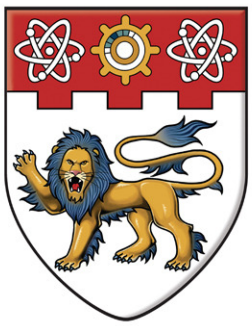


NANYANG TECHNOLOGICAL UNIVERSITY

HIERARCHICAL APPROACH TO STUDY WATER CLUSTER IN FIRST-PRINCIPLES CALCULATIONS

by

Nguyen Quoc Chinh

A thesis submitted in partial fulfillment for the degree of Doctor of Philosophy

in the

Division of Physics and Applied Physics
School of Physical and Mathematical Sciences

December 2010

NANYANG TECHNOLOGICAL UNIVERSITY

Abstract

Division of Physics and Applied Physics

School of Physical and Mathematical Sciences

Doctor of Philosophy

by Nguyen Quoc Chinh

First-principles calculation plays a central role in computational physics and chemistry in studying the properties of molecular systems. Nevertheless, its high computational cost limits the capability to carry out first-principles calculations for exploring the potential energy surface of molecular systems. In this work, a hierarchical method has been proposed to handle this issue in which we combine the semi-empirical methods and first-principles calculation in order to reduce the computational demand.

The proposed hierarchical method has been applied to study water clusters. The microscopic structures and properties of water clusters have attracted interest of theoreticians because they are extremely difficult to be revealed in experiments. In our studies, the potential energy landscape of protonated ($H^+(H_2O)_n$), deprotonated ($OH^-(H_2O)_n$) and neutral ($(H_2O)_n$) water clusters were thoroughly explored at quantum chemistry level. The distinct configurational isomers of different kinds of water clusters were uncovered and archived systematically by using a so-called archival memetic algorithm. The optimized geometries and relative stabilities of each system were analyzed afterward.

For studies on thermodynamics and structural transitions, harmonic superposition approximation has been used to investigate the thermodynamics at both

empirical and first-principles levels. The accuracy of harmonic superposition approximation has been tested by comparing the results with the ones predicted by Monte Carlo simulations. The finite temperature effects, the structural transitions as well as the thermodynamic profile of each kind of water for different sizes have been investigated systematically. For bridging the gap between theory and experiment as well as testing the accuracy of the simulation results, the vibrational spectra were simulated based on the calculated thermodynamic properties. The vibrational spectra were subsequently compared with recent experimental results. In addition, the effects of zero-point energy correction on the relative stabilities, thermodynamic properties and vibrational spectra of each system were discussed throughout the studies.

Last but not least, we reported the details on the development of the potential models for protonated hydrogen fluoride and deprotonated water clusters in order to extend the current work to other molecular systems and also to improve the efficiency of the hierarchical approach for our future studies.

Acknowledgements

First of all, I would like to acknowledge my supervisor, Asst. Prof. Kuo Jer-Lai. I owe my deepest gratitude for his encouragement, assistance, supervision throughout the years of my PhD project.

I am deeply grateful to thank Assoc. Prof. Ong Yew-Soon from School of Computer Sciences (SCE, NTU) who made a crucial supports in many ways. He offered invaluable assistance and helped me to collaborate with the people in computer science. His wide knowledge always inspires me and his suggestions contributed greatly to the growth in my research life.

It is a pleasure to thank all my lab buddies in Molecular Modeling lab in School of Physical and Mathematical Sciences (SPMS, NTU): Zhang Jingyun, Bing Dan, Fan Xiaofeng, Wu Hongyu, Xie Zhizhong and Guan Lixiu for the interesting working time, helpful ideas and useful discussions throughout the duration of my studies.

I would like to express my gratitude to the collaborators in Centre for Computational Intelligence (C2i) of School of Computer Sciences for their aids in computer relevances. In particular, I am greatly thankful to Harold Soh, Nguyen Quang Huy and Dudy Lim for their friendships and the pleasure time of collaboration throughout of my PhD project.

I would like to thank the colleagues in Institute of Atomic and Molecular Sciences, Academia Sinica for their helps during my exchange work in Taiwan.

I also grateful acknowledge the financial support of Nanyang Technological University.

At last, I would like express my deepest gratitude to my beloved family. I am deeply indebted to my mother for her endless love and support. I am under special obligations to my brother for his care to our family during my time abroad over the past few years.

Contents

List of Figures

List of Tables

Abbreviations

AMA	**A**rchiving **M**emetic **A**lgorithm
DOS	**D**ensity **O**f **S**tate
EA	**E**volution **A**lgorithm
C-HSA	**C**lassical **S**uperposition **A**pproximation
Q-HSA	**Q**uantum **H**armonic **S**uperposition **A**pproximation
GA	**G**enetic **A**lgorithm
HSA	**H**armonic **S**uperposition **A**pproximation
LJ	**L**ennard-**J**ones
MA	**M**emetic **A**lgorithm
MD	**M**olecular **D**ynamics
PT-MC	**P**arallel **T**empering **M**onte **C**arlo
PES	**P**otential **E**nergy **S**urface
SA	**S**uperposition **A**pproximation
ZPE	**Z**ero-**p**oint **E**nergy

This thesis is dedicated to my father. Gone but never forgotten. I will always miss him and love him forever.

Chapter 1

Introduction and Background

1.1 The importance of water clusters

Water receives much more attention of researchers than any other compounds. For decades, numerous studies have been published in order to reveal the structures, properties, dynamics of water.[3–19, 19–24, 24–27] Water is the most abundant substance on the Earth and plays a vital role in all kinds of life on the Earth. It is usually the major and essential component in most of living organisms. Covering over 70% of the Earth's surface, water is one of the key factors to regulate and maintain the temperature of the Earth's surface in the range favorable for life. Water is also the most popular solvent in industry and laboratory. It is capable of solvating most of substances. Hence, water participates in many important biological, chemical and physical processes.

Besides the importance of water, people are fascinated by its anomalous and unique properties[5, 28] which are hardly expected from the simple chemical formula of water molecules. For instance, water possesses an unusually high specific heat capacity and heat of vaporization. The density of water exhibits an anomalous temperature dependence in which the density reaches the maximum value at

4℃ and then decreases when the temperature decreases to 0℃. When frozen, the volume of water expands instead of shrinks so water in solid phase is lighter than in liquid phase. Moreover, water exhibits so many other unusual chemical, electrical and physical properties which have been presented in the variety of articles. Among them, readers are referred to the review works of Ref [5, 28].

The unusual properties of water are well-known to be essential for the existence of life. It has been commonly known that those properties of water come from the hydrogen bond network of water rather than the properties of a single water molecule. From chemical point of view, water should be described as a group of hydrogen bonded water molecules which is called "water clusters", $(H_2O)_n$, where n is an integer starting from 2. The existence of water clusters has been discovered experimentally with different sophisticated spectroscopy techniques in various forms: gas phase,[3, 6, 7, 9, 11, 29] ice[30] and aqueous solution.[31] Studying water clusters as well as the mixed forms with ions or protons ($H^+(H_2O)_n$, i.e.) is very essential in understanding the fundamental mechanisms of the hydrogen-bond characteristics and interactions of water.

1.2 Research in water clusters

1.2.1 A brief introduction to water clusters

A water molecule consists of one oxygen atom and two hydrogen atoms. Two hydrogen atoms can be covalently bonded with the oxygen atom with the bond length of approximately 0.96 Å. The covalent bonds are formed by sharing a single valence electron of the hydrogen atom and one of six outer-shell electrons of the oxygen atom. Four remaining electrons of oxygen atom are paired with each other to form other two lone pairs. In total, there are 4 pairs of electron surrounding one oxygen atom. Due to the repulsive forces between electron clouds, four pairs

of electron will arrange themselves in tetrahedral geometry. The oxygen nuclei is located at the center point of tetrahedron and the angle between any two bonding pairs is supposed to be 109.47°. Nevertheless, the angle of H-O-H is usually smaller than the expected value (109.47°) and is about 104° due to the strong repulsion of two lone pairs.

For inter-molecular interactions between water molecules, two lone pairs of the oxygen atom of a molecule can form hydrogen bonds with the hydrogen atoms of neighboring molecules. There has been no solid evidence to clarify the nature of hydrogen bonds yet. It is believed to rely on the attractive Coulomb and dipole-dipole interactions. If a water molecule forms a hydrogen bond, it is called a hydrogen acceptor. It can also be a hydrogen donor if any of its two hydrogen atoms participates in forming the hydrogen bond with neighboring molecules. In total, a water molecule can be bonded possibly with up to four other water molecules. Hence, there are so many possible ways to arrange the molecules in water clusters. It results in that the structural geometry of water clusters are usually complicated. It is different from most other hydrogen-bonded systems such as hydrogen fluoride clusters (HF_n) where each molecule (three lone pairs and one hydrogen atom) can only form two bonds. The structural geometry of hydrogen fluoride clusters is therefore much simpler where only the linear or ring forms can exist.

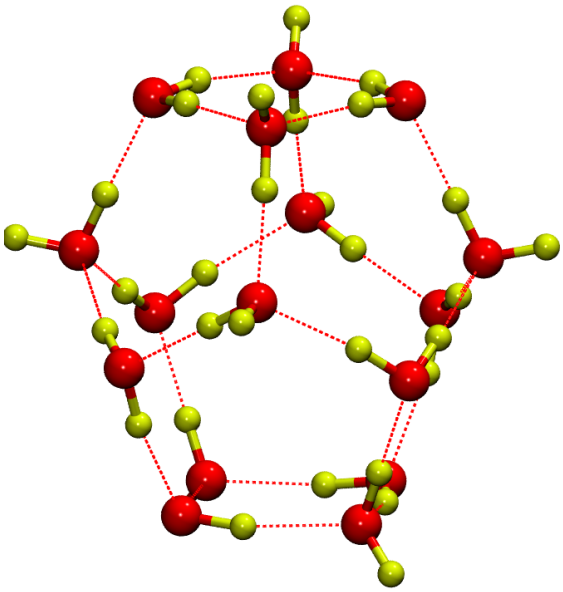


FIGURE 1.1: A decahedronal structures of $(H_2O)_{16}$. The dashed lines represent the hydrogen bonds.

The length of hydrogen bonds could not be defined precisely and supposed to vary in the range of 1.7 - 2.2 Å, depending on different situations. When a water molecule donates the hydrogen atoms, the electron density at its lone sites increases, encouraging the molecule to accept hydrogen atoms from the neighboring molecules. It is the so-called cooperative effect. In contrast, the anti-cooperative effect does the opposite. When accepting a hydrogen bond, the possibility of a water molecule to accept a new hydrogen bond is reduced but the possibility to donate hydrogen atoms to others is increased. The cooperative effect on a water molecule increases its own covalent bond (O-H) lengths and reduces the hydrogen bond (H ···O) distances whereas the anti-cooperative effect does the opposite things.

Although hydrogen bonds are considered stronger than Val der Waals interactions, they are still energetically weaker (from 5 - 30 kJ/mol and approximately 20 kJ/mol in liquid water) than covalent bonds in water molecules. Thus, the hydrogen bonds can be easily broken or reformed due to the thermal and vibrational motions of the molecules. The life time of a hydrogen bond is quite short at room temperature (in order of picoseconds). The structures of water clusters will vary with time, temperature and pressure, therefore, are very difficult to be predicted. In water systems, the number of hydrogen bonds is quite large so the hydrogen bonds are the dominant interactions. Thus, understanding the character of hydrogen bonds plays a significant role in studying the structures and dynamic properties of water.

1.2.2 Experimental studies

Investigating the microscopic structures and dynamics of water clusters are extremely difficult due to a couple of reasons. Firstly, the scattering area of hydrogen

atoms is very small which are actually unable to be observed with existing spectroscopic techniques. Thus, the geometries of water clusters can only be estimated indirectly. Secondly, the size of clusters is difficult to be selected and maintained precisely. Thirdly, it is not possible to control and monitor the temperature of clusters in experiments since the structures and dynamics of water are strongly affected by temperature as discussed in the previous section. Nevertheless, by engaging a lot of sophisticated experimental techniques, numerous experimental studies have been done in recent three decades. In this section, we only highlighted some typical works.

Dyke and coworkers[6, 7] determined the structural information of partially deutorated water dimer, the smallest water clusters,by using molecular beam electric resonance spectroscopy in 1977. A rigid-rotor model of water (the distance of covalent bonds and the angle of H-O-H were fixed) was used in conjunction with the microwave spectra to approximate the geometry of water dimers. The results of O-O distance and Euler angles of the water dimer displayed a quantitative agreement with the first-principles results derived afterward.

In 1982, Lee Y. T. et al.[8] used the crossed molecular beam technique to measure the infrared vibrational predissociation spectra of water dimer, trimer, tetramer, pentamer and hexamer in the range of 3000 - 3800 cm^{-1} for the first time. The sharp peaks of OH stretching of $(H_2O)_{n=3\text{-}6}$ clusters were detected and compared with those in liquid phase.

In 1992, Saykally and coworkers developed the far-infrared vibration-rotation-tunneling (FIR-VRT) spectroscopy.[10] This powerful method used the tunable FIR lasers to measure vibration-rotation-tunneling (VRT) bands of weak bonds such as hydrogen bonds. In the first application of FIR-VRT, the IR spectra of the cyclic water trimer was reported. Following that work, a series of experimental observations on the vibrational spectra and dynamics of water trimer,[9, 11]

tetramer,[32] pentamer,[33] and hexamer[34] have been done also by Saykally's group.

Recently, the vibrational predissociation spectra of $H^+(H_2O)_n$ for $n = 9 - 11$ were reported systematically by Lin and coworkers.[3] In this work, the size dependence of the spectra was witnessed. They found that the peaks of O-H stretching were shifted systematically to the higher frequency with increasing cluster sizes. Especially, the temperature of the clusters was approximated from the dissociation rate of water molecules for the first time. This approximated temperature was very useful for further studying the thermal effects of water clusters. In another work, Headrick et al.[35] measured the vibrational spectra of Argon-attached protonated water clusters. Afterward, the spectra of $H^+(H_2O)_{n=2\text{-}11}$ were analyzed carefully and the spectral signatures of hydrated proton vibrations was recognized.

More recently, Miyazaki et al.[36] probed the IR spectra of O-H stretching of $H^+(H_2O)_n$ with $n = 4 - 27$. From the spectral changes with cluster size, they concluded that the chain-like structures appearing at small sizes ($n < 10$) transform into two-dimensional net structures ($10 < n < 21$), and then into the cages ($n \geq 21$). A similar work was undertaken by Shin and coworkers[37] in which the OH stretching vibrational spectra of a series $H^+(H_2O)_n$ clusters with $n = 6 - 27$ were reported. Although using the method different from Miyazaki's group,[36] their experimental results somehow show a satisfactory agreement with each other.

1.2.3 Theoretical studies

Although a variety of techniques have been used in experiments, microscopic structures of water clusters are still somehow poorly understood. The information of geometry such as O-O distances could be estimated as in Ref. [11, 33], but the detailed configurations of water clusters cannot be identified experimentally. Also

no dynamic properties can be observed directly in the experiments. Therefore, theoretical calculations play a very important role in investigating the structures and properties of water clusters. To date, numerous number of theoretical works about water clusters have been carried out. In this section, we only emphasize on several interesting works about neutral water clusters. The current research on other kinds of water clusters such as protonated and deprotonated clusters will be discussed separately in Chapter 4 and 5, respectively.

There have been many extensive first-principles studies about water clusters so far. Most of the studies were limited to the small-sized clusters due to the highly computational cost of first-principles calculation. Ground-state structures and vibrational spectra of small-sized water clusters ($n < 30$) have been revealed at different levels of theory including Hartree-Fock (HF), density functional theory (DFT) and Møller-Plesset second order perturbation theory (MP2).[38–40] In these works, the ring structures were identified as the most energetically stable forms for $n = 3 - 5$. The morphological transition of $(H_2O)_n$ were recognized at $n = 6$ when the three-dimensional (3D) structures become the most stable ones. Several topologically distinct structures of water hexamer including book, cage, prism, ring were found to be energetically close to each other and their relative stabilities depend on the level of calculation theory.[12, 41] For water octamer, many studies with different levels of theory predicted two cubic isomers with S_4 and D_{2d} symmetry to be the ground-state structures which are energetically well separated from other isomers.[13, 14, 41] More recently, various properties including optimal structures, structural parameters, binding energies, hydrogen bonds, charge distributions, dipole moments of $(H_2O)_{n=2\text{-}34}$ have been computed at Hartree-Fock level.[42]

In the efforts to bridge the gap from clusters to bulk materials, large-sized water clusters ($n > 30$) have been investigated exclusively in literature. Numerous empirical models for water clusters have been developed, in conjunction

with advanced simulation methods such as MD or MC simulations, permitting to investigate large-sized clusters consisting of dozens[15, 16] to hundreds water molecules.[43] Lee et al.[16] applied simulated annealing method with the empirical potential function of Cieplak, Kollman, and Lybrand to optimize water clusters up to $n = 20$. Using basin hopping, Wales and coworkers studied the TIP4P[17] and TIP5P[44] potentials for $n \leq 21$, and compared the structures and formation energies obtained against the results of MP2 calculations. Along this line, Bandow and Hartke [15] developed a highly parallel evolutionary algorithm to examine water clusters up to $n = 34$ on TIP4P and TTM2-F potentials. In this study, they recognized the morphological transition at $n = 17$ in which the structure consisting of a single interior four-coordinate molecule at the center was found to be the global minimum. They also found two morphological trends for $n = 17 - 25$ where the odd-numbered clusters were the cages with a single interior molecule where as the even-numbered ones were the cube-and-pentarism structures. For studying huge-sized clusters, Martin Chapin[43] has recently proposed an icosahedral clustering model in which an 3-nanometer icosahedron containing 280 (H_2O) molecules was formed by arranging twenty regular 14-molecule tetrahedral clusters. The radial distributions of R(O-O) and R(O-H) derived from the model were quite consistent with X-ray diffraction data. In addition, the anomalous temperature-density behavior of water could also be explained partly from the results of this study.

1.3 Motivation and contributions of our work

To date, most of existing theoretical studies about water clusters have primarily focused on the ground-state structures without considering other meta-stable structures. Nevertheless, the ground-state structures themselves suffer from characterizing the properties of the water clusters. From the experimental observations, multiple stable conformations of water have been found to coexist. Moreover, the

relative stabilities of water isomers are usually influenced strongly by such entropy effects or zero-point-energy corrections. Hence, it is necessary to engage not only the ground-state structure but also the low-lying energetic isomers in studying water clusters. Nevertheless, identifying possible isomers is still challenging in theoretical calculations since the search space is huge and rough. The existing global optimization methods for molecular systems are only applicable with semi-empirical models and not feasible to be used with first-principles calculations due to the high computational cost. In this work, we proposed a hierarchical search method in order to explore thoroughly potential energy surface (PES) at first-principles calculations in which we not only identified the global minimum but also archived the local minima. In particular, the hierarchical search has been applied to neutral, protonated and deprotonated water clusters with different sizes. Numerous distinct configurational isomers have been uncovered and archived. The morphology of neutral and deprotonated water clusters have been found to be similar to each other and more compact than the protonated counterpart. The effect of ZPEs on the relative stabilities of each water species were analyzed afterward. The resultant archives were not only useful for studying geometries and relative stabilities but also for further thermodynamic analysis.

The vibrational spectra with increasing sizes have been measured in recent spectroscopic experiments in attempt to reveal the transition sizes and the signatures of structural changes of water clusters. Apparently, the structural transitions are strongly influenced by the temperature, therefore, finite temperature effects are necessary to be considered in those studies. However, the existing experimental techniques lack of the ability to control and monitor the temperature of clusters. Therefore, theoretical investigations are required to study the temperature dependence. To date, many theoretical simulations with different methods have been done, [2, 45–55]. Nevertheless, most of the simulations carried out empirical models and the amount of the studies at first-principles calculations still

remain limited. In this work, one of our goals is to investigate the thermodynamics and structural transitions of water clusters at first-principles calculations. To avoid a lot of problems such as quasi-ergodicity, slow convergence, high energy barrier crossing encountered in conventional Monte-Carlo simulations, superposition approximation approach has been carried out in which the density of states and other dynamic properties were acquired from the collected local minima. We showed the quantitative agreement in the thermal properties calculated by Monte Carlo simulation and harmonic superposition approximation (HSA), highlighting the reliability of HSA approach for studying water clusters. The thermodynamic behaviors of both neutral, protonated and deprotonated water clusters were subsequently examined on the same footing for a comfortable comparison.

The dynamic properties derived from theoretical calculations, however, cannot be measured experimentally. To bridge the gap between theory and experiment, the vibrational spectra were simulated based on the results of thermodynamics. Instead of looking at the spectra of individual isomers, we engaged a method in which all collected isomers are taken into account, in conjunction with the temperature effect, for simulating the total spectrum. This method was expected to offer a more realistic vibrational spectra than the conventional manner. Based on the comparison between experimental and theoretical spectra of protonated water clusters, we obtained a reasonable agreement.

Our method allows to investigate systematically from structures, dynamics to vibrational spectra at first-principles calculations. It has been applied to water clusters in this work but it is entirely possible to be used in other molecular systems. In our recent work,[56] it has been carried out to study the structures, stabilities and proton switch in protonated water and methanol mixed cluster, $H^+(CH_3OH)_m(H_2O)_n$ ($m+n=5$ and 6).

Last but not least, the collected isomers were used partly to parametrize new

empirical potentials. Two potentials for hydrogen fluoride and water clusters were developed. From our preliminary investigation, they could reproduce well the geometries, interactions and binding energies of high level MP2 calculations. The new potentials developed were expected to improve the efficiency of hierarchical approach for further exploration with larger cluster sizes in our future works.

1.4 Organization of the thesis

The thesis is organized as follows: in chapter 2, we describe the hierarchical approach in order to explore the potential energy landscape at quantum chemistry levels. The method for molecular global optimization based on Genetic Algorithms is described and the technical issues arising are addressed. In chapter 3, we present the details of harmonic superposition approximation for studying the thermodynamics and structural transitions from the archived isomers. In the following chapters (4, 5 and 6), we apply the proposed method to study extensively 3 species of water clusters including protonated, deprotonated and neutral water clusters, respectively. For each kind of water clusters, geometrical structures, thermodynamic properties and vibrational spectra are systematically investigated. The calculated results are subsequently compared with recent experiments. In chapter 7, we present the development of the potential models for hydrogen fluoride and water clusters in order to improve our methods for further study as well as to provide a good model for simulating hydrogen bonding clusters. In the final chapter, we summarize our works and discuss about the future studies.

Chapter 2

Hierarchical approach for exploring the potential energy surface

2.1 Hierarchical approach to study at quantum chemistry level

The potential energy surface can be considered as a hyperspace defined by the potential energy of a group of atoms. Exploration on the PES is very intricate since it is a high dimensional global optimization problem. For an N-atom system, the coordinate degrees of freedom is $3N$, so the computational scale is roughly $O(N^3)$. The most challenging thing, however, is the exponential increase in the number of stationary points according to the cluster size which has been demonstrated theoretically and empirically in literature.[57, 58] Besides that, the high energy barrier and ergodicity also make the exploration on PES very troublesome.

By carrying out first-principles calculations, we can yield accurate results which are close to the experiments. However, the computational scale of first-principles calculations is nearly $O(N^4)$ in most cases. Consequently, the applications of first-principles calculations are usually limited to small-sized clusters. It is not suitable to perform exploration or simulation directly on PES due to the high computational demand. Until now, the "on the fly" approaches have been used rarely on several molecular systems.[50] On the other hand, using empirical model is more favoured for such complicated systems like water clusters, because they are always faster than first-principles calculations. It is also helpful in studying the atomic interaction while leaving aside the issue of the coordinate system. By excluding the involvement of electrons, it always make obtaining the energy and the atomic forces much easier. Hence, global minimization and simulation can be performed for relatively larger systems up to hundreds of atoms.[17, 59] It is not only efficient and effective in locating ground-state structures, but also in sampling a large set of stationary points for further analysis.

On the other hand, empirical potentials are not accurate as compared to first-principles methods. To compensate for the weaknesses of two methods, we proposed an alternative approach that synergizes empirical models with first-principles calculations. Our aim is to benefit from the low computational cost of empirical potentials by coupling it with first-principles calculations to explore the PES of water clusters at the quantum mechanical level. Consequently, the search for isomers becomes an efficient highly accurate process. The main idea of the approach is sketched in Figure 2.1. Instead of searching directly on the PES of first-principles calculations (red curve) which is computationally very expensive, we first locate the isomers, denoted here as B_i, on the empirical PES (black curve) which serves a "pre-screening" stage. All distinct isomers B_i are subsequently refined to the nearest isomers A_i via first-principles optimizations. The low computational cost of the empirical potential allows possible extensive

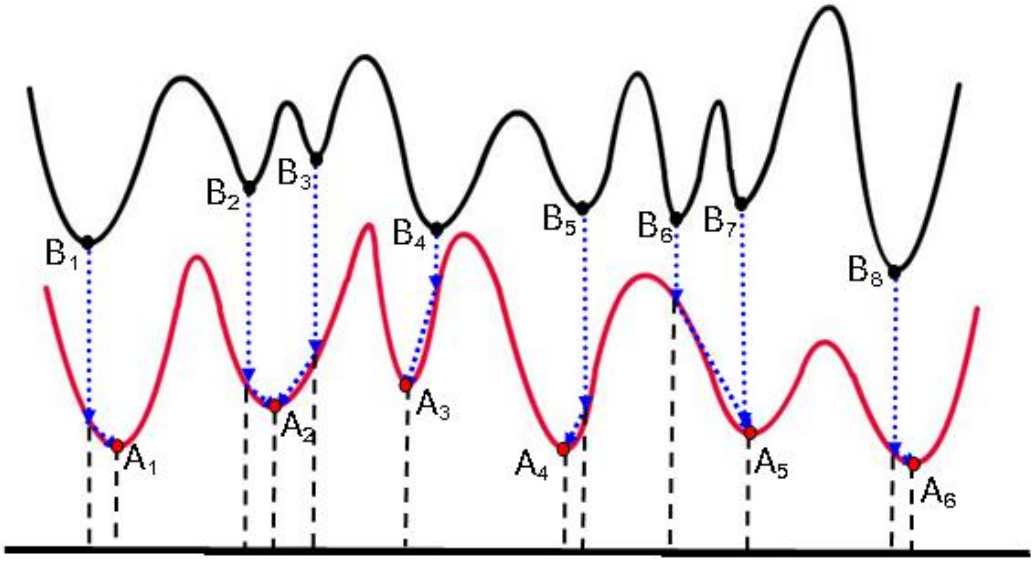


FIGURE 2.1: Sketch of hierarchical approach. A_i and B_i refer to the local minima of the empirical model (black curve) and first-principles calculations (red curve), respectively. The local minima B_i have been identified with the genetic algorithm to serve as appropriate starting points for further refinement or locally optimized using the *ab initio* calculations to arrive at the respective A_i

coverage and exploration of the PES for unique isomers. To speedup the entire isomers search process, the first-principles relaxation representing the most time consuming tasks, can be executed in parallel on multiple compute clusters. Our hierarchical approach has been previously proposed and applied to the studies of neutral and protonated water clusters.[58, 60]

2.2 Empirical potentials

Over the years, a lot of effort to construct potential models for water have been made. According to the review of Wallqvist et al.,[61] there are more than 50 empirical models including TIPS,[62] SPC, SPC/E,[63] TIP3P, TIP4P,[64] TIP4P-Ew,[65] KJ,[66] TIP5P,[67] TIP5P-E,[68] etc. Most of water models were designed for simulation in liquid and condensed phases. For studying water clusters, we

chose two sophisticated models, TTM2.1-F [18–20, 69] and OSS2,[70] as the first level of exploration since they were designed particularly for describing water clusters.

(a) **OSS2 potential:** is one of three versions of OSS family developed by Ojame and coworkers[70] to simulate water as participant in ionic chemistry. It was established from the interaction between polarizable O^{-2} anions and protons, with pair-wise and three-body terms and is suitable for studying $H^+(H_2O)_n$ system since it permits the disassociation and proton transfer in water clusters. The potential was parametrized by fitting to high level *ab initio* MP2 calculations and it can reproduce well the structures and binding energies of small-sized protonated water clusters including neutral ones. It is polarizable. However, unlike to other potentials such as TIP family, OSS2 does not use artificial charges to describe the polarizability of water molecules. Alternatively, screening functions were introduced to also model the overlap of electron clouds between oxygen and hydrogen atoms. OSS2 is a sophisticated potential with 40 parameters in which 24 ones are involved in the O-H interactions in order to describe correctly the vibrations of water molecules. Among OSS family, we chose OSS2 potential since it was demonstrated to offer the "best overall performance with regard to structure and energetics of larger neutral and protonated water clusters".[70] For details on the functional forms, stable structures and thermodynamic simulations of OSS2, the readers are referred to Ref. [2, 60, 70].

(b) **TTM2.1-F potential:** is the second sophisticated empirical potential considered in this work. It is a flexible, polarizable, Thole-type interaction potential developed by Bernham and coworkers[18–20, 69]. Although the model was parametrized using water dimer only, it was shown to reproduce the binding energies that are in close agreement to MP2 calculations for $(H_2O)_{n=2\text{-}6}$[12] and $(H_2O)_{20}$ clusters.[21] In this study, we first considered the TTM2.1-F

empirical model,[69] which is a revised version of the original TTM2-F, that reportedly resolves issues relating to dipole moment of individual water and short intermolecular interaction.

The close agreement of both TTM2.1-F and OSS2 with first-principles method on small-sized water clusters inspired the idea to use empirical models to enhance the exploration of PES at the quantum level. Because the PES landscape and the structures of water cluster isomers generated by these two models are sufficiently close to those of first-principles counterparts, it makes good sense to use their local minima as input to first-principles methods, thus reducing the computational efforts required.

2.3 First-principles calculations

All archived distinct isomers of the OSS2 model then undergo geometrical optimization using the Becke-'s three-parameter hybrid method[71] with the Lee, Yang, and Parr (B3LYP) functional[72] and 6-31+G* basis set. The convergence criteria for minimization are defined as root mean square and maximum component of gradient lower than 3.0×10^{-4} and 4.5×10^{-4} Hartree/Å, respectively. In several cases requiring a larger basis set such as $OH^-(H_2O)_n$ clusters, selected isomers of different topologies were re-optimized with the use of B3LYP/6-311+G(d,p) as well as Møller-Plesset second order perturbation theory (MP2). The effect of basis set on the structures and thermodynamics is also discussed throughout the study. To ensure the optimized structures are true isomers, vibrational analysis is subsequently performed and only those having non-imaginary frequencies are accepted for harmonic superposition approximation. Note that all calculations have been completed using the GAUSSIAN-03 package.[73]

2.4 Exploration of PES by using an Archiving Memetic Algorithm

2.4.1 Introduction

Exploration on multi-dimensional, ergodic PES for identifying the optimal solutions even with a simple objective potential function is not trivial. For decades, a lot of advanced heuristic search algorithms have been designed and developed for this purpose. In attempt to bridge the understanding gap from clusters to bulk materials, the size of investigated systems is larger and larger. The conventional methods such as simulated annealing, tabu search, basin hopping, Monte Carlo simulations are somehow less efficient for such non-convex, multi-modal, disjoint or noisy solution space. In the last decade, people started to apply genetic algorithm (GA), a technique of artificial intelligence, in molecular geometry optimization. Deaven and coworkers[59, 74] pioneered the application by carrying out GAs to locate the global minima of fullerene cluster structures up to C_{60} and Lennard-Jones clusters. For two recent decades, an abundance of studies have been done using GAs to investigate many kinds of systems from clusters[15, 58, 59, 74] to bulk materials.[75, 76]

GA was invented by John Holland in 1966. They borrow the ideas of genetics and biological evolution for solving real-world optimization problems. In fact, GA is inspired by Lamarkian spirit rather than Darwian one.[77, 78] Candidate solutions of the optimization problem evolve as individuals in a population. Each individual is characterized by its "genes" or "chromosome" which are usually implemented computationally as a bit string or an array of numbers. The "adaptive capability" to the "environment" of each individual is measured by a fitness function. The individuals create the "offspring" for the next generation via genetic

operators: selection, combination, mutation. The survival and reproduction possibilities of the individuals are based on the idea of "survival of the fittest" in the evolutionary theory where the individuals with higher fitness score will have a higher chance to survive and being reproduced. Due to this mechanism, new population of solutions biased towards better regions on PES will be generated.

The generality is one of biggest advantages which is demonstrated by the successes of GA in a wide variety of application domains: engineering, economics, finance, genetics, robotics, art, biology, physics and chemistry. GAs usually work well with many kinds of optimization problems since they generally do not require a prior-knowledge of PES neither make any assumption about the search space.

By working with a population of solutions, GA is able to cover PES well and enhance the search efficiency. Moreover, GAs can be implemented easily in embarrassingly parallel manner in which the population can be divided into many parts and each part can be independently handled by an individual processor. Consequently, the search capability of GAs is nearly proportional to number of processes in use which make GAs particularly well-suited in computer clusters.

By sampling the search space with a population of solutions rather than using local slope information from the fitness function, GA can mitigate the possibility to be trapped at local minima as encountered in other search algorithms. Nevertheless, it usually takes GA a relatively longer time to locate the exact local minima with a desired accuracy than using local search methods such as conjugate gradients, steepest decent or quasi-Newton. In order to achieve a better efficiency in exploration and exploitation, many algorithms[77, 79] combining the population based and local search procedures have been developed. They exploit the complementary advantages of GA (generality, robustness, global search efficiency) and problem-specific local search (exploiting application-specific problem structure, rapid convergence toward local minima). Such combinations of optimizers are

commonly known as memetic algorithms (MA). MA has been shown to be efficient and effective on a large number of testing and real-world problems.[15, 58, 60, 79–83] In our work, we are not only keen on identifying the global minima but also on local minima due to their essential role in analyzing PES (see Chapter 3). For this purpose, a so-called Archiving Memetic Algorithm (AMA) has been proposed. In the following sections, the components of AMA are presented in details.

2.4.1.1 Genetic operators

GA relies on operators to generate new solutions or "offspring" from existing solutions or "parents". These operators can be categorized into mutation and crossover operators. Mutation operators manipulate the chromosomes of single parents, usually through stochastic perturbation. Crossover operators combine features from the chromosomes of two or more parents to create offspring. The chromosomes are traditionally encoded as a bit string and both the mutations and crossovers are simply binary operators. Nevertheless, it was found that binary bit string GAs performed poorly in molecular optimization since operating on bit string encoded chromosomes usually destroys the cluster geometry and molecular constraints. Consequently, the clusters generated do not inherit the bond information of their parent structures and are usually worse than them. It leads to the fact that GA mostly operate on unphysical region of PES and the search becomes extremely slow and ineffective. To deal with this issue, our AMA works directly on the configuration space. The chromosomes are implemented as an array of real numbers based on the Cartesian coordinates of atoms. Additional, molecular mutation and crossover operators are designed in order to generate viable clusters that obey molecular constraints. This manner displays much more effective than the conventional counterpart as demonstrated in the previous studies.[15, 74]

(a) **Initialization:** The population of candidate solutions are initialized stochastically. The molecules are generated randomly within a sphere of radius of $(3/4N_{mol})^{1/3}R_E$, where N_{mol} is the number of molecules and R_E is the average equilibrium between two neighbouring molecules. The molecules are added one by one until the required cluster size is met. After initialization, all clusters are relaxed to their nearest local minima.

(b) **Crossover:** The crossover operator creates a single "child" cluster by merging two selected parent clusters, A and B. Firstly, both parent clusters are rotated randomly around an arbitrary axis passing through the clusters' centroids. Secondly, the clusters are cut into two sub-clusters using a random plane across their centers of mass. However, the numbers of molecules of the sub-clusters are usually not equal to each other if cut in that manner. To guarantee the sizes of the sub-clusters are similar, an array of the signed distances of all molecules to the cut plane is calculated and sorted ascendingly. The molecules corresponding to the first $N_{mol}/2$ array items of the "father" cluster are picked up to form the first sub-cluster. And the molecules corresponding to the last $N_{mol}/2$ array items of the "mother" cluster are picked up for the second one. Two sub-clusters are afterward swapped to form a new, possibly lower energy, cluster. The crossover should be constructive. The "child" cluster should inherit the characters of its parents. It means that the children are supposed to be quite similar to the "parent" ones. It has been tested and shown in our previous study[1] with 300 structures of $(H_2O)_{10}$ generated from the crossover operator. In average, the structure of the child cluster was found to be about 80% similar to the parent ones.

(c) **Mutation:** The child clusters created by crossover afterward mutate by one of three mutation operators: the perturbation, inner-molecule relocation (IMR) and outer-molecule relocation (OMR) operators as depicted in Figure 2.2. The perturbation operator is a standard operator used in previous research.[16, 83]

A couple of molecules selected randomly in a cluster are firstly rotated around the random axis through their mass centers by an arbitrary angle. Subsequently, they are translated by adding a random vector with the magnitude of ≤ 0.5 Å to the Cartesian coordinates of the atoms in the molecules. For more aggressive mutation to escape from deep minima, IMR and OMR operators are implemented. IMR relocates the molecule closest to a cluster's centroid to an arbitrary point on the cluster's surface, destroying the cluster's inner structure. OMR relocates the molecule furthest from the cluster's centroid to another arbitrary point on the cluster's surface, modifying the cluster's surface. To avoid generating unnatural geometries, constraints on bond lengths and bond angles are imposed on these operators. For example, O-O distances must be in the range of 2 - 3.4 Å.

2.4.1.2 Flow-chart of Archiving Memetic Algorithm

The conventional MA uses a generation-based scheme in which all individuals of the current population are replaced for every iteration. This scheme, however, encountered a bottle-neck since the new generation of individuals have to be synchronous. The populations have to wait until all the reproductions and local optimizations operating on each pair of individuals are finished. The overall efficiency will decrease significantly as a consequence.

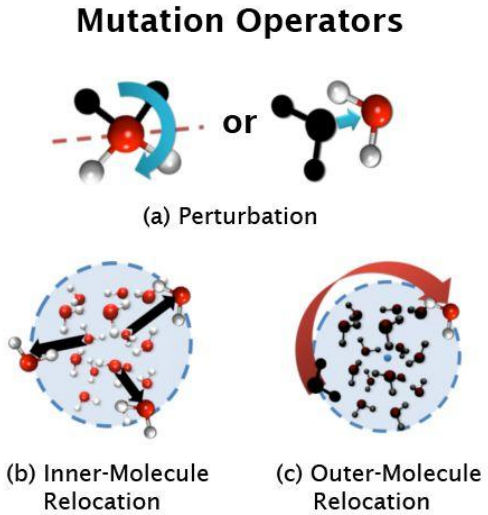


FIGURE 2.2: Three mutation operators[1].

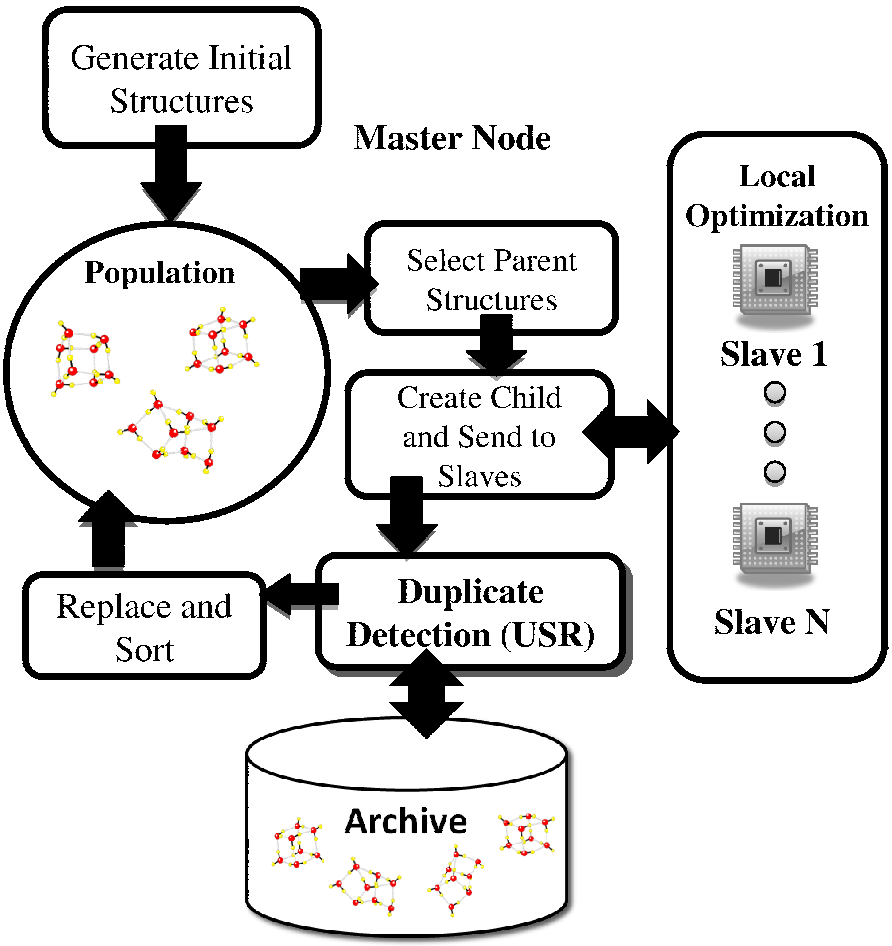


FIGURE 2.3: Flowchart of Genetic Algorithms

To avoid the bottle-neck issue, our real-coded AMA applied the asynchronous, non-generational based scheme which was successfully applied in previous studies.[15, 59, 74] The flowchart of our AMA is depicted in Figure 2.3. The master node manages the population of candidate structures and performs the genetic operators while the slave nodes locally optimize the structures sent by the master. After initialization, parent structures are rank-selected from the population before undergoing crossover and/or mutation to generate the offspring. Each individual offspring structure is sent to a slave node for local optimization. If the local optimization is successful, i.e. the root mean square of force is less than 3.0×10^{-4} Hartree/Å and the maximum component of force is less than 4.5×10^{-4} Hartree/Å, the resultant locally optimized structure is then inserted back into the

population of the master node to compete for reproductive opportunities in the spirit of Lamarckian learning.[78, 79] It is also archived for further thermodynamic and vibrational analysis. This entire process repeats until a maximum number of iterations or the wall-clock time limit is reached.

2.4.2 Ultra-fast shape recognition in archiving and preventing the premature convergence

In many cases, low-lying structures easily dominate the others in the population leading to premature search convergence.[84, 85] In the present context for example, the consequence is that the crossover operator fails to generate any further improvements when the population of structures becomes homogeneous. Hence, preserving the diversity in the population is very crucial in GAs. Unlike conventional GAs, it is not trivial for molecular GAs since the similarity of molecular geometry can not be evaluated implicitly by the fitness function. In addition, throughout the archiving process, a significant number of structures may be archived more than once. These duplicate structures not only enlarge the archive unnecessarily but also make the simulation results unreliable since the contribution by some isomers could be artificially amplified (see Chapter 3). A quantitative structure comparison method is extremely necessary in order to filter out the duplicate structures as well as maintain the diversity of population in AMA search.

A variety of methods have been developed to measure the (dis)similarity between molecules in three dimensions.[86–89] In superposition methods such as volume overlap,[86] grid point counts,[87] Gaussian approximations,[88] the molecular structures are aligned through a minimization process for measuring the shape similarity. Those methods provide accurately a quantitative measurement on molecular shape. Nevertheless, they are very costly in computing due to the use of optimization.

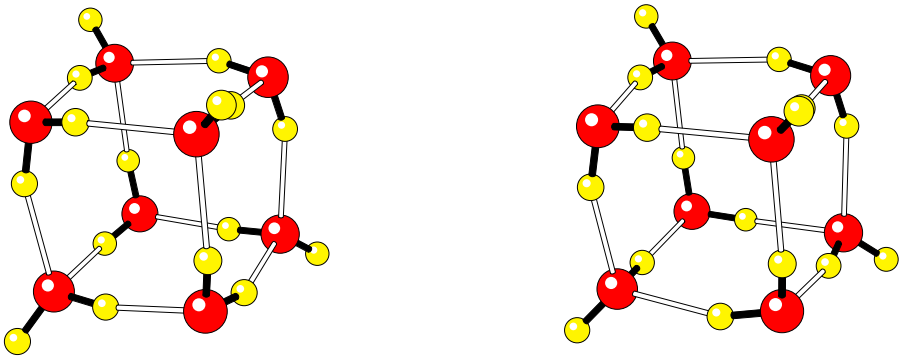
Instead, we used a computationally efficient non-superposition method with demonstrated accuracy: the Ultrafast Shape Recognition (USR) developed recently by Ballester et al.[90] Unlike the superposition methods described above, USR measures a molecular structure's shape using a signature vector of twelve atomic distance statistics, $U_{k=1-12}$. This signature captures the mean, standard deviation and asymmetry of the distances from each atom in the structure to four anchor points. The anchor points are a (the structures centroid), b (the atom closest to a), c (the atom furthest a) and d (the atom furthest from c). For details, the reader is referred to the original paper.[90]

This signature has nice properties in that it is invariant to translational and rotational symmetries. As such, we can easily define the similarity index, s_{v_i,v_j} between two molecular structures v_i and v_j as distances between the signatures. In this work, we follow the approach used by the original USR authors by using the inverse-scaled Manhattan distances between signatures:

$$s_{v_i,v_j} = \frac{1}{1 + \frac{1}{12}\sum_{k=1}^{12} |U_k^{v_i} - U_k^{v_j}|} \tag{2.1}$$

In this form, s_{v_i,v_j} is a number ranging from 0 to 1. A value of 0 indicates that two structures are totally dissimilar, whereas the other extreme represents a perfect match. Since the similarity index is constructed regardless of the number of atoms, USR can also compare two structures with different numbers of atoms.

Our tests with USR on water clusters indicated that it was effective at identifying duplicates and distinguishing dissimilar clusters. For example, two well-known water octamer isomers with the point-groups of S_4 and D_{2d} shown in Figure 2.4 are very difficult to be distinguished by naked eyes. The difference between them lies in the hydrogen-bond directions of two constituent tetramers (top and bottom surfaces). In S_4 octamer, they are both counter-clockwised (looking from the top) but in opposite directions in D_{2d}. The similarity of two isomers can be computed



a) S_4 b)D_2d

FIGURE 2.4: Structures of the S_4 and D_{2d} water octamer isomers. The similarity between them is 87% as computed by USR.

easily by USR signatures (87%). From a computational perspective, USR is highly efficient and at least three orders of magnitude faster than the previously most efficient method which is called Rapid Overlay of Chemical Structures.[91] The signature computations take at most $O(n)$ time for each pure water cluster of size n and need only be performed once for each cluster.[1]

AMA preserves diversity in the population by preventing duplication of structures in the population which have been implicated in the premature convergence.[92] Therefore, before a cluster is added to the population, it is checked against every population member. If the USR similarity to any existing population member is above 96% and the binding energy difference between the two clusters is less than 0.01 kcal/mol, the cluster is classified as a duplicate and is prevented from entering the population. The threshold values of 96% and 0.01 were chosen based on investigations performed in our works[58, 60] but can be easily modified for other studies. Moreover, to ensure that only unique structures are archived, USR algorithm is also used to sieve out potential duplicate structures.

Chapter 3

Harmonic superposition approximation

3.1 Introduction

Most of existing theoretical simulations nowadays carry out either molecular dynamics (MD) or Monte Carlo-based (MC) algorithms in predicting the structural transitions.[2, 45, 46, 93] Even though some techniques such as J-walk[44, 94] and parallel tempering[95, 96] are established as useful for resolving the issue of quasi-ergodicity, they remain to be plagued with problems of slow convergence, high energy barrier crossing, and poor sampling of the potential energy landscape, especially on large-scale or large-sized systems. An alternative approach to MD and MC is superposition approximation (SA) which acquires the diverse physical properties of a system from a collection of local minima and transition states instead of exploring directly on the potential energy landscape as like as MC or MD methods. To date, SA has been used widely in a variety of systems, from liquids to solids, glasses, clusters and etc. It was first considered by Stillinger et al. for studying hidden structures in liquids[97] and subsequently also liquid-solid

transitions.[98] Wales examined the coexistence of solid-like and liquid-like forms in a finite atomic cluster, with harmonic approximation employed in SA for the first time.[13] Calvo et al.[99] also used SA to calculate the physical and chemical properties of several atomic clusters and highlighted the close agreement between SA and MC simulations. Generally, SA is established to yield higher efficiency than conventional MC simulations in predicting thermal properties,[57, 99, 100] thus making SA feasible and practical for studying the thermodynamics of molecular systems at quantum chemistry level.

The underlying theory of SA is that the observed quantities of the system are approximated from a statistical set of local minima instead of performing sampling on the potential energy surface. The density of states ($\Omega(E)$) and the partition function ($Z(T)$) can be computed by summing up the contributions of all local minima found on PES [57, see Chapter 7, page 266] as follows

$$\Omega(E) = \sum_a \Omega_a(E) \text{ and } Z(T) = \sum_a Z_a(T), \tag{3.1}$$

where T is the temperature and $\Omega_a(E)$ and $Z_a(T)$ are the density of states and the partition function associated to a local minimum, a, respectively. In order to reduce the calculation complexity and computational efforts, the density of states or partition function of the local minima are normally approximated in some simple manners instead of being computed directly. Many schemes for approximation have been developed and applied in a lot of systems from clusters, liquid to bulk materials. [57, 97–99, 101] In several systems, only a simple approximation towards $\Omega_a(E)$ and $Z_a(T)$ produced a quantitatively good agreement with the experiments[60] and other simulation methods[99] with a low computational cost. As stated by D. J. Wales,[57] the advantage of SA is that the density of state or partition function of whole system can be broken down in terms of local

minima. It allows us to determine the contributions of a particular region of PES and to identify the major impact to thermodynamics of the whole system.

Another advantage of SA is the contribution of each local minimum to the DOS can be treated individually. Consequently, the calculation can be performed embarrassingly parallel in computer clusters making SA more efficient than the conventional MC methods where only one sampling move is carried out. SA requires to identify as many minima as possible so a sampling algorithm for effective and efficient discovery of true distinct isomers is crucial. Nevertheless, there is no restriction in how to sample the minima so a lot of advance algorithms and techniques can be utilized for this purpose. This advantage makes SA capable to overcome the difficulties such as slow convergence, high energy barrier crossing or ergodicity usually encountered in MC simulations as mentioned above.

The accuracy and success of SA depend strongly on how well the PES is explored, meaning how many isomers are involved for the calculation. However, this is not a trivial task due to the large numbers of existing local minima even for small-sized systems. The number of local minima usually increases exponentially with the cluster size as demonstrated in both theoretical and numerical studies.[57, 102] Fortunately, it can be seen from the formula that the major contribution to the density of states of the whole system comes from low-energy minima due to the Boltzmann factor, $\exp(-E/(k_B T))$. The natural bias toward the low-energy local minima suggests using only low-energy minima as the representative sample instead of a complete set of all minima which is impractical to be achieved. By limiting the search on the local minima whose energy are lower than an appropriate threshold, the number of considered minima can be reduced significantly, therefore, saving a lot of computational efforts in calculation without losing much accuracy. The biased-search algorithms such as basin hopping or GAs (described in Section 2.4) are very suitable for this purpose.

3.2 Harmonic approximation and thermodynamics

Harmonic superposition approximation (HSA) is one of the simple approximations toward the partition functions of basins. The underlying theory of HSA is to treat each local minimum as a harmonic and infinite basin characterized only by vibrational frequencies and relative energies as demonstrated in Figure 3.1. The contributions of rotation and translation to the density of states are neglected. The overlap between minima is also not taken into account. Although the approximation is relatively simple and many assumptions are proposed, HSA still represents an excellent choice for examining the equilibrium properties since its approximated results are significantly close to conventional MC-based algorithms' as demonstrated in our study of protonated water cluster[60] and many other works.[97–99, 101]

We assume that the partition function $Z_a(T)$ of each local minimum in Equation 3.1 is calculated only from the contribution of un-coupled vibrational mode. The effect of molecular rotation is not taken into account in the assumption that it has little impact. It means potential energy surface of each isomer is approximated as a hyper-ellipsoid. Based on this assumption, Hamiltonian in normal mode coordinates, Q_a of a single local minimum, a, can be written as

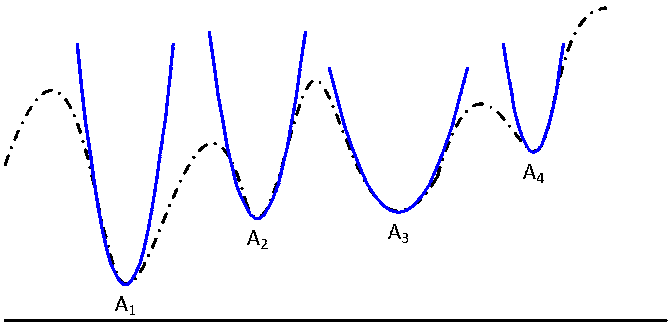


FIGURE 3.1: Schematic representation of HSA. X-axis is the reduced representation of the Cartesian coordinates. The local minima A_1, A_2, A_3, etc. are assumed to be harmonic and infinite with no overlap between them.

$$H = V_a + \frac{1}{2}\sum_{a=1}^{\kappa}(\dot{Q}_{a,\alpha}^2 + \omega_{a,\alpha}^2 Q_{a,\alpha}^2), \tag{3.2}$$

where $\omega_{a,\alpha}$ are vibrational frequencies, $\kappa = 3N - 6$ is the degree of freedom of vibration and V_a is potential energy. Each vibrational mode can be considered as a simple quantum harmonic oscillator whose energy spectrum is known as follows

$$E_{j,n} = \hbar\omega_j(n + \frac{1}{2}), n = 0, 1, 2, 3, ... \tag{3.3}$$

where j is an index of the vibrational mode, and n is the quantum number of the j^{th} vibrational mode. The canonical vibrational partition function, hence, is derived as

$$Z_{vib}(\beta) = \prod_j^{\kappa}\sum_{n=0}^{\infty} e^{-\beta E_{j,n}} = \prod_j^{\kappa}\sum_{n=0}^{\infty} e^{-\beta\hbar\omega_j(n+\frac{1}{2})} = \prod_j^{\kappa}\frac{e^{-\beta\hbar\omega_j/2}}{1 - e^{-\beta\hbar\omega_j}}, \tag{3.4}$$

where $\beta = 1/(k_B T)$, k_B is Boltzmann constant. The canonical partition function associated to a local minimum, a, is calculated as the product of vibrational contribution part,Z_{vib}, and $e^{-\beta V_a}$.

$$Z_a^Q(\beta) = e^{-\beta V_a}\prod_{\alpha=1}^{\kappa}\frac{e^{-\beta\hbar\omega_{a\alpha}/2}}{1 - e^{-\beta\hbar\omega_{a,\alpha}}}, \tag{3.5}$$

where Q denotes that the contribution of the vibrational modes are treated on the quantum-mechanical footing. For a specific configurationally distinct minimum of an N-atom system, the Hamiltonian is invariant to all permutation of equivalent nuclei and to inversion of all coordinates through a space-fixed origin (see Chapter 5.1 in Ref [57]). Hence, to derive the total canonical partition function, the degeneracy factor,

$$n_a = \frac{2N_A!N_B!N_C!...}{m_a}, \tag{3.6}$$

should be included in the formula of total partition function. Here, N_A, N_B, N_C, ... denote the number of atoms of atomic type A, B, C,.., respectively. m_a is the order of point-group symmetry of the local minimum a. The total canonical partition function in quantum regime (Q-HSA) is derived by taking summation over all configurationally distinct local minima,

$$Z^Q(\beta) = \sum_a n_a e^{-\beta V_a} \prod_{\alpha=1}^{\kappa} \frac{e^{-\beta\hbar\omega_{a,\alpha}/2}}{1 - e^{-\beta\hbar\omega_{a,\alpha}}}. \tag{3.7}$$

By considering the classical limit, $\hbar \longrightarrow 0$, the harmonic approximation of canonical partition function, $Z^C(\beta)$, in classical treatment (C-HSA), is derived as given below

$$Z^C(\beta) = \sum_a n_a e^{-\beta V_a} \prod_{\alpha=1}^{\kappa} \lim_{\hbar\to 0} \frac{e^{-\beta\hbar\omega_{a,\alpha}/2}}{1 - e^{-\beta\hbar\omega_{a,\alpha}}} = \sum_a n_a e^{-\beta V_a} \prod_{\alpha=1}^{\kappa} \frac{1}{\beta\hbar\omega_{a,\alpha}}. \tag{3.8}$$

From the total partition function, other thermodynamic quantities can be obtained easily as follows

The free energy is

$$A = -\frac{1}{\beta} \ln Z(\beta). \tag{3.9}$$

The average energy is

$$\langle E\rangle(\beta) = -\frac{\partial}{\partial\beta} lnZ(\beta). \tag{3.10}$$

The heat capacity (in k_B unit) is

$$C_v(\beta) = -\frac{1}{k_B}\frac{d\langle E\rangle}{dT} = -\beta^2\frac{\partial}{\partial\beta}\langle E\rangle. \tag{3.11}$$

By substituting the partition functions from Equation 3.8 and Equation 3.7 into above equations, we obtain easily the thermodynamics as follows.

In classical harmonic superposition approximation theory (C-HSA), the average energy and heat capacity are [57, see Chapter 7.1]

$$\langle E\rangle^C(\beta) = \frac{Z_1}{Z_0} \text{ and } C_v^C(\beta) = \beta^2\left(-\frac{Z_1^2}{Z_0^2} + \frac{Z_2}{Z_0}\right), \text{ respectively} \tag{3.12}$$

where

$$Z_p = \sum_a \frac{n_a V_a^p e^{-\beta V_a}}{\prod_\alpha \beta\hbar\omega_{a,\alpha}}, p = 0, 1, 2 \tag{3.13}$$

In quantum harmonic superposition theory (Q-HSA), the average energy and heat capacity are [57, see Chapter 7.1]

$$\langle E\rangle^Q(\beta) = \frac{Z_1^Q + Z_{11}^Q}{Z_0^Q} \quad C_v^Q(\beta) = \beta^2\left[\frac{Z_2^Q + 2Z_{11}^Q + Z_{02}^Q}{Z_0^Q} - \left(\frac{Z_1^Q + Z_0 1^Q}{Z_0^Q}\right)^2\right] \tag{3.14}$$

where

$$\begin{aligned}
Z_p^Q &= \sum_a \frac{n_a E_a^p e^{-\beta V_a}}{\prod_\alpha 1 - e^{-\beta\hbar\omega_{a,\alpha}}}, p = 0, 1, 2 \\
Z_{p1}^Q &= \sum_a \frac{n_a E_a^p e^{-\beta V_a}}{\prod_\alpha 1 - e^{-\beta\hbar\omega_{a,\alpha}}} \sum_\gamma \frac{\hbar\omega_{a,\gamma}}{e^{\beta\hbar\omega_{a,\gamma}} - 1}, p = 0, 1 \\
Z_{02}^Q &= \sum_a \frac{n_a E_a^p e^{-\beta V_a}}{\prod_\alpha 1 - e^{-\beta\hbar\omega_{a,\alpha}}} \left[\sum_\gamma \frac{(\hbar\omega_{a,\gamma})^2 e^{-\beta\hbar\omega_{a,\gamma}}}{(1 - e^{\beta\hbar\omega_{a,\gamma}})^2} + \left(\sum_\gamma \frac{\hbar\omega_{a,\gamma}}{e^{\beta\hbar\omega_{a,\gamma}} - 1}\right)^2\right],
\end{aligned} \tag{3.15}$$

where $E_a = V_a + \kappa\hbar\langle\omega\rangle/2$ and $\langle\omega\rangle = (\sum_\alpha \omega_{a,\alpha})/\kappa$ is the average of the frequencies.

In the following chapters, we apply HSA to derive thermodynamics of water clusters in both classical and quantum regimes. The structural transitions as well as the character of thermodynamic properties of three species of water clusters are simulated and studied thoroughly. The factors affecting on HSA results such as structures, relative stability, cluster sizes, etc., are also investigated. After that, the vibrational spectra are calculated based on the thermodynamic results and then compared with the experimental observation. Throughout the comparison, the reliability of HSA is tested and compared with other methods such as Monte Carlo simulations. The details will be discussed in subsequent chapters.

Chapter 4

Study of protonated water clusters

4.1 Introduction

Protonated water clusters have attracted numerous studies for a long period of time because of their important roles in ionic media and chemical reactions.[2, 3, 35, 36, 45–54, 70, 93, 103, 104] To date, significant progress has been made in experiments[3, 35–37, 103–106] and theoretical simulations.[2, 45–55, 107–109] In the latter, one of the core focuses is studying the dynamic structural transitions to reveal the complicated thermal behavior of water clusters. Using Monte Carlo simulations, Singer and coworkers identified the topological transitions of $H^+(H_2O)_8$ and $H^+(H_2O)_{16}$ to treelike structures at high temperature, using OSS2 model.[70] On the other hand, Christie and Jordan[93] with the use of MSEVB model[110] identified two sharp transitions of $H^+(H_2O)_8$ whereas that of $H^+(H_2O)_6$ is devoid of sharp structures. Kuo and Klein[2] carried out basin hopping[17, 53, 111] and parallel tempering algorithm to systematically examine the low-energy structures and structural transitions of $H^+(H_2O)_n$ for n up to 21. In a similar work,

James et. al.[44] employed a modified empirical valance bond potential to study the properties of selected small to medium-sized clusters. With the advance of computation power and methodology, several studies have been made in the attempt to study water clusters at *ab initio* levels.[51, 52, 94, 112, 113] Iyengar and coworkers proposed the *ab initio* atom-centered density matrix propagation method to investigate systematically the structures, dynamics and vibrational properties of $H^+(H_2O)_n$ for the magic size of $n = 21$[51, 52] as well as other systems.[114] Recently, Nakayama et. al. considered the use of an approximate potential to speed up the *ab initio* MC simulation on small protonated water clusters, $H^+(H_2O)_{n=1,2}$.[94]

Most of existing theoretical simulations have engaged either molecular dynamics (MD) or Monte Carlo-based (MC) algorithms in predicting the structural transitions of $H^+(H_2O)_n$.[2, 45, 46, 93] However, those methods face a lot of problems such as quasi-ergodicity, slow convergence, high energy barrier crossing causing the difficulties in sampling of the potential energy landscape, especially on large-scale or large-sized systems. An alternative approach to MD and MC is superposition approximation (SA) where the observed quantities of the system are approximated from a statistical set of basins instead of performing sampling and averaging of the entire PES.

Recently, Ohno and coworkers[48, 49] proposed the anharmonic downward distortion following (ADD) algorithm, a full first-principles based approach, to explore the PES and then examined the thermodynamics of $H^+(H_2O)_n$ for n up to 7 via HSA. The sets of 9, 24, and 131 isomers collected for $n = 5$, 6, and 7, respectively, at B3LYP/6-31+G** level were reasonably large. However, the requirement on second-order derivative calculations and serial searching regime limits the ADD from sampling widely for high-energy isomers on the quantum chemistry PES, especially for larger-sized systems.

In the present work, instead of a complete reliance on *ab initio* calculations, we employ a hierarchical methodology with the OSS2 model employed as a prescreening process to construct the archival of potential distinct isomers for $n = 5 - 9$ to be subsequently examined and refined by first-principles calculations at the B3LYP/6-31+G* level. The synergy between empirical model and first-principles method permits extensive and efficient exploration of the PES, dealing with large-sized systems (up to $n = 9$), and studying the systems at multiple distinctive levels of theory simultaneously. To deal with the issue of exponential growth on the number of isomers, a parallel asynchronous genetic algorithm is deployed for searching and archiving of distinct configurational isomers in $H^+(H_2O)_n$. The archival of isomers are then analyzed to investigate the thermodynamic properties and structural transitions of $H^+(H_2O)_{n=5\text{-}10}$, for both OSS2 model and *ab initio* calculations. From our obtained results, HSA is shown to be reliable on small-sized systems since it reproduces well the feature of heat capacity curves and structural transition of $H^+(H_2O)_n$ in MC simulations of OSS2 model. The structural transition trends of protonated water clusters at *ab initio* calculations also revealed close agreement with that of the OSS2 model. The vibrational spectra are derived upon on the results of HSA and compared with recent experimental study. The detailed results and discussions will be presented in the following sections.

4.2 Exploration on potential energy surface

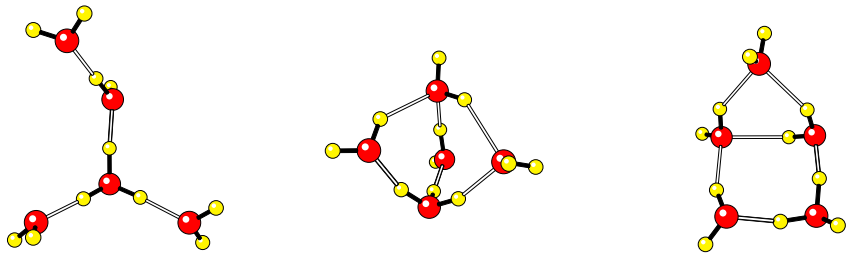
It is worth highlighting that our GA has reproduced successfully all the most stable structures of $H^+(H_2O)_{n=5\text{-}9}$ that were found by the basin-hopping algorithm reported in previous related works.[2] In addition to that, we have uncovered and archived large numbers of distinct isomers which are summarized in Table 4.1. Since a threshold value of 0.96 was employed in the USR technique to remove duplicate isomers automatically, an over-elimination of unique isomers may happen.

The resultant archive therefore may not represent the complete set of isomers on the PES. However, from our results the number of distinct isomers, $N_{isomers}$, is observed to increase exponentially with the number of atoms, N, in the range of $n = 5 - 9$ as expected in both theoretical and numerical studies.[57, 102] $N_{isomers}$ can be approximated as $Ae^{\alpha N}$ where both A and α are constants depending on the system considered. From the database of Lennard-Jones clusters, A and α have been approximated as 0.00397 and 0.9897, respectively, for $N \leq 16$.[115, 116] In the present work, A and α are approximated as 0.5849 and 0.3918, respectively. The significant larger A for $H^+(H_2O)_n$ might arise from the fact that water cluster system is less symmetric than atomic counterparts. This explains the larger number of isomers identified in the small-sized water clusters, hence a larger A value. Thus, A might represent as the measure of asymmetry in the system. In contrast, the smaller value of α of the water cluster system would imply its smaller degree of freedom due to the bonds.

The uncovered sets of archived OSS2 isomers subsequently serve as input structures that are optimized at B3LYP/6-31+G* level. At the end of each successful optimization, the local optima at B3LYP/6-31+G* undergo vibrational analysis to retrieve the vibrational frequencies. For $H^+(H_2O)_5$, $H^+(H_2O)_6$ and $H^+(H_2O)_7$, a total of 21, 135, 707 isomers have been uncovered and archived, respectively. For $H^+(H_2O)_8$ and $H^+(H_2O)_9$, because the number of isomers found using the OSS2 model is large, a stricter similarity threshold of 0.85 is used to

Cluster size	Number of isomers
5	218
6	1192
7	4700
8	11820
9	24693
10	60380

TABLE 4.1: Number of distinct isomers of $H^+(H_2O)_{n=5\text{-}10}$ with OSS2 potential.



a) 1.192 b) 3.739 c) 6.624

FIGURE 4.1: Several new isomers of $H^+(H_2O)_5$, the numbers denoted as relative energies in kcal/mol

filter out greater number of potentially duplicated OSS2 isomers before undergoing DFT optimization. Based on the given criterion, resultant sets of 422 and 877 isomers for $H^+(H_2O)_8$ and $H^+(H_2O)_9$, respectively, have been identified. In comparison to other recent works reported in the literature on similar studies,[48, 49] the archives represent the largest sets of isomers reported at quantum chemistry level to date. For instance, several isomers of $H^+(H_2O)_5$ that are missing in the recent reported work[49] are depicted in Figure 4.1.

The 10 lowest-energetic isomers of each cluster size have been sorted according to their binding energies and depicted in Figure 4.2. As observed, the ground-state structures of both $H^+(H_2O)_5$ and $H^+(H_2O)_6$ belong to the family of four-member ring as observed in OSS2 model. For $H^+(H_2O)_5$ cluster, several other families of isomer shapes is also found to exist as depicted in Figure 4.2. These include five-member ring, tree-like or cage isomers which are well separated at the energy level ($\sim$ 1kcal/mol). On the other hand, the family of four-member ring isomers is observed to dominate, with near iso-energetic at small energy gap of less than 0.2 kcal/mol. For $H^+(H_2O)_7$ and large-sized clusters, the structures tend to be more compact and all ground-state and low-energetic structures are of multi-ring shapes. The cubic isomers of $H^+(H_2O)_8$ are also relatively stable with a small energy gap of approximately 0.9 kcal/mol to the ground-state counterpart. Note that this is consistent with the observation reported by James et al.[44] where the

lowest cubic-like minimum of the modified MSEVB potential is approximately 0.8 kcal/mol higher than the most stable structure.

Figure 4.3 depicts the structures of $H^+(H_2O)_n$ sorted according to the energy after zero-point energy correction. It is worth noting that the resultant structures of all sizes are observed to be more open. For $H^+(H_2O)_5$ and $H^+(H_2O)_6$, tree-like isomers have lowest energy instead of single-ring, whereas for $H^+(H_2O)_7$, the single-ring structures dominate. For $H^+(H_2O)_8$ and $H^+(H_2O)_9$, the lowest-energetic isomers are no longer cage and few single-ring isomers are found. Besides the fact that open isomers are more favored, they are also more iso-energetic since the energy gaps between them are smaller (< 1 kcal/mol) than those without zero-point energy.

4.3 Thermodynamic transitions

4.3.1 Comparison with Parallel Tempering-Monte Carlo simulation

To observe the structural transformations, isomers of $H^+(H_2O)_n$ are classified into five categories of topological families, namely, multi-ring, double-ring, single-ring, tree-like and linear. The population or the canonical probability of the system to fall under topology A is calculated as $P_A(T) = \frac{\sum_{a \in A} Z_a(T)}{Z_{total}(T)}$. Other observables including specific heat capacity, free energy can also be easily derived based on the details available in Chapter 3.

In Figure 4.4, the temperature dependencies of heat capacity and the populations of five topological families obtained for $H^+(H_2O)_{n=5\text{-}10}$, are depicted. For

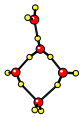
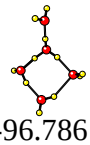
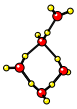
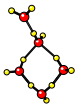
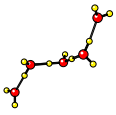
	W5+	W6+	W7+	W8+	W9+
I	 -99.184	 -113.854	 -129.765	 -145.381	 -159.635
II	 -97.981	 -113.834	 -129.379	 -144.657	 -159.127
III	 -97.920	 -113.602	 -129.365	 -144.511	 -158.400
IV	 -96.786	 -113.421	 -129.244	 -144.416	 -158.398
V	 -96.009	 -113.321	 -128.235	 -144.290	 -158.237
VI	 -95.445	 -113.293	 -128.225	 -144.116	 -158.105
VII	 -95.357	 -113.247	 -128.184	 -144.109	 -158.021
VIII	 -95.275	 -113.211	 -128.173	 -143.730	 -157.951
IX	 -94.998	 -113.132	 -128.172	 -143.558	 -157.920
X	 -94.807	 -112.977	 -128.166	 -143.550	 -157.769

FIGURE 4.2: Ten lowest-energetic isomers of $H^+(H_2O)_{n=5\text{-}9}$, found in B3LYP/6-31+G*, sorted with increasing order of binding energies (in kcal/-mol).

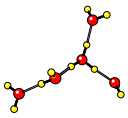
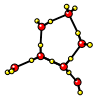
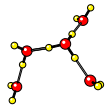
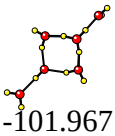
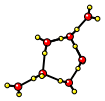
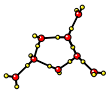
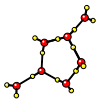
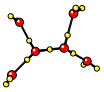
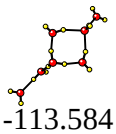
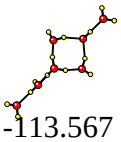
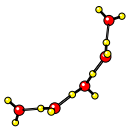
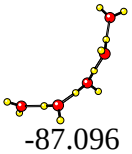
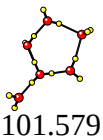
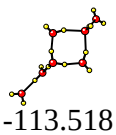
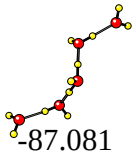
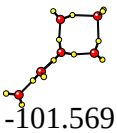
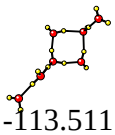
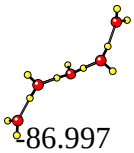
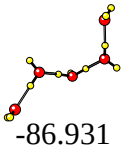
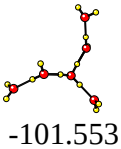
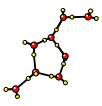
	W5+	W6+	W7+	W8+	W9+
I	 -89.645	 -102.271	 -114.160	 -125.739	 -136.998
II	 -89.354	 -101.967	 -113.689	 -125.091	 -136.835
III	 -89.312	 -101.953	 -113.657	 -125.070	 -136.714
IV	 -87.334	 -101.898	 -113.584	 -125.047	 -136.538
V	 -87.123	 -101.633	 -113.567	 -124.979	 -136.220
VI	 -87.098	 -101.606	 -113.562	 -124.881	 -136.216
VII	 -87.096	 -101.579	 -113.518	 -124.747	 -135.971
VIII	 -87.081	 -101.569	 -113.511	 -124.617	 -135.943
IX	 -86.997	 -101.554	 -113.397	 -124.542	 -135.801
X	 -86.931	 -101.553	 -113.365	 -124.509	 -135.795

FIGURE 4.3: Ten lowest-energetic isomers of $H^+(H_2O)_{n=5\text{-}9}$, found in B3LYP/6-31+G*, sorted with increasing order of binding energies with ZPE correction(in kcal/mol).

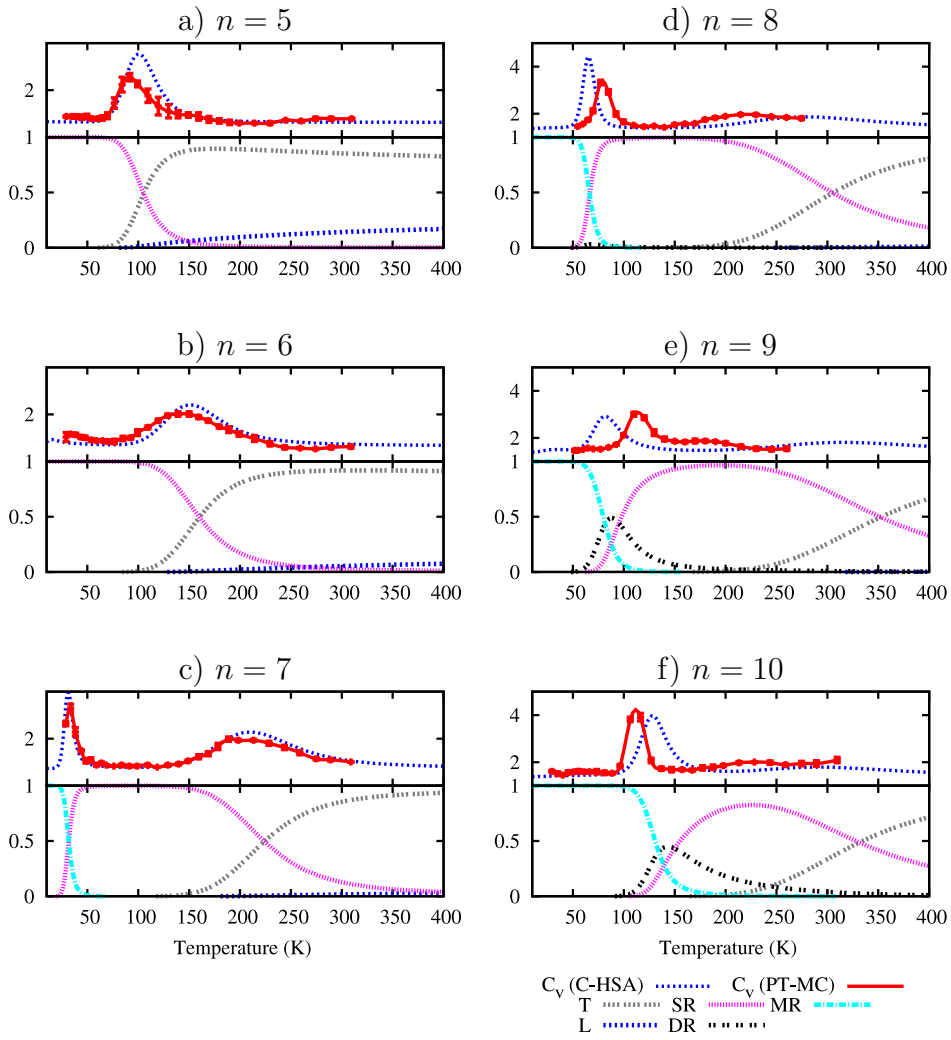


FIGURE 4.4: Canonical heat capacity Cv (upper panel) and population (lower panel) of five topologies for $H^+(H_2O)_{n=5\text{-}10}$, i.e., multi-ring (MR), double-ring (DR), single-ring (SR), linear (L) and tree-like (T), calculated using classical HSA (C-HSA) with the OSS2 model. Solid-red and dotted-blue lines in the upper panels represent heat capacity curves from previous parallel tempering Monte Carlo (PT-MC) simulation[2] and our C-HSA, respectively. For the sake of brevity, the heat capacity is plotted in dimensionless unit (Nk_B).

possible comparison, the heat capacity trends obtained by the parallel tempering Monte Carlo (PT-MC) simulations as reported in Ref [2] are reproduced in Figure 4.4 with the use of cubic-spline smoothing.

From Figure 4.4, HSA is observed to reproduce the transition peaks of the heat capacity trace well, for both sizes of 5 and 6. Both methods predicted similar transition temperatures in the proximity of 100 K and 150 K for $H^+(H_2O)_5$ and $H^+(H_2O)_6$, respectively. Analysis on the populations of topologies also indicated

the domination by single-ring (SR) structures dominate at low temperature and their transformations to tree-like (T) forms at high temperature. On the other hand, linear structures populating significantly in the archival appears only at high temperature with small probability. The similar observation was also recognized in PT-MC simulations.

For $n \geq 7$, when multi-ring structures posed as most stable, there are two maxima corresponding to two structural transitions: an exceptionally sharp peak for the rapid changes of multi-ring to single-ring while a broad and short peak for the gradual changes of single-ring to tree-like. In the case of $H^+(H_2O)_7$, the transition temperatures predicted by HSA are again in close agreement with those obtained by PT-MC, i.e., a significant rise at a low temperature of 40 K and a bump at higher temperature of about 230 K is observed in both. For $H^+(H_2O)_8$, HSA produces slight difference on the prediction of transition temperatures, i.e., the first peak (60 K) is lower than the counterpart in PT-MC (70 K) as expected, while the second peak shifts to higher temperature. Nevertheless, the overall feature and trend of the population and heat capacity obtained using HSA correlate well with those obtained by PT-MC.

For sizes 9 and 10, the transition temperatures become more difficult to be predicted accurately since both PT-MC and HSA fail to converge well. Note that the increase in transition temperatures due to cluster size in HSA matches that predicted by PT-MC previously. These results also revealed the transitions from single ring to treelike structures happening for all the cluster sizes investigated and the transition temperature shifting higher with growing cluster size. It is also worth noting the absence of double-ring structures for $n = 5$, 6, 7 and 8 in the population plots, even though double-ring populates significantly in the archival.

Except the first peaks of $H^+(H_2O)_8$ and $H^+(H_2O)_9$, HSA seems to overestimate the transition peak when compared to PT-MC and the shape of these peaks are

noted to be somewhat sharper. However, the overestimation is relatively small and regarded as acceptable for $n \leq 8$, although there is a slight increase for $n \geq 9$. We believe one of the core reasons is related to the underestimations of the contributions of the high-lying local minima, since our exploration of the PES for OSS2 using GA has placed greater emphasis on the low energy region than the high energy counterpart. Hence, the entropy of high-energy topologies in the form of tree and linear shapes has generally been underestimated. The resultant impact is a broadening and shifting of the high temperature transition peak to higher temperature, while a sharpening and shifting of the low temperature transition peak to lower temperature.

4.3.2 DFT calculation

Employing conventional Monte-Carlo methods to simulate thermodynamics at DFT calculation is impractical, even for small-sized systems, due to the high computational cost involved. The use of HSA, on the other hand, serves to be more appropriate due to its higher efficiency compared to conventional MC approaches. Furthermore, HSA has also been shown to yield quantitatively good agreement with Monte Carlo simulation on small-sized systems, as demonstrated in the previous sections for the OSS2 model. In what follows, we discuss the use of HSA to study protonated water clusters with DFT calculation in both classical (C-HSA) and quantum (Q-HSA) theory.

4.3.2.1 Classical theory

Using the archival of local minima, we applied classical HSA to simulate the thermodynamic transitions. It can be observed in Figure 4.5 that the heat capacity trend obtained using DFT calculations correlates well with those produced based on the OSS2 model for $n = 5 - 7$. In case of $H^+(H_2O)_5$ and $H^+(H_2O)_6$, a

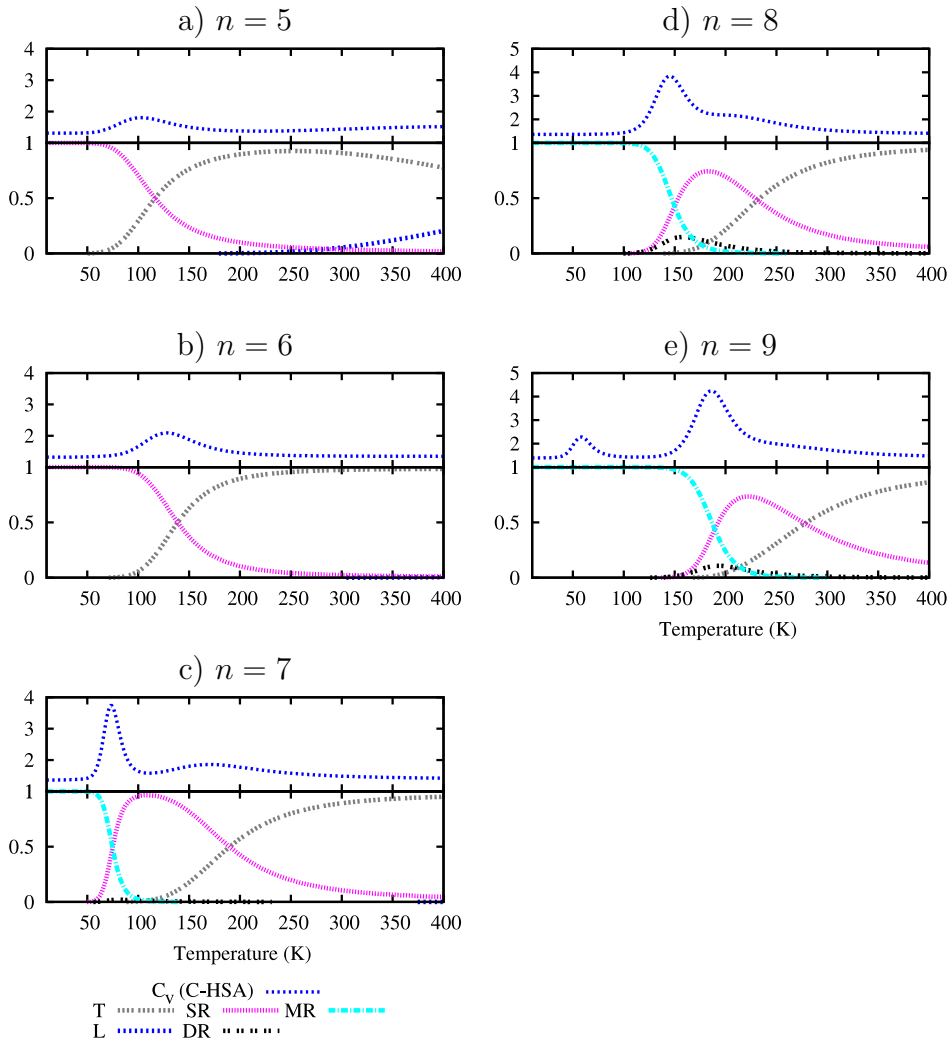


FIGURE 4.5: Canonical heat capacity Cv (upper panel) and population (lower panel) of five topologies of $H^+(H_2O)_{n=5\text{-}9}$, calculated using classical HSA (C-HSA) with B3LYP/6-31+G* calculation. The full details have been described in Figure 4.4

small maximum representing the transition from single-ring to treelike can be observed. The DFT results show a higher transition temperature with a phase change that is more gradual as reflected by a flat and shorter peak. For $H^+(H_2O)_7$ and $H^+(H_2O)_8$, on the other hand, two transition peaks found on DFT are similar to those of OSS2, except with the first occurring at a higher temperature than the OSS2 model, while the other at a lower temperature. Further, the two peaks of $H^+(H_2O)_8$ are relatively close and appear to merge as single peak in the case of $H^+(H_2O)_9$. Note also the small peak in the heat capacity trend of $H^+(H_2O)_9$ around 60 K which corresponds to the transition of the lowest-energetic isomers

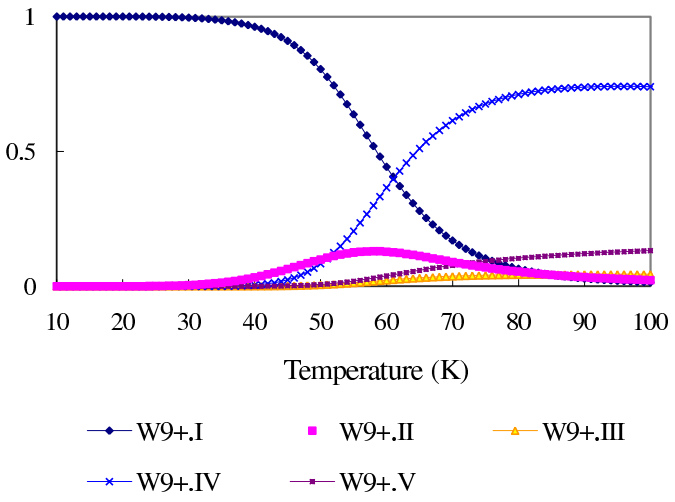


FIGURE 4.6: Population plot of five lowest energetic isomers of $H^+(H_2O)_9$ calculated using the classical HSA (C-HSA) with B3LYP/6-31+G* isomers. The transition from the lowest energetic isomer (W9+).I to the next lowest energetic ones resulted in the first heat capacity peak of Figure 4.5e

(denoted as W9+.I in Figure 4.2) to the second (W9+.II), third (W9+.III) and fourth (W9+.IV) lowest energetic isomer as depicted in Figure 4.6. Double-ring isomers appear to start from $H^+(H_2O)_8$ but only survive within a small range of temperature (around 150 K). The overall trend of the structural transitions agrees very well with the OSS2 model except on the precise positions of the transition temperatures. The size dependency of structural transitions is also consistent with OSS2 model. The shift to higher temperature from single-ring to treelike with increasing sizes is also observed for T_c = 120 K, 135 K, 170 K, 240 K and 275 K for $H^+(H_2O)_n$, n=5, 6, 7, 8 and 9, respectively.

4.3.2.2 Quantum theory

The key difference of quantum HSA from classical counterpart lies in the involvement of zero-point energy. As mentioned in the section 3.1, the inclusion of zero point energy changes the relative stability of structures and making open structures such as tree-like and linear forms more favorable. In this section, we discuss their impacts on thermodynamic properties.

The population and canonical heat capacity plots of quantum HSA are depicted in Figure 4.7. For $H^+(H_2O)_5$ and $H^+(H_2O)_6$, tree-like structures predominate the population in the range of 0-400 K instead of single-ring forms. As a result, there

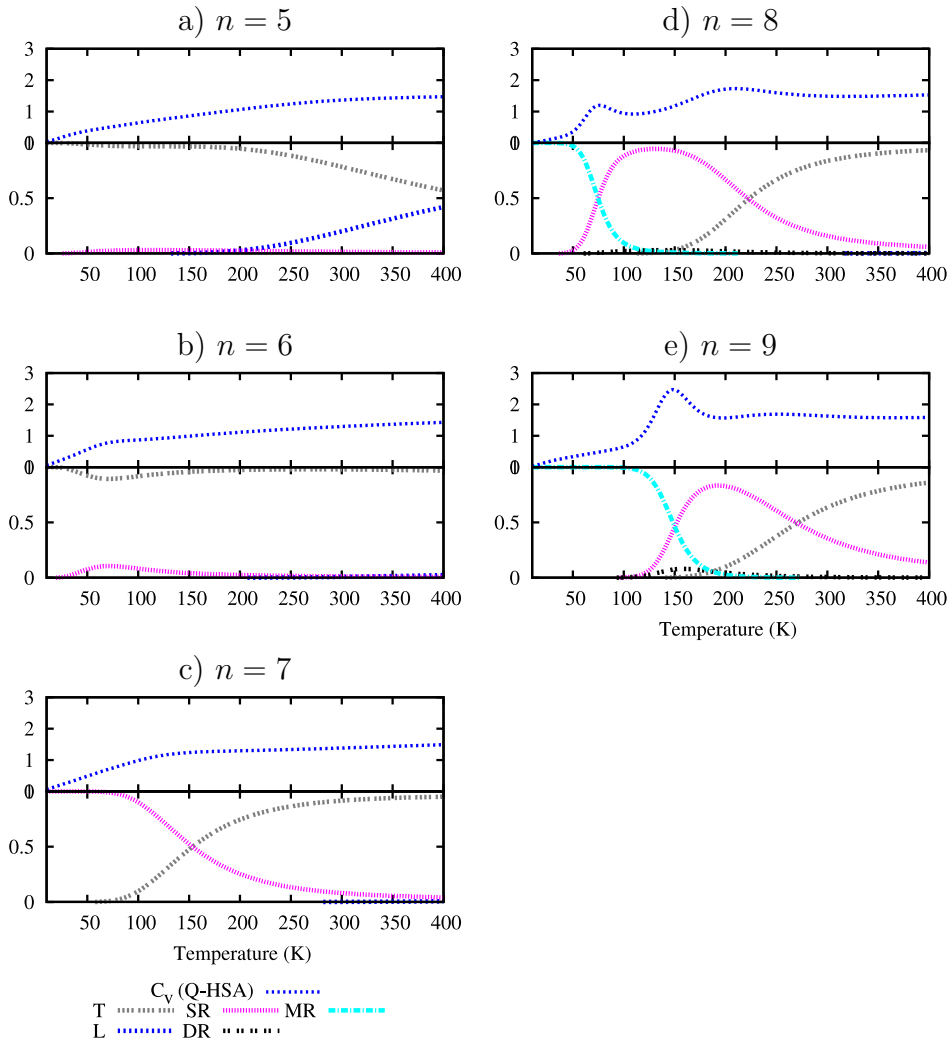


FIGURE 4.7: Canonical heat capacity Cv (upper panel) and population (lower panel) of five topologies of $H^+(H_2O)_n$, $n = 5 - 9$, calculated using quantum HSA (Q-HSA) with B3LYP/6-31+G* calculation. The full details have been described in Figure 4.4

is almost no structural transition as reflected in the heat capacity plots. For $H^+(H_2O)_7$, only one maximum of the transformation from single-ring to tree-like structures arises at around 150 K and in contrast to classical HSA, there is zero contribution by multi-ring structures. It is worth noting that the observations made are consistent with the recent work reported by Luo et al.[49]

For $H^+(H_2O)_8$ and $H^+(H_2O)_9$, the transition characteristics are similar to classical HSA. Two main transitions, namely, from multi-ring to single-ring and from single-ring to tree-like structures, displayed small contributions of double-ring at

around 150 K. In addition, the transitions seem to occur at lower temperatures than classical HSA. This makes sense since the zero-point energy correction has the effect of decreasing the entropy of compact structures. As explained earlier in section Section 4.3.1, this causes the shift in transitions to lower temperatures.

4.4 Vibrational spectra in free OH-stretching region

From the results of vibrational analysis, the IR spectrum $I_a(\omega)$ of an isomer, a, is approximated using Lorentz line shape. The total IR spectrum $I_{total}(\omega, T)$ is then calculated as the weighted sum of $I_a(\omega)$ with the canonical probability $p_a(T)$ of the isomer a derived from thermodynamic simulations which is given by:

$$I_{total}(\omega, T) = \sum_a I_a(\omega) p_a(T) \tag{4.1}$$

In Figure 4.8, the IR spectra calculated using both classical (C-HSA) and quantum HSA (Q-HSA) in the free OH-stretching region are plotted together with the experimental results which are reproduced from the work of Lin and coworkers.[3] For consistency, the spectra are simulated at temperature $T = 170K$ which lies within the range of cluster temperature used in Lin"s work[3] as deduced from the Arrhenius equation.[117] Besides, several relevant works[48, 49] have typically considered a temperature of 170K in their investigations on protonated water clusters. Note that all calculated frequencies are also scaled by a factor of 0.973 as suggested in a recent related work.[118]

Four groups of peaks in the free-OH stretching region are highlighted in Figure 4.8, namely, symmetric free-OH, asymmetric free-OH of one-coordinate (H_2O), free-OH of three-coordinate H_2O and free-OH of two-coordinate (H_2O), which are

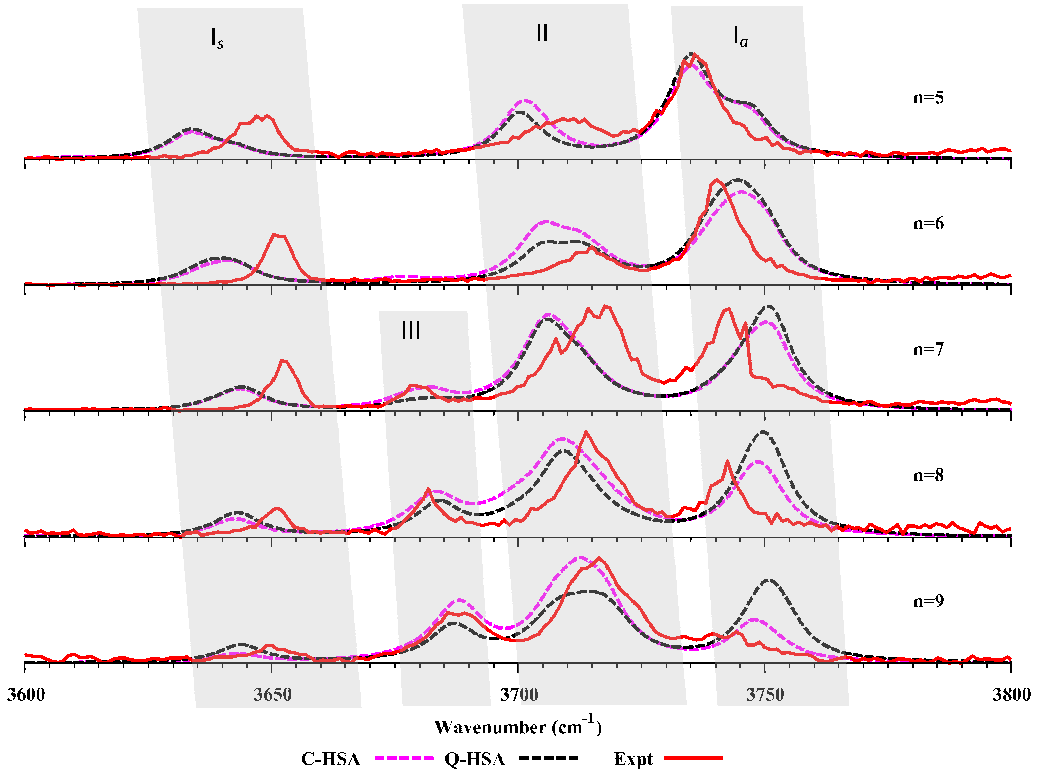


FIGURE 4.8: Vibrational spectra of free OH-stretching bands for $H^+(H_2O)_{n=5\text{-}9}$, calculated using classical (C-HSA) and quantum (Q-HSA) theory of harmonic superposition approximation. The experimental results of Ref [3] are also plotted in solid-red curve for the purpose of comparison.

denoted as I_s, I_a, II and III, respectively. As inferred from the figure, the theoretical spectra derived from both C-HSA and Q-HSA also indicated size dependency as observed in experimental research. The intensity of peaks I_s (around 3650 cm^{-1}) and peaks I_a (around 3750 cm^{-1}) decrease with increasing cluster size. Nevertheless, the relative decrease of peaks I_a are underestiated in our results, especially for $H^+(H_2O)_9$, with C-HSA showing better approximation accuracy than Q-HSA, even though the prediction is higher than in experimental research. The calculated spectra also agree with the experimental observation that the peaks III associated to the free-OH stretching band of three-coordinate (H_2O) has appeared at $n = 7$ and the intensity continue to rise consistently when n increases. As mentioned in the study of Lin,[3] this trend can serve as evidence of structural transition to ring isomers. Note that this agrees with the thermodynamic results discussed in the previous section where the ring isomers started to predominate the population

at $n = 7$. The blue-shift of the whole free-OH stretching bands witnessed in the experiment is also observed in our calculated spectra.

4.5 Conclusion

In this work, we have considered a hierarchical approach to thoroughly explore the PES of $H^+(H_2O)_{n=5\text{-}10}$, with OSS2 potential model and $n = 5 - 9$ at the B3LYP/6-31+G* level. The distinct isomer sets uncovered using our hierarchical methodology is the largest-ever archival found to date. The archival are subsequently used for the investigating the thermodynamic and structural transitions of $H^+(H_2O)_n$ at two theoretical levels, namely OSS2 model and B3LYP/6-31+G* level using the harmonic superposition approximation approach. In comparison with the parallel tempering Monte Carlo simulation (PT-MC) involving OSS2 model, our results revealed good quantitative agreement between HSA and PT-MC. The consistency with PT-MC results in structural transition and features of capacity curve are good indications of the HSA reliability. Further, the simulations with *ab initio* method also revealed the size dependency of $H^+(H_2O)_n$ in both thermal behaviors and vibrational spectra. The calculated vibrational spectra in free-OH stretching band when compared to recent experimental results also arrived at good agreements.

Chapter 5

Study of deprotonated water clusters

5.1 Introduction

Both deprotonated water clusters and their protonated companions have played essential roles in the studies of ionic media and acid-base reactions. The former have been investigated at a very early stage in the theoretical work of Bernal and Fowler.[119] In the last decade, numerous studies[50, 120–126] have been made to reveal the optimal structures, vibrational frequencies,[120] role of the first-shell solvation of OH^- anions[50, 125] and proton transfer.[121, 125] In particular, Xantheas[120] reported the geometries, vibrational frequencies and enthalpies of $OH^-(H_2O)_{n=1\text{-}3}$ at MP2 level of theory with aug-cc-pVDZ basis set and the decrease of OH^- bond length due to hydration was also addressed. Wei et al.[124] examined in detail the electronic structures of $OH^-(H_2O)_{n=1\text{-}3}$ using different methods. Del Valle and coworkers[122] performed tests to show the reliability of B3LYP functional against MP2 method and study the various properties of

$OH^-(H_2O)_{n=1\text{-}3}$. Vegiri et al.[125] studied systematically the hydration shell structure of clusters up to 15 water molecules using the canonical Monte-Carlo simulation with an empirical model function. In several experimental works, researchers have attempted to reveal the structures and H-bonding patterns of $OH^-(H_2O)_n$ via infra-red spectroscopic measurements.[127–130] Yang et al.[127] produced the large-sized clusters of $OH^-(H_2O)_{n=0\text{-}59}$ for the first time and uncovered the "magic numbers" at n = 11, 14, 17 and 20. Chaudhuri and coworkers,[128] on the other hand, systematically examined the infra-red spectra of $OH^-(H_2O)_{n=1\text{-}5}$ both experimentally and using DFT calculations to reveal the existence of several isomers. More recently, Robertson et al.[130] reported the vibrational spectra of $OH^-(H_2O)_n$ and identified the signatures of open solvation shell for hydroxide anion.

In recent years, the increased intensive activities on accurate experimental studies of small-sized clusters have led to significant theoretical effort in this field. Nevertheless, to date the amount of simulation study that brings insights into the morphological transitions of deprotonated water clusters remains limited. In the present work, we endeavor to fill in this gap by investigating the equilibrium structures, relative stability and vibrational spectra of $OH^-(H_2O)_n$ from small to mid sizes of n = 4 - 7 systematically at quantum chemistry levels. In our study, numerous isomers have been identified at different levels of first-principles calculations. Thermodynamic characteristics were subsequently derived in conjunction with vibrational analysis. The vibrational spectra were then calculated and compared with the recent experimental results. The influence of zero-point energy (ZPE) correction on the structures and properties of $OH^-(H_2O)_n$ was also examined.

From the fact that studying deprotonated cluster systems requires a large basis set than the protonated counterparts, different levels of first-principles calculations have been employed in the present work. The effect of basis set on the structures and thermodynamics are also discussed throughout the study.

One of the central issues in deprotonated cluster system is whether hydroxide anion can be treated as "proton hole". The scheme of proton hole has been discovered and investigated carefully in proton transfer process of solvated solution[131] and biological systems.[132, 133] In order to study how the scheme works in water cluster systems, the structures and thermodynamic properties of deprotonated water clusters in this work are compared to those of the protonated companions[60] in detail.

5.2 Isomers of deprotonated water clusters

The sets of 23, 70, 172 and 713 distinct isomers of $OH^-(H_2O)_n$ for n = 4 - 7, respectively, are identified at B3LYP/6-31+G* using the hierarchical search approach described in Chapter 2. According to the "compactness" of topological geometries, all isomers are classified into five structural categories, namely, multi-ring (MR), double-ring (DR), single-ring (SR), tree-like (T) or linear (L) forms. The isomers of MR, DR, SR are separated from the rest based on the numbers of smallest rings (n_{rings}), evaluated simply using the Cauchy formula: $n_{rings} = n_{edges} - n_{vertices} + 1$, where vertices and edges, in this study, are the oxygen atoms and hydrogen bonds, respectively. Several most stable isomers (up to ten) of each category are subsequently selected to be optimized further at the B3LYP/6-311+G(d,p) level. The utmost stable isomers of each category of B3LYP/6-31+G* and B3LYP/6-311+G(d,p), in terms of electronic energies, are depicted in Figure 5.1 and Figure 5.2, respectively.

For $OH^-(H_2O)_4$, the pyramid-shape isomer, $OH^-(H_2O)_4$-DR-I, depicted in Figure 5.1 is found to be the most stable structure. Nevertheless, upon ZPE corrections, two symmetric four-membered ring isomers, $OH^-(H_2O)_4$-SR-II and $OH^-(H_2O)_4$-SR-I, turned out to be the utmost and second most stable structures,

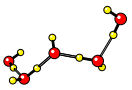
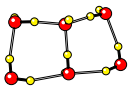
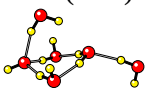
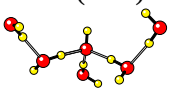
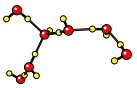
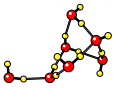
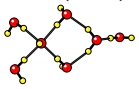
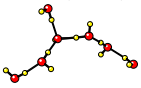
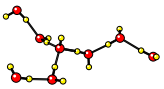
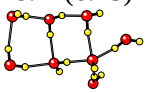
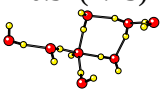
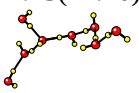
n		MR	DR	SR	T	L
4	I		 0.00(0.67)	 0.61(0.03)	 1.48(1.84)	 7.95(4.97)
	II			 0.71(0.00)	 3.08(2.37)	
5	I	 0.00(0.00)	 1.84(0.50)	 3.10(2.27)	 7.05(3.74)	
	II	 0.01(0.00)	 2.12(0.80)	 3.37(1.04)	 7.08(3.23)	
6	I	 0.00(0.69)	 3.73(2.13)	 5.91(3.72)	 9.16(5.80)	
	II	 0.23(0.00)	 3.76(2.24)	 6.47(3.63)	 11.65(7.25)	
7	I	 0.00(0.00)	 8.11(6.15)	 10.92(7.75)	 17.13(11.10)	
	II	 0.17(0.27)	 8.54(6.56)	 11.24(7.31)	 18.14(12.71)	

FIGURE 5.1: Low-energy isomers of $OH^-(H_2O)_{4\text{-}7}$, grouped according to their topologies, namely from left to right, MR (multi-ring), DR (double-ring), SR (single-ring), L (linear) and T (tree) and their relative electronic energies (E_0) optimized at B3LYP/6-31+G*. The numbers enclosed in parentheses denote the relative energies with ZPE correction. All values are shown in the unit of kcal/mol.

respectively. Consistent with previous calculations in Ref [128], the tree-like isomer ($OH^-(H_2O)_4$-T-I) with fully solvated shell is among the predicted low-energy isomers uncovered in the present study. Beside that, we found another tree-like isomer deviating from the former, merely in the direction of the free OH bonds. The existence of a linear isomer which has not been reported elsewhere, was also uncovered at high relative electronic energies of 7.95 kcal/mol.

For $OH^-(H_2O)_5$, in contrast to previous study,[128] the $OH^-(H_2O)_5$-SR-II isomer is no longer the most stable structure. To be more precise, it is well separated from the lowest-energertic isomer ($OH^-(H_2O)_5$-DR-I) with an energy gap of 3.37 kcal/mol. The ground-state isomers are the two MR structures which are almost iso-energetic since they differ from each other only in the free-OH directions. Similarly, two most stable isomers of DR form which also differs in the free-OH directions are quite close to each other with an energy gap of < 3 kcal/mol. The stable tree-like isomers of $n = 5$ become 3-coordinated instead of 4-coordinated as those of $n = 4$.

For $OH^-(H_2O)_6$, the geometry of the ground-state isomer is ($OH^-(H_2O)_6$-MR-I) which is consistent with the results computed with B3LYP/6-311++G**(sp) in Ref [134]. For $OH^-(H_2O)_7$, in contrast to the protonated counterpart of $H_3O^+(H_2O)_7$, the cubic isomer turns out to be the ground-state structure. Also the stable isomers of SR form are now five-membered ring instead of four-membered ring for $n < 7$.

In comparison to the $H_3O^+(H_2O)_n$ counterparts,[60] overall the isomers of $OH^-(H_2O)_n$ are generally observed to be more compact with multi-ring isomers emerging as utmost stable for smaller cluster size ($n = 5$) and none of the linear isomer exists except for $n = 4$. This implies that the length of the hydrogen bond in $OH^-(H_2O)_n$ is likely to be shorter than that of $H_3O^+(H_2O)_n$. It is worth noting that most of the relative energies with ZPE correction (i.e., those enclosed in parentheses in the figure) are lower than those without ZPE correction. This is inline with the observation obtained in the previous study on $H_3O^+(H_2O)_n$ clusters.[60]

Figure 5.2 depicts the stable isomers of different topologies similar to those in Figure 5.1, but optimized at a higher basis set of B3LYP/6-311+G(d,p). Overall, the geometry of isomers are similar to those found at B3LYP/6-31+G*, especially

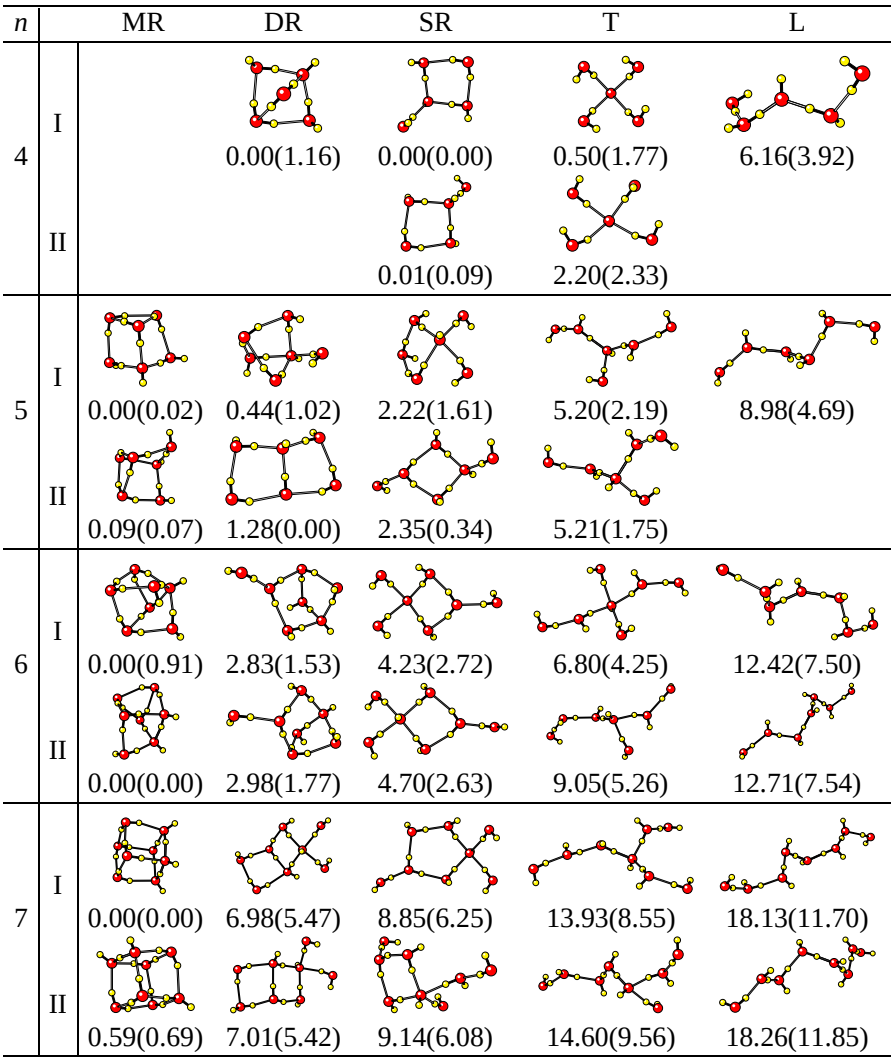


FIGURE 5.2: Low-energy isomers of $OH^-(H_2O)_{n=4\text{-}7}$ optimized at B3LYP/6-311+G(d,p). The explanations are provided in Figure 5.1.

for the ground-state structures. The isomers are noted to become more open with the existence of linear structures across the different sizes. It appears that the 6-311+G(d,p) basis set has the effect of weakening the strength of hydrogen bonds and increasing the bond lengths. Besides the observations on favoring of open isomers, they are also more iso-energetic as most of the relative energies are lower than those computed at B3LYP/6-31+G*. From the two figures of isomers, it is noted that the coordination number of anions are mostly 3 especially on the compact isomers (MR, DR), whereas 4-coordinated ones are only observed to exist in tree-like isomers.

5.3 Relative stability

The relative stability of isomers optimized at B3LYP/6-31+G* and B3LYP/6-311+G(d,p) are summarized in Figure 5.3 and Figure 5.4, respectively. The isomers are classified into 5 categories (MR, DR, SR, T, L) as discussed in the previous section. In these figures, each subplot, from left to right, depicts electronic energies (E_0), zero-point energies (E_{ZPE}), and total energies ($E_0 + E_{ZPE}$), respectively. For the sake of brevity, E_0, E_{ZPE} and ($E_0 + E_{ZPE}$) are shifted by taking the minimum as reference zero of the energy scale. All the values of energies are reported in the unit of kcal/mol.

Depicted in Figure 5.3, the electronic energy (E_0) generally displays a trend that favored the compact structures. The energy gaps between different topologies tend to widen with increasing cluster size. For instance, the gap between double-ring (DR) and single-ring (SR) increases from less than 1 kcal/mol ($n = 4$) to more than 2 kcal/mol ($n = 7$). In addition, the gap between MR and DR increases more rapidly, from around 2 kcal/mol ($n = 5$) to 4 kcal/mol ($n = 6$) and up to 8 kcal/mol ($n = 7$). In contrast, the ZPEs displayed an opposite trend where the open structures are more favored. This is in accord with the observation that the compact isomers are stronger in bondings, thereby possessing higher vibrational frequencies. Nevertheless, the energy gaps of ZPEs between the topologies are smaller than those in the electronic energies (E_0). Hence, ZPEs do not change the relative orders of topologies in terms of total energies which are displayed in the right-most column of Figure 5.1. It can be seen that except the case of $n = 4$, the others also displayed the "compact-favoring" trends as mentioned above. From this trend, the inclusion of the ZPE correction is found to have less significant influence on the structural transitions of deprotonated water clusters. However, ZPE correction generate the effect of lowering the transition temperature since it reduces the energy gaps between different topologies.

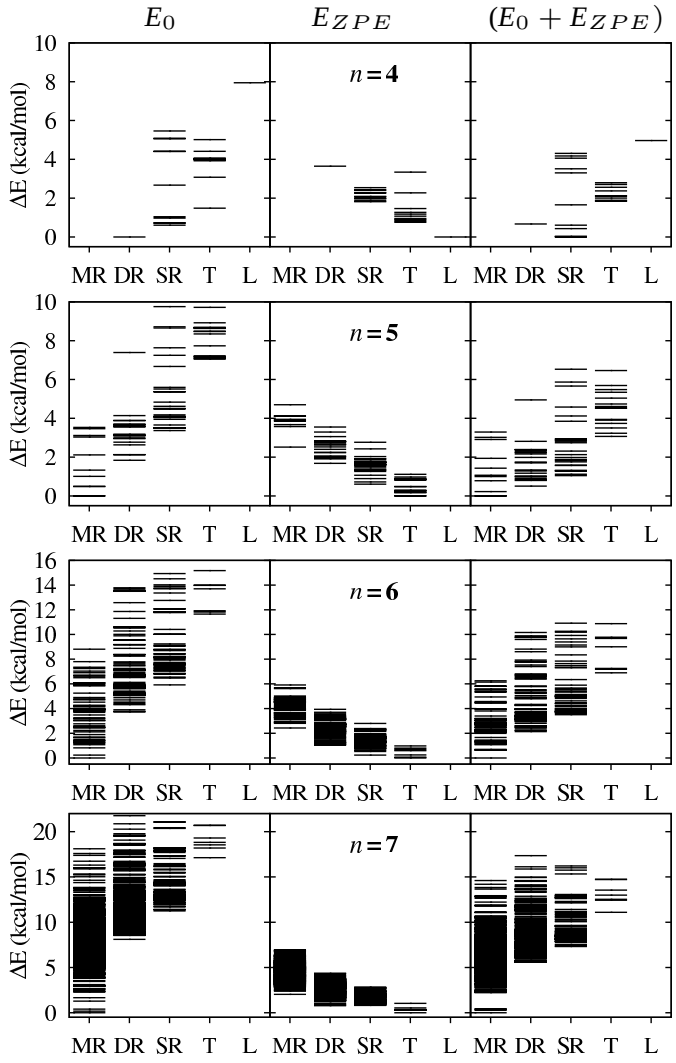


FIGURE 5.3: Relative stability diagrams of $OH^-(H_2O)_{n=4\text{-}7}$ isomers computed at B3LYP/6-31+G* with ascending order of water cluster size and from top to bottom. Each subplot, from left to right, depicts the electronic energies (E_0), zero-point energies (E_{ZPE}), and total energy ($E_0 + E_{ZPE}$), respectively. For details on the notations, the reader is referred to Figure 5.1.

The relative stability of isomers computed at the larger basis set of B3LYP/6-311+G(d,p) are plotted in Figure 5.4. Note that the diagram is observed to be more sparse since only ten isomers of B3LYP/6-31+G* were selected to be re-optimized at B3LYP/6-311+G(d,p). Overall, the "compact-favoring" trend is still conserved where the favored order in topology is MR, DR, SR, T and L as remarked above. However, the gaps between topologies, in terms of electronic energies, decrease significantly and the ZPEs have larger influence than those computed at B3LYP/6-31+G*. This indirectly results in more iso-energetic distinct-topological

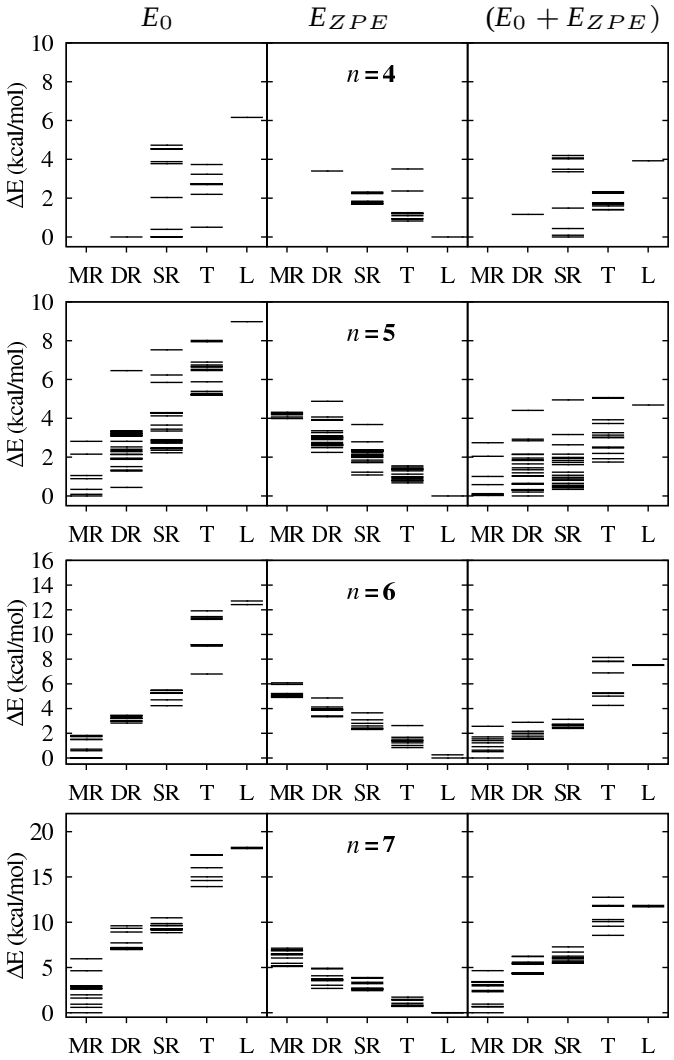


FIGURE 5.4: Relative stability diagrams of $OH^-(H_2O)_{n=4\text{-}7}$ isomers computed at B3LYP/6-311+G(d,p) with ascending order of water cluster size and from top to bottom. The explanations are provided in Figure 5.3.

isomers. For $n = 4$, $\Delta(E_0 + E_{ZPE})$ of SR/T and T/L are around 1.77 kcal/mol and 2.1 kcal/mol, respectively, whereas these values with B3LYP/6-31+G* are approximately 1.84 kcal/mol and 4.97 kcal/mol, respectively. For $n = 5$, the three most stable isomers of MR, DR, SR forms differs at 0.3 kcal/mol while the counterparts of B3LYP/6-31+G* are separated from each other by > 0.5 kcal/mol. Moreover, DR isomers turn out to be the ground-state structure instead of MR, as expected. The reduction in energy gaps can also be seen for $n = 6$, 7 although they are less significant compared to the smaller cluster size s ($n = 4$, 5). From the study, the equilibrium properties simulated with B3LYP/6-311+G(d,p) are

expected to be similar to the B3LYP/6-31+G* counterpart since the topological order is largely conserved. It is also expected that the transition temperatures are lower since the energy gaps between different topologies are smaller, that is similar to the effect of ZPEs mentioned in the last paragraph.

5.4 Thermal behaviors

The thermal behaviors of deprotonated water clusters are investigated by calculating the canonical heat capacity in both C-HSA and Q-HSA regimes.[60] The population or canonical probability of each topology is subsequently calculated as a function of temperature when studying the structural transitions. Figure 5.5 shows the HSA simulation results of $OH^-(H_2O)_{n=4\text{-}7}$ based on the archival of B3LYP functional isomers. The C-HSA result of $OH^-(H_2O)_4$ is depicted in Figure 5.5a. There are two maxima corresponding to the transitions of double-ring to single-ring (DR-SR) and single-ring to treelike (SR-T). In the results of Q-HSA counterpart (the right panels of Figure 5.5), only one maximum transition of SR-T is observed. This is consistent with the observation obtained in Section 5.3 since the DR isomers are no longer most stable upon undergoing ZPE correction. For $OH^-(H_2O)_{n=5\text{-}7}$, the characteristics of C_v is that the MR isomers predominate the population at low temperature which is then followed by the emergence of DR and SR isomers. The transition of SR-T occurs at slightly higher temperature. Because the transition points are quite close to one another, the overlapping transitions resulted in a broad and short peak in the heat capacity curves. The peak becomes sharper and shifts to higher temperature for increasing cluster sizes. In detail, the critical temperatures, T_c, of $OH^-(H_2O)_n$ for n = 5 - 7 are (175 K, 220 K, 260 K) and (110 K, 180 K, 240 K) for C-HSA and Q-HSA, respectively, as indicated in the figure. Individually, the shift can also be observed in each transition

of MR-(DR-SR) and SR-T. The size-dependency is analogous to the observations made in the study of $H_3O^+(H_2O)_n$.[2]

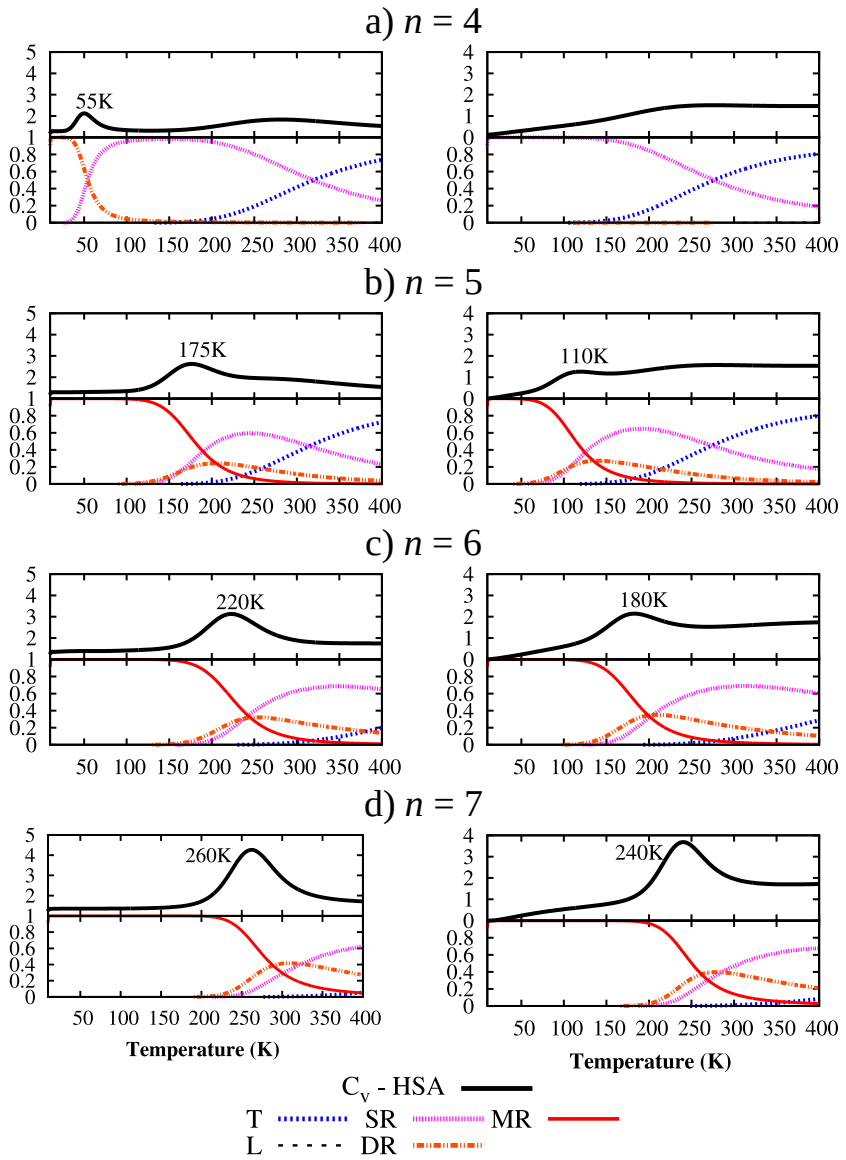


FIGURE 5.5: Canonical heat capacity C_v (upper panel) and population (lower panel) of five topologies of $OH^-(H_2O)_{n=4\text{-}7}$ calculated using classical HSA (C-HSA) in the left column and quantum HSA (Q-HSA) in the right column with B3LYP/6-31+G* calculation. The peak positions representing the transitions in heat capacity profile are reported above the peaks.

The difference between Q-HSA and C-HSA arises partly from the ZPE correction used in the former. As mentioned in Section 5.3, the inclusion of the ZPE correction does not change the relative stability among topologies except

the case of $OH^-(H_2O)_4$. This is further reflected in the thermodynamic simulations depicted in Figure 5.5. For $OH^-(H_2O)_4$, the ZPE correction changes the relative order of stability with single-ring replacing the double-ring structures as most stable isomers. This explains the different thermal behaviors of $OH^-(H_2O)_4$ for Q-HSA and C-HSA. For $OH^-(H_2O)_{5\text{-}7}$, however, the topological transitions of C-HSA and Q-HSA are very similar. The transition temperatures of Q-HSA are lower than the C-HSA counterpart as observed in Figure 5.5. This can be explained by using the relative stability diagrams in Figure 5.3 and Figure 5.2 where the ZPE correction reduces the energy gaps between different topologies, thus reducing the entropy of low-energy topologies and resulting in the shifting of the transition peak to lower temperature. Generally, this results the peaks in Q-HSA to occur at lower temperature.

In comparison to the study in Ref [2], the thermodynamic characteristics of both protonated and deprotonated water clusters are relatively similar, i.e., the transitions of MR-(SR-DR) and SR-T are observed in both systems. The size-dependency is also a common feature. Nevertheless, there exists some differences. The DR isomers of $OH^-(H_2O)_n$ exist with a greater proportion in the population. The topological transitions of $OH^-(H_2O)_n$ also happened at higher temperature than the companions of $H_3O^+(H_2O)_n$. This is due to the fact that the compact isomers in $OH^-(H_2O)_n$ are more favored. In contrast to $H_3O^+(H_2O)_n$, the ZPE correction has little or almost no influence on the thermodynamics since the energy gaps between the compact (MR,DR) and open (SR, L , T) topologies are much larger as noted in Section 5.3. This observation would be useful for future studies on large-sized systems since calculations of ZPE involving the vibration analysis are significantly more expensive in computation.

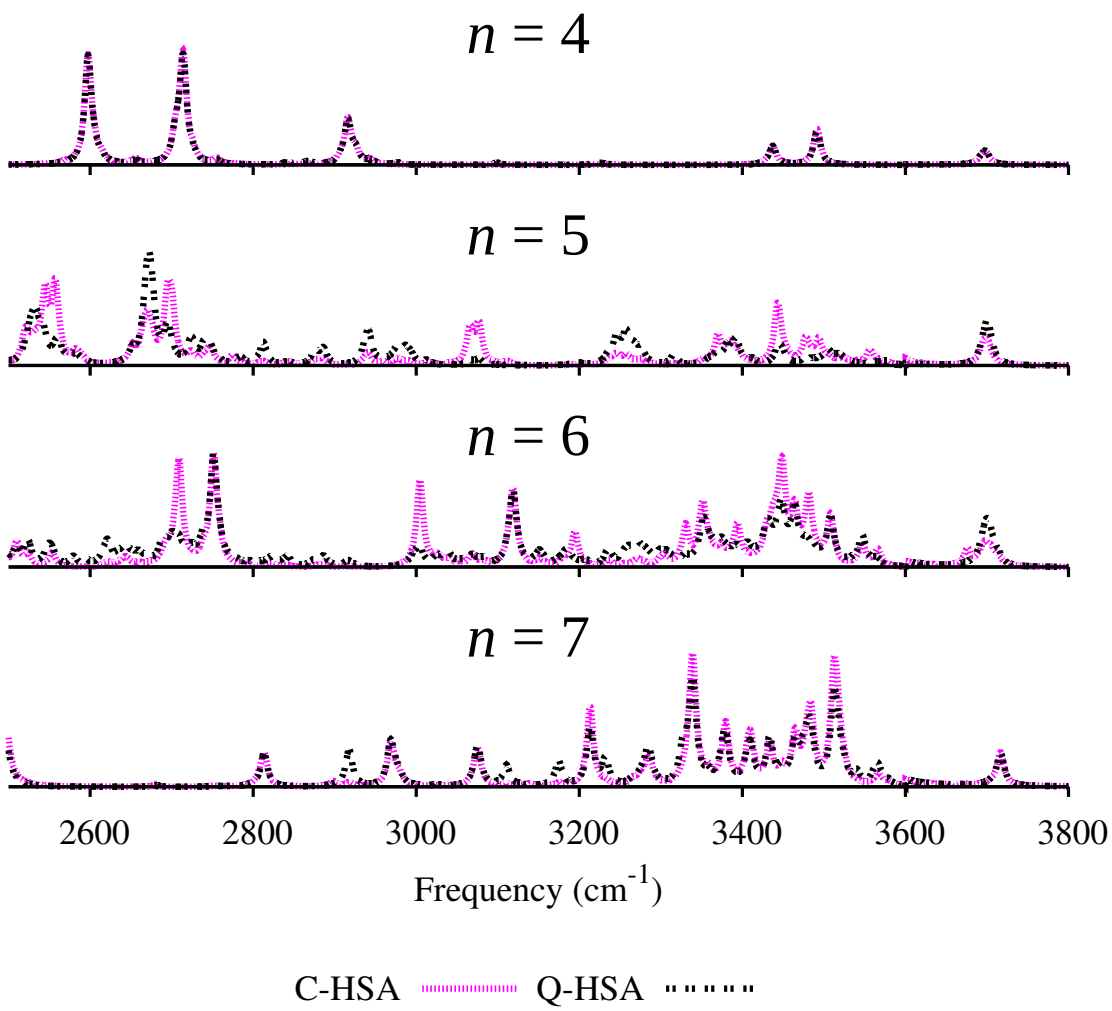


FIGURE 5.6: Vibrational spectra of $OH^-(H_2O)_{n=4\text{-}7}$ calculated using C-HSA and Q-HSA with B3LYP/6-31+G* calculations

5.5 Vibrational spectra

The vibrational spectrum, $I_{total}(\omega)$, is calculated as the sum of the spectrum of each individual isomer weighted by its corresponding canonical probability as follows:

$$I_{total}(\omega) = \sum_a I_a(\omega) p_a(T)$$

,

where the sum is taken for all identified isomers, $I_a(\omega)$ is the spectrum of isomer a derived from the vibrational analysis and $p_a(T)$ denotes the canonical probability

computed in both classical and quantum HSA as described in Section 5.4. Lorentz line-shape approximation is then applied with the half-width wavelength of 10 cm^{-1} for retrieving the continuous spectra of $OH^-(H_2O)_n$ which are depicted in Figure 5.6. To be consistent with the experimental results[128–130] and those of protonated counterpart,[60] the simulations are also carried out at the temperature $T = 170$ K in the present study. Note that the intensities of spectra in free OH stretching band (3600 - 3800 cm^{-1}) are magnified by 5 times for the sake of brevity. The vibrational frequencies in the figure are calibrated by scaling down with a factor of 0.973 as suggested in Ref [118].

In the free OH-stretching band, it is clear that the vibrational spectra of $OH^-(H_2O)_{4\text{-}7}$ are featured with a broad peak around 3700 cm^{-1}. Note that this is a deviation from the protonated counterpart in previous experimental [3] and theoretical[60] studies, since the spectra of $H_3O^+(H_2O)_n$ are well resolved and shown to depend on the coordination numbers. Nevertheless, it is in accord with the observations made in the experiment for $n = 4$, 5 in Ref [128–130]. And this feature is also observed to persist for larger cluster sizes ($n = 6$, 7). As shown in the previous analysis,[128] the broad peaks in the spectra are due to an overlapping of the normal modes in this band that are particularly too close to be resolved from one another. It may arise from the co-existences of many distinct-topological isomers at $T = 170$ K which are observed in the thermodynamic simulations. In addition, the peaks are noted to shift slightly to higher frequency with increasing cluster size, which is analogous to the size-dependency of the thermodynamics. At present, there remains a lack of experimental results to validate or confirm this observation in $OH^-(H_2O)_n$ system. However, the size-dependency has also been witnessed in both theoretical and experimental studies of protonated water clusters.[3] In the region corresponding to hydrogen bond vibration, the spectra of $OH^-(H_2O)_{n=5\text{-}7}$ contains a significant number of peaks. Note that this is consistent with the observation obtained in the experiment of Chaudhuri et al.[128]

where the spectra in this band were also severely broadened. In comparison with those of $H_3O^+(H_2O)_n$, the spectra of $OH^-(H_2O)_n$ are more complicated with more peaks observed. This suggests the possibilities of more distinct-topological isomers in $OH^-(H_2O)_n$ coexisting at $T = 170$ K than in $H_3O^+(H_2O)_n$, such as the involvement of DR isomers beside the MR and SR structures as noted in the previous section. Nuclear quantum effect has a minor effect on the spectra since the overall spectra of Q-HSA are very similar to those of C-HSA as depicted in the figure. This observation can be inferred from the small influences of ZPE correction on relative stability and the thermal behaviors as discussed in the previous sections.

5.6 Conclusion

In this work, we have systematically studied the structures, relative stability, thermal behaviors and vibrational spectra of $OH^-(H_2O)_n$ for cluster sizes of $n = 4-7$ using DFT calculations. The comparison against protonated counterparts has highlighted the structures of $OH^-(H_2O)_n$ to be more compact since the formation of 3D structure occurred at smaller sizes and the energy gaps between compact and open structures have been found to become larger. The dissimilarity in the ground-state structures and structural orders indicates that the structures of $OH^-(H_2O)_n$ could not be inherited simply from protonated water clusters through flipping OH bonds and considering OH^- anion as "proton hole". The thermodynamic simulations derived via both classical and quantum HSA displayed similar trends of topological transitions, despite some minor deviations between them. In detail, for $n = 7$, the cubic structures were reported as the utmost stable structures and dominated the population at temperatures $T < 200K$. The ZPE correction has also been investigated and concluded to have little impact on relative stability, structural transitions as well as on the vibrational spectra of $OH^-(H_2O)_n$. The

size-dependency in thermal behaviors and vibrational spectra, which are analogous to those of $H_3O^+(H_2O)_n$ was also recognized.

Chapter 6

Study of neutral water clusters

6.1 Introduction

Neutral water clusters have been investigated extensively for a long period because they provide important understanding of properties of water molecules in aqueous media.[4, 12–19, 19–24, 24–27, 108, 135] A number of empirical potential models have been developed for estimating the interaction energies and reproducing the ground state structures of first-principles calculations.[12, 70] To date, many researchers have focused on using the global minima to validate and compare the different potential models. Lee et al.[16] applied simulated annealing method with the empirical potential function of Cieplak, Kollman, and Lybrand to optimize water clusters up to $n = 20$. Using basin hopping, Wales and coworkers studied the TIP4P[17] and TIP5P[44] potentials for $n \leq 21$, and performed a comparison of the structures and formation energies obtained against MP2 calculation. More recently, Bandow and Hartke [15] developed a highly parallel evolutionary algorithm to study water clusters up to $n = 34$ on TIP4P and TTM2-F potentials.

It is worth noting that these earlier methods generally focus on global optimization with empirical models for large-sized systems to bridge the gap from single molecule to bulk materials.

Recently, Maeda et al.[4] proposed a full first-principles based approach for studying the structures and thermodynamic transitions of water clusters by searching for isomers instead of focusing only on the global minima. They have developed a scaled hypersphere search (SHS) method that employs an uphill-walking technique to locate the isomers sequentially. The set of 168 identified isomers for $(H_2O)_8$ at B3LYP/6-31+G(d,p) level was considered to be relatively large. The shortcoming of the proposed method however is its complete reliance on first-principles and second order derivative calculations. The high computational demand of first-principles calculations renders this method computationally infeasible for exploring the quantum chemistry PES with a large number of local minima.

In this study, we applied an alternative approach that synergizes empirical model with first-principles calculations. Our aim is to benefit from the low computational cost of empirical models by coupling it with first-principles calculation to explore the PES of water clusters at the quantum mechanical level. Consequently, the search for isomers becomes an efficient and highly accurate process. We demonstrated the approach with two sophisticated and flexible potential models, OSS2 and TTM2.1-F. The isomers of $(H_2O)_n$ in the range of $n = 4-8$ of these models were collected extensively using an asynchronous parallelized genetic algorithm and subsequently refined using B3LYP/6-31+G*. The details are described in the next section, followed by the results and discussions.

Size	OSS2			TTM2.1-F			B3LYP/6-31+G*
	I	II	III	I	II	III	
4	38	21	55%	14	10	71%	21
5	186	68	37%	33	23	70%	76
6	736	332	45%	154	95	62%	379
7	2700	1175	44%	639	406	64%	1443
8	5521	2455	44%	2331	1429	61%	3563

TABLE 6.1: Numbers of distinct isomers of OSS2, TTM2.1-F and B3LYP/6-31+G* for $(H_2O)_{n=4\text{-}8}$. The success rate (shown in column III) is the ratio between the number of successfully located isomer using B3LYP/6-31+G* (column II) to the initial number of isomers used in the empirical model (column I).

6.2 Structures of water clusters

The number of distinct isomers is an important index to reflect the complexity of PES. While there is no strict rules on the number of isomers, N_{isomer}, with system size, it was demonstrated in LJ clusters and TIP models that N_{isomer} would grow exponentially with system size n, that is $N_{isomer} \sim e^{\alpha n}$, where α is a system dependent constant.[14, 24, 57, 102] Here, the number of isomers identified using our GA search on both OSS2 and TTM2.1F models are tabulated in Table 6.1. Note that since a value of 0.96 is used as the similarity threshold index to remove duplicate isomers, it is possible for over-elimination of unique isomers to happen. Hence our results may be regarded as a lower bound for the number of unique isomers that exist. Nevertheless, the number of isomers we collected increases rapidly from dozens to thousands in the range of $n = 4 - 8$. The exponential dependence of isomer size n, is depicted clearly in Figure 6.1a and by linear regression we estimated the values of α for OSS2 and TTM2.1F to be 1.26 and 1.32, respectively.

To evaluate the efficiency of using empirical potential models to guide the exploration of PES at the quantum level, the success rate, the ratio between the number of successfully located isomer using B3LYP/6-31+G* to the initial number

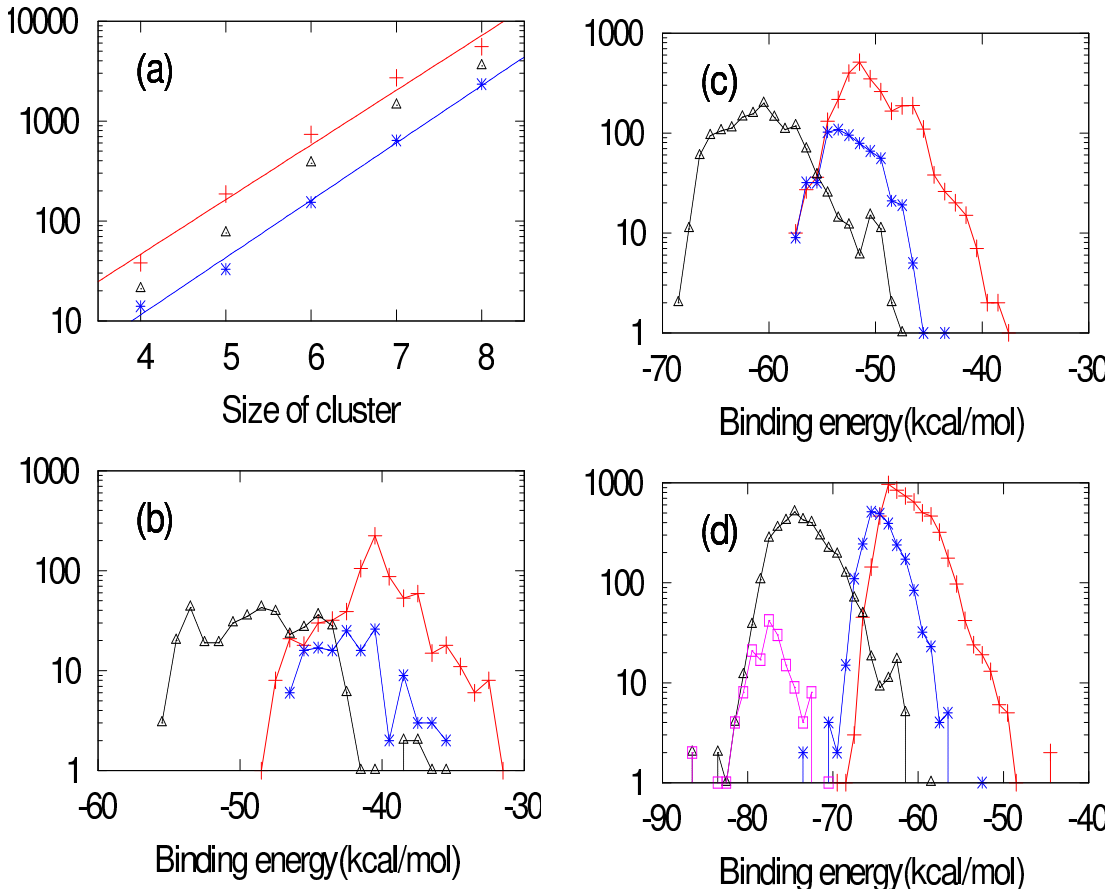


FIGURE 6.1: The number of isomers of water clusters in the range of $n = 4-8$ of OSS2 (red plus), TTM2.1-F (blue star) and B3LYP/6-31+G* (black triangle) versus the cluster size. The density of state (DOS) of $(H_2O)_6$, $(H_2O)_7$ and $(H_2O)_8$ in terms of binding energy (BE) (kcal/mol) are shown in (b), (c), and (d), respectively. In subplot (d), the DOS of $(H_2O)_8$ originated from Ref[4] is included for comparison (shown in pink square).

of isomers used as starting geometries, was computed. As shown in Table 6.1, upon relaxation using DFT, approximately 50 % of the isomers found in OSS2 successfully converged to equivalent-topology isomers. On the other hand, using TTM2.1-F displays a higher efficiency of $> 60\%$. It is worth noting that this also highlights the rugged PES of OSS2 over TTM2.1-F on $n = 4 - 8$. Nevertheless, by cross-referencing both empirical models simultaneously, the exploration on the PES of first-principles method has led to the identification of far more isomers than would be possible when using only any one of the empirical models in isolation. For example, the set of $(H_2O)_7$ isomers found increased to 1443 isomers compared to 1175 and 406 on OSS2 and TTM2.1F alone, respectively.

The density of states (DOS) for $(H_2O)_6$, $(H_2O)_7$, $(H_2O)_8$ are illustrated in the Figure 6.1b, c and d, respectively, to reflect the distribution of isomers on the energy scale. Note that calculations based on B3LYP/6-31+G* lead to a shift in DOS to larger binding energy in comparison with OSS2 and TTM2.1-F. This is as expected since both OSS2 and TTM2.1F have been parametrized based on MP2 binding energies, and it is commonly acknowledged that B3LYP method tends to overestimate the interaction energies. On the other hand, the overall features of DOS for all three models are similar. For example, in $(H_2O)_8$ the energy gaps separating the most stable cubic structure from the others is of similar value, with a Gaussian distribution of the DOS span of about 20 kcal/mol. This suggests thermal simulation of all three models would be comparable. Maeda et al. [4] independently examined the energy landscape of $(H_2O)_8$ and found 164 isomers. Here, we re-optimized their isomers using the B3LYP/6-31+G* method and the resultant DOS is depicted in Figure 6.1d. The figure indicates that Maeda and coworkers have generated a fairly good coverage of the low-energy region. Nevertheless, our results highlight a significant more complex energy landscape of $(H_2O)_8$ than was revealed earlier, since we discovered at least 2093 isomers.

The structures of the five most stable isomers for OSS2, TTM2.1-F and B3LYP/6-31+G* are depicted in Figures 6.2 and 6.3. For $n = 4$ and 5, the most stable forms for all three models are ring structures. For $n = 6$, the global minimum of OSS2 is also of ring form which is similar to the TIP5P model; the TTM2.1F global minimum has a cage-like form similar to TIP4P model whereas B3LYP/6-31+G* is of two-ring-membered form. For $n = 7$, TTM2.1-F and DFT calculations predict the cube-like structure with a missing corner as the most stable state while OSS2 retains the prediction of double ring structure. Overall, it can be concluded that OSS2 tends to favor more open structures rather than compact forms, while TTM2.1F displays an opposite trend. From the binding energies (by OSS2, TTM2.1F, B3LYP/6-31+G* and MP2/aug-cc-pvDZ) of these low-energy isomers,

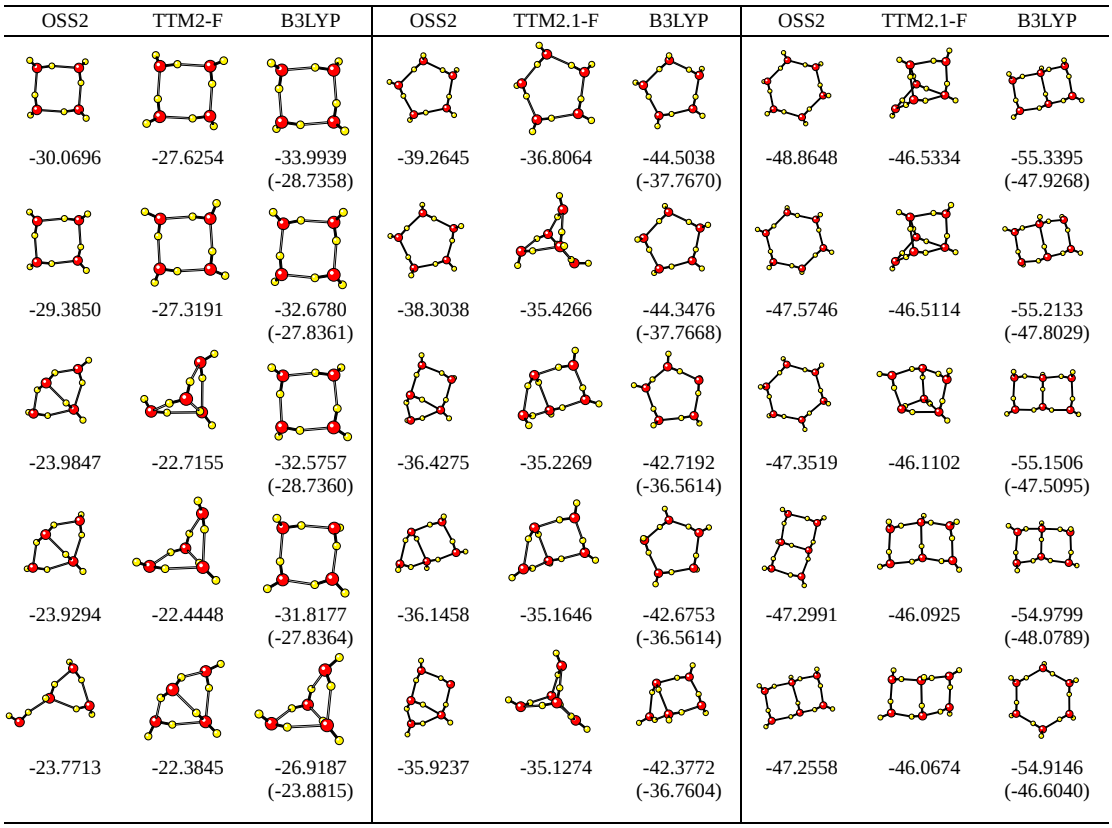


FIGURE 6.2: Molecular structures of the five most stable isomers in OSS2, TTM2.1-F and B3LYP/6-31+G*,at $n = 4, 5, 6$. The unit of the binding energy is reported in kcal/mol. Note that the binding energies of the B3LYP/6-31+G* isomers re-optimized using MP2/aug-cc-pvDZ are in brackets.

it is clear that the PES of water clusters are characterized by many iso-energetic isomers and it is beyond the scope of this work to make an extensive comparison of the empirical models. Nevertheless, the isomers we have archived will serve as a good starting effort to span greater investigation on using a multi-scale approach to explore the potential energy surface of water cluster. Furthermore, we have also provided the 5 most stable structures in this paper with the aim of facilitating future studies in the field.

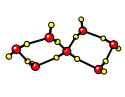
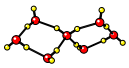
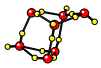
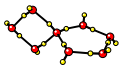
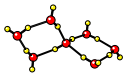
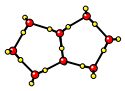
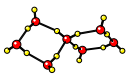
OSS2	TTM2-F	B3LYP	OSS2	TTM2.1-F	B3LYP
					
-57.7660	-57.8342	-68.5069 (-60.9106)	-69.1409	-73.3289	-86.7434 (-77.3861)
					
-57.7409	-57.5960	-68.0896 (-60.4691)	-68.6715	-73.2965	-86.6625 (-77.3663)
					
-57.5125	-57.5447	-67.7464 (-60.9106)	-67.1448	-70.9212	-83.2212 (-74.3787)
					
-57.4999	-57.3696	-67.7345 (-60.2692)	-67.0946	-70.8902	-83.2055 (-74.3560)
					
-57.4541	-57.1882	-67.5632 (-59.7587)	-67.0613	-70.1057	-82.0980 (-73.4505)

FIGURE 6.3: Molecular structures of the five most stable isomers in OSS2, TTM2.1-F and B3LYP/6-31+G*,at $n = 7, 8$. The unit of the binding energy is in kcal/mol. Note that the binding energies of the B3LYP/6-31+G* isomers re-optimized using MP2/aug-cc-pvDZ are in brackets.

6.3 Thermodynamics

From the sets of B3LYP/6-31+G* isomers, we applied HSA (see Chapter 3) to investigate thermal behavior of $(H_2O)_{n=5\text{-}8}$. To study structural transitions, isomers of $(H_2O)_n$ are also classified into five categories of topological families as similar to the studies of protonated and deprotonated water clusters (Chapter 4 and 5). The calculated heat capacity and population (canonical probability) of each topological families in both C-HSA and Q-HSA are shown in Figure 6.4.

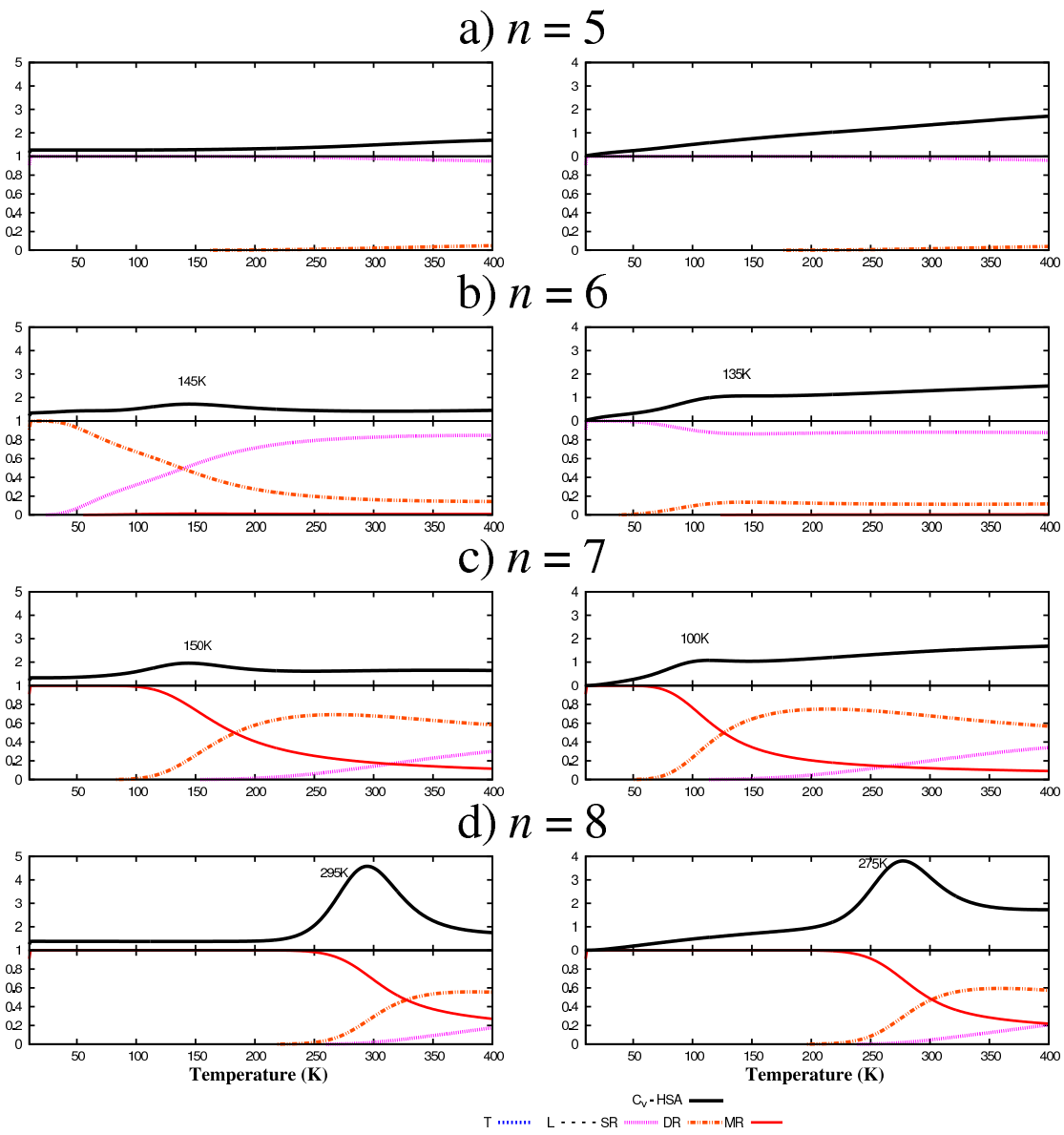


FIGURE 6.4: Canonical heat capacity C_v (upper panel) and population (lower panel) of five topologies of $(H_2O)_{n=5\text{-}8}$ calculated using classical HSA (left column) and quantum HSA (right column) with B3LYP/6-31+G* calculation.

At first, we discuss about the C-HSA results depicted in the left-hand column subfigures of Figure 6.4. For $n = 5$, there is no structural transition and SR isomers dominate the population. It is explained from the fact that SR isomers are well separated energetically from the rest in both DFT (2.2 kcal/mol) and MP2 calculations ($\sim$ 1 kcal/mol) as seen from Figure 6.2. For $n = 6$, DR is the most stable structure but the energetic separation between SR-DR isomers is relatively small (<0.3 kcal/mol in DFT, <0.007 kcal/mol in MP2). Therefore, DR

isomers dominate the population at low temperature due to the smaller average of frequencies (see Equation 3.8). We then observe a gradual change from DR to SR at about 145 K. For $n = 7$ and 8, three-dimensional isomers appear to be the most stable structures. As seen from the figure, MR isomers predominate the population which is then followed by the rising of DR and SR isomers. For $n = 8$, the transition of MR-DR occurs at 295 K implying that the MR isomers are thermodynamically very stable. This is consistent with other experimental and theoretical observations.[14, 17, 24]

Q-HSA results are shown in the right-hand subfigures of Figure 6.4. Overall, the characteristics of heat capacity and population are similar to C-HSA counterpart except that for $n = 6$, SR isomers are the dominant instead of DR ones. It is explained from the fact that ZPE correction is involved in the partition function of Q-HSA (Equation 3.7), and with the inclusion of ZPE correction, SR turn to be more stable than DR. For the same reason, the peaks on heat capacity plots representing the structural transitions generally occur at lower temperature than those in C-HSA. Similar to $H^+(H_2O)_n$ and $OH^-(H_2O)_n$, we can also observe the size-dependency of the thermodynamic properties of $(H_2O)_n$ where the critical temperatures of MR-(DR-SR) and SR-T transitions are shifted to the higher temperature with increasing cluster size.

6.4 Vibrational spectra

Similar to $H^+(H_2O)_n$ and $OH^-(H_2O)_n$ counterparts, vibrational spectra of $(H_2O)_{n=5\text{-}8}$ are also calculated by using Equation 4.1. To be consistent with the previous results, the cluster temperature of $T = 170$ K is chosen. To correct the anharmonicity of vibrational motion of water, the vibrational frequencies are calibrated by scaling down with a factor of 0.973 as suggested in Ref [118]. Note that the

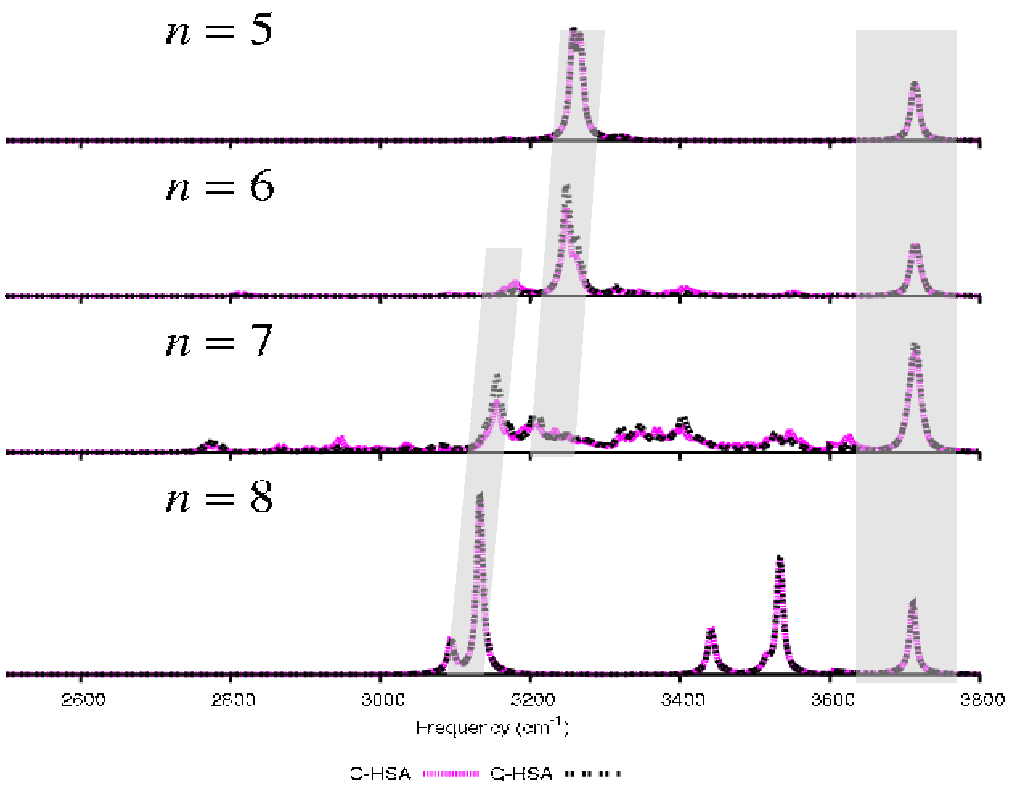


FIGURE 6.5: Vibrational spectra of $(H_2O)_{n=5\text{-}8}$ calculated using C-HSA and Q-HSA with B3LYP/6-31+G* calculations

intensities of spectra in free O-H stretching band (3600 - 3800 cm^{-1}) are magnified by 5 times for the sake of brevity.

In Figure 6.5, the free OH-stretching band gives a rise to an absorption peak at approximately 3750 cm^{-1}. The peak's position remains unchanged for all cluster sizes, $n = 5-8$ which is consistent with the results of other studies.[136–138] Unlike in the O-H stretching spectra of protonated water clusters,[3, 60] the peaks seem to be unaffected by the coordination numbers of water molecules participating in hydrogen bonds; but this feature is quite similar to the one of deprotonated water clusters described in Chapter 5.

In the range of hydrogen-bond vibration (2600-3600 cm^{-1}), the spectrum of $(H_2O)_5$ is characterized by a sharp peak originated from the hydrogen bond vibration of SR isomers. Since all water molecules of SR isomers are 2-coordinated, the peak can be assigned to the signature of hydrogen bonds in 2-coordinated

molecules. This peak is still persistent for the larger size ($n = 6$), where the configurations of $(H_2O)_n$ transform from 2-dimensional to 3-dimensional geometries. Besides that, a number of other small peaks corresponding to hydrogen bonds in 3-coordinated molecules appear in the spectrum of $(H_2O)_6$. It is related to the fact that DR isomers consisting of 3-coordinated water molecules are found to coexist with SR ones at $T = 170$ K, according to the thermodynamic analysis in Section 6.3. Therefore, the vibration of hydrogen bonds in three-coordinated molecules is involved in the spectra. For $n = 7$, SR isomers only occupy a small portion of the population whereas MR and DR isomers are dominant. Thus, the intensity of 3-coordinated hydrogen bond vibration increases relative to the intensity of 2-coordinated counterpart. For $n = 8$, the spectrum come from the cubic isomers where all molecules are three-coordinated; therefore, the peak of H-bonds in two-coordinated molecules is not observed. Another trend that can be observed is the red-shift of the peak's positions with increasing cluster size. It is highlighting that the strength of hydrogen bonds is reduced when the cluster's size becomes larger.

6.5 Conclusion

In this work, we demonstrated a multi-scale approach for exploring the PES of water clusters at the quantum chemistry level in the range of $n = 4 - 8$. Two flexible models TTM2.1F and OSS2 were used in the pre-screening process for identifying the probable locations of the isomers in the PES of B3LYP/6-31+G* calculation. The numbers of distinct isomers found using the present methodology are much larger than those reported in the literature highlighting the complexity of the PES of water clusters. The thermal behavior and vibrational spectra are simulated from the collected isomers. The spectral signature of coordination number of hydrogen bonds is analyzed. In this study, the moderate 6-31+G* basis

set has been chosen due to its high computational efficiency. Currently, we are extending the present work to a closely coupled multi-scale optimization approach where the higher level calculations are directly coupled to get rid of the basis set superposition error which results from the limited basis set method.

Chapter 7

Development of empirical models for water and hydrogen fluoride

7.1 Introduction

In the previous chapters, structures, thermodynamics and vibrational spectra of water clusters by using HSA have been studied. As mentioned, the accuracy of simulation results strongly depends on the number of isomers sampled on PES. And yet, the large databases of water isomers have been constructed by using hierarchical approach described in Chapter 2. As explained in Figure 2.1, the efficiency of hierarchical approach relies on the quality of the empirical potential used in the first level of exploration. The close agreement between empirical potentials and *ab initio* calculations will make the relaxation to the true local minima of *ab initio* calculations much faster. In this chapter, our aim is to develop new empirical potentials in order to improve the efficiency of hierarchical approach for further exploration with larger cluster sizes.

An empirical potential consists of two parts: a functional form and a set of parameters. We are particularly interested in OSS2 functional form (see Section 2.2

and Ref [70]) for three reasons. Firstly, it achieves a good agreement with first-principles calculations for protonated and pure water clusters as shown in the previous chapters. Secondly, the overlap of electronic charge clouds is modelled by screening functions instead of using artificial charges as in other models. Thirdly, it is an atomic model where there is no bonding constraints. Hence, it permits the disassociation of molecules. Particularly, it does not require two separated forms for modeling Zundel($H_2O_5^+$)[139] and Eigen (H_3O^+)[140] ions as others do such as Kozack-Jordan (KJ) potential.[66] In addition, it is designed for protonated water clusters but it surprisingly works well for pure water clusters as shown by original authors. According to our preliminary investigation, the functional form can be potentially used to model other systems such as deprotonated water, hydrogen fluoride clusters or mixed systems. For each molecular systems, we only sought to re-parametrize to retrieve a suitable set of parameters. It not only reduces the efforts for designing functional form but also provides a convenient way to compare the molecular interactions between different kinds of molecules and atoms since they are treated on the same footing as Lennard Jones potential which can be applied for many kinds of liquid and inert gas.

Parametrization to obtain suitable parameters is not straightforward especially with the complicated OSS2 potential with more than 40 parameters. To test our parametrization method, we firstly construct the potential model for a simpler system, protonated hydrogen fluoride clusters ($H^+(HF)_n$). Then, the parametrization for the potential of water clusters is performed. The hydrogen bond network and the morphology of $H^+(HF)_n$ clusters are simpler than the water counterparts due to the fact that the coordination number of fluorine molecules is 2 instead of 4 like water clusters. It is not only for testing the parametrization methodology but also to investigate whether OSS2 functional form is capable of modeling a system other than water clusters. The details of the work will be presented in the subsequent sections.

7.2 Potential model for protonated hydrogen fluoride clusters

7.2.1 Introduction

Hydrogen fluoride is established as one of the most important raw materials in the industry due to its crucial role in fluorine production. In research, it serves as a very popular reactant in laboratory and study. It has been long acknowledged as a key component of superacids.[141, 142] For instance, fluoroantimonic acid (SbF_5-HF), a 1:1 combination of hydrogen fluoride and antimony pentafluoride, has been recognized as the strongest superacid to date. The H_2F^+ species plays a significant role in the strong acidity of SbF_5-HF[143–147] as well as other mixtures including AsF_5-HF, NbF_5-HF, PF_5-HF, etc.[148] However, despite the importance in studying superacids, a limited number of experimental works on hydrogen fluoride have been reported due to its toxicity and corrosive chemical instinct.[149–153]

In theory, only few studies of hydrogen fluoride have been considered in the last decade.[148, 154] One of the core reasons lies on the lack of empirical potential models designed for modelling hydrogen fluoride. Empirical potentials server as more efficient and effective to simulate a molecular system as compared to first-principles calculations. They are also useful for studying the interactions within a system, which is comforta seem asusually very difficult for studying with first-principles calculation due to the complexity involved. To date, there are only few empirical models developed for HF clusters[155–159] although hundreds potentials have been developed for other H-bonding systems such as water clusters. Furthermore, for saving computational costs, most of existing empirical potentials developed are rigid models where the bond distances and angles are fixed as constants and an additional charged site is used to simulate the polarizability of

hydrogen fluoride molecules.[155, 156] Particularly, all the empirical potentials focus on neutral HF clusters, anionic HF clusters, HF-water mixture and HF liquid. From the best of our knowledge, to date, no polarizable potential model developed for protonated HF clusters.

Our goal in this study is to create a flexible, polarizable and disociable potential model for $H^+(HF)_n$ clusters. The new potential is supposed to accurately reproduce the geometry-optimized structures and formation energies at high-level *ab initio* leve for small to large-sized clusters. The potential model focuses on describing the proton transfer process in $H^+(HF)_n$ clusters since the high mobility of the proton in $H^+(HF)_n$ chains is the most important for studying the acidity of the hydrogen fluoride mixtures.[148] The great interest of the potential lies in the sophisticated functional form of OSS2 potential which is designed for protonated water clusters. Only the parameters are altered. So the success of this potential is prominent in developing a more sophisticated potential for a more complicated system such as $HF-H_2O$ mixtures which are crucial in ionic chemistry.

7.2.2 Generation of parametrization data

From sampling the potential energy surface of protonated hydrogen clusters, 9 sets of structural configurations for fitting are prepared and subsequently evaluated using the *ab initio* MP2 method with 6-311+G(d,p) basis set. All MP2 evaluations are conducted based on the Gaussian 03 package.[73] In total, approximately 3000 structural configurations are generated in various ways which are described in what follows.

To represent the intermolecular interactions, a set of configurations are obtained to parametrize the OSS2 potential using a procedure similar to that used in Ref [70]. Following this procedure, the potential energy landscape is sampled along normal modes or their linear combinations of each local minimum. First, the

ground-state structures of HF monomer and $H^+(HF)_n$ clusters with $n = 1-6$ are identified. This process is quite straightforward since these structures have been established in the form of chain-like morphology, with the HF^+ cation located at the center of the clusters. Second, harmonic vibrational analysis is performed on each structure to retrieve the normal mode displacement vectors. Subsequently, the new configuration points are generated from the ground-state structures by translating their atoms along those displacement vectors or the linear combinations of up to three displacement vectors. The translations are conducted on both positive and negative directions with a sufficient maximum displacement to guarantee the minimum energy span of approximately 2 kcal/mol. About 3 to 4 points are sampled along each direction depending on the cluster sizes. The configurations collected from this procedure are denoted as the sets of 1-7 tabulated in Table 7.1. Overall, a total of 2383 configurations is obtained.

In addition, significant number of configurations are generated to represent molecular interactions, including F-F interactions, torsion angle interactions, etc., that are not fully characterized in molecular vibrations. For F-F interactions, the configurations are generated by moving HF molecules of HF_2 clusters along F-F axis at an interval displacement of 0.01 Å from the equilibrium. The similar precedure is performed for $H^+(HF)_2$ cluster, where an additional 300 configurations is attained. In total, there are 371 configurations for F-F interaction profile, as denoted by the data sets 8 and 9 in Table 7.1. To represent the proton transfer in $H^+(HF)_2$ cluster, 61 configurations are generated by moving the center hydrogen atom of $H^+(HF)_2$ cluster from the equilibrium position along the F-F axis with an interval displacement of 0.2 Å. For torsion interaction, the torsion angle of $H^+(HF)_2$ is varied from 0° (equilibrium) to 90° with an interval of 1°. For the details, the reader is referred to the summary in Table 7.1.

7.2.3 Parametrization

The so-called HF-OSS2 potential developed in this study possesses the functional form inherited from the OSS2 potential model[70] which was originally designed to simulate solvated protons in water. OSS2 potential has been widely used in various studies of water[58, 60, 70]. For detailed descriptions on the OSS2 functional form, the reader is referred to Ref [70]. In this section, we present the modifications made to adapt the original OSS2 potential for modelling hydrogen fluoride clusters, $H^+(HF)_n$.

Data set	Description	Number
1	HF vibration	51
2	$H^+(HF)$ vibration	126
3	$H^+(HF)_2$ vibration	585
4	$H^+(HF)_3$ vibration	730
5	$H^+(HF)_4$ vibration	231
6	$H^+(HF)_5$ vibration	297
7	$H^+(HF)_6$ vibration	363
8	(HF)-(HF) interaction	71
9	(HF)-(H^+HF) interaction	300
10	(HF-H-HF) transfer	61
11	$H^+(HF)_2$ torsion	361
Total		2815

TABLE 7.1: Numbers of configurations of the data sets generated for fitting.

In our design, the charge of oxygen atoms (-2) in the electrostatic energy is replaced by the charge of fluorine atoms (+1). The polarizability of water monomer, α, is modified from 1.444 Å^3 to that of hydrogen fluoride monomer at 0.88 Å^3, by taking the value from experimental studies.

Two-body and three-body interactions remain unchanged in the parameterization process except that the parameters of r_0, and θ_0 denoting F-H bond length and H-F-H angle, respectively, are fixed at 0.96 Å and 1.9523 *radians*, respectively (see Table 7.2). The values of these two parameters are taken from optimized H^+(HF) monomer with MP2/6-311+G(d,p) calculations. These parameters are fixed as constants reference points as claimed in Ref [70]. They are redundant

due to the presence of the k_1 term. The rest of parameters are then fitted by using the Levenberg-Marquardt nonlinear least square algorithm[160] where thhe difference between the potential and *ab initio* MP2 calculation in binding energies are minimized. The objective function which is the root mean square of ($E^k_{OSS2}(p)$ - E^k_{MP2}) can thus be derived as

$$f_{RMS}(p) = \sqrt{\frac{1}{M}\sum_{k=1}^{M}(E^k_{HF-OSS2}(p) - E^k_{MP2})^2} \tag{7.1}$$

where p is the candidate parameter, M is total number of configurations, $E^k_{HF-OSS2}(p)$ and E^k_{MP2} are the binding energies of the configuration k obtained from HF-OSS2 and *ab initio* MP2 calculations, respectively. Here we considered the C++ implementation of Levenberg-Marquardt nonlinear least square algorithm[160] in Ref[161].

It is noted that the most computationally intensive potion of the HF-OSS2 potential lies in the three-body interactions which is a function of the forth-order polynomial series of bond angles and distances between each H-F-H combination multiplied by a cutoff term as follows[70]

$$\begin{aligned} V_{HFH}(r_1, r_2, \theta) =& f_{cutoff}(r_1, r_2, \theta)(k_1 + k_2(\Delta r_1 + \Delta r_2) + k_3\Delta\theta + k_4(\Delta r_1^2 + \Delta r_2^2) \\ &+ k_5\Delta r_1\Delta r_2 + k_6\Delta\theta^2 + k_7(\Delta r_1 + \Delta r_2)\Delta\theta + k_8(\Delta r_1^3 + \Delta r_2^3) \\ &+ k_9(\Delta r_1^2\Delta r_2 + \Delta r_1\Delta r_2^2) + k_{10}\Delta\theta^3 + k_{11}(\Delta r_1^2 + \Delta r_2^2)\Delta\theta \\ &+ k_{12}\Delta r_1\Delta r_2\Delta\theta + k_{13}(\Delta r_1 + \Delta r_2)\Delta\theta^2 + k_{14}(\Delta r_1^4 + \Delta r_2^4) \\ &+ k_{15}\Delta r_1^2\Delta r_2^2 + k_{16}\Delta\theta^4) \end{aligned} \tag{7.2}$$

where $\Delta r = r - r_0$ and $\Delta\theta = \theta - \theta_0$, with r_1, r_2 are the F-H distances and θ is the H-F-H angle. r_0 and θ_0 denote the equilibrium F-H distances and -H angles,

respectively. The number of H-F-H combinations is $N_F N_H(N_H - 1)/2$. Hence, the three-body interaction becomes extremely computationally inpensive on large-sized clusters. Nevertherless, it can be established that cutoff term, controlling the interaction range of three body interactions,

$$f_{cutoff}(r_1, r_2, \theta) = e^{-(m_0(\Delta r_1^2+\Delta r_2^2)+m_1\Delta\theta^2+m_2(\Delta r_1^2+\Delta r_2^2)\Delta\theta^2)} \tag{7.3}$$

decays exponentially to zero if the F-H distances are larger than r_0. In the bid to reduce the computatuonal effort involved, the f_{cutoff} term is computed first so that if it becomes small significantly ($< 10^{-20}$), this polynomial series can be neglected. In this way, the computational cost of energy evaluations is significantly reduced and scales from $O(N^3)$ to $O(N^2)$.

7.2.4 Results and Discussions

7.2.4.1 Binding energy correlation

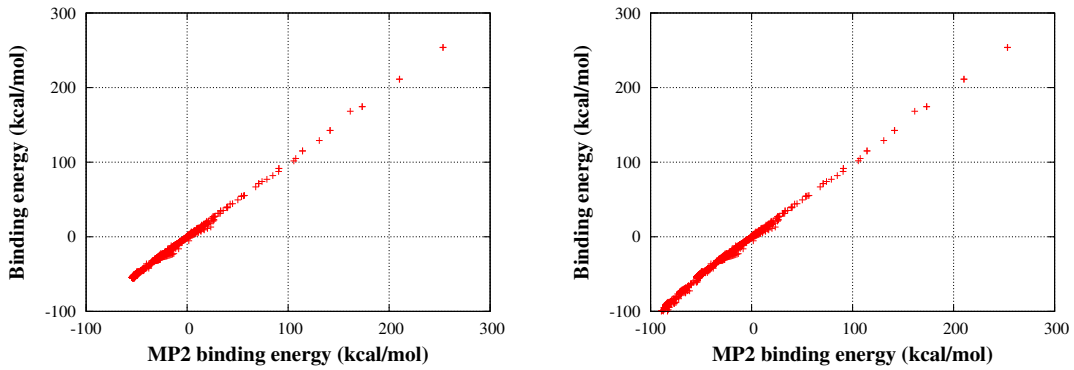
The multi-start search approach is used to optimize the HF-OSS2 potential. All parameters are first initialized randomly and subsequently optimized without any bound constraint by using Levenberg-Marquardt algorithm.[160] A total of 5 independent runs have been performed and all runs reached the convergences successfully within 1000 iterations. The best value of $f_{RMS}(p)$ obtained after fitting was less than 0.4 kcal/mol. This result was achieved twice over the 5 runs. The very small value of $f_{RMS}(p)$ can also be noticed in the chart of binding energy correlation between HF-OSS2 and *ab initio* MP2/6-311+G(d,p) calculations depicted in the left hand subplot of Figure 7.1. The data points almost lie exactly on the diagonal line as seen from the chart. We tabulate the fitted parameters in Table 7.2 with the same notation with OSS2 potential. For detailed descriptions of the parameters, the reader is referred to the study of Ojamae et. al[70].

Parameter	Value	Unit	Parameter	Value	Unit
r_0	0.960000	Å	θ_0	1.952300	rad
a_1	3.052465	Å^2	a_2	1.347178	Å^{-1}
b_1	332.285292	Å^2	b_2	2.846835	Å^{-1}
c_1	1.095686	Å^2	c_2	0.000002	Å^{-1}
h_1	0.002444		h_2	3.494190	Å
h_3	0.710774	Å^{-1}	h_4	3.936791	Å^{-1}
h_5	0.010923		o_1	39.023696	
o_2	1.138219	Å^{-1}	o_3	-41.187846	
o_4	1.155520	Å^{-1}	o_5	10.990913	
o_6	0.618724	Å^{-2}	o_7	-0.690755	Å
k_1	-0.020092		k_2	0.106009	Å^{-1}
k_3	-0.021725	rad^{-1}	k_4	-0.134639	Å^{-2}
k_5	-0.171111	Å^{-2}	k_6	-0.015599	rad^{-2}
k_7	0.036413	Å^{-1}rad^{-1}	k_8	0.102691	Å^{-3}
k_9	0.330649	Å^{-1}	k_{10}	0.036372	rad^{-3}
k_{11}	-0.120433	Å^{-2}rad^{-1}	k_{12}	-0.111249	Å^{-2}rad^{-1}
k_{13}	0.037841	Å^{-1}rad^{-2}	k_{14}	-0.019087	Å^{-4}
k_{15}	0.326415	Å^{-4}	k_{16}	-0.114364	rad^{-4}
m_1	1.310153	Å^{-2}	m_2	1.352019	rad^{-2}
m_3	0.242257	Å^{-2}rad^{-2}	α	0.880000	Å^3

TABLE 7.2: The parameters of the potential model for $H^+(HF)_n$ clusters obtained from fitting.

To investigate the performance of the optimized parametersof the HF-OSS2 potential on large-size clusters, further study involving the additional of more than 1500 configurations of $H^+(HF)_{n=7\text{-}9}$ clusters sampled via translating along the normal mode vectors as described in the previous section is conducted. The resultant correlation plot is depicted on the right hand plot of Figure 7.1b. As observed in the figure, the agreement between the HF-OSS2 potential and *ab initio* MP2 calculation is fairly good with the root mean square less than 2 kcal/mol. As expected, the deviation is largest in the low binding-energy range which is associated to large-size clusters ($n = 7 - 9$) but still acceptable since the data points still lie relatively close to the diagonal line. Note that the configurations of $H^+(HF)_{n=7\text{-}9}$ are not included in the data used for optimizing the parameters. This highlights the robustness of the HF-OSS2 potential in reproducing *ab initio* MP2 calculations resonably well for $H^+(HF)_{n=7\text{-}9}$ even though only the configurations

of $H^+(HF)_{n=1\text{-}6}$ are used in parametrizing the potential model.



a) Fitting b) Testing

FIGURE 7.1: Binding energy correlation of $H^+(HF)_n$ between HF-OSS2 potential and *ab initio* MP2/6-311+G(d,p) calculations with fitting and testing data sets.

7.2.4.2 Optimized geometries

Several basin hopping runs are also performed on each cluster size starting from random structures to identify the ground state structures of hydrogen fluoride clusters using HF-OSS2 potential with the optimized parameters. After running, all ground-state structures are identified as the linear chain morphology as expected. For comparison, the bond distances, angles and torsion angles of $H^+(HF)_n$ and HF clusters derivied using both *ab initio* MP2 calculation and HF-OSS2 are displayed in Figure 7.2. Note that only the bond distances of intramolecular covalent HF (D(F-H)), hydrogen bonds (D(F–H) and angles involved in the left-hand branches of the linear $H^+(HF)_n$ chains are presented in the figure since the cluster geometries are absolutely planar and symmetric.

Further, as observed in Figure 7.2, the bond parameters of small-sized clusters ($n = 1 - 3$) with the HF-OSS2 potential are significantly close to those obtained using *ab initio* MP2 calculations. The differences are small, with an average percentage deviations (on bond lengths and angles) of less than 1 % and 5%, respectively. Note that the bond angles seem to deviate more than the bond

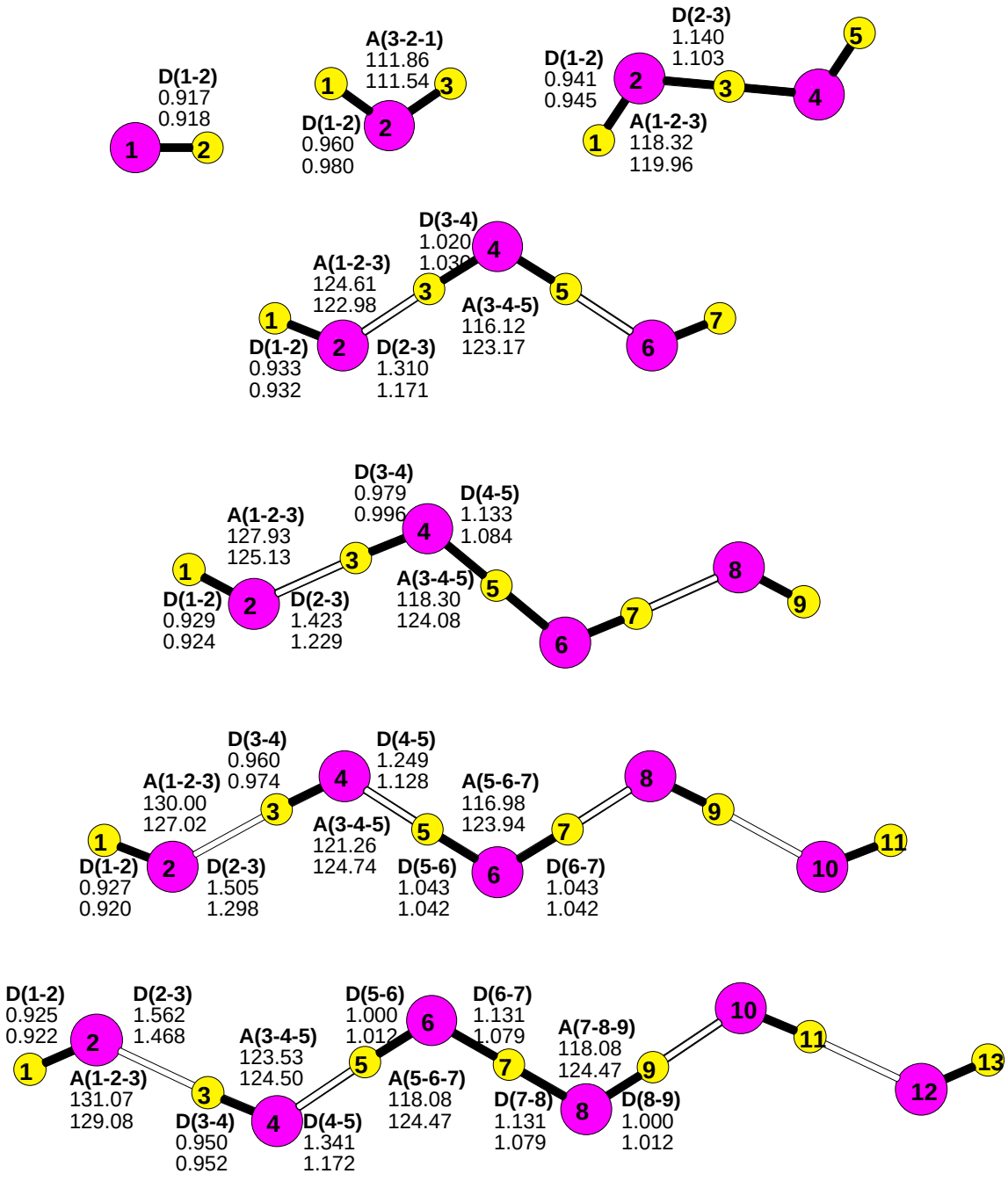


FIGURE 7.2: Optimized geometries of $H^+(HF)_n$ and HF clusters. The numerical data are the bond distances and angles calculated using MP2/6-311+G(d,p) calculations (top) and HF-OSS2 potential (bottom). The purple big circles are fluorine atoms whereas the yellow smaller ones are hydrogen.

lengths, but this is unsurprising as the functional form of HF-OSS2 emphasizes bond lengths involving in both two-body (F-F and F-H) and three-body (H-F-H) interaction terms, whereas the angles only form part of the three-body interactions. Hence, there are more configurations produced for bond distances in the data used to parametrize the potential. For larger-sized clusters, the difference between the results using the HF-OSS2 potential and MP2 calculations are observed to be larger than the small-sized clusters. Nevertheless, the covalent F-H distances are generally in close agreement with the MP2 results with an error of less than 0.1 Å while the hydrogen bond distances are overestimated slightly.

In the basin hopping simulations, other well-known morphologies of $H^+(HF)_n$ including ring or ring with tails appearing at $n > 5$ have also been observed. Above these known morphologies, a unique morphology of $H^+(HF)_{n=5}$ that has not been observed before (according to our knowledge) has been discovered in the present study. This newly discovered morphology as depicted in Figure 7.3 is relaxed under *ab initio* MP2/6-311+G(d,p) calculations with a root mean square gradient of less than 0.0001 Hartree/Å . It is then validated as a true stationary point since all eigen-values of the Hessian matrix are positive. It is energetically well separated from the ground state level with a relative energy of 5 kcal/mol. For larger-sized clusters ($n = 75$), the similarly morphological isomers were also discovered with high relative energy above the ground-state level. This morphology is particularly interesting due to the strange molecular arrangement where the top fluorine atom appears to be bonded weakly

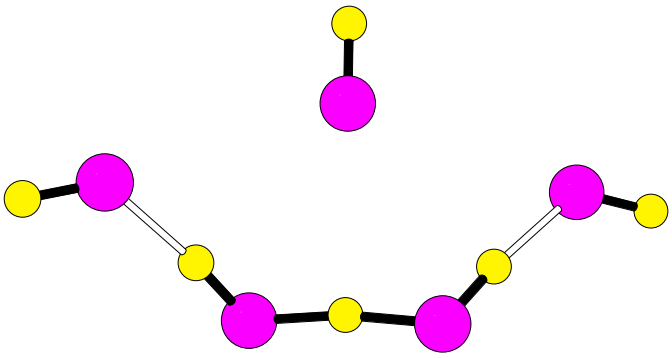
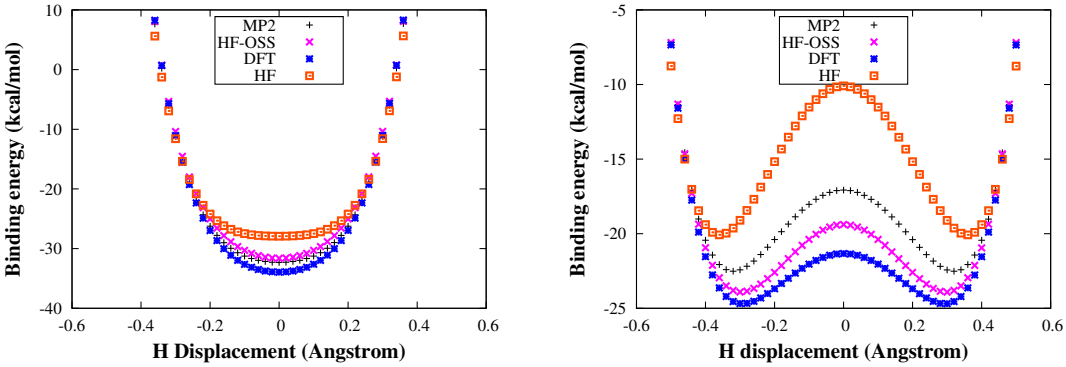


FIGURE 7.3: New morphology of $H^+(HF)_5$ is found under HF-OSS2 potential.



(a) D(F-F)=2.28 Å (b) D(F-F)=2.68 Å

FIGURE 7.4: Proton transfer energy profile in $H^+(HF)_2$ cluster where the proton is translated along F-F axis where two F atoms are placed in the equilibrium distance of 2.28 Å (a) and separated from each other by 2.68 Å (b). The x and y axes represent the relative displacement of moving proton from the middle point of O-O axis and binding energies, respectively.

to the surrounding fluorine atoms as observed from Figure 7.3. Although the bond distances between them exceed the range of regular hydrogen bonding ($\sim$ 1.5 - 2.2 Å), the fluorine atom can still be considered as solvated with 4 other ones instead of 1 or 2 as usual. From this discovery, HF-OSS2 potential is remarkable to in leading to the prediction an unknown morphology without any use of prior expert knowledge since the potential is simply parametrized using example configurations of linear-chain clusters.

7.2.4.3 Proton transfer profile

One of our interests in this work is to create HF-OSS2 potential that is deem as suitable for studying the proton transfer process in $H^+(HF)_n$ clusters. For this purpose, a subset of fitting data was created by translating the center proton in $H^+(HF)_2$ ions along F-F axis at equilibrium positions as discussed in the previous sectionSection 7.2.2. The energy profile with HF-OSS2 potential is depicted in Figure 7.4a. For comparison, the results of both *ab initio* calculations using MP2, restricted Hartree-Fock (RHF/6-311+G(d,p)) and density functional theory (DFT/6-311+G(d,p)) methods are also included. From the figure, the agreement

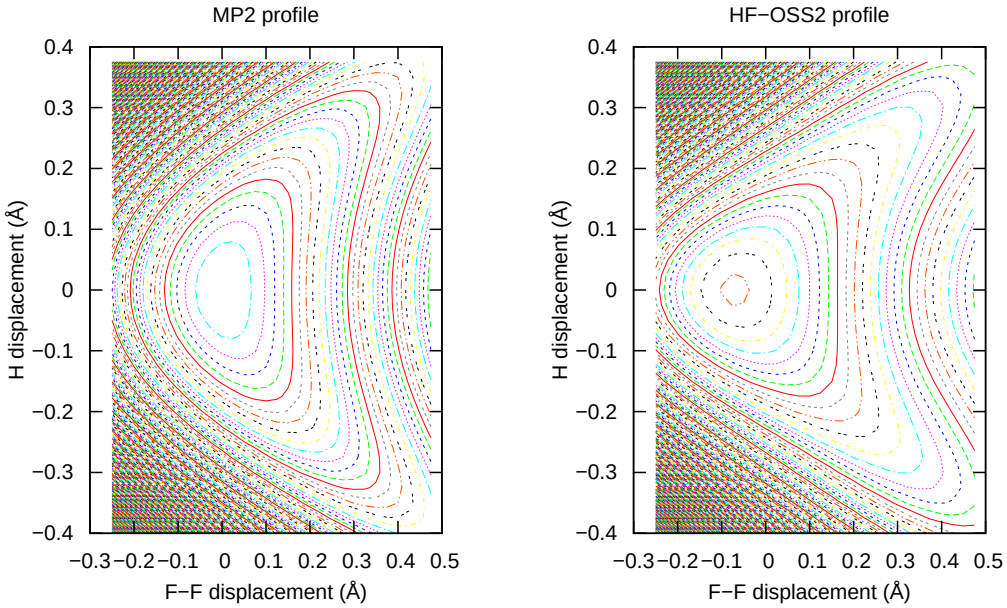


FIGURE 7.5: Contour plots of proton transfer profile of $H^+(HF)_2$ calculated in *ab initio* MP2 calculations (left) and HF-OSS2 potential (right). The x and y axes are the relative displacement of F-F distance and the position of the proton, respectively. The contour interval is 0.001 Hartree ($\sim$ 0.63 kcal/mol).

between the results obtained using HF-OSS2 potential and MP2 calculations are generally satisfactory since both predicted the depth of the potential curve as 32 kcal/mol. The DFT value of around 35 kcal/mol is also quite close to the MP2 value while the HF value is underestimated by approximately 5 kcal/mol.

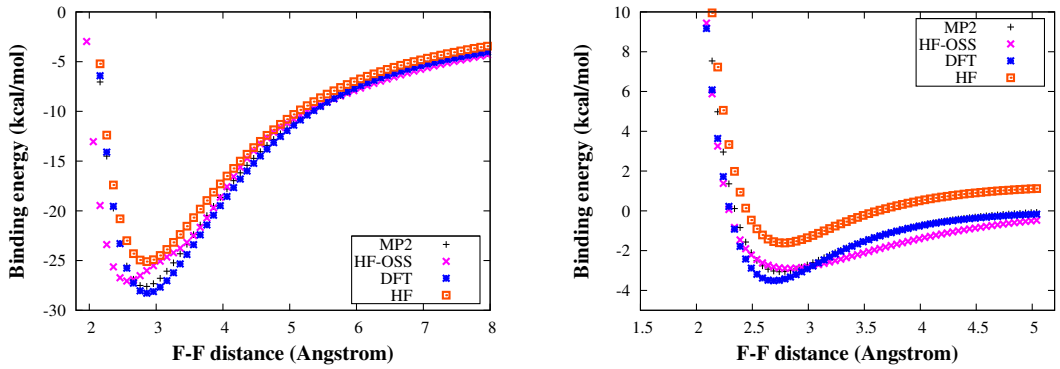
For further investigation, another set of configurations has been constructed in which the points were sampled in a similar procedure but two HF molecules were separated from each other by 2.68 Å (away from the equilibrium distance by 0.4 Å). The corresponding energy profile is depicted in Figure 7.4b. Unlike the previous case, there are energy barriers for proton transfer in $H^+(HF)_2$ ions. The agreement with MP2 calculations is not as high as previously due to the lack of those configurations to parametrize the model. The potential curves of HF-OSS2 potential and MP2 calculations are separated from each other by approximately 2 kcal/mol. Nevertheless, the height of energy barrier predicted using HF-OSS2

potential is consistent with those based on *ab initio* MP2 calculations. Both have value around 4 kcal/mol. Further, the result of HF-OSS2 potential remains good relative to those attained using DFT and RHF methods. The energy barrier predicted using DFT method (3 kcal/mol) is generally biased downwards whereas those by RHF (10 kcal/mol) is biased upwards.

It is crucial to test the performance of HF-OSS2 potential at the larger range of distances. For this purpose, we constructed a dense grid of nearly 1000 configurations where both the F-F distance and the position of proton of $H^+(HF)_2$ clusters are varied simultaneously. Figure 7.5 depicts the contour plots of the potential surface obtained using *ab initio* MP2 calculations and HF-OSS2 potential. The difference in elevation between successive lines is approximately 0.001 Hartree ($\sim$ 0.63 kcal/mol). The (0,0) origin corresponds to the equilibrium arrangement of the MP2-optimized $H^+(HF)_2$ clusters. However, the optimized F-F distance obtained from HF-OSS2 potential is smaller than MP2 value by 0.08 Å (see Figure 7.2). Hence the contour plot computed with HF-OSS2 potential is left-hand shifted relatively to the one computed with MP2 calculations. On the other hand, the shallowest directions referring to the reaction pathway of proton transfer in both two contour plots are consistent with each other. The potential energy surface with HF-OSS2 potential is slightly shallower. Overall, HF-OSS2 potential agrees well with *ab initio* MP2 calculations.

7.2.4.4 Interaction energy profile

Figure 7.6 shows the variation of the energy interaction respect with the F-F distance in two cases: between an $H^+(HF)$ ion and a HF molecule and between two HF molecules. The sampled configurations in both cases are involved in the fitting data as mentioned in Section 7.2.2. Hence, a very satisfactory agreement between the results obtained from HF-OSS2 potential and MP2 calculations have



a)H^+(HF) and HF b) HF and HF

FIGURE 7.6: The interaction energy between (a) an $H^+(HF)$ ion and a HF molecule and (b) two HF molecules. The x axis represents the distance of F-F.

been observed in the charts. In Figure 7.6a, the depths of potential wells computed using HF-OSS2 potential and MP2 calculations have roughly the same value of around 27 kcal/mol. The deviation between two curves is mainly at short F-F distances (within 2 - 2.5 Å). This is a result of the shorter equilibrium distance of F-F predicted using HF-OSS2 potential is shorter than the MP2 value as discussed in the previous section. Moreover, the interaction energies vary dramatically at short F-F distances, which makes the parametrizaton of the potential harder. Although, a straightforward measure is to sample sample more points in this region or assigning higher weights in Equation 7.1 to favour points in this region during the parametrization process. However, the final RMS obtained will be larger meaning that the agreement between HF-OSS2 potential and MP2 calculations become worse in overall. Figure 7.6b depicts the interaction profile of two HF molecules. The potential wells are shallower where the potential depths are only about 2 kcal/mol. As seen in the figure, the potential curve of HF-OSS2 potential matches nicely with the MP2 curve. However, the curve of HF-OSS2 potential is slightly underestimated compared to the counterpart of MP2 calculations when F-F distance is large. Comparing DFT results to HF ones, it is found that DFT calculations also achieve a quite close agreement with MP2 calculations. However, the curves of HF calculations are too far off.

Figure 7.7 displays the torsion or dihedral interaction profile where two fluorine atoms and the proton involved in the MP2 optimized $H^+(HF)_2$ cluster are placed at the equilibrium position while the dihedral angle of H-F···F-H are adjusted from 0° to 180°. As seen from the figure, the potential curves predicted by HF and DFT methods are too far off to MP2 calculation while the agreement between HF-OSS2 potential and MP2 calculation is very good in overall. The deviation between MP2 and HF-OSS2 curves appears only when the dihedral angle is around 0°. However, the maximum deviation is still less than 1 kcal/mol. On the other hand, the two curves match perfectly with each other at large dihedral angle.

7.2.5 Summary

In this work, a potential model to simulate protonated hydrogen fluoride clusters has been proposed. The potential is based on OSS2 functional form which was originally designed for protonated water clusters by Ojamäe and coworkers.[70] It acquires the characters from OSS2 model such as polarizable and disassociable. Several modifications have been introduced to adapt to $H^+(HF)_n$ clusters. The potential was then derived by fitting to the electronic surface of small-sized $H^+(HF)_{n=1\text{-}6}$ clusters at *ab initio* MP2/6-311+G(d,p) calculations. Our results show that the fitted potential reproduces correctly ground-state structures and binding energies at *ab initio* MP2 calculations for n up to 9. It is demonstrated to describe correctly the interaction in proton transfer process of protonated hydrogen fluoride dimers. The test on intramolecular interactions also

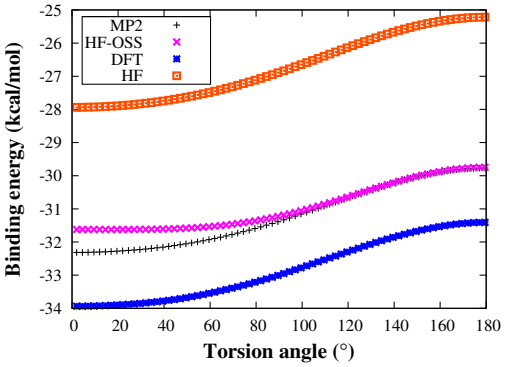


FIGURE 7.7: Dihedral interaction profile of $H^+(HF)_2$ cluster. The x axis represents the dihedral angle of H-F···F-H.

reveal a good agreement of the potential with MP2 calculations. The good results open a possibility to extend the model to be capable of simulating other hydrogen bonding systems such as the mixed clusters of hydrogen fluoride and water in the future.

7.3 Potential model for deprotonated water clusters

7.3.1 Introduction

A tremendous number of water models have been developed for decades. Nevertheless, most of attempts have focused on neutral clusters or bulk water due to their central roles in chemical and biological processes while only limiting number of potential models have been designed to simulate ionic clusters. For computational efficiency and simplicity, water molecules are usually treated as rigid in most of potential models. In several flexible potential models, bond distances and angles of water molecules can vary but they are still unable to allow dissociations of water molecules.[18–20, 69] The lack of this feature makes studying the proton transfer and dissociations and reassociations of water clusters difficult. To fulfill this requirement, several dissociable potentials have been created.[162, 163] Among those attempts, Ojamae and coworkers have developed a family of dissociable, polarizable potentials (OSS1, OSS2, OSS3) for simulation of water with excess proton. Those potentials were originally designed for protonated water clusters. However they can be used for neutral counterparts because they are purely atomic models without using any molecular constraints. In fact, OSS potentials have been engaged in several studies of protonated[2, 3, 60] as well as neutral water clusters[1, 58], showing a relatively good performance.

We realized that OSS potentials can also be used for deprotonated water clusters. In Chapter 5, OSS2 potential have been used in the AMA to identify distinct local minima of $OH^-(H_2O)_n$. However, the performance of this potential for $OH^-(H_2O)_n$ was not good as for $H^+(H_2O)_n$ since the original parameters of OSS2 potential were fitted to the potential energy surface of $H^+(H_2O)_2$. Although the ground-state structures of small-sized $OH^-(H_2O)_n$ clusters optimized by *ab initio* MP2 calculation still can be reproduced by OSS2 potential but the bond distances and angles of water molecules are quite far from those generated by *ab initio* MP2 calculation. As claimed by the authors,[70] OSS2 potential with current parameters usually exhibits too large H-O-H angles compared to *ab initio* MP2 results. It suggests the necessity to refine the current parameters to be used for $OH^-(H_2O)_n$. Another important motivation is that to our best knowledge, there has been no potential model particularly designed for deprotonated water clusters until now. In this work, we try to fill this gap by creating a new set of parameters suitable for deprotonated water clusters. OSS2 potential model is reparameterized to the geometry configurations from *ab initio* MP2 calculation. After that, the performance of the model with new parameters will be tested and reported in details.

7.3.2 Generation of parametrization data

Seven data sets of the structural configurations for parametrization are constructed by following the similar sampling strategy described in Section 7.2.2. In total, there are 1708 configurations generated in different manners. The detailed numbers of configurations of the data sets are summarized in Table 7.3.

More than 1000 configurations are obtained by sampling along the vibrational normal mode displacement vectors. In detail, we performed vibrational analysis on (H_2O), OH^- and $OH^-(H_2O)$ clusters to yield their normal mode displacement

vectors. Subsequently, we translated the atoms of the clusters from the equilibrium positions along the normal mode vectors with the maximum displacement of 0.2 Å. Similar to the OSS2 parametrization for protonated water clusters in Ref [70], the potential energy surface of deprotonated water dimer was particularly focused. In addition to sampling along the normal mode vectors, we also sampled along their linear combinations of up to three vectors to obtain totally 913 points for $OH^-(H_2O)$.

For O-O interactions, one water molecule of optimized $(H_2O)_2$ cluster was translated along O-O axis with the step size of 0.05 Å while keeping other atoms fixed to obtain about 80 points. In the same fashion, 60 additional points were obtained by translating OH^- anion along O-O axis from the optimized geometry of $OH^-(H_2O)$. For torsion angle interaction, 360 points were generated by gradually changing the torsion angle of H-O···O-H of $OH^-(H_2O)$ from the optimized value by 1 ° per step. For describing the proton transfer, 182 configurations are generated by moving the center hydrogen atom of $OH^-(H_2O)$ cluster from the equilibrium position along O-O axis with the interval displacement of 0.01 Å.

Set	Description	Number
1	(H_2O) vibration	33
2	OH^- vibration	80
3	$OH^-(H_2O)$ vibration	913
4	OH^--(H_2O) interaction	60
5	(H_2O)-(H_2O) interaction	80
6	$OH^-(H_2O)$ torsion	360
7	(HO-H-OH) transfer	182
Total		1708

TABLE 7.3: Numbers of configurations of the data sets generated for fitting.

All sampled configurations underwent *ab initio* electronic structure calculations at Møller-Plesset second order perturbation theory (MP2) with the "augmented"

Parameter	Value	Unit	Parameter	Value	Unit
r_0	0.961400	Å	θ_0	1.817338	rad
a_1	8.562466	Å^2	a_2	1.776328	Å^{-1}
b_1	332.285292	Å^2	b_2	2.846835	Å^{-1}
c_1	1.095686	Å^2	c_2	0.000002	Å^{-1}
h_1	0.057677		h_2	4.222943	Å
h_3	0.193396	Å^{-1}	h_4	2.204317	Å^{-1}
h_5	0.033396		o_1	376.731836	
o_2	0.406371	Å^{-1}	o_3	-377.264771	
o_4	0.407325	Å^{-1}	o_5	34100.405473	
o_6	0.354191	Å^{-2}	o_7	-3.940189	Å
k_1	-0.048393		k_2	0.166865	Å^{-1}
k_3	0.013679	rad^{-1}	k_4	-0.292374	Å^{-2}
k_5	-0.651854	Å^{-2}	k_6	0.006988	rad^{-2}
k_7	-0.043325	Å^{-1}rad^{-1}	k_8	0.399625	Å^{-3}
k_9	1.148092	Å^{-1}	k_{10}	0.057061	rad^{-3}
k_{11}	-0.173036	Å^{-2}rad^{-1}	k_{12}	0.438720	Å^{-2}rad^{-1}
k_{13}	-0.089550	Å^{-1}rad^{-2}	k_{14}	-0.228694	Å^{-4}
k_{15}	0.000000	Å^{-4}	k_{16}	-0.029092	rad^{-4}
m_1	6.250000	Å^{-2}	m_2	0.003772	rad^{-2}
m_3	6.250000	Å^{-2}rad^{-2}	α	1.444000	Å^3

TABLE 7.4: The parameters of the potential model for $OH^-(H_2O)_n$ clusters obtained from fitting.

correlation consistent polarized valence basis sets (aug-cc-pVDZ). All MP2 energy evaluations were performed by Gaussian 03 package.[73]

7.3.3 Parametrization

The functional form of OSS2 potential model described in Ref [70] is used in this study without any change. The parameter representing the polarizability of water monomer, α, was still set to the experimental value of 1.444 Å^3[164] and fixed as a constant in the fitting process. Those parameters of m_1, m_2 and m_3 governing the localization of three-body interaction (H-O-H) are also fixed. Similar to the parametrization of HF-OSS2 potential (Section 7.2), the remaining parameters are also fitted by using Levenberg-Marquardt nonlinear least square algorithm.[160] The objective function to be minimized is the root mean square of

the difference of the potential and *ab initio* MP2 calculation in binding energies given in Equation 7.1

7.3.4 Results and Discussions

7.3.4.1 Binding energy correlation and optimized geometry

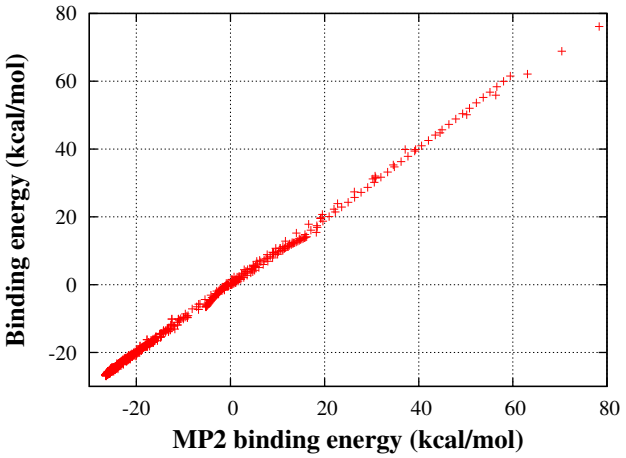


FIGURE 7.8: Binding energy correlation between OSS2(depW) potential and *ab initio* MP2 calculation.

We conducted ten independent fitting runs starting with different initial guesses. All of the runs converged successfully within a few thousands of iterations. In fact, different runs lead to different final results. In this work, we only report the best result attained from those fitting in which the final value of root mean square in binding energy, $f_{RMS}(p)$, was less than 0.5 kcal/mol. The final parameters are listed in Table 7.4. For the detailed explanation of the parameters, the reader is referred to Ref [70]. For the sake of brevity, OSS2 potential with the fitted parameters is denoted as OSS2(depW).

The binding energies computed by OSS2(depW) potential versus the *ab initio* MP2 energies are plotted in Figure 7.8. The energy span of fitting data of nearly 100 kcal/mol is very large. Nevertheless, the correlation between the computed energies and reference MP2 ones was achieved very well as seen from the figure.

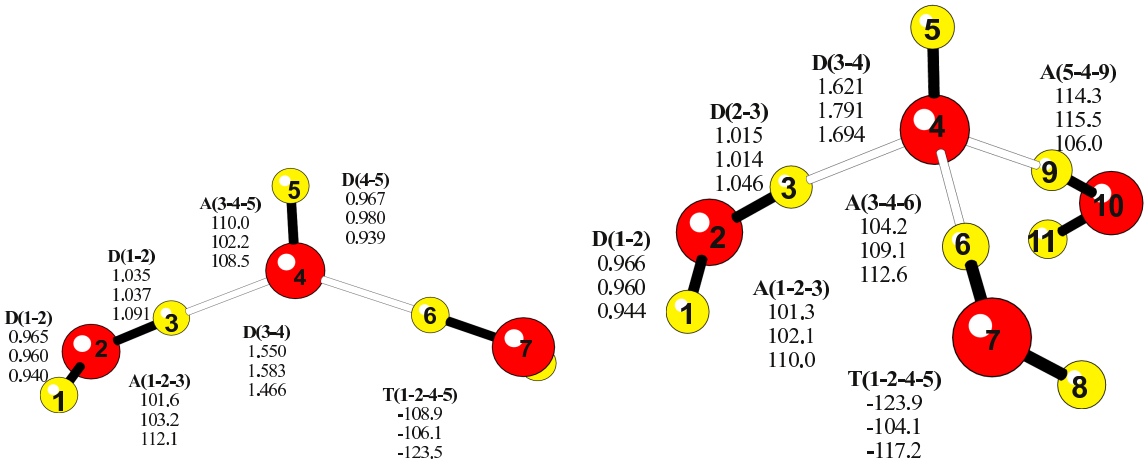


FIGURE 7.10: Optimized structure of $OH^-(H_2O)_3$ (C_2 symmetry) and $OH^-(H_2O)_4$(C_3 symmetry). The structural parameters for *ab initio* MP2, OSS2(depW) and OSS2 potentials are shown from top to bottom, respectively .

The deviation of the data points from the diagonal line is very small, indicating the good convergence of the fitting process.

The structures of deprotonated water monomer, dimer, trimer and tetramer optimized with *ab initio* MP2 calculation were re-optimized using OSS2 potential with both original and fitted parameters and depicted in Figure 7.9 and 7.10. For comparison, the internal coordinates for each models are shown together in the figures. Because of the symmetry of the cluster, only necessary bond parameters are represented in these figures.

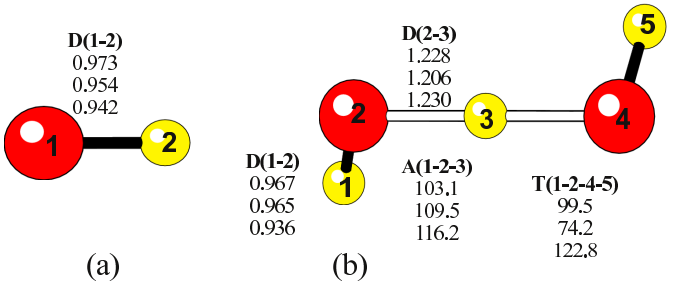


FIGURE 7.9: Optimized structures of OH^- and $OH^-(H_2O)$ (C_2 symmetry). The structural parameters for *ab initio* MP2, OSS2(depW) and OSS2 potentials are shown from top to bottom, respectively .

In Figure 7.9, it is seen that the bond distances of O-H obtained with OSS2(depW) are much closer to MP2 results than OSS2 potential. However, the distance of

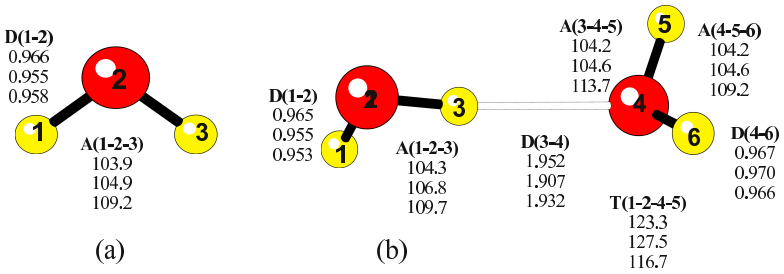
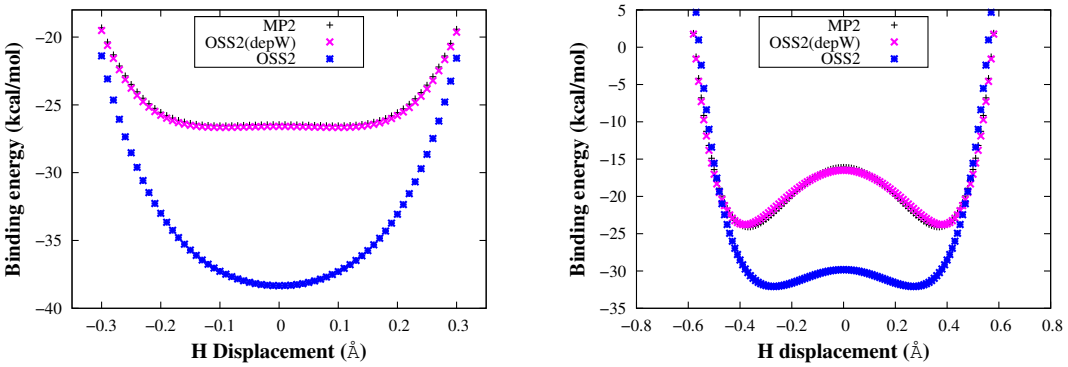


FIGURE 7.11: Optimized structures of (H_2O) and $(H_2O)_2$ (C_s symmetry). The structural parameters for *ab initio* MP2, OSS2(depW) and OSS2 potentials are shown from top to bottom, respectively .

hydrogen bond in $OH^-(H_2O)$ of OSS2(depW) potential is too short by about 0.02 Å while OSS2 potential predicts a quite close value. On the other hand, the bond angle of H-O-H is improved a lot with OSS2(depW). As seen from the figure, the OSS2 bond angle of 116.2° deviates quite far from the MP2 result of 103.1° comparing to OSS2(depW) result of 109.5°. In Figure 7.10, a similar situation is observed. The O-H bond distances for OSS2 potential are underestimated by roughly 0.03 Å while the values predicted by OSS2(depW) always lie within 0.01 Å of the results of MP2 calculation. The bond angles of H-O-H obtained with OSS2(depW) also achieve a better agreement with MP2 results. The same thing can be concluded from Figure 7.11 when examining the structures of neutral water monomer and dimer.

In Ref [70], it was stated that both OSS1 and OSS2 potentials did not reproduce correctly the bond angle of H-O-H and suggested to use the sister model, OSS3. However, OSS3 potential with the inclusion of dipole coupling in the three body interaction and 54 parameters is a too complicated potential. Hence, energy and derivative evaluations with OSS3 potential are not only difficult to be implemented but also computationally expensive. In this study, the new set of parameters for OSS2 potential makes a quite good improvement in reproducing the bond angle with MP2 calculation as shown above.



(a) D(O-O)=2.46 Å (b) D(O-O)=2.86 Å

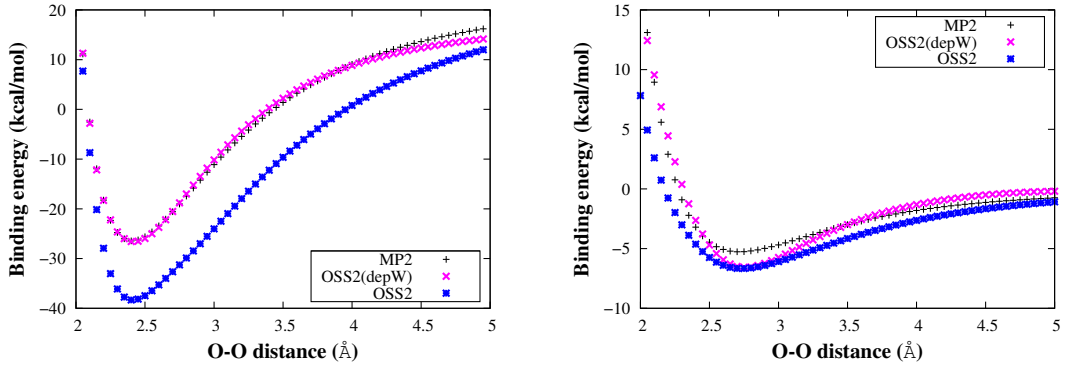
FIGURE 7.12: Proton transfer energy profile in $OH^-(H_2O)$ cluster where the proton is translated along O-O axis where two O atoms are placed in the equilibrium distance of 2.46 Å (a) and separated from each other by 2.86 Å (right). The x and y axes represent the relative displacement of moving proton from the middle point of O-O and binding energies, respectively.

7.3.4.2 Proton transfer profile

For a fair comparison with OSS2 potential for protonated water clusters, the proton transfer energy profile for $OH^-(H_2O)$ was created in the same fashion as described in Ref [70]. The center proton of $OH^-(H_2O)$ was translated along O-O axis from middle point of O-O where the rest atoms are placed at equilibrium positions. The energy profile is depicted in Figure 7.12a. Another one depicted in Figure 7.12b is for the O-O separation of 2.86 Å. In both cases, OSS2(depW) curves match very well with MP2 calculation. OSS2 potential in Figure 7.12a also has a relative good agreement in the shape of potential curve. However, the offset between the curves of OSS2 potential and MP2 calculation of about 20 kcal/mol is very large. In the second subplot, the energy barriers of proton transfer is underestimated for OSS2 potential by approximately 5 kcal/mol.

7.3.4.3 Interaction energy profile

For O-O interaction, we examined the interaction energy in two cases: between OH^- anion and H_2O molecule and between two H_2O molecules. The interaction



a) OH^- and (H_2O) b) $(H_2O)_1$ and $(H_2O)_2$

FIGURE 7.13: The interaction energy between (a) OH^- anion and H_2O molecule, and (b) two H_2O molecules. The x axis represents the distance of O-O.

energy versus O-O distance is shown in Figure 7.13. In the subplot (a), it is seen that the agreement between the potential curves of OSS2(depW) potential and MP2 calculation is quite satisfactory. For OSS2 potential, the potential well is overestimated by nearly 15 kcal/mol. In the subplot (b), the potential wells is shallower since the potential depth is less than 10 kcal/mol. Both OSS2 and OSS2(depW) potential achieve a relatively good agreement with MP2 calculation. The deviation from the MP2 curve is only about 1 kcal/mol.

For testing the performance of new potential in dihedral interactions, the interaction profile was created by changing the dihedral angle of H-O···O-H in $OH^-(H_2O)$ from the MP2 optimized geometry. The dependence of binding energy on the dihedral angle is depicted in Figure 7.14. It is clearly seen from the figure that the potential curve of OSS2 does not agree well with MP2 one. Comparing the results of OSS2(depW) with the counterparts of *ab initio* MP2 calculation, the agreement is fairly good in overall. There is only a deviation of 0.5 kcal/mol between them around 250 °. Notice that the interaction energy span regarding to the dihedral angle is only 1 kcal/mol, meaning that the potential well is very shallow. Thus, reproducing correctly the equilibrium dihedral angle for MP2 calculation is very difficult. In Section 7.3.4.1, the dihedral angles predicted by OSS2(depW) and OSS2 potentials do not agree well with MP2 results.

7.3.5 Summary

In this work, we reparameterized OSS2 model in order to obtain a more suitable potential model to simulate deprotonated water clusters. The new parameters were fitted to the potential energy surface of deprotonated water dimer at high level *ab initio* MP2 calculation. Several tests on the performance of reparameterized OSS2 model have been performed, showing that the bond angles and distances were improved significantly. The energy profile in proton transfer and intermolecular interaction obtained from reparameterized OSS2 model shows a satisfactory agreement with *ab initio* MP2 calculation. This potential would be very useful in simulating the dynamics of deprotonated water clusters in future works.

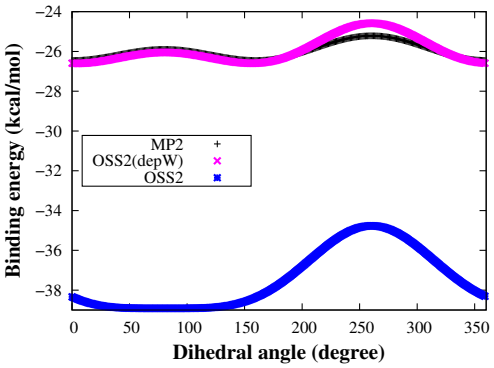


FIGURE 7.14: Dihedral interaction profile of $OH^-(H_2O)_2$ cluster. The x axis represents the dihedral angle of H-O···O-H.

Chapter 8

Conclusion and Future works

8.1 Conclusion

The thesis contains three parts. In the first part, we proposed a hierarchical approach combining empirical potential models and first-principles methods in order to explore PES at quantum chemistry level. We implemented an archive memetic algorithm to locate the global minima as well as to archive all possible local minima. In the second part of thesis, we present a systematic study on structures, thermodynamics and vibrational spectra of different systems: protonated, deprotonated and pure water clusters. For each kind of systems, a collection of isomers of different sizes are gathered via a hierarchical approach at first-principles calculation. Subsequently, thermodynamic properties and vibrational spectra are yielded by using HSA from the identified isomers. In the third part, the collected isomers are used to parametrize new potential models for water and hydrogen fluoride clusters.

8.1.1 Studies of water clusters

In the study of protonated water clusters, the PES of $H^+(H_2O)_n$, $n = 5-10$ with OSS2 potential model and $n = 5-9$ with B3LYP/6-31+G* calculation, have been explored thoroughly. The archival of distinct isomer have been uncovered and subsequently used for investigating the thermodynamic and structural transitions of $H^+(H_2O)_n$ with both OSS2 and DFT calculation. In comparison with PT-MC simulations involving OSS2 potential model, the results computed by HSA achieved a quantitative agreement with PT-MC. The consistency with PT-MC results is the good indication of the reliability and accuracy of HSA. From the results of simulation, the size dependency of $H^+(H_2O)_n$ in both thermal behaviors and vibrational spectra have been revealed. The simulated vibrational spectra are in accord with the recent experimental results.

In the study of deprotonated water clusters, we also studied the structures, relative stability, thermal behaviors and vibrational spectra of $OH^-(H_2O)_n$ for cluster sizes of $n = 4-7$ using DFT calculations. In addition, the comparison against protonated counterparts has been made throughout the study. Our results emphasized that the structures of $OH^-(H_2O)_n$ were more compact than $H^+(H_2O)_n$ in which the formation of 3D structure occurred at smaller sizes and the energy gaps between compact and open structures were found to become larger. The dissimilarity in the ground-state structures and structural orders indicates that the structures of $OH^-(H_2O)_n$ could not be inherited simply from protonated water clusters through flipping OH bonds and considering OH^- anion as "proton hole". The thermodynamic simulations derived via both C-HSA and Q-HSA displayed similar trends of topological transitions, despite some minor deviations between them. The ZPE correction has also been investigated and concluded to have little impact on relative stability, structural transitions as well as on the vibrational

spectra of $OH^-(H_2O)_n$. The size-dependency in thermal behaviors and vibrational spectra, which are analogous to those of $H^+(H_2O)_n$ was also recognized.

In the study of neutral water clusters, we have explored the PES of water clusters in the range of $n = 4 - 8$ with TTM2.1-F and OSS2 potential as well as DFT calculations. Two potential models were used in the pre-screening process for identifying the probable locations of the isomers in the PES of B3LYP/6-31+G* calculation. The numbers of distinct isomers found using the present methodology are much larger than those reported in the literature highlighting the complexity of the PES of water clusters. The thermal behaviors were simulated from the collected isomers and the size-dependency in the thermodynamics were also observed as similar as protonated and deprotonated clusters. The vibrational spectra of $(H_2O)_n$ were derived afterward. From the spectra, the spectral signatures of coordination number of hydrogen bonds were identified.

8.1.2 Potential models for protonated hydrogen fluoride and deprotonated water clusters

In the second part, the development of two potential models for protonated hydrogen fluoride (HF-OSS2) and deprotonated water clusters (OSS2(depW)) have been reported consecutively. Both two potential models used the functional form of OSS2 model which was originally designed for protonated water clusters. We realized that three body interaction in the functional form was one of the most computationally expensive parts. Thus, a simple improvement for three body interaction has been made to reduce the computational scale of the potential from $O(N^3)$ to $O(N^2)$. For parametrization, large sets of configurations at high level *ab initio* MP2 calculations were generated in different manners. In fitting process, the parameters were optimized by Levenberg-Marquardt algorithm. On assessing the performance of new potentials, it was shown that the interaction energies and

optimized geometries of small-sized clusters with *ab initio* MP2 calculation were reproduced correctly. In particular, from a basin hopping run with HF-OSS2 potential, we uncovered a new morphology of $H^+(HF)_5$ which has not been reported in literature. To our best knowledge, the potentials developed in this study are first ones particularly designed for protonated hydrogen fluoride and deprotonated water clusters. They are probably very useful for further studying those systems.

8.2 Future work

8.2.1 Representative sampling

In this thesis, we focused on the small-sized clusters, $n < 10$. For larger-sized ones, the high computational cost of first-principles calculations and the high dimensional search space make sampling all possible local minima impractical. A sampling strategy capable to cover the features of PES is necessary. We are considering to use a representative sampling method in which the isomers gathered at first exploration level by using AMA with empirical model will be classified into different topologies as similar as we have done in the previous studies of water clusters. At the second level, only ten or hundreds utmost isomers of each topology are selected as the representative samples to be further re-optimized instead of re-optimizing all isomers. By this way, the computational effort can be reduced significantly.

8.2.2 Anharmonicity correction

In SA, the harmonic approximation imposed on each basin to calculate the partition function was fairly good for the small-sized clusters since the thermodynamic

results of HSA were in accord with PT-MC as shown in Section 4.3.1. Nevertheless, the deviation between the results of HSA and PT-MC tends to be larger with increasing cluster size. Although the deviation is quite acceptable for $n < 10$, it probably becomes large for larger sizes. One of the reasons is that the harmonic approximation is no longer suitable for large-sized clusters. The anharmonicity should be taken into account for correcting the approximation. To handle the anharmonicity, we are considering to use one of the popular methods including Morse potential approximation,[165] vibrational self-consistent field approach,[166] or perturbation theory.[167]

8.2.3 Broadening of spectral lines

In this study, we used the Lorentzian line shape approximation to simulate the homonegous broadening of the vibrational spectra. The current value of half-width of spectral lines was approximated from the experimental IR spectra. To improve the accuracy of the calculated spectra, the half-width should be determined in other fashion in which life-time broadening contribution, temperature dependence should be taken into account for further studies.

Bibliography

[1] Soh, H.; Ong, Y.-S.; Nguyen, Q. C.; Nguyen, Q. H.; Habibullah, M.; Hung, T.; Kuo, J.-L. *IEEE Trans. Evol. Comput.* **2010**, *14*(3), 419 –437.

[2] Kuo, J. L.; Klein, M. L. *J. Chem. Phys.* **2005**, *122*(2), 1–9.

[3] Lin, C. K.; Wu, C. C.; Wang, Y. S.; Lee, Y. T.; Chang, H. C.; Kuo, J. L.; Klein, M. L. *Phys. Chem. Chem. Phys.* **2005**, *7*(5), 938–944.

[4] Maeda, S.; Ohno, K. *J. Phys. Chem. A* **2007**, *111*(20), 4527–4534.

[5] Ludwig, R. *Angew. Chem. Int. Ed.* **2001**, *40*(10), 1808–1827.

[6] Dyke, T. R.; Mack, K. M.; Muenter, J. S. *J. Chem. Phys.* **1977**, *66*(2), 498–510.

[7] Odutola, J. A.; Dyke, T. R. *J. Chem. Phys.* **1980**, *72*(9), 5062–5070.

[8] Vernon, M. F.; Krajnovich, D. J.; Kwok, H. S.; Lisy, J. M.; Shen, Y. R.; Lee, Y. T. *J. Chem. Phys.* **1982**, *77*(1), 47–57.

[9] Viant, M. R.; Cruzan, J. D.; Lucas, D. D.; Brown, M. G.; Liu, K.; Saykally, R. J. *J. Phys. Chem. A* **1997**, *101*(48), 9032–9041.

[10] Pugliano, N.; Saykally, R. *Science* **1992**, *257*(5078), 1937–1940.

[11] Liu, K.; Loeser, J. G.; Elrod, M. J.; Host, B. C.; Rzepiela, J. A.; Pugliano, N.; Saykally, R. J. *J. Am. Chem. Soc.* **1994**, *116*(8), 3507–3512.

[12] Xantheas, S. S.; Burnham, C. J.; Harrison, R. J. *J. Chem. Phys.* **2002**, *116*(4), 1493–1499.

[13] Wales, D. J.; Ohmine, I. *J. Chem. Phys.* **1993**, *98*(9), 7245–7256.

[14] Tsai, C. J.; Jordan, K. D. *J. Chem. Phys.* **1991**, *95*(5), 3850–3853.

[15] Bandow, B.; Hartke, B. *J. Chem. Phys. A* **2006**, *110*, 5809–5822.

[16] Lee, C.; Chen, H.; Fitzgerald, G. *J. Chem. Phys.* **1995**, *102*(3), 1266–1269.

[17] Wales, D. J.; Hodges, M. P. *Chem. Phys. Lett.* **1998**, *286*(1-2), 65–72.

[18] Burnham, C. J.; Xantheas, S. S. *J. Chem. Phys.* **2002**, *116*(4), 1500–1510.

[19] Burnham, C. J.; Xantheas, S. S. *J. Chem. Phys.* **2002**, *116*(4), 1479–1492.

[20] Burnham, C. J.; Xantheas, S. S. *J. Chem. Phys.* **2002**, *116*(12), 5115–5124.

[21] Lagutschenkov, A.; Fanourgakis, G. S.; Niedner-Schateburg, G.; Xantheas, S. S. *J. Chem. Phys.* **2005**, *122*(19), 194310–194319.

[22] Tsai, C. J.; Jordan, K. D. *J. Chem. Phys.* **1993**, *99*(9), 6957–6970.

[23] Tsai, C. J.; Jordan, K. D. *J. Phys. Chem.* **1993**, *97*(43), 11227–11237.

[24] Tsai, C. J.; Jordan, K. D. *J. Phys. Chem.* **1993**, *97*(20), 5208–5210.

[25] James, T.; Wales, D. J.; Rojas, J. H. *J. Chem. Phys.* **2007**, *126*(5), 054506–13.

[26] Wales, D. J. *Mol. Phys.* **2004**, *102*(9), 891–908.

[27] Tsai, C.; Jordan, K. *Chem. Phys. Lett.* **1993**, *213*(1-2), 181 – 188.

[28] Ball, P. *Nature* **2008**, *452*(7185), 291–292.

[29] Steel, E. A.; Merz, K. M.; Selinger, A.; Castleman, A. W. *J. Phys. Chem.* **1995**, *99*(19), 7829–7836.

[30] Gruenloh, C. J.; Carney, J. R.; Arrington, C. A.; Zwier, T. S.; Fredericks, S. Y.; Jordan, K. D. *Science* **1997**, *276*(5319), 1678–1681.

[31] Sobott, F.; Wattenberg, A.; Barth, H.-D.; Brutschy, B. *Int. J. Mass spectrom.* **1999**, *185-187*, 271 – 279.

[32] Cruzan, J. D.; Braly, L. B.; Liu, K.; Brown, M. G.; Loeser, J. G.; Saykally, R. J. . *Science* **1996**, *271*(5245), 59–62.

[33] Liu, K.; Brown, M. G.; Cruzan, J. D.; Saykally, R. J. *Science* **1996**, *271*(5245), 62–64.

[34] Liu, K.; Brown, M. G.; Carter, C.; Saykally, R. J.; Gregory, J. K.; Clary, D. C. *Nature* **1996**, *381*, 501–503.

[35] Headrick, J. M.; Diken, E. G.; Walters, R. S.; Hammer, N. I.; Christie, R. A.; Cui, J.; Myshakin, E. M.; Duncan, M. A.; Johnson, M. A.; Jordan, K. D. *Science* **2005**, *308*(5729), 1765–1769.

[36] Miyazaki, M.; Fujii, A.; Ebata, T.; Mikami, N. *Science* **2004**, *304*(5674), 1134–1137.

[37] Shin, J.-W.; Hammer, N. I.; Diken, E. G.; Johnson, M. A.; Walters, R. S.; Jaeger, T. D.; Duncan, M. A.; Christie, R. A.; Jordan, K. D. *Science* **2004**, *304*(5674), 1137–1140.

[38] Xantheas, S. S. *J. Chem. Phys.* **1994**, *100*(10), 7523–7534.

[39] Xantheas, S. S. *J. Chem. Phys.* **1995**, *102*(11), 4505–4517.

[40] Xantheas, S. S.; Dunning, T. H. *J. Chem. Phys.* **1993**, *99*(11), 8774–8793.

[41] Kim, K. S.; Dupuis, M.; Lie, G. C.; Clementi, E. *Chem. Phys. Lett.* **1986**, *131*(6), 451 – 456.

[42] Qian, P.; Song, W.; Lu, L.; Yang, Z. *Int. J. Quantum Chem.* **2010**, *110*(10), 1923–1937.

[43] Chaplin, M. F. *Biophys. Chem.* **2000**, *83*(3), 211 – 221.

[44] James, T.; Wales, D. J.; Hernandez-Rojas, J. *Chem. Phys. Lett.* **2005**, *415*(4-6), 302–307.

[45] Singer, S. J.; McDonald, S.; Ojamäe, L. *J. Chem. Phys.* **2000**, *112*(2), 710–716.

[46] Mella, M.; Kuo, J.-L.; Clary, D. C.; Klein, M. L. *Phys. Chem. Chem. Phys.* **2005**, *7*, 2324–2332.

[47] Svanberg, M.; Pettersson, J. B. C. *J. Phys. Chem. A* **1998**, *102*(10), 1865–1872.

[48] Luo, Y.; Maeda, S.; Ohno, K. *J. Phys. Chem. A* **2007**, *111*(42), 10732–10737.

[49] Luo, Y.; Ohno, S. M. K. *J. Comput. Chem.* **2008**, *30*(6), 952–961.

[50] Tuckerman, M.; Laasonen, K.; Sprik, M.; Parrinello, M. *J. Chem. Phys.* **1995**, *103*(1), 150–161.

[51] Iyengar, S. S. *J. Chem. Phys.* **2007**, *126*, 216101–216103.

[52] Iyengar, S. S.; Petersen, M. K.; Day, T. J.; Burnham, C. J.; Teige, V. E.; Voth, G. A. *J. Chem. Phys.* **2005**, *123*(8), 084309–084318.

[53] Hodges, M. P.; Wales, D. J. *Chem. Phys. Lett.* **2000**, *324*(4), 279–288.

[54] Wei, D.; Salahub, D. R. *J. Chem. Phys.* **1997**, *106*(14), 6086–6094.

[55] Christie, R. A.; Jordan, K. D. *J. Phys. Chem. A* **2001**, *105*(32), 7551–7558.

[56] Bing, D.; Hamashima, T.; Nguyen, Q. C.; Fujii, A.; Kuo, J.-L. *J. Phys. Chem. A* **2010**, *114*(9), 3096–3102.

[57] Wales, D. J. *Energy landscape;* Cambridge University Press, 2003.

[58] Nguyen, Q. C.; Ong, Y. S.; Soh, H.; Kuo, J. L. *J. Phys. Chem. A* **2008**, *112*(28), 6257–6261.

[59] Deaven, D. M.; Ho, K. M. *Phys. Rev. Lett.* **1995**, *75*(2), 288–291.

[60] Nguyen, Q. C.; Ong, Y. S.; Kuo, J.-L. *J. Chem. Theory and Comput.* **2009**, *5*, 2629–2639.

[61] Wallqvist, A.; Mountain, R. D. *Molecular Models of Water: Derivation and Description;* 2007.

[62] Jorgensen, W. L. *J. Am. Chem. Soc.* **1981**, *103*(2), 335–340.

[63] Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*(24), 6269–6271.

[64] Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*(2), 926–935.

[65] Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. *J. Chem. Phys.* **2004**, *120*(20), 9665–9678.

[66] Kozack, R. E.; Jordan, P. C. *J. Chem. Phys.* **1993**, *99*(4), 2978–2984.

[67] Mahoney, M. W.; Jorgensen, W. L. *J. Chem. Phys.* **2000**, *112*(20), 8910–8922.

[68] Rick, S. W. *J. Chem. Phys.* **2004**, *120*(13), 6085–6093.

[69] Burnham, C. J.; Li, J.; Xantheas, S. S.; Leslie, M. *J. Chem. Phys.* **1999**, *110*(9), 4566–4581.

[70] Ojamäe, L.; Shavitt, I.; Singer, S. J. *J. Chem. Phys.* **1998**, *109*(13), 5547–5564.

[71] Becke, A. D. *Phys. Rev. A* **1988**, *38*(6), 3098–3100.

[72] Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*(2), 785–789.

[73] Gaussian 03, Revision D.01. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. **2004**.

[74] Deaven, D. M.; Tit, N.; Morris, J. R.; Ho, K. M. *Chem. Phys. Lett.* **1996**, *256*(1-2), 195–200.

[75] Chua, A. L. S.; Benedek, N. A.; Chen, L.; Finnis, M. W.; Sutton, A. P. *Nat. Mater.* **2010**, *9*(5), 418–422.

[76] Hooper, J.; Hu, A.; Zhang, F.; Woo, T. K. *Phys. Rev. B* **2009**, *80*(10), 104117–104125.

[77] Ong, Y.-S.; Lim, M.-H.; Zhu, N.; Wong, K. W. *IEEE. Trans. Syst., Man, Cybern. B* **2006**, *36*(1), 141–152.

[78] Turner, G. W.; Tedesco, E.; Harris, K. D. M.; Johnston, R. L.; Kariuki, B. M. *Chem. Phys. Lett.* **2000**, *321*(3-4), 183–190.

[79] Ong, Y. S.; Keane, A. *IEEE Trans. Evol. Comput.* **2004**, *8*(2), 99–110.

[80] Hasan, S. M. K.; Sarker, R.; Essam, D.; Cornforth, D. *Memetic Computing* **2009**, *1*, 69–83.

[81] Nguyen, Q. H.; Ong, Y.-S.; Lim, M. H. *IEEE Trans. Evol. Comput.* **2009**, *13*(3), 604–623.

[82] Zhu, Z.; Ong, Y.; Zurada, J. *IEEE/ACM Trans. Comput. Biol. Bioinf.* **2010**, *7*(2), 263–277.

[83] Guimaraes, F. F.; Belchior, J. C.; Johnston, R. L.; Roberts, C. *J. Chem. Phys.* **2002**, *116*(19), 8327–8333.

[84] Smith, S. S.-F. *Comput. Ind.* **2004**, *54*(1), 35–49.

[85] Leung, Y.; Gao, Y.; Xu, Z.-B. *IEEE Trans. Neural Networks* **1997**, *8*, 1165 – 1176.

[86] Masek, B. B.; Merchant, A.; Matthew, J. B. *J. Med. Chem.* **1993**, *36*(9), 1230–1238.

[87] Meyer, A. Y.; Richards, W. G. *J. Comput.-Aided Mol. Des.* **1991**, *5*, 427–439.

[88] Grant, J. A.; Pickup, B. T. *J. Chem. Phys.* **1995**, *99*, 3503–3510.

[89] Good, A. C.; Richards, W. G. *Perspect. Drug Discovery Des.* **1998**, *9*, 321–338.

[90] Ballester, P. J.; Richards, W. G. *J. Comput. Chem.* **2007**, *28*(10), 1711–1723.

[91] Nicholls, A.; MacCuish, N. E.; MacCuish, J. D. *J. Comput.-Aided Mol. Des.* **2004**, *18*, 451–474.

[92] Ronald, S. *IEEE International Conference on Evolutionary Computation* **1998**, pages 793 –798.

[93] Christie, R. A.; Jordan, K. D. *J. Phys. Chem. B* **2002**, *106*(33), 8376–8381.

[94] Nakayama, A.; Seki, N.; Taketsugu, T. *J. Chem. Phys.* **2009**, *130*(2), 024107–024117.

[95] Swendsen, R. H.; Wang, J. . *Phys. Rev. Lett.* **1986**, *57*(21), 2607–2609.

[96] Earl, D. J.; Deem, M. W. *Phys. Chem. Chem. Phys.* **2005**, *7*(23), 3910–3916.

[97] Stillinger, F. H.; Weber, T. A. *Phys. Rev. A* **1982**, *25*(2), 978–989.

[98] Stillinger, F. H.; Weber, T. A. *Science* **1984**, *225*(4666), 983–989.

[99] Calvo, F.; Doye, J. P. K.; Wales, D. J. *Chem. Phys. Lett.* **2002**, *366*(1-2), 176–183.

[100] Sharapov, V. A.; Meluzzi, D.; Mandelshtam, V. A. *Phys. Rev. Lett.* **2007**, *98*, 105701–105705.

[101] Wales, D. *Mol. Phys.* **1993**, *78*, 151–171.

[102] Stillinger, F. H. *Phys. Rev. E* **1999**, *59*(1), 48–51.

[103] Jiang, J. C.; Wang, Y. S.; Chang, H. C.; Lin, S. H.; Lee, Y. T.; Niedner-Schatteburg, G.; Chang, H. C. *J. Am. Chem. Soc.* **2000**, *122*(7), 1398–1410.

[104] Wu, C. C.; Lin, C. K.; Chang, H. C.; Jiang, J. C.; Kuo, J. L.; Klein, M. L. *J. Chem. Phys.* **2005**, *122*, 074315–074323.

[105] Wei, S.; Shi, Z.; A. W. Castleman, J. *J. Chem. Phys.* **1991**, *94*(4), 3268–3270.

[106] Asmis, K. R.; Pivonka, N. L.; Santambrogio, G.; Brummer, M.; Kaposta, C.; Neumark, D. M.; Woste, L. *Science* **2003**, *299*(5611), 1375–1377.

[107] Zwier, T. S. *Science* **2004**, *304*(5674), 1119–1120.

[108] Laasonen, K.; Klein, M. L. *J. Phys. Chem.* **1994**, *98*(40), 10079–10083.

[109] Karthikeyan, S.; Park, M.; Shin, I.; Kim, K. S. *J. Phys. Chem. A* **2008**, *112*(41), 10120–10124.

[110] Schmitt, U. W.; Voth, G. A. *J. Chem. Phys.* **1999**, *111*(20), 9361–9381.

[111] Li, Z. Q.; Scheraga, H. A. *Proc. Nat. Acad. Sci. U.S.A.* **1987**, *84*, 6611–6615.

[112] Iyengar, S. S.; Day, T. J. F.; Voth, G. A. *Int. J. Mass spectrom.* **2005**, *241*(2-3 SPEC.ISS.), 197–204.

[113] Termath, V.; Sauer, J. *Mol. Phys.* **1997**, *91*(5), 963–975.

[114] Li, X.; Teige, V. E.; Iyengar, S. S. *J. Phys. Chem. A* **2007**, *111*(22), 4815–4820.

[115] Wales, D. J.; Doye, J. P. K. *J. Chem. Phys.* **2003**, *119*(23), 12409–12416.

[116] Bogdan, T. V.; Wales, D. J.; Calvo, F. *J. Chem. Phys.* **2006**, *124*(4), 1–13.

[117] Lovejoy, E. R.; Bianco, R. *J. Phys. Chem. A* **2000**, *104*(45), 10280–10287.

[118] Wang, Y.-S.; Chang, H.-C.; Jiang, J.-C.; Lin, S. H.; Lee, Y. T.; Chang, H.-C. *J. Am. Chem. Soc.* **1998**, *120*(34), 8777–8788.

[119] Bernal, J. D.; Fowler, R. H. *J. Chem. Phys.* **1933**, *1*(8), 515–548.

[120] Xantheas, S. *J. Am. Chem. Soc.* **1995**, *117*(41), 10373–10380.

[121] Tuckerman, M.; Marx, D.; Klein, M.; Parrinello, M. *Science* **1997**, *275*(5301), 817–820.

[122] del Valle, C. P.; Novoa, J. J. *Chem. Phys. Lett.* **1997**, *269*(5-6), 401–407.

[123] Turki, N.; Milet, A.; Rahmouni, A.; Ouamerali, O.; Moszynski, R.; Kochanski, E.; Wormer, P. E. S. *J. Chem. Phys.* **1998**, *109*(17), 7157–7168.

[124] Wei, D.; Proynov, E.; Milet, A.; Salahub, D. *J. Phys. Chem. A* **2000**, *104*(11), 2384–2395.

[125] Vegiri, A.; Shevkunov, S. *J. Chem. Phys.* **2000**, *113*, 8521–8530.

[126] Shevkunov, S. V. *Colloid J.* **2008**, *70*(6), 784–795.

[127] Yang, X.; Castleman Jr., A. *J. Phys. Chem.* **1990**, *94*(23), 8500–8502.

[128] Chaudhuri, C.; Wang, Y. S.; Jiang, J. C.; Lee, Y. T.; Chang, H. C.; Niedner-Schatteburg, G. *Mol. Phys.* **2001**, *99*(14), 1161–1173.

[129] Price, E. A.; Robertson, W. H.; Diken, E. G.; Weddle, G. H.; Johnson, M. *Chem. Phys. Lett.* **2002**, *366*, 412–416.

[130] Robertson, W. H.; Diken, E. G.; Price, E. A.; Shin, J.-W.; Johnson, M. A. *Science* **2003**, *299*(5611), 1367–1372.

[131] Ando, K.; Hynes, J. T. *Adv. Chem. Phys.* **1999**, *110*, 381–430.

[132] Schultz, B. E.; Chan, S. I. *Ann. Rev. Biophys. Biomol. Struct.* **2001**, *30*, 23.

[133] Riccardi, D.; Koenig, P.; Prat-Resina, X.; Yu, H.; Elstner, M.; Frauenheim, T.; Cui, Q. *J. Am. Chem. Soc.* **2006**, *128*(50), 16302 – 16311.

[134] Lee, H.; Tarkeshwar, P.; Kim, K. *J. Chem. Phys.* **2004**, *121*, 4657–4664.

[135] Shi, Z.; Ford, J. V.; Wei, S.; Castleman, Jr., A. W. *J. Chem. Phys.* **1993**, *99*(10), 8009–8015.

[136] Lenz, A.; Ojame, L. *J. Phys. Chem. A* **2006**, *110*(50), 13388–13393.

[137] Chang, H.-C.; Wu, C.-C.; Kuo, J.-L. *Int. Rev. Phys. Chem.* **2005**, *24*(3/4), 553 – 578.

[138] Singh, N. J.; Park, M.; Min, S. K.; Suh, S. B.; Kim, K. S. *Angew. Chem. Int. Ed.* **2006**, *45*(23), 3795–3800.

[139] Zundel, G.; Metzger, H. *Z. Phys. Chem.* **1968**, *58*, 225–245.

[140] Eigen, M. *Angew. Chem. Int. Ed.* **1964**, *3*(1), 1–19.

[141] Sommer, J.; Bukala, J.; Hachoumy, M.; Jost, R. *J. Am. Chem. Soc.* **1997**, *119*, 3274–3279.

[142] Minkwitz, R.; Schneider, S. *Angew. Chem. Int. Ed.* **1999**, *38*(5), 714–715.

[143] Kim, D.; Klein, M. L. *J. Phys. Chem. B* **2000**, *104*(43), 10074–10079.

[144] Esteves, P. M.; Ramirez-Solis, A.; Mota, C. J. A. *J. Phys. Chem. B* **2001**, *105*(19), 4331–4336.

[145] Esteves, P. M.; Ramirez-Solis, A.; Mota, C. J. A. *J. Am. Chem. Soc.* **2002**, *124*(11), 2672–2677.

[146] Raugei, S.; Klein, M. L. *J. Phys. Chem. B* **2001**, *105*(34), 8212–8219.

[147] Raugei, S.; Klein, M. L. *J. Chem. Phys.* **2002**, *116*(16), 7087–7093.

[148] Gillespie, R. J.; Liang, J. *J. Am. Chem. Soc.* **1988**, *110*(18), 6053–6057.

[149] Deraman, M.; Dore, J. C.; Powles, J. G.; Holloway, J. H.; Chieux, P. *Mol. Phys.* **1985**, *55*(6), 1351–1367.

[150] Pfleiderer, T.; Waldner, I.; Bertagnolli, H.; Todheide, K.; Fischer, H. E. *J. Chem. Phys.* **2000**, *113*(9), 3690–3696.

[151] McLain, S. E.; Benmore, C. J.; Siewenie, J. E.; Urquidi, J.; Turner, J. F. C. *Angew. Chem. Int. Ed.* **2004**, *43*(15), 1952–1955.

[152] McLain, S. E.; Benmore, C. J.; Siewenie, J. E.; Molaison, J. J.; Turner, J. F. C. *J. Chem. Phys.* **2004**, *121*(13), 6448–6455.

[153] Kreitmeir, M.; Heusel, G.; Bertagnolli, H.; Todheide, K.; Mundy, C. J.; Cuello, G. J. *J. Chem. Phys.* **2005**, *122*(15), 154511–154518.

[154] Karpfen, A.; Yanovitskii, O. *J. Mol. Struct.* **1994**, *307*, 81 – 97.

[155] Pártay, L.; Jedlovszky, P.; Vallauri, R. *J. Chem. Phys.* **2006**, *124*(18), 184504–184515.

[156] Jedlovszky, P.; Vallauri, R. *J. Chem. Phys.* **1997**, *107*(23), 10166–10176.

[157] Della Valle, R. G.; Gazzillo, D. *Phys. Rev. B* **1999**, *59*(21), 13699–13706.

[158] Wierzchowski, S. J.; Kofke, D. A.; Gao, J. *J. Chem. Phys.* **2003**, *119*(14), 7365–7371.

[159] Izvekov, S.; Voth, G. A. *J. Phys. Chem. B* **2005**, *109*(14), 6573–6586.

[160] Marquardt, D. W. *SIAM J. Appl. Math.* **1963**, *11*(2), 431–441.

[161] Levenberg-marquardt nonlinear least squares algorithms in c/c++. Lourakis, M. http://www.ics.forth.gr/~lourakis/levmar, **2000**.

[162] Kumagai, N.; Kawamura, K.; Yokokawa, T. *Mol. Simul.* **1994**, *12*(3), 177–186.

[163] Stillinger, F. H.; Weber, T. A. *J. Chem. Phys.* **1982**, *76*(8), 4028–4036.

[164] Eisenberg, D.; Kauzmann, W. *The Structure and Properties of Water;* Oxford University Press Inc., 2005.

[165] Yukhnevich, G. *Dokl. Phys.* **2000**, *45*(5), 201–204.

[166] Bihary, Z.; Gerber, R. B.; Apkarian, V. *J. Chem. Phys.* **2001**, *115*, 2695–2701.

[167] Patnaik, P. K. *Phys. Rev. D* **1986**, *33*(10), 3145–3146.