed.; Springer: New York, 2007.

[197] Ingold, C. K. *Structure and Mechanism in Organic Chemistry*, 2nd ed.; Cornell University Press: Ithaca, 1986.

[198] Eberson, L. *Electron-Transfer Reactions in Organic Chemistry*; Springer–Verlag: New York, 1987.

[199] Olah, G. A.; Kuhn, S. J.; Flood, S. H. *J. Am. Chem. Soc.* **1961**, *83*, 4571–4580.

[200] Schofield, K. *Aromatic Nitration*; Cambridge University Press: Cambridge, 1980; pp 44–54, 104–128.

[201] Attinã, M.; Cacace, F. *Gazz. Chim. Ital.* **1988**, *118*, 241–250.

[202] Attinã, M.; Cacace, F.; Ricci, A. *Tetrahedron* **1988**, *44*, 2015–2022.

[203] Attinã, M.; Cacace, F.; Yanez, M. *J. Am. Chem. Soc.* **1987**, *109*, 5092–5097.

[204] Bernardi, F.; Hehre, W. J. *J. Am. Chem. Soc.* **1973**, *95*, 3078–3080.

[205] Hehre, W. J.; Stewart, R. F.; Pople, J. A. *J. Chem. Phys.* **1969**, *51*, 2657–2664.

[206] Hehre, W. J.; Ditchfield, R.; Stewart, R. F.; Pople, J. A. *J. Chem. Phys.* **1970**, *52*, 2769–2773.

[207] Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257–2261.

[208] Binkley, J. S.; Pople, J. A.; Hehre, W. J. *J. Am. Chem. Soc.* **1980**, *102*, 939–947.

[209] Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650–654.

[210] Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; DeFrees, D. J.; Pople, J. A. *J. Chem. Phys.* **1982**, *77*, 3654–3665.

[211] Raghavachari, K.; Trucks, G. W.; Head-Gordon, M.; Pople, J. A. *Chem. Phys. Lett.* **1989**, *157*, 479–483.

[212] Stanton, J. F.; Gauss, J.; Watts, J. D.; Nooijen, M.; Oliphant, N.; Perera, S. A.; Szalay, P. G.; Lauderdale, W. J.; Gwaltney, S. R.; Beck, S.; Balkova, A.; Bernholdt, D. E.; Baeck, K. -K.; Sekino, H.; Rozyczko, P.; Huber, C.; Bartlett, R. J. ACES II program is a product of the Quantum Theory Project, University of Florida. Integral packages included are VMOL (Almlof and Taylor), VPROPS (Taylor) and a modified version of the ABACUS integral derivative package (Helgaker, Jensen, Olsen, Jorgensen and Taylor). E-mail: aces2@qtp.ufl.edu.

[213] Taketsugu, T.; Gordon, M. S. *J. Chem. Phys.* **1995**, *103*, 10042–10049.

[214] Singer, S. J.; Chandler, D. *Mol. Phys.* **1985**, *55*, 621–625.

[215] Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187–217.

[216] MacKerell, A. D.; Bashford, D.; Bellott,; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.

[217] Jorgensen, W. L. *J. Am. Chem. Soc.* **1981**, *103*, 335–340.

[218] Berendsen, H. J. C.; Postma, J. P. M.; vanGunsten, W. F.; Hermans, J. In *Intermolecular Forces*; Pullman, B., Ed.; Reidel Publishing Company: Dortrecht, 1981.

[219] Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269–6271.

[220] Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657–1666.

[221] Yu, H. A.; Roux, B.; Karplus, M. *J. Chem. Phys.* **1990**, *92*, 5020–5033.

[222] Yu, H. A.; Karplus, M. *J. Chem. Phys.* **1988**, *89*, 2366–2379.

[223] Lebedev, V. I. *Zh. Vychisl. Mat. Mat. Fiz.* **1975**, *15*, 48–54.

[224] Lebedev, V. I. *Zh. Vychisl. Mat. Mat. Fiz.* **1976**, *16*, 293–306.

[225] Lebedev, V. I. *Sibirsk. Mat. Zh.* **1977**, *18*, 132–142.

[226] Lebedev, V. I.; Skorokhodov, A. L. *Russian Acad. Sci. Dokl. Math.* **1992**, *45*, 587–592.

[227] Lebedev, V. I. *Russian Acad. Sci. Dokl. Math.* **1995**, *50*, 283–286.

[228] Lebedev, V. I.; Laikov, D. N. *Dokl. Math.* **1999**, *59*, 477–481.

[229] T. H. Dunning, J. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

[230] Kendall, R. A.; Thom H. Dunning, J.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796–6806.

[231] Clough, S. A.; Beers, Y.; Klein, G. P.; Rothman, L. S. *J. Chem. Phys.* **1973**, *59*, 2254–2259.

[232] Lovas, F. J. *J. Phys. Chem. Ref. Data* **1978**, *7*, 1445–1750.

[233] Xantheas, S. S.; Thom H. Dunning, J. *J. Chem. Phys.* **1993**, *99*, 8774–8792.

[234] Xantheas, S. S. *J. Chem. Phys.* **1994**, *100*, 7523–7534.

[235] Xantheas, S. S. *J. Chem. Phys.* **1995**, *102*, 4505–4517.

[236] Gregory, J. K.; Clary, D. C.; Liu, K.; Brown, M. G.; Saykally, R. J. *Science* **1997**, *275*, 814–817.

[237] M. Yang, P. S. C. V. A. *Int. J. Quant. Chem.* **2005**, *101*, 535–542.

[238] Moro, R.; Rabinovitch, R.; Xia, C.; Kresin, V. V. *Phys. Rev. Lett.* **2006**, *97*, 123401.

[239] Bukowski, R.; Szalewicz, K.; Groenenboom, G. C.; van der Avoird, A. *Science* **2007**, *315*, 1249–1252.

[240] Rai, D.; Kulkarni, A. D.; Gejji, S. P.; Pathak, R. K. *J. Chem. Phys.* **2008**, *128*, 034310.

[241] Lodi, L.; Tolchenov, R. N.; Tennyson, J.; Lynas-Gray, A. E.; Shirin, S. V.; Zobov, N. F.; Polyansky, O. L.; Császár, A. G.; van Stralen, J. N. P.; Visscher, L. *J. Chem. Phys.* **2008**, *128*, 044304.

[242] Badyal, Y. S.; Saboungi, M.-L.; Price, D. L.; Shastri, S. D.; Haeffner, D. R.; Soper, A. K. *J. Chem. Phys.* **2000**, *112*, 9206–9208.

[243] Coulson, C. A.; Eisenberg, D. *Proc. Roy. Soc. London A* **1966**, *291*, 445–453.

[244] Batista, E. R.; Xantheas, S. S.; Jónsson, H. *J. Chem. Phys.* **1998**, *109*, 4546–4551.

[245] Tunón, I.; Martins-Costa, M. T. C.; Millot, C.; Rivail, M. F. R.-L. J. L. *J. Comp. Chem.* **1996**, *17*, 19–29.

[246] Bernardo, D. N.; Ding, Y.; Krogh-Jespersen, K.; Levy, R. M. *J. Phys. Chem.* **1994**, *98*, 4180–4187.

[247] Gao, J.; Xia, X. *Science* **1992**, *258*, 631–635.

[248] Silvestrelli, P. L.; Parrinello, M. *Phys. Rev. Lett.* **1999**, *82*, 3308–3311.

[249] Dang, L. X.; Chang, T.-M. *J. Chem. Phys.* **1997**, *106*, 8149–8159.

[250] Chalmet, S.; Ruiz-López, M. F. *J. Chem. Phys.* **2001**, *115*, 5220–5227.

[251] Rocha, W. R.; Coutinho, K.; de Almeida, W. B.; Canuto, S. *Chem. Phys. Lett.* **2001**, *335*, 127–133.

[252] Tu, Y.; Laaksonen, A. *Chem. Phys. Lett.* **2000**, *329*, 283–288.

[253] Jansen, G.; Colonna, F.; Ángyán, J. G. *Int. J. Quant. Chem.* **1996**, *58*, 251–265.

[254] Laasonen, K.; Sprik, M.; Parrinello, M.; Car, R. *J. Chem. Phys.* **1993**, *99*, 9080–9089.

[255] Millot, C.; Costa Cabral, B. J. *Chemical Physics Letters* **2008**, *460*, 466–469.

[256] Sprik, M. *J. Chem. Phys.* **1991**, *95*, 6762–6769.

[257] Brancato, G.; Rega, N.; Barone, V. *J. Chem. Phys.* **2008**, *128*, 144501.

[258] Poulsen, T. D.; Ogilby, P. R.; Mikkelsen, K. V. *J. Chem. Phys.* **2002**, *116*, 3730–3738.

[259] Jensen, L.; van Duijnen, P. T.; Snijders, J. G. *J. Chem. Phys.* **2003**, *118*, 514–521.

[260] M. Yang, P.; Alsenoy, S. C. V. *Int. J. Quant. Chem.* **2005**, *101*, 535–542.

[261] Tu, Y.; Laaksonen, A. *J. Chem. Phys.* **1999**, *111*, 7519–7525.

[262] Watanabe, K.; Klein, M. L. *Chem. Phys.* **1989**, *131*, 157–167.

[263] Röntgen, W. C. *Sitzungsberichte der Würzburger Physikalis chmedicinis chen Gesellschaft*, 1895, Also in *Annalen der Physik und Chemie*, **1898**, *64*, 1; Engl. transl. *Nature*, **1896**, *53*, 274.

[264] Röntgen, W. C. *Sitzungsberichte der Würzburger Physikalis chmedicinis chen Gesellschaft*, 1896, Also in *Annalen der Physik und Chemie*, **1898**, *64*, 12;.

[265] Röntgen, W. C. *Sitzungsber. Preuss. Akad. Wiss. (Berlin)*, 1987, Also in *Annalen der Physik und Chemie*, **1898**, *64*, 18;.

[266] Thiessen, W. E.; Narten, A. H. *J. Chem. Phys.* **1982**, *77*, 2656–2662.

[267] Soper, A. K.; Phillips, M. G. *Chemical Physics* **1986**, *107*, 47–60.

[268] Nieto-Draghi, C.; Avalos, J. B.; Rousseau, B. *J. Chem. Phys.* **2003**, *118*, 7954–7964.

[269] Sorenson, J. M.; Hura, G.; Glaeser, R. M.; Head-Gordon, T. *J. Chem. Phys.* **2000**, *113*, 9149–9161.

[270] Jedlovszky, P.; Vallauri, R. *J. Chem. Phys.* **2001**, *115*, 3750–3762.

[271] Izvekov, S.; Voth, G. A. *J. Chem. Phys.* **2002**, *116*, 10372–10376.

[272] Krack, M.; Gambirasio, A.; Parrinello, M. *J. Chem. Phys.* **2002**, *117*, 9409–9412.

[273] Eggert, J. H.; Weck, G.; ; Loubeyre, P. *J. Phys. Condens. Mat.* **2002**, *14*, 1138511394.

[274] Netzloff, H. M.; Gordon, M. S. *J. Chem. Phys.* **2004**, *121*, 2711–2714.

[275] Leung, K.; Rempe, S. B. *Phys. Chem. Chem. Phys.* **2006**, *8*, 2153–2162.

[276] Teplukhin, A. *J.Struc. Chem.* **2008**, *49*, 270–277.

[277] Lu, H.; Wang, Y.; Wu, Y.; Yang, P.; Li, L.; Li, S. *J. Chem. Phys.* **2008**, *129*, 124512–124517.

[278] Mitsuo, K. *Bull. Chem. Soc. Jpn* **1973**, *46*, 384–387.

[279] Koichi Itoh, T. S. *Biopolymers* **1967**, *5*, 921–930.

[280] Fillaux, F.; Baron, M. H. *Chem. Phys.* **1981**, *62*, 275–285.

[281] Trabelsi, S.; Bahri, M.; Nasr, S. *J. Chem. Phys.* **2005**, *122*, 024502.

[282] Kearley, G. J.; Johnson, M. R.; Plazanet, M.; Suard, E. *J. Chem. Phys.* **2001**, *115*, 2614–2620.

[283] Dixon, D. A.; Dobbs, K. D.; Valentini, J. J. *J. Phys. Chem.* **1994**, *98*, 13435–13439.

[284] Han, W.-G.; Suhai, S. *J. Phys. Chem.* **1996**, *100*, 3942–3949.

[285] Mirkin, N. G.; Krimm, S. *J. Am. Chem. Soc.* **1991**, *113*, 9742–9747.

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

[287] Tannor, D. J.; Marten, B.; Murphy, R.; Friesner, R. A.; Sitkoff, D.; Nicholls, A.; Honig, B.; Ringnalda, M.; Goddard, W. A. *J. Am. Chem. Soc.* **1994**, *116*, 11875–11882.

[288] Langley, C. H.; Allinger, N. L. *J. Phys. Chem. A* **2003**, *107*, 5208–5216.

[289] Guo, H.; Karplus, M. *J. Phys. Chem.* **1992**, *96*, 7273–7287.

[290] Guo, H.; Karplus, M. *J. Phys. Chem.* **1994**, *98*, 7104–7105.

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

[292] Perera, S. A.; Bartlett, R. J. *Magnet. Res. Chem.* **2001**, *39*, S183–S189.

[293] Yang, Z.-Z.; Qian, P. *J. Chem. Phys.* **2006**, *125*, 064311.

[294] Hirst, J. D.; Hirst, D. M.; Brooks, C. L. *J. Phys. Chem. A* **1997**, *101*, 4821–4827.

[295] Serrano-Andrés, L.; Fülscher, M. P. *J. Am. Chem. Soc.* **1996**, *118*, 12190–12199.

[296] Besley, N. A.; Hirst, J. D. *J. Phys. Chem. A* **1998**, *102*, 10791–10797.

APPENDIX

PARTIAL ATOMIC CHARGES AND SOLVATED FOCK MATRIX

As mentioned in Chapter 6, partial atomic charges are determined so as to reproduce the electrostatic potential around the solute with a least squares fitting procedure. In this appendix the detailed procedure is described following the paper by Ten-no et al. [70].

The electrostatic potential at a random point ($\mathbf{r}$) near a molecule can be written as

$$U(\mathbf{r}) = U_N(\mathbf{r}) + U_e(\mathbf{r}), \tag{A.1}$$

and

$$U_N(\mathbf{r}) = \sum_{A=1}^{N} \frac{Z_A}{\mathbf{r} - \mathbf{R}_A}, \tag{A.2}$$

$$U_e(\mathbf{r}) = -Tr\left(\mathbf{PA}(\mathbf{r})\right), \tag{A.3}$$

where $U_N(\mathbf{r})$ is the potential from the nuclei, and $U_e(\mathbf{r})$ is the potential from the electrons. $\mathbf{P}$ is the one particle density matrix, as described in Eq. (2.52),

$$P_{\mu\nu} = 2\sum_{i}^{N/2} C_{\mu i} C^*_{\nu i}, \tag{A.4}$$

where N is number of electrons, and C is the MO coefficients. $\mathbf{A}(\mathbf{r})$ stands for the three-center one-electron integrals,

$$\left(\mathbf{A}(\mathbf{r})\right)_{\mu\nu} = \int \frac{\chi^*_\mu(\mathbf{r}')\chi_\nu(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \mathrm{d}\mathbf{r}', \tag{A.5}$$

where $\chi(\mathbf{r}')$ is a set of basis function. For practical purposes the two contributions are independently approximated by the field at the grid points. The filed is induced by a set of partial charges assigned $\left(\left\{q_\alpha^{(N)}\right\}, \left\{q_\alpha^{(e)}\right\}\right)$ to the same interaction sites,

$$\mathscr{U}_N(\mathbf{r}) = \sum_{\alpha=1}^{n} \frac{q_\alpha^{(N)}}{|\mathbf{r} - \mathbf{R}_\alpha|}, \tag{A.6}$$

$$\mathscr{U}_e(\mathbf{r}) = \sum_{\alpha=1}^{n} \frac{q_\alpha^{(e)}}{|\mathbf{r} - \mathbf{R}_\alpha|}, \tag{A.7}$$

where n is the number of interaction sites. $\left(\left\{q_\alpha^{(N)}\right\}, \left\{q_\alpha^{(e)}\right\}\right)$ are determined with the standard procedure of minimizing target functions subject to the constraints preserving the correct total number of electrons.

$$\frac{\partial}{\partial q_i^{(N)}}\left\{\sum_{k=1}^{l} w_k \left[U_N(\mathbf{r}_k) - \mathscr{U}_N(\mathbf{r}_k)\right]^2 + 2\lambda_N \left(\sum_{j=1}^{n} q_j^{(N)} - (\text{constant})\right)\right\} = 0, \quad \text{(A.8)}$$

$$\frac{\partial}{\partial q_i^{(e)}}\left\{\sum_{k=1}^{l} w_k \left[U_e(\mathbf{r}_k) - \mathscr{U}_e(\mathbf{r}_k)\right]^2 + 2\lambda_e \left(\sum_{j=1}^{n} q_j^{(e)} - (\text{constant})\right)\right\} = 0, \quad \text{(A.9)}$$

where l is the number of grid points, w_k are the weight coefficients, and λ are Lagrange multipliers. For the nuclear part $q_\alpha^{(N)}$ are solved to be,

$$\mathbf{q}^{(N)} = \begin{pmatrix} q_1^{(N)} \\ q_2^{(N)} \\ \vdots \\ q_n^{(N)} \end{pmatrix} = \mathbf{a}^{-1}\mathbf{a}'\mathbf{Z} - \lambda_N \mathbf{a}^{-1}\mathbf{1}, \quad \text{(A.10)}$$

where $\mathbf{1}$ is a column vector whose elements are all one, and $\mathbf{a}'$ and $\mathbf{a}$ are $n \times N$ and $n \times n$ matrices given by

$$(\mathbf{a}')_{Ai} = \sum_{k=1}^{l} \frac{w_k}{r_{kA} r_{ki}}, \quad \text{(A.11)}$$

$$(\mathbf{a})_{ij} = \sum_{k=1}^{l} \frac{w_k}{r_{ki} r_{kj}}. \quad \text{(A.12)}$$

By using the total charge of the nucleus, $N_N = \mathbf{1}^t \cdot \mathbf{Z}$, the Lagrange multiplier λ_N can be wrttien as

$$\lambda_N = \frac{\mathbf{1}^t \mathbf{a}^{-1}\mathbf{a}'\mathbf{Z} - N_N}{\mathbf{1}^t \mathbf{a}^{-1}\mathbf{1}}. \quad \text{(A.13)}$$

Now, $\mathbf{q}^{(N)}$ is

$$\mathbf{q}^{(N)} = \mathbf{a}^{-1}\mathbf{a}'\mathbf{Z} - \frac{\mathbf{1}^t \mathbf{a}^{-1}\mathbf{a}'\mathbf{Z} - N_N}{\mathbf{1}^t \mathbf{a}^{-1}\mathbf{1}} \mathbf{a}^{-1}\mathbf{1}, \quad \text{(A.14)}$$

When the set of interaction sites is identical to that of the nuclei, $\mathbf{a} = \mathbf{a}'$, and Eq. (A.14) becomes, $\mathbf{q}^{(N)} = \mathbf{Z}$.

The electronic contribution can be estimated from

$$\mathbf{q}^{(e)} = -\mathbf{a}^{-1} Tr\left(\mathbf{PB}\right) - \lambda_e \mathbf{a}^{-1}\mathbf{1}, \tag{A.15}$$

and

$$\mathbf{q}^{(e)} = -\mathbf{a}^{-1} Tr\left(\mathbf{PB}\right) - \frac{\mathbf{1}^t\mathbf{a}^{-1} Tr\left(\mathbf{PB}\right) - N_e}{\mathbf{1}^t\mathbf{a}^{-1}\mathbf{1}}\mathbf{a}^{-1}\mathbf{1}, \tag{A.16}$$

where $\mathbf{B}$ is a column supermatrix having the elements

$$\left(\mathbf{B}\right)_{\mu,\nu,i} = \sum_{k=1}^{l} \frac{w_k}{r_{ki}} A_{\mu\nu}\left(\mathbf{R}_k\right) \tag{A.17}$$

N_e is the total numbers of electrons. We can rearrange $\mathbf{q}^{(e)}$ as

$$\mathbf{q}^{(e)} = Tr\left\{\mathbf{PD}\right\} \tag{A.18}$$

where,

$$\mathbf{D} = \left[\left(\mathbf{a}^{-1}\mathbf{B} - \frac{\mathbf{a}^{-1}\mathbf{1}}{\mathbf{1}^t\mathbf{a}^{-1}\mathbf{1}}\left(\mathbf{1}^t\mathbf{a}^{-1}\mathbf{B} - \mathbf{S}\right)\right]\right(\tag{A.19}$$

where $\mathbf{S}$ is the overlap integrals between basis functions. Now we can write the partial atomic charges as the following equation:

$$\mathbf{q} = \mathbf{q}^{(N)} + \mathbf{q}^{(e)}. \tag{A.20}$$

The solvated Fock matrix is the sum of the isolated Fock matrix and an additional term coming from the solvent contribution

$$F_{\mu\nu} = F^0_{\mu\nu} + F^{sol}_{\mu\nu}. \tag{A.21}$$

The required integrals for the solvent contribution to the Fock matrix are already discussed in Section 7 and can be expressed as

$$F_{\mu\nu}^{sol} = -\mathbf{V}^t \cdot \left[\left(\mathbf{a}^{-1}\mathbf{B} + \frac{\mathbf{a}^{-1}\mathbf{1}}{\mathbf{1}^t\mathbf{a}^{-1}\mathbf{1}}\left(\mathbf{1}^t\mathbf{a}^{-1}\mathbf{B} - \mathbf{S}\right)\right]\right(\tag{A.22}$$

where $\mathbf{V}$ is the electrostatic potential from Eq. (6.14).