First Principles Studies on the Vibrational Spectra of Cationized Amino Acids and Hydrated Anions

SUN, Shoutian

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy
in
Chemistry

The Chinese University of Hong Kong
March 2014

Thesis/Assessment Committee

Professor Chai Mei YU (Chair)

Professor Zhifeng LIU (Thesis Supervisor)

Professor Man Chor CHAN (Committee Member)

Abstract of thesis entitled:

First Principles Studies on the Vibrational Spectra of Cationized Amino Acids and Hydrated Anions

Submitted by SUN, Shoutian

for the degree of Doctor of Philosophy in Chemistry

at the Chinese University of Hong Kong in March 2014

The combination of the FEL with FTICR makes it a powerful tool in researching the biochemical molecules and solvation dynamics of small hydrated ions. And theoretical methods are often employed in the determination of the molecular structures as auxiliary tools. In this thesis, density functional theory and *ab initio* molecular dynamics method have been applied to study the experimental infrared multiphoton dissociation (IRMPD) processes for cationized amino acids and hydrated anions.

The second Chapter is the AIMD simulated IR spectra for Asparagine cationized by H^+, Li^+, Na^+, K^+, and Cs^+. The previous work simulated these type clusters with harmonic approximation in which the dynamics and anharmonicity effects were missed. My result shows that M^+(Asn) complexes are very flexible and they could convert into each other at finite temperatures by torsional rotation and stretching the distances between the ion and N, O atoms in Asn. The broaden features and shift of

the relative intensities of the vibrational profiles for these conformers are dependent on the structural flexibility of the complexes.

The third Chapter is the IRMPD spectra of the hydrated bicarbonate anion clusters for n=1-7. For some clusters, as mentions in the former harmonic results, comparing the experimental and calculated spectra does not always identify a single isomer as being responsible for the spectrum. By using AIMD technique, I found discrepancies between the harmonic and experimental spectra, especially for the absence of water wagging and rocking mode in the experimental result for n=1-5. By calculating the thermal energies of various isomers of each cluster, I found that the less dominate conformer also shows its large contributions to the vibrational profiles and this could explain the peak intensity trend as the cluster size increase.

The vibrational spectra for $(HO)_2PO_2^-(H_2O)_{n=1\text{-}4,\ 6,\ \text{and}\ 8}$ are presented in the last two chapters. It is the first time we collaborate with experiment group. In these works, I have done all the theoretical simulations. For these clusters, both the low and high (n=1-4) frequency regions of IRMPD spectra are measured. Interestingly for n=1, both the two stable isomer cannot alone represent both the low and high frequency profiles. Base on the AIMD simulations, I found that the existence of an unstable transient state shows it absorption profile in the high frequency region due to the zero point energy and tunneling effect of hydrogen atom. Similar to n=1, the IRMPD spectra of larger clusters (n>1) also show evidence for isomerization at cryogenic temperatures which leads to a complete breakdown of the harmonic approximation.

香港中文大學
化學課程哲學博士學位
論文摘要

第一性原理研究陽離子化氨基酸和水和陰離子

孫守田
二零一四年三月

自由電子雷射器和傅裡葉變換質譜儀的組合使之成為一種研究生物分子和一些水和離子的溶解動力學的有力工具。而理論方法作為實驗的輔助手段常用于判斷一些分子的結構。在本論文中，我運用密度泛函理論和第一性分子動力學方法研究了陽離子化的氨基酸和一些水合陰離子的實驗紅外多光子吸收譜。

第二章研究了氫離子和鹼金屬離子化的天冬醯胺酸的紅外多光子吸收譜。以前有文獻用簡諧近似的方法研究過這個團簇，不過這種方法忽略了分子的動力學和非諧效應。我用第一性分子動力學方法重新研究了這個體系，結果顯示這種陽離子化的氨基酸的分子結構非常柔軟，在有限溫度下可以通過二面角的轉動和離子與氨基酸中的氧和氮原子距離的振動可以相互的轉化。而這些構型的相互轉化可以解釋實驗紅外吸收峰的展寬和相對強度變化趨勢。

第三章研究了水合碳酸氫根陰離子的紅外多光子吸收譜。以前的文獻指出，對於這種團簇，僅通過對比實驗和簡諧近似的紅外譜圖不能夠判斷是否只有一種異構體存在。應用第一性分子動力學方法，我發現簡諧近似紅外譜和實驗譜的不吻合之處，特別是水分子個數小於等於 5 時，水分子中氫原子的左右搖擺振動和來回搖擺振動吸收峰在實驗譜中的缺失。通過計算一系列異構體的熱能，我發現

一些含量較少的異構體仍然對吸收峰有比較大的貢獻，考慮了這些結構的貢獻我們就可以理解為什麼隨著團簇尺寸的增加，一些吸收峰的強度有一定的變化趨勢。

最後兩章研究了水合磷酸二氫根的紅外多光子吸收譜。這是第一次我們和實驗組合作成果。在這個專案裡，我負責分析和全部的理論計算。對於這種團簇，實驗上測量了紅外譜的低頻和高頻區(n=1-4)。有趣的是，特別是體系只包含一個水分子時，兩個穩定的異構體均不能單獨解釋低頻和高頻的實驗結果。基於分子動力學的研究發現，有一個過渡態的吸收峰出現在了高頻區。其原因是氫原子的零點能和隧穿效應。與 n=1 類似,對於存在多個水分子的體系，實驗譜也揭示了低溫下的異構體效應，而這個效應使得簡諧效應完全失效。

ACKNOWLEDGMENTS

I would like to give my sincere acknowledgments to my supervisor Prof. LIU Zhi-Feng, Department of Chemistry of the Chinese University of Hong Kong (CUHK) for his kindly guidance and instruction of my research work and the writing of this thesis.

Special thanks are given to my colleagues in room 225A, Science Center, CUHK: LIU Jian-Wen, YUAN Qing-Hong, ZHANG Han, CHAN Qiu-Bo, NG Kwok-Hung, XIA Guang-Jie, LI Hui-Yan and YUAN Lu-Yan for their helpful discussions in the research process.

I also thank the staffs of Department of Chemistry of CUHK, for helping submit and post thesis. Special thanks are given to the National Supercomputing Center in Shenzhen for the generous allocation of computer time.

Finally, I would like to thank my parents, Mr. SUN Fa-Ke, Ms. CHEN Wen-Ying and my wife XU Ying for their strong supports.

TABLE OF CONTENTS

Chapter One
General introduction and methodology

1.1 Introduction

1.1.1 Introduction to the infrared multiphoton dissociation (IRMPD) process

With recent progresses of experimental techniques, such as the "soft ionization" method[1] and free electron lasers with a wide range of IR frequency, direct probe of charged biomolecules and micro-hydrated clusters in the gas phase has been made possible by measuring the infrared multiphoton dissociation (IRMPD) processes. Cationized amino acids and hydrated ions have been among the systems investigated by IRMPD since they are models for understanding the electrostatic and solvation interaction underlying many biological and atmospherical processes. The spectroscopic signatures not only provide structural information, but also insights into the solvation dynamics.

IRMPD spectroscopy was originally designed and implemented in Lee group for investigation of reactive ionic species in the gas phase.[2,3] The basic idea was to stabilize floppy reactive ions by solvating them with molecules, such as H_2O and H_2, and storing them in an ion trap. As such solvent molecules were only weakly bound to the ions; absorption of IR photons could easily lead to their dissociation. The advantage for this technique is that the detection for the changes in ion signals was much more sensitive than the direct measurement of photon attenuation. This design has been gradually improved over the years, to measure the infrared spectra for a wide variety of solvated ions.[4,5] With the introduction of the "soft ionization" methods such as electrospray or matrix-assisted laser desorption ionization[1], it has become possible in the last few years to measure fragile molecules, such as charged amino acids and hydrated ions, in the gas phase. Fourier-transform ion cyclotron resonance mass spectrometers have greatly increased the ion storage and analysis capacity.[6-8]

And the free electron laser facilities have expanded the wavelength range to measure the important vibrational modes[9], providing information complementary to the measurement of stretching mode vibrations by tabletop lasers.[10]

The phenomenon of infrared multiphoton dissociation has been studied theoretically in the early days after the first experimental work by Letokhov[11] and Quack.[12,13] And more details on this topic can be found in reviews[11,14-17] for the infrared laser photochemistry and the infrared multiphoton excitation process. Here, we will concentrate the IRMPD process suggested by Quack and others.

The most general description of the IRMPD process is to use the solution of the time dependent Schrodinger equation in the presence of intense laser radiation, treated as a classical electromagnetic wave. However, accurate solution of this problem is only possible for small molecules for which the density of vibrational states is relatively low. However, in typical IRMPD experiments, the numbers of the vibrational states are enormous, even for relatively small polyatomic molecules such as CF_3I and SF_6.[17] Quack proposed to use a statistical mechanical approach for the description of the IRMPD process.[12] It provides a realistic model that can be used for the reasonable calculations of the multiphoton excitation dynamics if complete information about the spectroscopy of the molecule is available.

A semi-empirical model is often used if complete molecule information is absent, as proposed in Ref.[11]. This model does not give a complete physical description of the IRMPD process, but offers a simplified and intuitive way to understanding the multiphoton absorption excitation phenomenon in strong IR laser field.[11] The total IRMPD process is divided into four steps and corresponding to four zones as shown in Figure 1.1. Zone I is the lower vibrational levels. The excitation of the molecule in this region is a resonant process due to the low vibrational density of states. The vibrational anharmonicity effect will lead to the detuning of successive vibrational states from the fixed IR photon frequency which is known as anharmonic bottleneck. This effect can be partially compensated by changes in the molecule's rotational energy by splitting or broadening the excited vibrational states. And in case of an intense laser field, more IR photons can be absorbed. However, the absorption cross

section of this zone is low. The excitation in this region is the key factor that limits the dissociation efficiency. The vibrational states in zone II is quasi-continuum and the mutual interaction vibrational energy levels are high in density. For a fixed frequency IR photon, the molecule can absorb successively in this region and accumulate the amount of energy to the dissociation threshold if the laser fluence is high enough and relaxation of the absorbed energy is slow. Zone III describes the accumulated energy above the molecule dissociation threshold. This process has a statistical character conditioned by energy flow between each vibrational state. The collision relaxation will compensate the dissociation efficiency and the IRMPD dissociation can only occur faster than the relaxation. Zone IV represents the multiphoton excitation of the dissociation fragment. It is only possible when the dissociation fragments are a polyatomic molecule, in resonance between the IR laser and one of its vibrational modes.

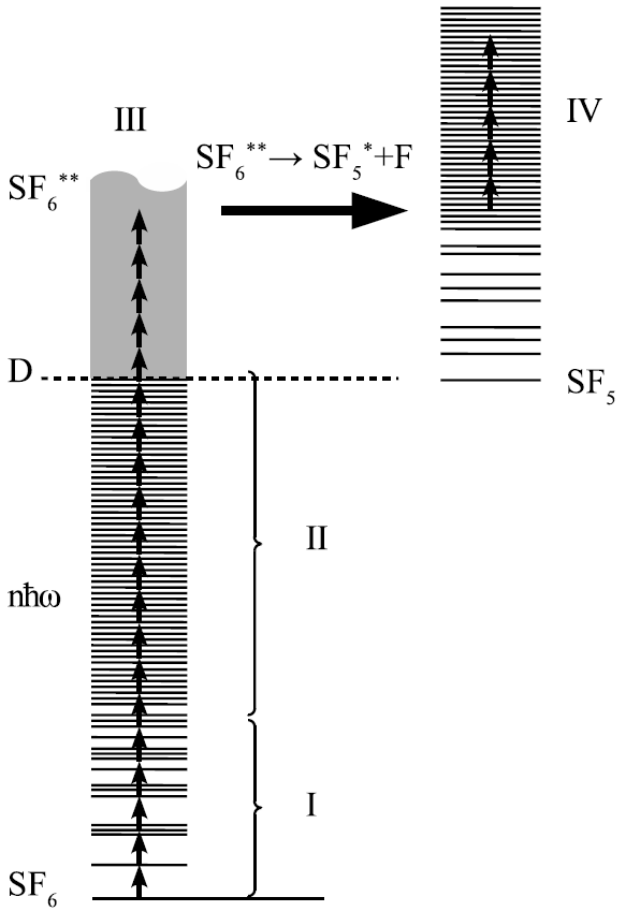

Figure 1.1: Semi-empirical model for the dissociation process of a molecule in a strong IR laser field, showing the four major steps of IRMPD.[11]

1.1.2 IRMPD technique with FTICR mass spectrometers (MS) and Free Electron Laser (FEL)

FT-ICR Mass spectrometry shows great advantages in analyzing biological molecules at atto-mole quantities over X-ray crystallography (XRD) and nuclear magnetic resonance (NMR) which are of limited when sample quantities are low. The developments of MS techniques yield accurate conformer information of biomolecules and begin to challenge XRD and NMR as a key technology in life sciences. FT-ICR provide ideal "test tubes" for the study of gas phase reactions in absence of solvent or other environmental effects, given 1). high vacuum operation (normally 10^{-9} mbar) which can minimize gas phase collisions and allow relatively fast reactions; 2). relatively long ion trapping time (typically seconds to hours), permitting kinetic study of relatively slow processes; 3). high mass resolution (10^6) and accuracy (<ppm), allowing differentiation of ionic species identical in mass.[18]

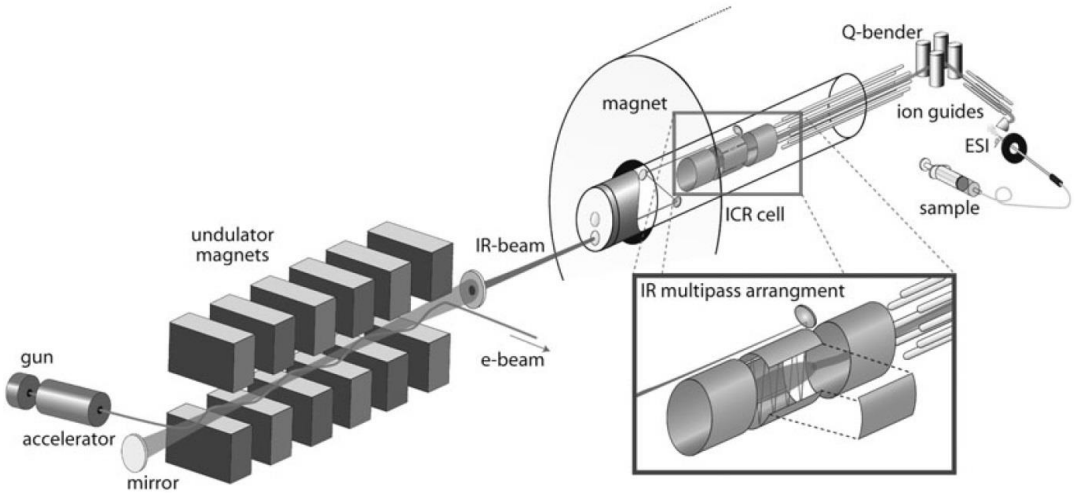

Figure 1.2 The graph for experimental apparatus: ESI-FTICR mass spectrometer (right) coupled to the infrared beam line of FELIX (left). Typically, the length of the FEL from gun to undulator is more than 20 m and the length of the superconducting FTICR magnet is about 1 m.[18]

Infrared spectroscopy for the molecular ions stored in an MS ion trap can be measured by infrared multiple photon dissociation (IRMPD) technique.[19] When IR photons are in resonance/quansi-resonance with an absorption band of a molecular ion, the ion can undergo dissociation. By monitoring the yield of the residue molecular

ions in the ion trap as the function of the laser wavelength, the IRMPD spectrum can be obtained. Normally, powerful laser resource is needed to induce dissociations which require absorption many IR photons. Historically, CO_2 and CO discharge lasers[20-24] were used. However, the wavelength of these types of lasers is limited and discontinuous. For example, the frequency of CO_2 laser roughly falls into the 1500-1900 cm^{-1} region, which limits its application in IR spectroscopy. In 2000, an infrared free electron laser (FEL) was equipped with FTICR at FOM Rijnhuizen[25] which can provide wide and continuous tunable free electron laser for infrared experiments (FELIX).

The combination of the FEL with FTICR makes it a powerful tool in studying the charged molecules. The graph for the apparatus is shown in Fig. 1.2. In the FEL, electrons are accelerated in a linear accelerator to relativistic speed and guided into an undulator[26,27] which is a periodic magnetic apparatus that wriggles the electrons by Lorentz force field and emits synchrotron radiation. The radiation wavelength is dependent on the B-field, which can be used to adjust the wavelength of the free electron laser. Normally, the FTICR-MS is coupled with an electrospray (Micro-mass Z-Spray) ion source to facilitate ionization and vaporization of bimolecular samples.[28] The precursor ions are accumulated in an rf hexapole ion trap for about 500 ms and then injected into the IRC cell via a quadrupole ion deflector and a long rf octopole ion guide. The precursor ions, once the ions are captured in the ICR cell, can be mass isolated by a stored waveform inverse Fourier transform (SWIFT) excited pulse ejecting all other massed from the trap.[29,30] The reactions happen in the ICR cell and the residual precursor ions and the product ions are monitored. And the entire process is repeated while the FEL is scanned. Then the IRMPD spectrum is obtained.

1.1.3 Problems in theoretical calculations

Understanding the spectroscopic signatures obtained by IRMPD requires theoretical calculations, for the identification of structures and for the assignment of observed features. The normal routine is to locate the most stable structures and calculate the harmonic frequencies. Comparison between experimental and theoretical

results leads to the identification of the dominant isomer. However, it has long been realized that the ground state potential surfaces for the ions are not flat and contain many local minima separated by small barriers. It is also known that in solvated clusters, the most stable structures in terms of energy is not necessarily the most favored in terms of free energy. A flat potential surface and inter-conversions among various isomers are important considerations. For understanding the vibrational spectra, harmonic analysis may not be good enough, and anharmonic effects could play a significant role. For understanding the relative stabilities, energies obtained by structural optimizations at 0 K are not the sole concern, as entropy effects obtained by dynamic simulations at finite temperature could also be important.

IRMPD spectroscopy requires multiple vibrational excitations to produces the detected signals. Obviously, knowledge about the vibrational density of sates at higher temperature should be very helpful, although such information is neglected in the harmonic analysis. Ab initio molecular dynamics (AIMD) method is well suited for the study of IRMPD processes.[31,32] With the potential energy and atomic forces calculated from first principles (typically by DFT), it can simulate the dynamics of a floppy ion in real time and at finite temperature. And the vibrational density of states can be easily obtained by the Fourier transform of the time correlation function for velocity.[33,34] And the intensity of vibrational peaks as a function of frequency can be obtained by Berry's phase method[35] and Fourier transform of the dipole time correlation function, with dynamic and anharmonic effects naturally included.[36,37]

1.2 General background

1.2.1 Schrödinger equation

In quantum chemistry, the key problem is to solve the Schrödinger equation for many-electron systems, which is formulated in 1926 by Austrian physicist Erwin Schrödinger.[38] The general form of the time-dependent Schrödinger equation can be expressed as

$$-i\hbar\frac{\partial\Psi(x,t)}{\partial t}=\hat{\mathrm{H}}\Psi(x,t) \tag{1.1}$$

where $\Psi(x,t)$ is the wave-function, dependent on time t and space coordinate x. $i\hbar\frac{\partial}{\partial t}$ is the energy operator. And $\hat{\mathrm{H}}$ is the Hamiltonian operator of the system. $\hat{\mathrm{H}}$ can be written as

$$\hat{\mathrm{H}} = -\frac{\hbar^2}{2m}\nabla_x^2 + V(x,t) \tag{1.2}$$

For the stationary state, the time-dependent Schrödinger equation can be separated into time-dependent part and time-independent part (space part),

$$\Psi(x,t) = \Phi(x)f(t) \tag{1.3}$$

$$E\Phi(x) = \hat{\mathrm{H}}\Phi(x) \tag{1.4}$$

with (1.4) being the time-independent Schrödinger equation.

For a molecule system containing N electrons, the Hamiltonian can be written as

$$\hat{\mathrm{H}} = -\sum_A \frac{\hbar^2}{2M_A}\nabla_A^2 - \sum_i \frac{\hbar^2}{2m_i}\nabla_i^2 + \sum_{A<B}\frac{Z_A Z_B e^2}{\left|\vec{R}_A - \vec{R}_B\right|} - \sum_{A,i}\frac{Z_A e^2}{\left|\vec{R}_A - \vec{r}_i\right|} + \sum_{i<j}\frac{e^2}{\left|\vec{r}_i - \vec{r}_j\right|} \tag{1.5}$$

where $\vec{R}_A$ is the position of nuclei and $\vec{r}_i$ is the position of *ith*-electron, M_A is the mass and Z_A the atomic number of nucleus A. The Laplacian operators ∇_A^2 and ∇_i^2 involve differentiation with respect to the coordinates of the Ath nucleus and the ith electron. The first two terms are the kinetic energy operators of the nucleus and the electrons, respectively. The last three terms are the potential energy operators for coulomb interactions between nucleus-nucleus, nucleus-electron and electron-electron, respectively. This equation cannot be solved exactly for complex systems and additional approximations must be made to simulate chemical systems.

1.2.2 Born-Oppenheimer approximation

Born-Oppenheimer approximation makes it possible to compute the wavefunction in electronic structure calculation in consecutive steps. The approximation was proposed by Born and Oppenheimer in 1927[39] which was based on the fact that the nuclei were much heavier than the electrons and hence they move much more slowly than the electrons. Using Born-Oppenheimer approximation, the

wavefunction and Hamiltonian can be separated into an electronic and a nuclear part. The final form of Schrödinger equation can be written as

$$\Phi_{molecule}(\vec{R},\vec{r}) = \Phi_{\text{nuclei}}(\vec{R})\Phi_{\text{elec}}(\vec{R},\vec{r}) \tag{1.6}$$

$$\hat{\mathrm{H}} = \hat{\mathrm{H}}_{elec} + \hat{\mathrm{H}}_{nuclei} \tag{1.7}$$

$$\hat{\mathrm{H}}_{\text{elec}}(\vec{R})\Phi_{\text{elec}}(\vec{R},\vec{r}) = E_{\text{elec}}\Phi_{\text{elec}}(\vec{R},\vec{r}) \tag{1.8}$$

Even with this approximation, solving the N-electron system is still very expensive and further approximation must be made.

1.3 Density functional theory (DFT)

Density functional theory (DFT), developed since 1964[40], has become popular due to its computational efficiency and high accuracy. The modern DFT originates from the famous Hohenberg-Kohn theorems proposed by Hohenberg and Kohn. The first theorem provides the proof of DFT stating that the ground state electronic energy E_0 is a functional of the ground state electron density $\rho_0 : E_0 = E_0[\rho_0]$. Therefore the properties of the system are formally determined by the ground state density. The second one is the functional which delivers the ground state energy of the system, delivers the lowest energy if and only if the input density is the true ground state density ρ_0. In this theory, the interaction between electrons is described by its ground state electron density instead of its 3N variables wave functions. There are only three variables for the density whereas for a wave function there are 3N degrees of freedom.

1.3.1 Kohn-Sham (KS) scheme

Kohn and Sham[41] formulated a method for the implementation of DFT. They introduced the concept of a non-interacting system built from KS orbitals $\{\chi_i^{\mathrm{KS}}\}$ with $\rho(\vec{r}) = \sum_i \sum_{\omega=\alpha,\beta} \chi_i^{\mathrm{KS}}(\vec{r}_i, w)$. The ground state energy of the system in the Hohenberg-Kohn theorem can be written as

$$E_0 = \min_{\rho \to N}(T[\rho] + \int \rho(\vec{r}) V_{\mathrm{Ne}} d\vec{r}) \tag{1.9}$$

where the kinetic energy functional $T[\rho]$ is

$$T_s[\rho(\vec{r})] = -\frac{1}{2}\sum_i^N \left\langle \chi_i^{\mathrm{KS}} \middle| \nabla_i^2 \middle| \chi_i^{\mathrm{KS}} \right\rangle \tag{1.10}$$

For a real system, the ground state electronic energy can be expressed as

$$E_0 = \min_{\rho \to N}(F[\rho] + \int \rho(\vec{r}) V_{\mathrm{Ne}} d\vec{r}) \tag{1.11}$$

$$F[\rho(\vec{r})] = T[\rho(\vec{r})] + J[\rho(\vec{r})] + E_{\mathrm{ncl}}[\rho(\vec{r})] \tag{1.12}$$

where, $T[\rho(\vec{r})]$ is the kinetic energy term, $J[\rho(\vec{r})]$ is the classical electrostatic electron-electron repulsion term and $E_{\mathrm{ncl}}[\rho(\vec{r})]$ is the sum of other non-classical electrostatic interaction term. For the kinetic energy functional $T[\rho(\vec{r})]$, Kohn and Sham suggested to use the exact kinetic energy functional $T_s[\rho(\vec{r})]$ of the non-interacting system with the same density as the real interacting one $T[\rho(\vec{r})]$, and put the difference between $T_s[\rho(\vec{r})]$ and $T[\rho(\vec{r})]$ into the exchange-correlation term E_{XC}. Then the real system energy can be written as

$$E[\rho(\vec{r})] = T_s[\rho(\vec{r})] + J[\rho(\vec{r})] + E_{\mathrm{XC}}[\rho(\vec{r})] + \int \rho(\vec{r}) V(\vec{r}) d\vec{r} \tag{1.13}$$

and E_{XC} is define as

$$\begin{aligned} E_{\mathrm{XC}}[\rho(\vec{r})] &= (T[\rho(\vec{r})] - T_s[\rho(\vec{r})]) + (E_{\mathrm{ee}}[\rho(\vec{r}) - J[\rho(\vec{r})]) \\ &= T_C[\rho(\vec{r})] + E_{\mathrm{ncl}}[\rho(\vec{r})] \end{aligned} \tag{1.14}$$

The exchange-correlation functional includes two terms. One is the correction term for kinetic energy $T_C[\rho(\vec{r})]$ and the other is the non-classical electrostatic $E_{\mathrm{ncl}}[\rho(\vec{r})]$ contribution.

The orthogonal KS orbitals can be determined by KS equations:

$$-\frac{1}{2}\nabla_i^2 \chi_i^{\mathrm{KS}}(\vec{r}_1) + V_s(\vec{r}_1)\chi_i^{\mathrm{KS}}(\vec{r}_1) = \varepsilon_i \chi_i^{\mathrm{KS}}(\vec{r}_1) \tag{1.15}$$

The effective potential of electrons

$$V_s(\vec{r}_1) = \int d\vec{r}_2 \frac{1}{|\vec{r}_1 - \vec{r}_2|} \left|\chi_j^{\mathrm{KS}}(\vec{r}_2)\right|^2 + V_{XC}(\vec{r}_1) - \sum_I \frac{Z_I}{\left|\vec{R}_I - \vec{r}_1\right|} \tag{1.16}$$

where the exchange-correlation potential is

$$V_{\mathrm{XC}}(\vec{r}_1) = \frac{\partial E_{\mathrm{XC}}[\rho(\vec{r}_1)]}{\partial \rho(\vec{r}_1)} \tag{1.17}$$

The KS energies could be obtained by the minimization of the total energy expressed as a functional of the electron density based on the variational principle.

Usually the two terms E_{XC} and V_{XC} could not be known exactly. Therefore, the central problem in DFT is to find good approximations to E_{XC} and V_{XC}. Currently the widely used schemes are local density approximation (LDA), generalized gradient approximation (GGA) and hybrid functional (B3LYP and etc.).

1.3.2 Local density approximation (LDA)

Local density approximation (LDA) is the simplest way to approximate the exchange-correlation functional E_{XC}. This approximation is based on the hypothesis of uniform electron gas which is fairly good for the simple metals such as sodium. In this method, the exchange-correlation functional E_{XC} can be expressed as

$$E_{\mathrm{XC}}^{\mathrm{LDA}}[\rho(\vec{r})] = \int \rho(\vec{r}) \varepsilon_{\mathrm{XC}}[\rho(\vec{r})] d\vec{r} \tag{1.18}$$

where $\varepsilon_{\mathrm{XC}}[\rho(\vec{r})]$ is the exchange-correlation energy per particle of a uniform electron gas with density $\rho(\vec{r})$. Furthermore, the term $\varepsilon_{\mathrm{XC}}[\rho(\vec{r})]$ can be split into an exchange and a correlation part:

$$\varepsilon_{\mathrm{XC}}[\rho(\vec{r})] = \varepsilon_{\mathrm{X}}[\rho(\vec{r})] + \varepsilon_{\mathrm{C}}[\rho(\vec{r})] \tag{1.19}$$

where $\varepsilon_{\mathrm{X}}[\rho(\vec{r})]$ is the exchange energy of an electron in a uniform electron gas of a particular density, which is also equal to the HF exchange functional, originally derived by Bloch and Dirac[42]:

$$\varepsilon_{\mathrm{X}}[\rho(\vec{r})] = -\frac{3}{4}\sqrt[3]{\frac{3\rho(\vec{r})}{\pi}} \tag{1.20}$$

$$E_{\mathrm{X}}[\rho] \cong C_{\mathrm{X}} \int \rho(\vec{r})^{\frac{4}{3}} d\vec{r} \tag{1.21}$$

In this equation, the exchange functional is only dependent on the electron density at position $\vec{r}$. Therefore, it was named local density approximation (LDA).

For the correlation functional part ε_{C}, the most widely used formula is developed by Vosko, Wilk and Nusair in 1980[43] and by Perdew and Wang in 1992[44]. And currently, the most popular LDA schemes are SVWN and VWN5.

The LDA approximation is developed on the basis of uniform charge density and produces significant errors in the non-uniform charge density, which is prevalent in quantum chemistry and condensed matter physics, in which the non-local corrections are important.

1.3.3 Generalized gradient approximation (GGA)

In generalized gradient approximation (GGA), electron density gradient is added to correct the E_{XC}, written as

$$E_{\mathrm{XC}}^{\mathrm{GGA}}[\rho(\vec{r})] = \int \rho(\vec{r}) \varepsilon_{\mathrm{XC}}[\rho(\vec{r})] d\vec{r} + F_{\mathrm{XC}}[\rho(\vec{r}), \nabla\rho(\vec{r})] d\vec{r} \tag{1.22}$$

Some popular exchange functions are developed by Becke in 1988[45] (B) and 1986 (B86), by Perdew in 1986 (P), and by Perdew, Burke and Ernzerhof in 1996 (PBE).[46,47] There are also a number of gradient corrected correlation functionals such as P86, PW91,[44] and LYP[48].

1.3.4 Hybrid functionals

GGA functionals have their own problems for example, overestimate of hydrogen bond strength and inadequacy for van der Waals interactions. But they provide a large improvement in the energy calculations, and can usually produce results with reasonable chemical accuracy. Further improvement in accuracy can be obtained by combining the HF exchange integrals with a DFT exchange-correlation functional. Such a mixture is generally known as a hybrid functional.

Currently, the most popular hybrid functional is B3LYP[49] which can be expressed as

$$E_{\mathrm{XC}}^{\mathrm{B3LYP}} = (1-a)E_{\mathrm{X}}^{\mathrm{LSD}} + aE_{\mathrm{X}}^{\mathrm{HF}} + bE_{\mathrm{X}}^{\mathrm{B88}} - cE_{\mathrm{C}}^{\mathrm{LYP}} + (1-c)E_{\mathrm{C}}^{\mathrm{LSD}} \tag{1.23}$$

The three parameters *a*, *b* and *c* are be obtained by fitting. $E_{\mathrm{X}}^{\mathrm{LSD}}$ is the extended LDA in unrestricted case (local spin-density approximation). $E_{\mathrm{X}}^{\mathrm{HF}}$ is the exchange functional from HF approximation. $E_{\mathrm{X}}^{\mathrm{B88}}$ refers to exchange functional developed by Becke at 1988.[45] $E_{\mathrm{C}}^{\mathrm{LYP}}$ is the correlation functional developed by Lee, Yang and Parr.[48]

1.4 *Ab initio* molecular dynamics (AIMD)

1.4.1 Molecular dynamics

Molecular dynamics is a powerful technique to investigate the equilibrium and transport properties of many–body systems. The nuclear motion of the particles follows the laws of classical mechanics. This is a very good approximation since quantum effects are only important at low temperature and mostly for the lightest atom, H. The Hamiltonian of N-particle system can be written as

$$\hat{\mathrm{H}}(\{\vec{R}_I\},\{\vec{P}_I\}) = \sum_I \frac{\vec{P}_I^2}{2M_I} + U(\{\vec{R}_I\}) \tag{1.24}$$

The forces acted on the particles can be derived from the potential:

$$\vec{F}_I(\{\vec{R}_I\}) = M_I\ddot{\vec{R}}_I = -\frac{\partial U(\{\vec{R}_I\})}{\partial \vec{R}_I} \tag{1.25}$$

The equations of motion can be expressed as

$$\dot{\vec{R}}_I = \frac{\partial \hat{\mathrm{H}}}{\partial \vec{P}_I} = \frac{\vec{P}_I}{M_I} \tag{1.26}$$

$$\dot{\vec{P}}_I = -\frac{\partial \hat{\mathrm{H}}(\{\vec{R}_I\},\{\vec{P}_I\})}{\partial \vec{R}_I} = -\frac{\partial U(\{\vec{R}_I\})}{\partial \vec{R}_I} = \vec{F}_I(\{\vec{R}_I\}) \tag{1.27}$$

Also, the equations of motion can be written by using the Lagrange formalism with Euler-Lagrange equation:

$$L(\{\vec{R}_I\},\{\dot{\vec{R}}_I\}) = \sum_I \frac{1}{2} M_I \dot{\vec{R}}_I^2 - U(\{\vec{R}_I\}) \tag{1.28}$$

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\vec{R}}_I} = \frac{\partial L}{\partial \vec{R}_I} \tag{1.29}$$

which leads to the same result. The two formulations expressed above are equivalent.

The traditional approach followed in molecular dynamics is to determine the potentials in advance. The full potential can be broken up into two-body and many-body interactions, long-rang and short-range terms, electrostatic and non-electrostatic interactions, etc.. Empirical potentials can describe many physical processes, such as thermal motions, harmonic vibration, and phase changes. But for chemical reactions, which involve the formation and breaking of chemical bonds, empirical potentials are generally inadequate, since the energy change induced by electron rearrangement cannot be easily broken into two-body or three-body potentials.

1.4.2 *Ab initio* molecular dynamics (AIMD)

The development of AIMD method makes it possible to simulate the dynamic processes involving chemical reactions. The seminal contributions of Roberto Car and Michele Parrinello came in the mid-1980s.[50] They unified molecular dynamics and electronic structure theory in a method often dubbed as "Car–Parrinello molecular dynamics". In such a scheme, the forces are described in the framework of quantum mechanics at the level of DFT. The dynamics is captured in terms of classical and statistical mechanics. Due to its inherent advantages, AIMD is currently a very popular and useful computational tool to study physical, chemical, and biological phenomena. Most important physical properties can be calculated by AIMD simulations, such as optical, IR, Raman, NMR, dipole moments, localized orbitals, and effective atomic charges.[51]

The basic idea of the Car-Parrinello method is to exploit the quantum-mechanical adiabatic time-scale separation of fast electronic and slow nuclear motion by transforming it into classical-mechanical adiabatic energy scale separation in the

framework of dynamical systems theory.[13] Car and Parrinello postulated the following class of Lagrangian:

$$L_{CP} = \sum_I \frac{1}{2} M_I \dot{R}_I^2 + \sum_i \frac{1}{2} \mu_i \langle \dot{\psi}_i \mid \dot{\psi}_i \rangle - \langle \Psi_0 \mid H_e \mid \Psi_0 \rangle + \text{constraints} \tag{1.30}$$

The corresponding Newtonian equations of motion are obtained from the associated Euler-Lagrange equations:

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{\vec{R}}_I} = \frac{\partial L}{\partial \vec{R}_I} \tag{1.31}$$

$$\frac{d}{dt} \frac{\delta L}{\delta \dot{\psi}_i^*} = \frac{\delta L}{\delta \psi_i^*} \tag{1.32}$$

The generic Car-Parrinello equations of motion can be found to be the form:

$$M_I \ddot{R}_I(t) = -\frac{\partial}{\partial R_I} \langle \Psi_0 \mid H_e \mid \Psi_0 \rangle + \frac{\partial}{\partial R_I} \{\text{constraint}\} \tag{1.33}$$

$$\mu_i \ddot{\psi}_i(t) = -\frac{\delta}{\delta \ddot{\psi}_i^*} \langle \Psi_0 \mid H_e \mid \Psi_0 \rangle + \frac{\delta}{\delta \psi_i^*} \{\text{constraint}\} \tag{1.34}$$

where $\mu_i (= \mu)$ are the "fictitious mass" assigned to the orbital degrees of freedom; the units of the mass parameter μ are energy times a squared time of reasons of dimensionality. Such an approach cuts down the computational cost. The electronic subsystem is optimized close to the instantaneous minimum energy or to the exact Born-Oppenheimer surface. Thus, a ground-state wavefunction optimized for the initial configuration of the nuclei will stay close to its ground state during the time evolution. One drawbacks of this approach is that adiabaticity between nuclei and electronic subsystem cannot be guaranteed during long time MD run. And recently developed wave function extrapolation method makes the Born-Oppenheimer molecular dynamics faster than the traditional Car-Parrinello molecular dynamics.

An alternative approach for including the electronic structure in molecular dynamics simulations is to solve the static electronic structure problem at each MD step, given a set of fixed nuclear positions. Thus, the propagating of the nuclei can be reduced to solving the classical molecular dynamic while the electronic structure part

is to solving a time independent quantum problem. The Born-Oppenheimer MD method can be defined as (1.35) and (1.36) for the electronic ground state.

$$M_I \ddot{R}_I(t) = -\nabla_I \min_{\Psi_0} \{\langle \Psi_0 \mid H_e \mid \Psi_0 \rangle\} \tag{1.35}$$

$$E_0 \Psi_0 = H_e \Psi_0 \tag{1.36}$$

The equations of motion can be formulated to be effective one-particle Hamiltonians. This might be the Hartree-Fock approximations defined by the variational minimum of the energy expectation value $\langle \Psi_0 \mid H_e \mid \Psi_0 \rangle$ given a single Slater determinant $\Psi_0 = \det\{\psi_i\}$. As derived in standard textbooks, the equations of motion corresponding to Eqs. (1.35)-(1.36) read

$$M_I \ddot{R}_I(t) = -\nabla_I \min_{\{\psi_i\}} \{\langle \Psi_0 \mid H_e^{HF} \mid \Psi_0 \rangle\} \tag{1.37}$$

$$0 = -H_e^{HF} \psi_i + \sum_j \Lambda_{ij} \psi_j \tag{1.38}$$

for the Hartree-Fock case. A similar set of equations is obtained if Hohenberg-Kohn-Sham density functional theory is used, where H_e^{HF} has to be replaced by Kohn-Sham one-article Hamiltonian H_e^{KS}.

1.4.3 Basis set

Wavefunction is often expressed as a linear combination of the basis set functions. Plane wave (PW) basis set is a good choice for periodic systems. The periodic parameters produce a periodic potential, which imposes the same periodicity on the density. Plane waves are used for the expansion of orbitals because they are complete and orthogonal functions. Plane waves can be written as

$$f_G^{\mathrm{PW}}(\vec{r}) = \frac{1}{\sqrt{\Omega}} \exp\left[i\vec{G} \cdot \vec{r}\right] \tag{1.39}$$

where Ω is the volume of the lattice (box). $\vec{G}$ is the reciprocal space vectors which also satisfies the particular periodic boundary conditions.

In Bloch's theorem, a periodic wave function can be expanded with plane wave basis set as:

$$\psi(\vec{r}) = \psi(\vec{r} + \vec{L}) = \frac{1}{\sqrt{\Omega}} \sum_{\vec{G}} \psi(\vec{G}) \exp\left[i\vec{G} \cdot \vec{r}\right] \tag{1.40}$$

where $\psi(\vec{r})$ and $\psi(\vec{G})$ are related by a 3D Fourier transform. $\vec{L}$ is the lattice vectors.

The periodic extended Kohn-Sham orbitals in the plane wave basis form can be written as:

$$\Phi(\vec{r}) = \frac{1}{\sqrt{\Omega}} \sum_{\vec{G}} c_i(\vec{G}, \vec{k}) \exp[i(\vec{G} + \vec{k}) \cdot \vec{r}] \tag{1.41}$$

where $\vec{k}$ is a reciprocal vector in the first Brillouin zone and $c_i(\vec{G}, \vec{k})$ is complex coefficient.

There are both advantages and disadvantages for using plane wave basis. One advantage is that the plane wave basis is independent on the position of nuclei and no Pulay force[52] exists. Secondly, wave functions can be efficiently solved via Fast Fourier Transform (FFT). One disadvantages is the unbiased property of plane wave basis set is not suitable for description biased regions where more functions are required than others regions. Another one is that it is difficult to expand electron wave functions in the rapidly oscillation core regions. And to overcome this problem, pseudopotentials are usually used.

Gaussian function is another widely used basis set for the expansion of the electron wave functions. The electron density $n(\vec{r})$ can also be expressed based on an expansion in atom centered, contracted Gaussian functions:

$$n(\vec{r}) = \sum_{uv} P^{uv} \varphi_u(\vec{r}) \varphi_v(\vec{r}) \tag{1.42}$$

where P^{uv} is a density matrix element and $\varphi_u(\vec{r}) = \sum_i d_{iu} g_i(\vec{r})$ with primitive Gaussian functions $g_i(\vec{r})$ and the corresponding contraction coefficients is d_{iu}. Different types of Gaussian basis set are widely used, such as 6-31G, 6-311++G**, and etc. One disadvantage of Gaussian basis set is that it is non-periodic and has limitations in implementations in periodic system. But a hybrid Gaussian and plane

wave density function scheme (GPW)[53] could combine the advantages of both plane waves and Gaussian basis set together.

In the hybrid Gaussian and plane wave density function scheme (GPW), plane waves are employed as an auxiliary basis and given by[54]

$$\tilde{n}(\vec{r}) = \frac{1}{\Omega}\sum_{\vec{G}} \tilde{n}(\vec{G}) \exp(i\vec{G}\cdot\vec{r}) \tag{1.43}$$

where Ω is the volume of the unit cell, and $\vec{G}$ are the reciprocal lattice vectors. The expansion coefficients $\tilde{n}(\vec{G})$ are such that $\tilde{n}(\vec{r})$ is equal to $n(\vec{r})$ on a regular grid in the unit cell. This choice allows for a rapid conversion between $n(\vec{r})$, $\tilde{n}(\vec{r})$ and $\tilde{n}(\vec{G})$ using an efficient mapping procedure and fast Fourier transforms (FFT).

Using the hybrid Gaussian and plane wave density function scheme, the Kohn-Sham DFT energy expression employed is written as

$$\begin{aligned} E[n] &= E^T[n] + E^V[n] + E^H[n] + E^{XC}[n] + E^{II}[n] \\ &= \sum_{uv} P^{uv} \left\langle \varphi_u(\vec{r}) \middle| -\frac{1}{2}\nabla^2 \middle| \varphi_v(\vec{r}) \right\rangle \\ &+ \sum_{uv} P^{uv} \left\langle \varphi_u(\vec{r}) \middle| V_{loc}^{PP}(\vec{r}) \middle| \varphi_v(\vec{r}) \right\rangle \\ &+ \sum_{uv} P^{uv} \left\langle \varphi_u(\vec{r}) \middle| V_{nl}^{PP}(\vec{r},\vec{r}') \middle| \varphi_v(\vec{r}') \right\rangle \\ &+ 2\pi\Omega \sum_{G} \frac{\tilde{n}^*(\vec{G})\tilde{n}(\vec{G})}{G^2} + \int e^{XC}(\vec{r})dr \\ &+ \frac{1}{2}\sum_{I\neq J} \frac{Z_I Z_J}{\left|\vec{R}_I - \vec{R}_J\right|} \end{aligned} \tag{1.44}$$

where $E^T[n]$ is the electronic kinetic energy, $E^V[n]$ is the electronic interaction with ionic cores, $E^H[n]$ is the electronic Hartree energy and $E^{XC}[n]$ is the exchange-correlation energy. The interaction energies of ionic cores with charges Z_I and positions R_I are denoted by $E^{II}[n]$. $E^V[n]$ is the norm-conserving Goedecker, Teter, and Hutter (GTH) type pseudopotentials,[55,56] split into local part $V_{loc}^{PP}(\vec{r})$ and a fully non-local part $V_{nl}^{PP}(\vec{r},\vec{r}')$.

1.4.4 Pseudopotentials (PP)

As a large number of plane waves are required to describe the core electrons, the computational cost of an all-electron-plane-wave calculation is huge. The pseudopotential approximation is to replace the true nuclear potential $V_{\text{ion}}(\vec{r})$, which is due to the core electrons and the nuclei, with a weaker pseudopotentials $V_{\text{ion}}^{\text{PP}}(\vec{r})$. The wavefunction ψ is described by a pseudo wavefunction $\psi_{\text{ion}}^{\text{PP}}$. $\psi_{\text{ion}}^{\text{PP}}$ should be smooth and nodeless close to the nuclear core region and identical to the all-electron wavefunction outside a chosen cutoff radius, showing the same scattering properties as illustrated in Figure 1.3.[57] The approximation is based on the fact that the valence electrons play a more important role than the tightly bound core electrons in determining the physical and chemical properties of molecules.

Two most commonly used pseudopotentials are Vanderbilt's Ultra-Soft pseudopotentials (US-PP)[58] and Projector Augmented Wave potential (PAW).[59,60]

US-PP was proposed by Vanderbilt in the early 1990s.[58] Such pseudopotentials attain much smoother/softer pseudo-wavefunctions while using considerably fewer plane waves for calculations of the same accuracy. The pseudo-wavefunctions in the core region are soft to make it possible for a drastic reduction in the cutoff kinetic energy. The material properties of complex ionic systems consisting of many different kinds of atoms can be well reproduced in calculations with US-PP.

The projector augmented wave potential (PAW) potential was introduced by Blöchl in 1994.[59] The wavefunction is divided into three parts, atomic-like partial waves $\{\phi_i\}$, auxiliary partial waves $\{\widetilde{\phi}_i\}$ expanded within atomic sphere regions, and auxiliary wavefunction $\{\widetilde{\psi}_i\}$ expanded outside the spheres. Generally the PAW potentials are more accurate than the US-PP types pseudopotentials.

For the implementation of the GPW basis set, Goedecker-Teter-Hutter type pseudopotential (GTH) was developed in 1996 by Goedecker, Teter, and Hutter.[55,56] It is accurate and transferable, and has an analytic form that allows for an efficient

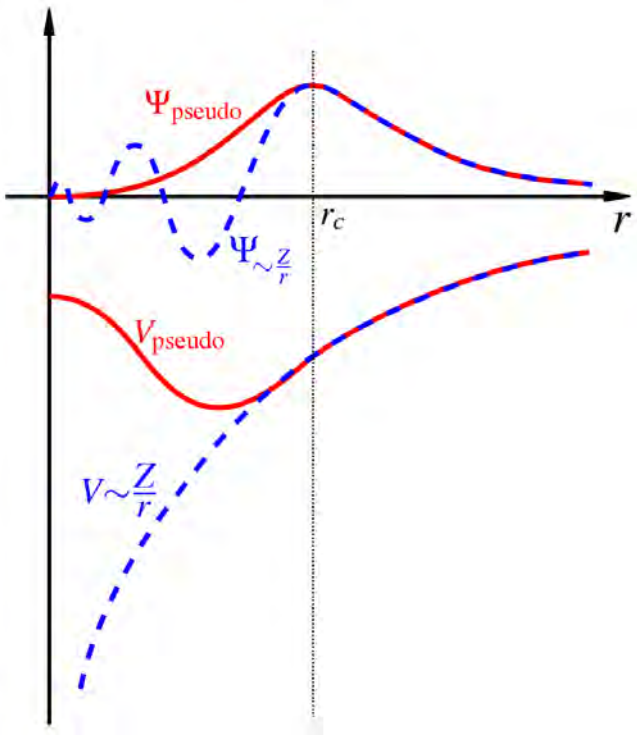

Figure1.3: Illustrations of the full all-electronic wavefunction and electronic potentials (blue dashed lines) plotted against distance *r* from the atomic nucleus. The corresponding pseudo-wavefunction and pseudo-potential are plotted in red lines. Outside the cutoff radius r_c the all electron and pseudo electron values should be same.[57]

treatment of all energy functional terms within the GPW method. The norm-conserving, separable, dual-space GTH pseudopotentials consist of a local part including a long-ranged (LR) and a short-ranged (SR) term:

$$\begin{aligned} V_{loc}^{PP}(\vec{r}) &= V_{loc}^{LR}(\vec{r}) + V_{loc}^{SR}(\vec{r}) \\ &= -\frac{Z_{ion}}{r} erf(\alpha^{PP}\vec{r}) + \sum_{i=1}^{4} C_i^{PP} \left(\sqrt{2}\alpha^{PP}\vec{r}\right)^{2i-2} \\ &\times \exp[-(\alpha^{PP}\vec{r})^2] \end{aligned} \tag{1.45}$$

with

$$\alpha^{PP} = \frac{1}{\sqrt{2} r_{loc}^{PP}}$$

and a non-local part

$$V_{nl}^{PP}(\vec{r},\vec{r}') = \sum_{lm}\sum_{ij} \left\langle \vec{r} \mid \vec{p}_i^{lm} \right\rangle h_{ij}^{l} \left\langle \vec{p}_j^{lm} \mid \vec{r}' \right\rangle \tag{1.46}$$

with the Gaussian-type projectors

$$\left\langle \vec{r} \mid \vec{p}_i^{lm} \right\rangle = N_i^l Y^{lm}(\hat{r}) r^{l+2i-2} \exp[-\frac{1}{2}\left(\frac{\vec{r}}{\vec{r}_l}\right)^2] \tag{1.47}$$

where N_i^l are the normalization constants and $Y^{lm}(\hat{r})$ is spherical harmonics.

1.5 Dipole moment and infrared spectrum

Wannier functions is a powerful tool[61] in the study of the condensed matters. It is equivalent to localized molecular orbitals[62] and provides an insightful picture of the nature of chemical bonding, which is missing from the Bloch extended orbitals. By transforming the occupied electronic manifold into a set of maximal localized Wannier functions, it is possible to obtain the chemical coordination and bonding properties. In most cases, Wannier function analyses can be carried out in the pseudopotential framework.

Within density functional theory, the polarization of a system can be written as:

$$\vec{M}_\alpha^{el} = \frac{2e}{|\vec{G}_\alpha|} \operatorname{Im} \ln \det \vec{Z}_\alpha \tag{1.48}$$

α denotes the Cartesian components of the electronic contribution M^{el} to the dipole moment vector M of the total system. The matrices Z_α are defined in terms of matrix elements of Kohn-Sham orbitals:

$$(Z_\alpha)_{kl} = \left\langle \Phi_k \left| \exp\left[-i\vec{G}_\alpha \cdot \vec{r}\right] \right| \Phi_l \right\rangle \tag{1.49}$$

where the indices *k* and *l* run over all occupied orbitals. The nuclear contributions obtained from the positions and effective core charges of nuclei according to the pseudopotentials can be added to get the total dipole moment of the system.

By carrying out a suitable unitary transformation U, the localized Wannier functions with a set of occupied Bloch Kohn-Sham orbitals can be expressed as

$$\omega_k(\vec{r}) = \sum_l U_{kl} \phi_l(\vec{r}) \tag{1.50}$$

The expectation value of position operator is

$$\vec{r}_k = -\frac{L}{2\pi}\,\mathrm{Im}\ln\vec{z}_k \tag{1.51}$$

which is the center of a given maximally localized Wannier function. k labels the different Wannier orbitals/electrons. With this definition the electronic part dipole moment of the whole system can be approximated by $M^{el} = -2e\sum_k \vec{r}_k$, where the sum includes all occupied states, and molecular dipole moments can be defined as

$$\vec{\mu}_I^{el} = -2e\sum_{k\in I}\vec{r}_k \tag{1.52}$$

where the sum runs over all Wannier function centers that can be associated with molecule ***I***. This is possible as Wannier function centers for closed shell molecules tend to be located close to atoms or bond centers and can therefore be attributed uniquely to a molecule. In the Wannier representation, as a simple sum of position vector $\vec{r}_k$ multiplied by the appropriate charges (i.e. -1e and -2e in spin-polarized and spin-restricted calculations, respectively), like for classical point particles.

The infrared absorption spectra can be calculated by the time-correlation function [61] with for a set of total dipole moments obtained at each molecular dynamics step:

$$\alpha(\omega) = \frac{2\pi\beta\omega^2}{3n(\omega)cV}\int_{-\infty}^{+\infty} dt\left\langle \vec{M}(t)\bullet\vec{M}(0)\right\rangle \exp(i\omega t) \quad \textbf{\textit{Error! Reference source not found.}} \tag{1.53}$$

where**Error! Reference source not found.** $\beta = 1/kT$, **Error! Reference source not found.**is the refractive index, c is the speed of light in vacuum, and V is the volume. **Error! Reference source not found.** is the total dipole moment of the system. In this formula, a quantum correction factor in the form of **Error! Reference source not found.** is multiplied into the classical line shape, which was shown to provide the most accurate results on the calculated IR amplitudes[62]. The assignment of each active band into individual atomic displacements or vibrational modes is done by using the vibrational density of states (VDOS), which is obtained by the Fourier form of the atomic velocity autocorrelation functions[61]:

Error! Reference source not found. (1.54)

Where i runs over all toms of the system.

1.6 Reference

[1]. H. B. Oh, C. Lin, H. Y. Hwang, H. L. Zhai, K. Breuker, V. Zabrouskov, B. K. Carpenter, and F. W. McLafferty, J. Am. Chem. Soc. **2005**, 127, 4076.

[2]. M. Okumura, L. I. Yeh, and Y. T. Lee, J. Chem. Phys. **1985**, 83, 3705.

[3]. M. Okumura, L. I. Yeh, J. D. Myers, and Y. T. Lee, J. Phys. Chem. **1990**, 94, 3416.

[4]. J. M. Lisy, J. Chem. Phys. **2006**, 125, 132302.

[5]. M. A. Duncan, Int. Rev. in Phys. Chem. **2003**, 22, 407.

[6]. A. Gapeev and R. C. Dunbar, J. Am. Chem. Soc. **2001**, 123, 8360.

[7]. O. P. Balaj, C. Kapota, J. Lemaire, and G. Ohanessian, Int. J. Mass Spectrometry **2008**, 269, 196.

[8]. A. Simon, L. MacAleese, P. Maitre, J. Lemaire, and T. B. McMahon, J. Am. Chem. Soc. **2007**, 129, 2829.

[9]. C. Kapota, J. Lemaire, P. Maitre, and G. Ohanessian, J. Am. Chem. Soc. **2004**, 126, 1836.

[10]. M. F. Bush, J. T. O'Brien, J. S. Prell, R. J. Saykally, and E. R. Williams, J. Am. Chem. Soc. **2007**, 129, 1612.

[11]. V. S. Letokhov, Nonlinear Laser Chemistry, Springer-Verlag, Berlin, Heidelberg, New York, **1983**.

[12.] M. Quack, J. Chem. Phys. **1978**, 69, 1282.

[13]. M. Quack, Adv. Chem. Phys. **1982**, 50, 395.

[14]. V. S. Letokhov, Progress in optics **1978**, 16.

[15]. M. Quack, Infrared Phys. Technol. **1995**, 36, 365.

[16]. M. Quack, Infrared Phys. **1989**, 29, 441.

[17]. M. Quack, Encyclopedia of Computational Chemistry, Vol. 3 and John Wiley & Sons Chichester, UK, **1998**, pp. 1775.

[18]. N. C. Polfer and J. Oomens, Phys. Chem. Chem. Phys. **2007**, 9, 3804

[19]. D. P. Little, J. P. Speir, M. W. Senko, P. B. O'Connor and F. W. McLafferty, Anal. Chem. **1994**, 66, 2809.

[20]. R. N. Rosenfeld, J. M. Jasinski and J. I. Braumann, J. Am. Chem. Soc. **1982**, 104, 658.

[21]. S. K. Shin and J. L. Beauchamp, J. Am. Chem. Soc. **1990**, 112, 2066.

[22]. D. M. Peiris, M. A. Cheeseman, R. Ramanathan and J. R. Eyler, J. Phys. Chem. **1993**, 97, 7839.

[23]. D. A. Kirkwood and A. J. Stace, J. Am. Chem. Soc. **1998**, 120, 12316.

[24]. S. R. Atrill and A. J. Stace, Phys. Chem. Chem. Phys. **2000**, 2, 823.

[25]. J. Oomens, A. J. A. van Roij, G. Meijer and G. von Helden, Astrophys. J. **2000**, 542, 404.

[26]. H. P. Freund and T. M. Antonsen, Jr, Principles of Free Electron Lasers, Chapman & Hall, London, **1992**.

[27]. P. Luchini and H. Motz, Undulators and Free-Electron Lasers, Clarendon Press, Oxford, **1990**.

[28]. J. B. Fenn, M. Mann, C. K. Meng and C. M. Whitehouse, Science. **1989**, 246, 64.

[29]. A. G. Marshall, T. C. L. Wang and T. L. Ricca, J. Am. Chem. Soc. **1985**, 107, 7893;

[30]. S. H. Guan and A. G. Marshall, Int. J. Mass Spectrom. Ion Processes **1996**, 158, 5.

[31] M. E. Tuckerman, P. J. Ungar, T. Vonrosenvinge, and M. L. Klein, J. Phys. Chem. 100,

12878 (**1996**).
[32] D. Marx and J. Hutter, Modern methods and algorithms of quantum chemistry, in Ab initio molecular dynamics: Theory and Implementation, edited by J. Grotendorst, volume 1 of NIC Series, page 301, Juelich, **2000**, John von Neumann Institute for Computing.
[33]. D. W. Boo, Z. F. Liu, A. G. Suits, J. S. Tse, and Y. T. Lee, Science **1995**, 269, 57.
[34]. A. L. L. East, Z. F. Liu, C. McCague, K. Cheng, and J. S. Tse, J. Phys. Chem. A **1998**, 102, 10903.
[35] R. Resta, Rev. Mod. Phys. **1994**, 66, 899.
[36] M. S. Lee, F. Baletto, D. G. Kanhere, and S. Scandolo, J. Chem. Phys. **2008**, 128, 214506.
[37] P. L. Silvestrelli, M. Bernasconi, and M. Parrinello, Chem. Phys. Lett. **1997**, 277, 478.
[38]. Schrödinger, E. Annalen der Physik **1926**, 385, 437.
[39]. Born, M.; Oppenheimer, R. Annalen der Physik **1927**, 389, 457.
[40]. Hohenberg, P.; Kohn, W. Physical Review B **1964**, 136, B864.
[41]. Kohn, W.; Sham, L. J. Physical Review **1965**, 140, 1133.
[42]. Dirac, P. A. M. Mathematical Proceedings of the Cambridge Philosophical Society **1930**, 26, 376.
[43]. Vosko, S. H.; Wilk, L.; Nusair, M. Can. J. Phys. **1980**, 58, 1200.
[44]. Perdew, J. P.; Wang, Y. Physical Review B **1992**, 45, 13244.
[45]. Becke, A. D. Physical Review A **1988**, 38, 3098.
[46]. Perdew, J. P.; Burke, K.; Ernzerhof, M. Phys. Rev. Lett. **1996**, 77, 3865.
[47]. Perdew, J. P.; Burke, K.; Ernzerhof, M. Phys. Rev. Lett. **1997**, 78, 1396.
[48]. Lee, C. T.; Yang, W. T.; Parr, R. G. Physical Review B **1988**, 37, 785.
[49]. Becke, A. D. J Chem. Phys. **1993**, 98, 5648.
[50]. Car, R.; Parrinello, M. Phys. Rev. Lett. **1985**, 55, 2471.
[51]. Marx, D. and Hutter, J. AB INITIO MOLECULAR DYNAMICS: BASIS THEORY AND ADVANCED METHODS, Cambridge University Press, **2009**.
[52]. Pulay, P. Mol. Phys. **1969**, 17, 197.
[53]. Lippert, G. Hutter, J. and Parrinello, M. MOLECULAR PHYSICS **1997**, 92, 477
[54]. Vondele, J. V. Krack, M. Mohamed, F. Parrinello, M. Chassaing, T. and Hutter, J. Computer Physics Communications **2005**, 167, 103
[55]. S. Goedecker,M. Teter, J. Hutter, Phys. Rev. B **1996**, 54, 1703.
[56]. C. Hartwigsen, S. Goedecker, J. Hutter, Phys. Rev. B **1998**, 58, 3641.
[57]. http://en.wikipedia.org/wiki/Pseudopotential
[58]. Vanderbilt, D. Physical Review B **1990**, 41, 7892.
[59]. Blochl, P. E. Physical Review B **1994**, 50, 17953.
[60]. Kresse, G.; Joubert, D. Physical Review B **1999**, 59, 1758.
[61]. McQuarrie, D. A., Statistic Mechanics; Harper-Collins Publishers; New York, **1976**.
[62]. M. P. Gaigeot, R. Vuilleumier, M. Sprik, and D. Borgis J. Chem. Theory and Computation **2005**, 1, 772.

Chapter Two
IRMPD Spectroscopy of Cationized Asparagine: Effects of Temperature and Anharmonicities

2.1 Introduction

Asparagine (Asn) is one of the natural amino acids, and its interaction with a metal cation is of fundamental importance to understand its biological function. In the P-type Na^+-K^+-ATPase enzyme, which regulates the volume and potential in animal cells, the biological functionality depends on the high affinity of Na^+ and of K^+ for a few amino-acids, including asparagines, aspartic acid, glutamic acid, and glutamine.[1,2] Together with glutamine, asparagines also play an important role in the deamidation of proteins, with their amide group on the side chain converted to a carboxylate or carboxylic acid under physiological conditions.[3] Such reactions cause time-dependent changes in charge and conformation of proteins, and are hypothesized to serve as a molecular clock that controls aging and the development of certain diseases, such as Alzheimer, Parkinsons, and various form of cancers.[4] The reaction rate is again dependent on the presence of H^+ and metal cations.

The metal complexes of amino-acids provide the basic models for the pairwise interaction between a metal cation and an amino-acid. In the case of asparagine, Armentrout and co-workers have experimentally measured its complexes with alkali metal cations and H^+, using both guided ion beam mass spectrometry[5-9] and infrared multiple photon dissociation spectroscopy[10] (IRMPD) techniques.

With one additional amide group and carbonyl group on its side chain, asparagine binds a metal ion in a tridendately or bidentately, which produces a significant binding energy. The cation affinity at 0 K, as measured by collision-induced dissociation (CID), is 1.6 eV for K^+[5], and 2.2 eV for Na^+.[6] Interestingly, such collisions could also induce deamidation for the complexes,[9] with loss of NH_3, and

in the case of Na^+-asparagine complex, the reaction energy is 1.6 eV.[8] For protonated asparagine, deamidation (with the loss of NH_3 or NH_4^+) and the loss of H_2O+CO are observed as the main dissociation channels, and the threshold dissociation energy is around 1.5 eV.[7] For comparison with these data, the conformers for these complexes and their dissociation channels have been systematically mapped out with density functional theory (DFT) calculations, and found in good agreement with experiment.[5-9]

With the IRMPD technique, the structure and vibrational modes in a complex ion is directly probed. And for asparagines complexes, a systematic study[10] has been reported, with the list of cation including Li^+, Na^+, K^+, Rb^+, and Cs^+, as well as Ba^{2+} and H^+. Gradual changes in both peak positions and relative intensities are observed in the IRMPD spectra as the metal cation changes through the alkali series, from Li^+ to Cs^+, while the spectra for H^+(Asn) and for Ba^{2+}(Asn) are quite distinct both from the alkali metal cation complexes and from each other. Based on DFT calculations, the spectral features have been assigned based on the most stable conformers, and their calculated harmonic spectra.

In this report, we examine another aspect in the IRMPD process, the vibrational profile of the cation-asparagine complexes at finite temperature. Signals measured in the IRMPD are due to the dissociation of the ionic complexes. For example, in the case Na^+(Asn), the lowest energy required for dissociation is 1.6 eV, as measured by CID experiment.[8,9] At 2000 cm^{-1}, it would require the absorption of six photons to push the complex over the dissociation threshold. Its IRMPD spectrum should contain contributions both from the vibrational profiles both at low and elevated temperatures as photons are absorbed. Such effects can not be captured by the harmonic analysis of conformer structures obtained at 0 K. By using *ab initio* molecular dynamics (AIMD) method, we systematically simulate the dynamics at successive temperatures and calculate the corresponding vibrational profiles by the time-correlation function for the dipole moment, which accounts for both the harmonic and anharmonic effects. Our simulations indicate significant variations in the internal dynamics of the complexes, as the cation is varied from Li^+, Na^+, K^+, to

Cs^+, while in terms of backbone flexibility, H^+(Asn) is the most flexible and Ba^{2+}(Asn) the most rigid. Such changes are also important for understanding the features in the IRMPD spectra.

2.2 Methods

Two methods were employed in our study. The optimization of static structure and the calculation of harmonic frequency were performed by the G03[11] package, with the B3LYP exchange-correlation functional and a 6-311+G(d,p) basis set, [12,13] similar to the level of previous calculations reported in reference [10]. For Cs^+, the basis set was LANL2DZ-ECP,[14] with an effective core potential. The calculated harmonic frequencies were scaled by a factor of 0.975.

Ab initio molecular dynamics (AIMD) simulations at finite temperature were performed with Quickstep in CP2K package.[15,16] The total energy and atomic forces were calculated within the framework of density functional theory, using PBE exchange and correlation functional.[17] The electron density was represented by double-ζ Gaussians and auxiliary plane waves with an electronic density cutoff of 400 Ryd., based on the mixed Gaussian and plane wave scheme (GPW),[18] while the wave functions were represented by Gaussians. Atomic core regions were described by Goedecker-Teter-Hutter type pseudopotentials (GTH).[19-21] Corrections for dispersion were also added via Grimme's scheme at D3 level.[22] Our test calculations showed that PBE+D3 provided band positions in better agreement with experimental values than BLYP+D3.[23,24] A cluster ion was put in a cubic box, with a length of 16 Å for H^+(Asn), Li^+(Asn), Na^+(Asn) and 18 Å for the larger K^+(Asn) and Cs^+(Asn), to mimic the gas-phase condition. The decoupling technique developed by Martyna and Tuckerman[25] was employed to eliminate the effect of the periodical charge density images. The convergence criterion for the SCF procedure was set to 10^{-7} a.u. at each time step to guarantee the conservation of total energy. In the dynamic part, an equilibration period of 1.0~2.5 ps was first simulated with the temperature controlled by rescaling. It was then followed by data collection (>15 ps) in the NVE ensemble with a time step of 1.0 fs for the simulations at 100 K and 0.5 fs for the rest.

For each AIMD trajectory, the IR absorption coefficient was calculated by the Fourier transform of the dipole time correlation function (1.53), derived from linear response theory.[26,27]

2.3 Results and Discussion

Conformers for the Asn complexes have been extensively searched and explored in previous studies,[10] which are reproduced and verified with excellent agreement by our structural optimizations. We shall follow the same scheme for naming these conformers, each defined first by the cation binding sites in a bracket, and followed by a description of the ligand orientation based on four dihedral angles starting from the carboxylic acid hydrogen of the backbone (or the NH_2 in zwitterionic conformer), and ending with the terminal side chain nitrogen. Each dihedral angle is labeled as c (cis) for a value between 0^o and 45^o, g (gauche) between 45^o and 135^o, and t (trans) between 135^o and 180^o, respectively.

2.3.1. The dynamics of M^+(Asn) with M=Li, Na, K, and Cs

For Li^+(Asn), Na^+(Asn), and K^+(Asn), the most stable conformer is the [N,CO,CO]tggt structure, as shown in Figure 2.1, with the metal ion coordinated with the N4 of the amide group, and the two carbonyl oxygen atoms, O2 and O3. For Na^+(Asn), the structure is stable at 500 K. The fluctuation for Na–N4 and Na–O3 (side chain carbonyl) distances is within 1 Å, while the fluctuation for Na–O2 distance (the carbonyl on the carboxyl group) is more significant.

When the simulation temperature is raised to 1000 K, structural transformation is observed a few times during the 16 ps period of AIMD simulation. The weak link in Na^+(Asn) is the Na^+–N4 bond, which is periodically extended to values above 4 Å. The two Na^+–O bonds are more stable, although Na^+–O2 bond is also broken for a short while between 1-2 ps. When the Na^+–N4 bond is broken, the structure is transformed into a bidentate [CO,CO] conformer, which is the second most stable, in both previous[10] and our own calculations, with an energy of more than 5 kcal/mol above the tridentate [N,CO,CO] conformer.

The dynamics of Li^+(Asn) is similar to that of Na^+(Asn). At 500 K, the [N,CO,CO] structure is maintained, with the fluctuations in Li^+–N and Li^+–O distances smaller than those for Na^+(Asn). At 1000 K, the tridentate structure is transformed into bidentate structures, [N,CO] or [CO,CO], with the Li^+–O3 (on the side chain carbonyl) bond being maintained during most of the simulation period, similar to Na^+(Asn) (as shown in the appendix).

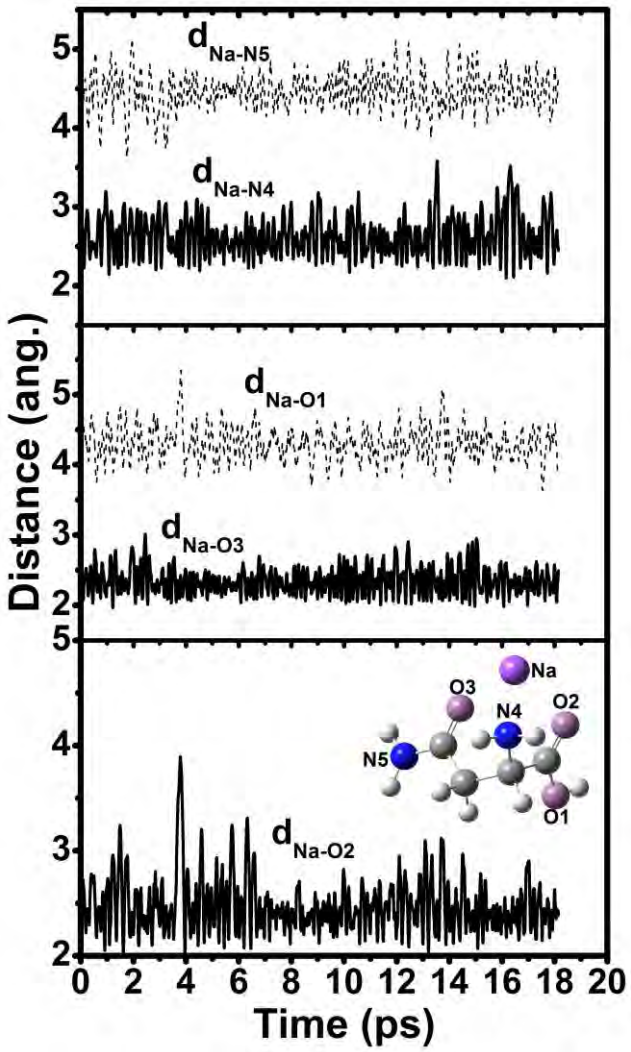

Figure 2.1. Fluctuation of the distance from Na^+ to the N and O atoms in Na^+(Asn) at 500 K. The Na^+ ion is coordinated to N4, O3 and O2, and the structure is maintained during the 16 ps simulation, despite fluctuations in bond distances.

For K^+(Asn) and Cs^+(Asn), the bond distance fluctuation indicates that the metal complex becomes more flexible in structure. In both cases, extensive structural transformation is already observed at 500 K. For Cs^+(Asn), the breaking of Cs^+–O2

bond and big fluctuations in the Cs^+−N4 distance are observed even at 300 K, as shown in Figure 2.3.

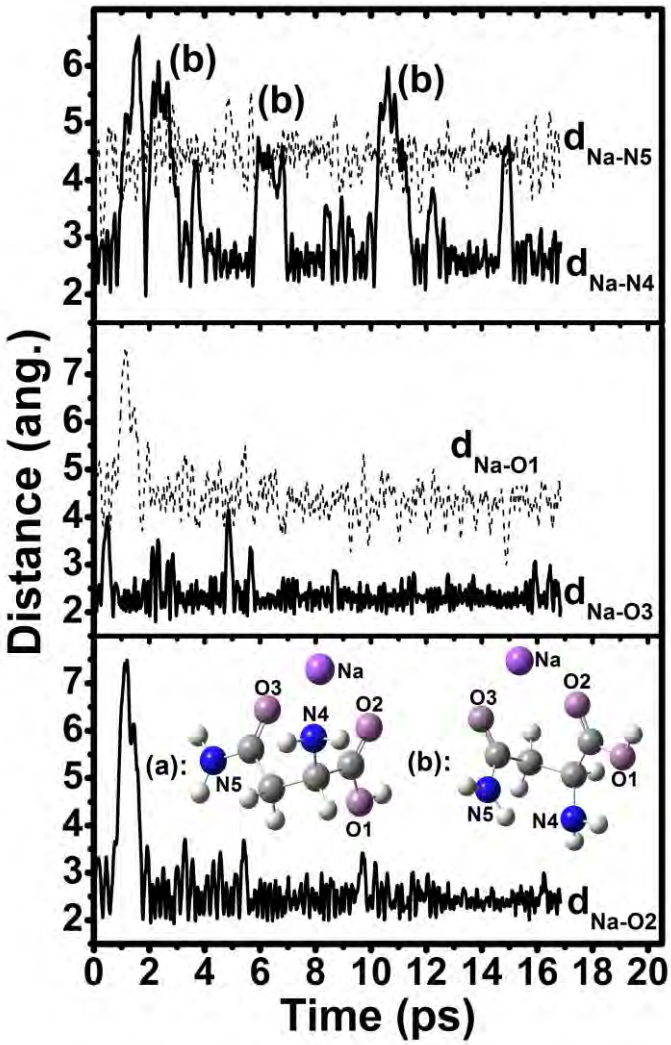

Figure 2.2. Significant fluctuation for the distances from Na^+ to the N and O atoms in Na^+(Asn) at 1000 K. The [N,CO,CO]tggt structure shown in (a) is periodically transformed into the [CO,CO] structure shown in (b).

These comparisons show a clear trend: the metal complex becomes progressively more flexible as the metal cation changes from Li^+, to Na^+, K^+, and finally to Cs^+. This is obviously due to the relative strength of the binding between the metal cation and Asn. We have also calculated the barrier for flipping the carboxyl group, which brings the OH group to metal ion by replacing the C=O group. As expected, its value increases from 5.3 kcal/mol for Cs^+(Asn), to 8.4 kcal/mol for K^+(Asn), 9.9 kcal/mol for Na^+(Asn), and 10.7 kcal/mol for Li^+(Asn). It is also interesting to notice that

despite all these transformations, the bond between the metal cation and the carbonyl oxygen (O3) on the side chain is maintained in these simulations, indicating that its coordination interaction with the metal cation is the strongest.

It should be further pointed out that for Cs^+(Asn) the most stable conformer is no longer the tridentate [N,CO,CO], but the bidentate [COOH,CO], although the energy gap between the two optimized conformers is less than 1 kcal/mol. Such a bidentate [COOH,CO] structure is also optimized for K^+(Asn), but in this case it is higher in energy than the tridentate [N,CO,CO] by ~2 kcal/mol. At 300 K, K^+ in the [COOH,CO] K^+(Asn) is still solvated by the two carbonyl oxygen atoms, O2 and O3, although the hydroxyl oxygen O1 does approach K^+ periodically. For Cs^+(Asn), the Cs^+−O3 bond is broken during more than half of the simulation period, and Cs^+ is solvated by the two oxygen atoms on the carboxyl group, O2 and O1. Within such a structure, the acidic hydrogen H17, originally bounded to O1, jumps to the amid group, producing a zwitterion. In other words, there is a dynamic presence of the zwitterion at room temperature in Cs^+(Asn).

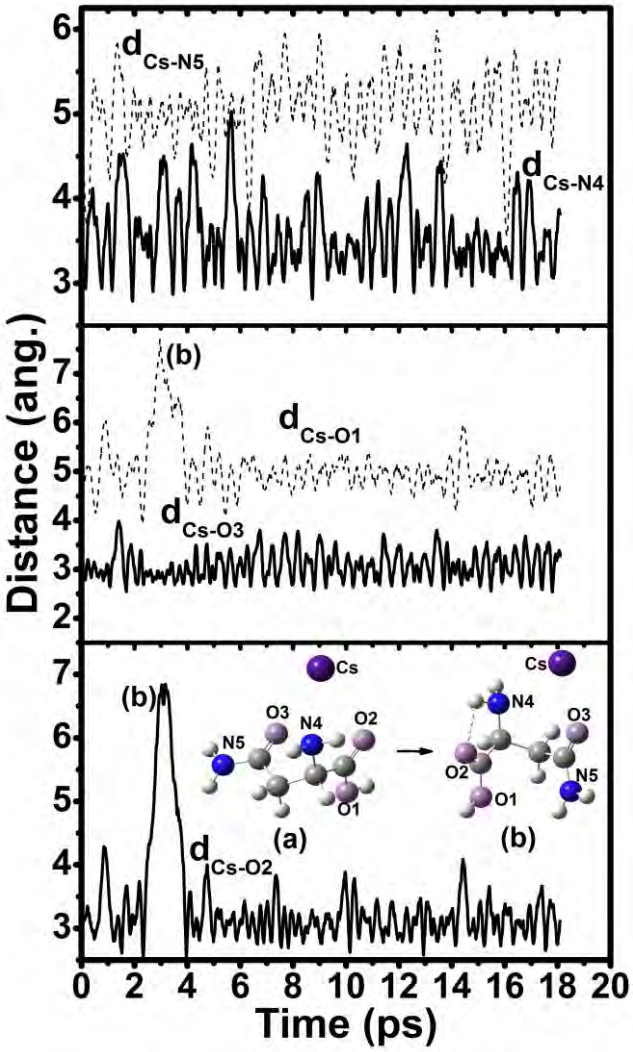

Figure 2.3. Fluctuation of the distance from Cs^+ to the N and O atoms in for the [N,CO,CO]tggt conformer of Cs^+(Asn) at 300 K.

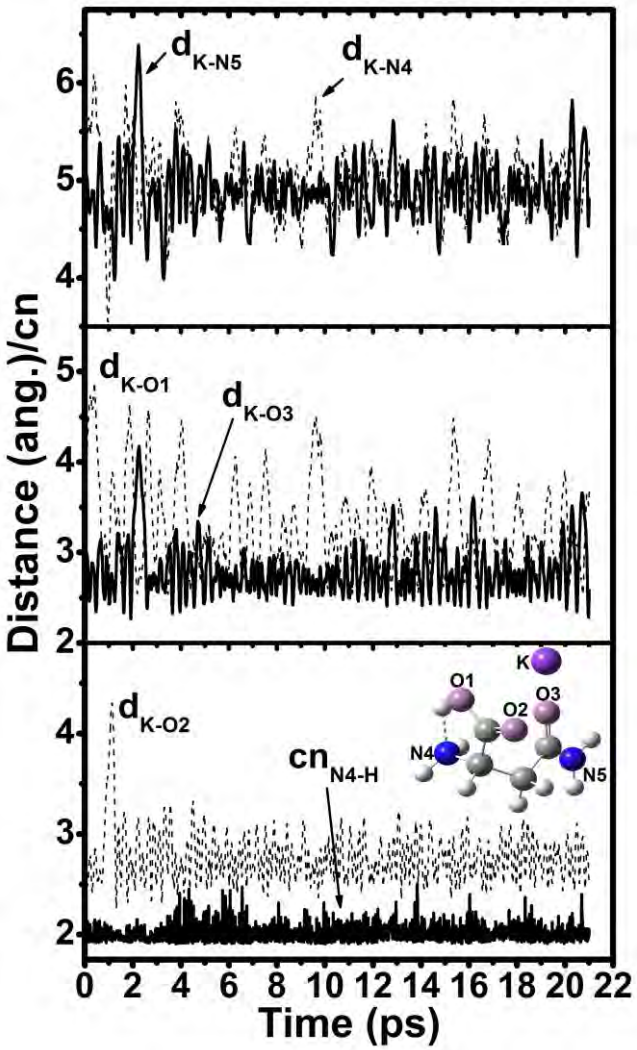

Figure 2.4. Fluctuation of the distance from K^+ to the N and O atoms in the [COOH,CO]ctgt conformer of K^+(Asn) at 300 K. The bottom panel also shows the time evolution of the hydrogen coordination number around N4.

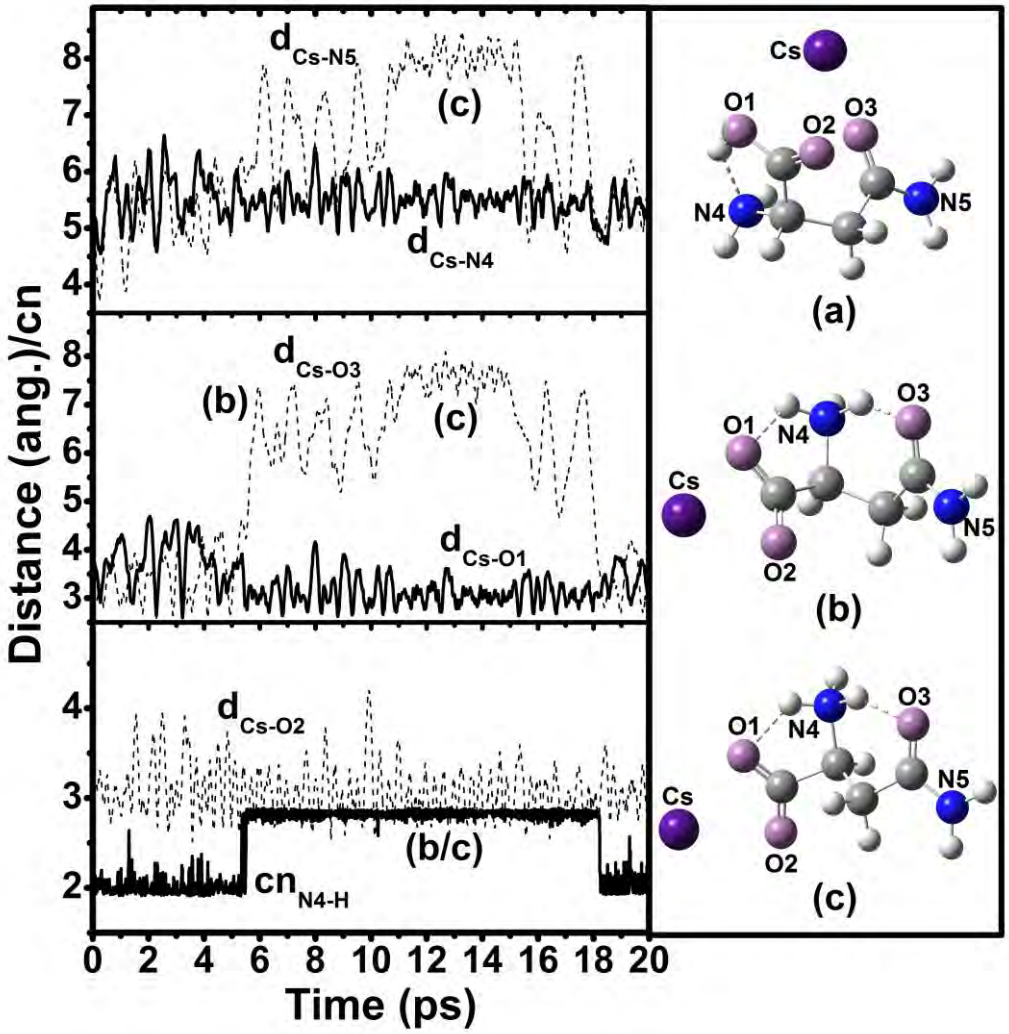

Figure 2.5. Fluctuation of the distance from Cs^+ to the N and O atoms in the [COOH,CO]ctgt conformer of Cs^+(Asn) at 300 K. The bottom panel also shows the time evolution of the hydrogen coordination number around N4.

2.3.2. The dynamics of H^+(Asn)

H^+(Asn) differs from the metal complex M^+(Asn) in structure. In M^+(Asn), the M^+ cation is stabilized by interactions with the electron negative N and O atoms by bidentate or tridentate binding. In H^+(Asn), the amide group is protonated in the most stable structure, which is then further solvated by oxygen atoms in the carboxyl group and in the side chain carbonyl group. The dominant interaction in H^+(Asn) is hydrogen bonding, which could again lead to a number of possible configurations, with varying dihedral angles. These conformers have been carefully mapped out before,[10] and for the sake of discussion, they are shown again in Figure 2.6. As the

conformation of H^+(Asn) is hold up by internal hydrogen bonds, it is more flexible than the metal complexes discussed above.

The two most stable conformers for H^+(Asn) are [N,CO]tgtt and [N,CO]tcgt, as shown in Figure 2.6, with an energy difference of only 0.4 kcal/mol in favor of the former. The most notable structural difference between these two conformers is in the orientation of the two H atoms, H11 and H12, in the CH_2 group. However, their dynamics differs considerably from each other, as shown in Figure 2.7 and 2.8 respectively.

While [N,CO]tgtt (Figure 2.6a) is the most stable conformer at 0 K, it is dynamically less robust than [N,CO]tcgt (Figure 2.6b). Even at 200 K (Figure 2.7), there are considerable fluctuations in the hydrogen bond between the protonated amide group and C=O on the carboxyl group (O1). It is actually broken during part of the simulation, when the carboxyl group rotate along the C6-C8 axis and a hydrogen bond is formed between the hydroxyl group (O3) and the protonated amide, leading to the formation of [N,CO]tttt (Figure 2.6d). At 300 K, the fluctuation is more prominent and the structure is transformed to another conformer, [N,CO]ttgt (Figure 2.6c) after 6 ps, staying there for the rest of the simulation. At 500 K, the structure is transformed into [N,CO]tcgt (Figure 2.6b).

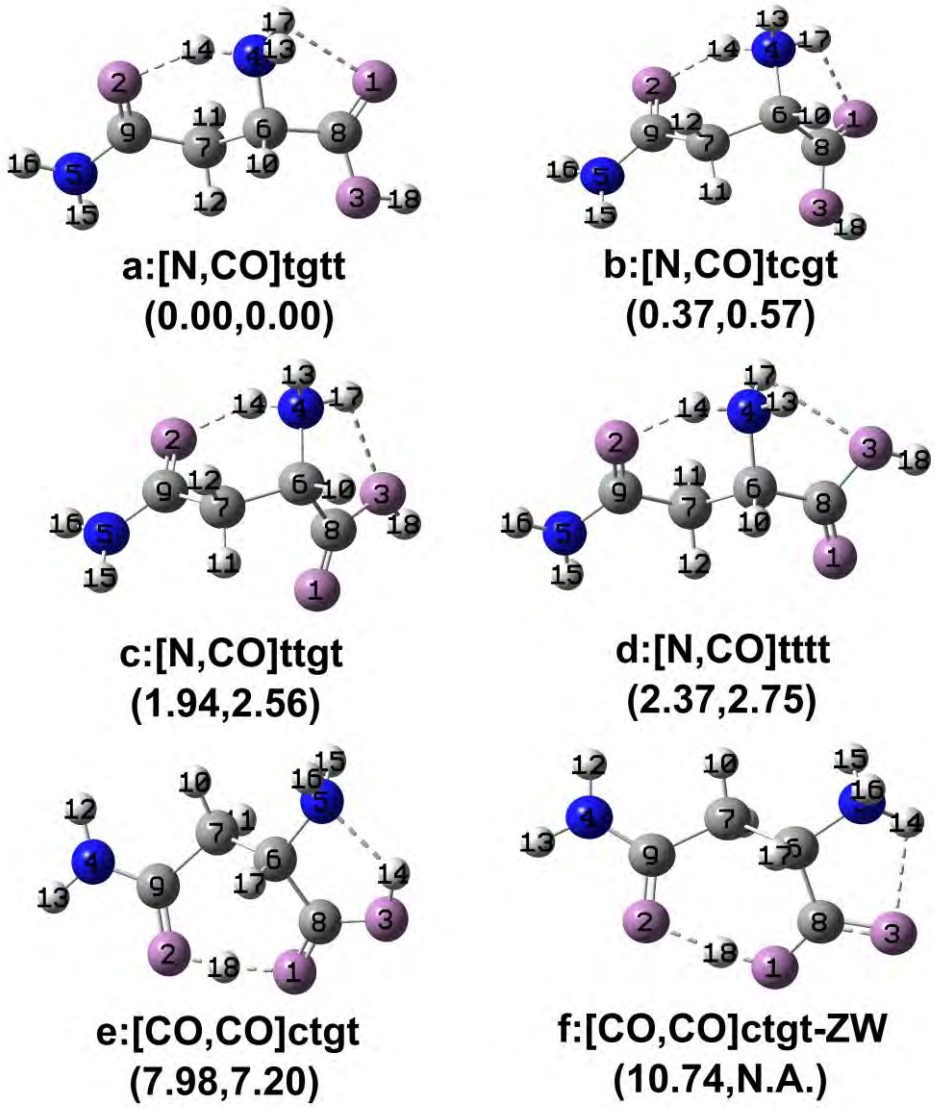

Figure 2.6. The structure and relative energy of H^+(Asn). The two numbers in parenthesis are the relative energy in kcal/mol as calculated in our study (first number) and as reported in reference [10].

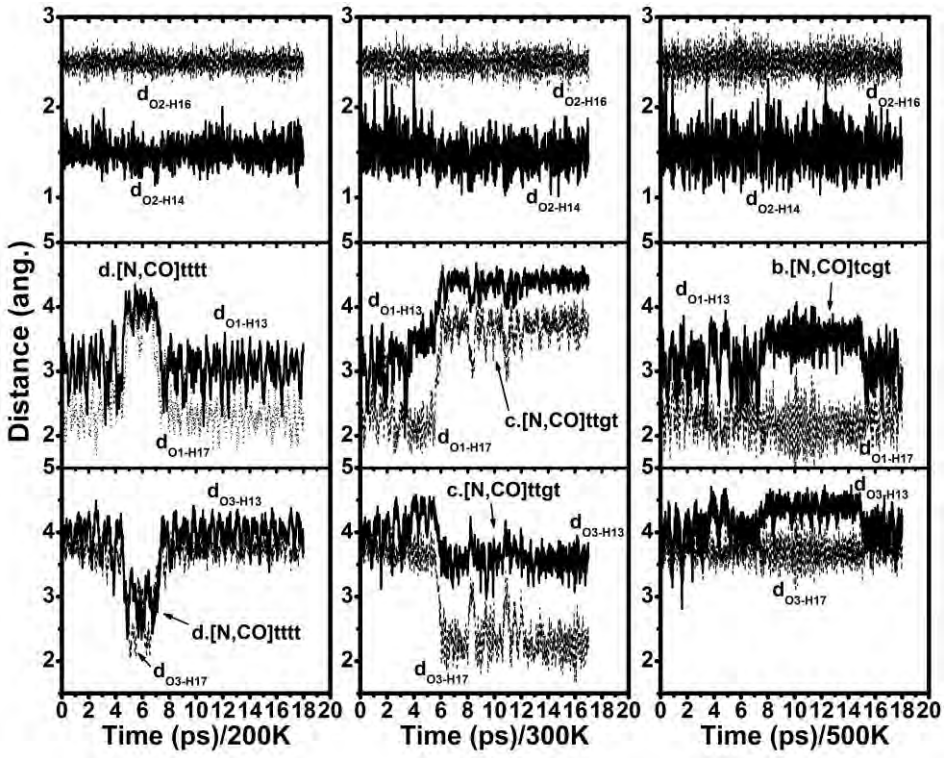

Figure 2.7. Fluctuation of atomic distances for the [N,CO]tgtt conformer of H^+(Asn), at 200 K, 300 K and 500 K. For the labeling of atoms, see Figure 2.6a.

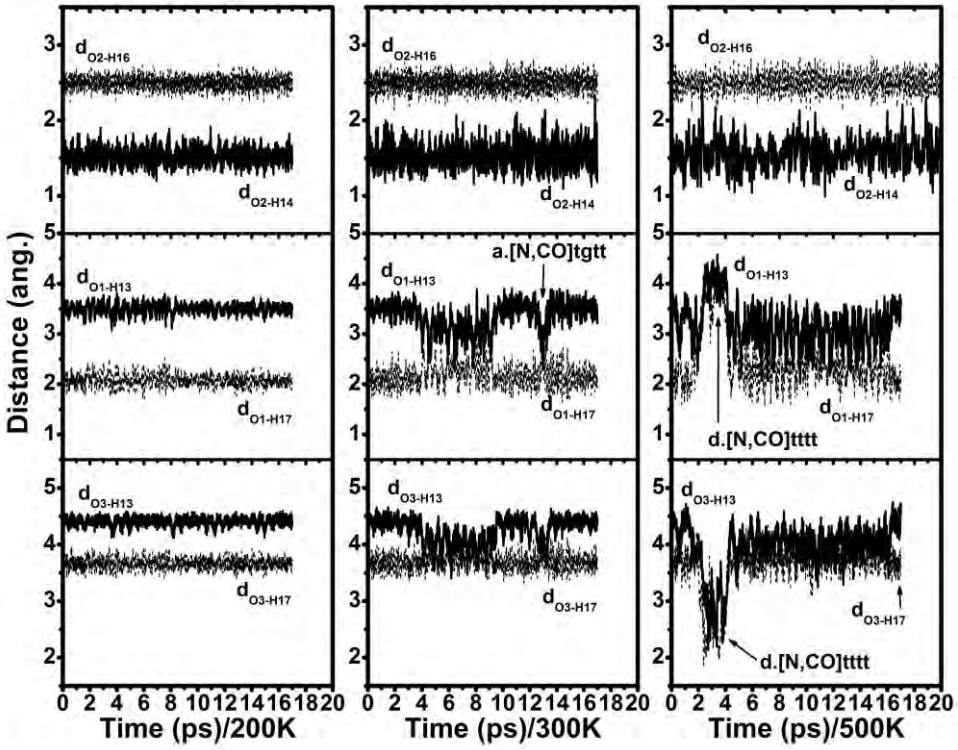

Figure 2.8. Fluctuation of atomic distances for the [N,CO]tcgt conformer of H^+(Asn), at 200 K, 300 K and 500 K. For the labeling of atoms, see Figure 2.6b.

In contrast, [N,CO]tcgt (Figure 2.6b) is maintained all the way through the simulation at 200 K (Figure 2.8), and the fluctuation in bond distances is considerably less than that for [N,CO]ttgt. The structure is also almost maintained at 300 K, although for a short while, it becomes [N,CO]ttgt. Even at 500 K, the structure is basically maintained. But the fluctuation in distances does increase significantly and from 2ps to 4ps during the simulation, the structure is transformed to [N,CO]tttt.

These results show that inter-conversion among the various conformers of H^+(Asn) is easier than that in the metal complexes M^+(Asn). For the [N,CO] conformers, it is the hydrogen bond between the protonated amide and the side chain carbonyl (O2) that is maintained throughout the simulations, while the hydrogen bond between the amide group and the carboxyl group is more flexible. The barriers separating H^+(Asn) conformers can be overcome at 300 K and even 200 K. Considering the fact that the atomic motion is treated by Newtonian mechanics in our simulations, which does not take into consideration the zero point vibration, H^+(Asn) could be a mixture of various conformers even at low temperature.

For H^+(Asn), there are also [COOH,CO] conformers, with the proton solvated by two carbonyl oxygen atoms (O1 and O2 in Figure 2.6e-h), similar to the [COOH,CO] conformer for the metal complexes discussed above. Their energies are at least 8 kcal/mol higher than the [N,CO] conformers. AIMD simulations for these conformers show that these conformers can also converting to each other upon heating. However, the proton stays between O1 and O2 through out the simulation, and basic [COOH,CO] structure is maintained (see appendix). It is also interesting to notice that the hydrogen on the hydroxyl group actually jumps back and forth between the hydroxyl and the amide groups, to form zwitterion during part of the simulation, especially at 500 K. However, these structures are not expected to be abundant due to their higher energies, relative to the [N,CO] conformers.

2.3.3. Vibration of M^+(Asn) and H^+(Asn) complexes at finite temperature

The reported experimental IRMPD spectra for the metal complex M^+(Asn) were measured between 600-1800 cm^{-1}, and as the metal ion changed from Li^+, to Na^+, K^+, and Cs^+, their spectra were broadly similar to each other and most features can be satisfactorily assigned based on harmonic analysis.[10] With AIMD trajectories, the temperature effects on the IRMPD spectra could be simulated by the Fourier transformation of the dipole time correlation function, as shown in Figure 2.9 and 2.10, which lend further supports to the previous assignment.

For Li^+(Asn) as shown in Figure 2.9, the 100 K spectrum, calculated by the dipole time-correlation function method, is basically the same as the harmonic spectrum based on optimized [N,CO,CO]tggt structure. And both spectra are in general agreement with the experimental IRMPD spectrum. As temperature rises, the sharp peaks become broadened. This is especially true in the low frequency region around 600 cm^{-1}, where the OH and NH wagging peaks merge into one broad feature, as observed experiment. For peaks above 1000 cm^{-1}, the temperature must be raised to 1000 K to observe significant broadening.

A word of caution is needed here regarding the temperature. Experimentally, the temperature of the measured ionic complexes was not reported.[10] In our simulations, the atomic motions are treated by Newtonian mechanics, with no consideration of the zero-point vibrations. A simulation temperature of 100 K could actually be cooler than what is achievable in experiment at 0 K. Furthermore, IRMPD signals are produced when ions are brought to the dissociation threshold by absorbing several photons successively. As estimated in the introduction, around than 6 photons at 2000 cm^{-1} would be needed for these ionic Asn complexes, and the resulting IRMPD spectrum is the overlap of vibrational profiles at several temperatures. As such, an observed feature is produced when photons of a certain frequency are absorbed in all these profiles. In other words, the basic shape in the experimental spectrum should already be present in the vibrational profile for the absorption of the first photon. By comparison between the experimental and simulated spectra, this first profile should correspond approximately to the simulated spectrum at 500 K. It is in the 500 K spectrum, when the two NH_2 bending peaks around 1600 cm^{-1}, one for the backbone

NH_2 and the other for the side-chain NH_2, are merged into one peak (**D**), and the low frequency features below 1000 cm^{-1} are smeared into weak and broad features (especially **A**), while at the same time, peaks **B**, **C**, **E**, and **F** are still well preserved. Similar analysis is also applicable to the spectra for Cs^+(Asn) shown in Figure 2.10 (and the spectra for Na^+(Asn) and K^+(Asn) shown in appendix).

For Cs^+(Asn), the [COOH,CO] is the most stable conformer and should make an important contribution to the IRMPD spectrum. Its presence has been identified previously as the reason for the appearance of **C'** and the increase in the height of peak **C** (relative to peak **B**),[10] both of which are due to a combination of several backbone modes, such as CH_2 bending (C7-H11-H12), C-C-C stretching (C7-C6-C8) and N-C-H bending (N4-C6-H10). The significant intensities of **C** and **C'** and the absence of **B** and **B'** in the AIMD simulated spectra for the [COOH,CO] conformers at 300 and 500 K support this assignment, as these vibrational profiles not only enhance band **C** and **C'**, but also suppress band **B** and **B'**. However, there is a significant presence of **B** and **B'** which indicates the presence of the [N,CO,CO] conformer. In any case, the conversion between [COOH,CO]/[CO,CO] and tridentate [N,CO,CO] conformers becomes easier when the metal ion is Cs^+, as discussed above.

The other trend in the IRMPD spectra of Asn complexes is the gradual emergence of a small band **B'** around 1100 cm^{-1} as the metal ion changes from Li^+, to Na^+, K^+, and Cs^+. Based on harmonic analysis, band **B** is due to the bending of the C-OH group, and band **B'** is due to the C-N stretching on the backbone (C6-N4), although backbone modes such as CH_2 bending and C-C-C stretching are mixed into both bands. The presence of **B'** is predicted for the tridentate [N,CO,CO] conformer for all four metal complexes by harmonic analysis, and its position shifts to the blue, from 1097 cm^{-1} for Li^+(Asn), to 1113 cm^{-1} for Na^+(Asn), to 1115 cm^{-1} for K^+(Asn), and to 1118 cm^{-1} for Cs^+(Asn), together with a slight increase in intensity. The blue shift is due to a decrease in the interaction between N and the metal ion, which strengthens the C-N bond. In the AIMD simulated spectra, **B'** is clearly resolved at 300 K and 500 K for both K^+(Asn) and Cs^+(Asn) (Figure 2.10 and Supporting Information), but smeared out at 300 K for Li^+(Asn) (Figure 2.9), in good agreement with the

experimental observations. In the case of M^+(Asn), a more flexible backbone seems to enhance the intensity of **B'** (C-N stretching) and makes it more prominent even at raised temperatures. A similar trend is observed for a band around 1100 cm^{-1} in the cases of Threonine,[28] Serine,[29] Histidine,[30] and Cysteine[31] complexes. It would be interesting to further investigate whether such a peak is correlated with a more flexible backbone in those cases.

The experimental IRMPD spectrum for H^+(Asn) differs from the spectra of the metal complexes in that the observed bands are further broadened. Previously, the H^+(Asn) spectrum was interpreted as containing the contribution from the two most stable conformers, [N,CO]tgtt and [N,CO]tcgt. Their harmonic spectra, shown in Figure 2.11 and 2.12, are actually quite similar to each other, and a simple mixture of the two should not produce the very broadened features such as **B** and **C**. Based on the AIMD simulations discussed above, we know that H^+(Asn) is more flexible and undergoes considerable structural changes. The simulated spectra show considerable broadening even at 200 K. Such internal dynamics is the reason for the broad features observed experiment. While a hint of **B'** is still observable, **C'** is merged into **C**, which produces raised baseline from 1000 cm^{-1} all the way up to 1800 cm^{-1}. For the cationic Asn complexes study in this report, H^+(Asn) provides the most dramatic example of how the internal motion and dynamics could affect the experimentally observed vibrational spectrum.

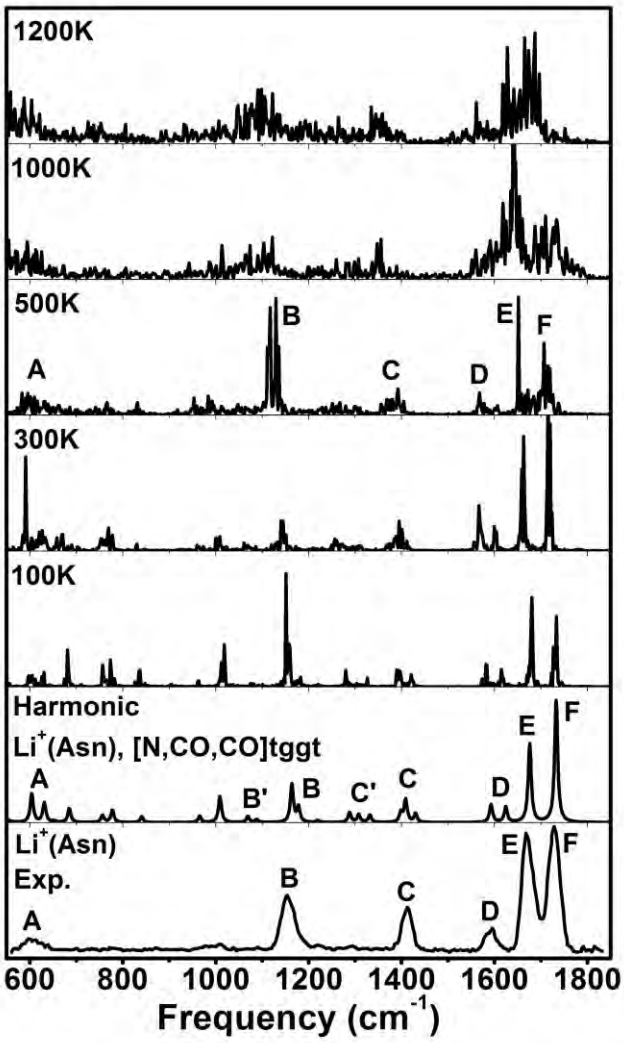

Figure 2.9. Vibrational profiles for Li^+(Asn),[N,CO,CO]tggt at finite temperatures obtained by AIMD simulations, together with the harmonic and experimental spectrum (from Ref.[10]).

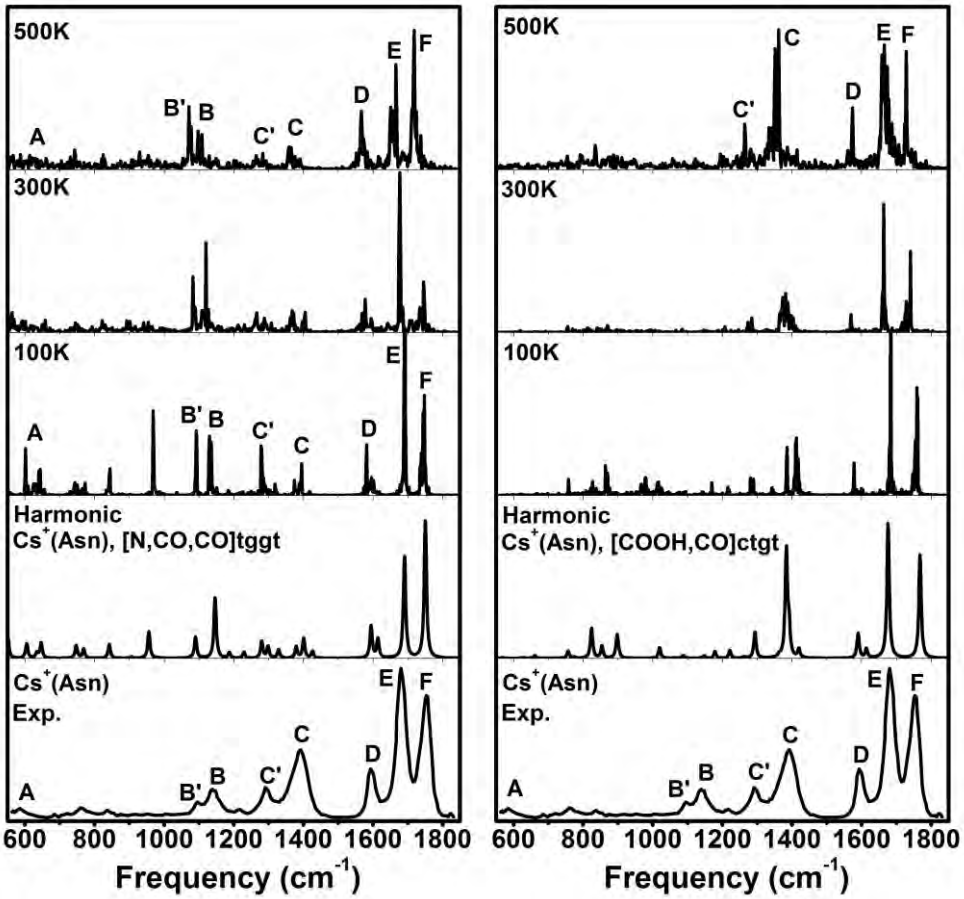

Figure 2.10. Vibrational profiles for Cs^+(Asn) at finite temperatures obtained by AIMD simulations. Left: [N,CO,CO]tggt conformer; Right: [COOH,CO]ctgt conformer. Both in comparison with the harmonic and experimental spectrum (from Ref.[10]).

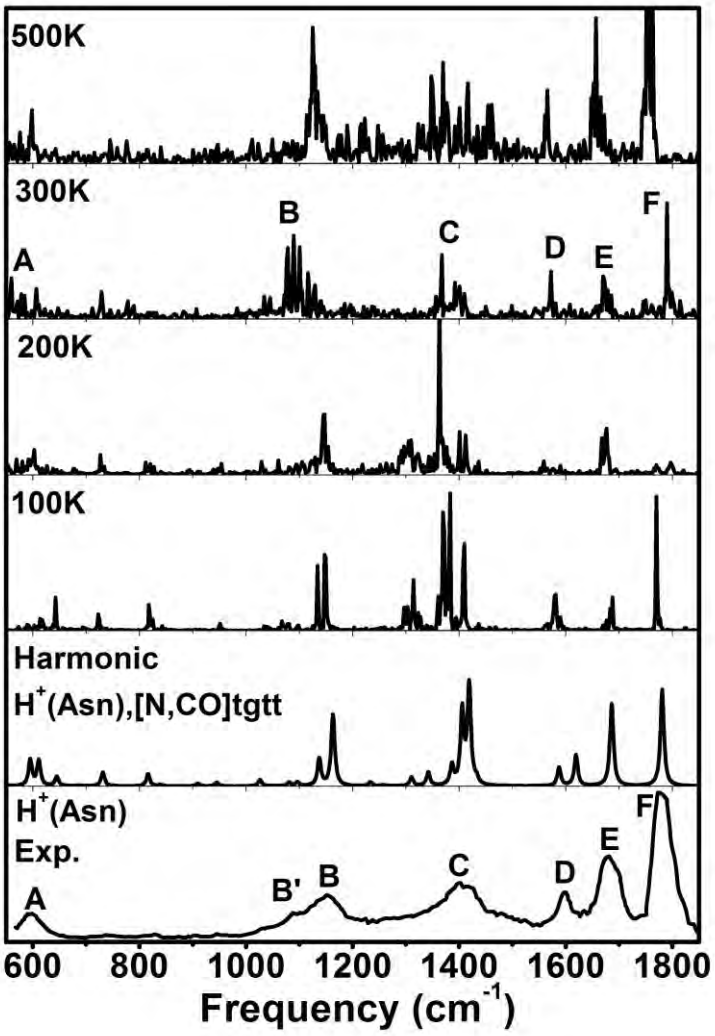

Figure 2.11. Vibrational profiles for H^+(Asn),[N,CO]tgtt conformer, at finite temperatures, obtained by AIMD simulations, together with the harmonic and experimental spectrum (from Ref.[10]).

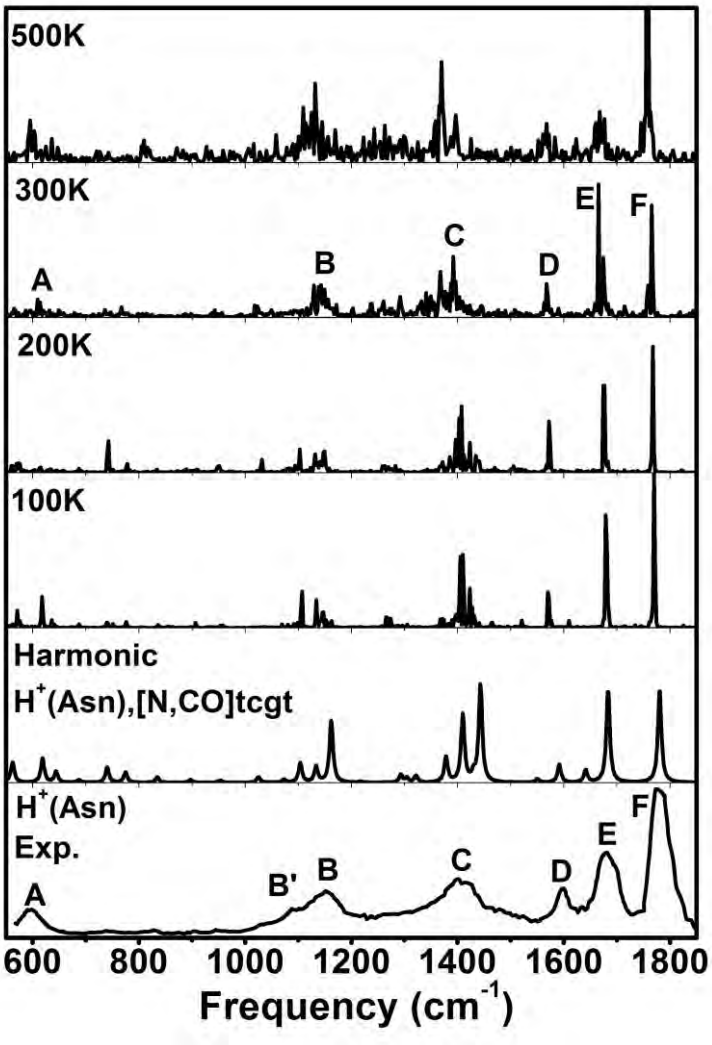

Figure 2.12. Vibrational profiles for H^+(Asn),[N,CO]tcgt conformer, at finite temperatures, obtained by AIMD simulations, together with the harmonic and experimental spectrum (from Ref.[10]).

2.4 Conclusions

AIMD simulations indicate that the structures of M^+(Asn) complexes are flexible, and there could be extensive conversions among various conformers at finite temperatures, by torsional rotation and stretching of N-M^+ and O-M^+ distances, which change the dihedral angles and the interactions between the cation and the N and O atoms in Asn. The flexibility depends on the cation. While Li^+ and Na^+ complexes are relatively rigid and go through conversions at elevated temperature, K^+ and Cs^+ complexes are more flexible, even at room temperature. The protonated H^+(Asn) is the most flexible, and conversion between conformers is observed even at 200 K.

The vibrational profiles obtained from the Fourier transformed dipole time correlation function lend support to previous assignment of the IRMPD spectra. The structural conversion implies that an IRMPD spectrum should contain contributions from various isomers. But since the vibrational profiles for these conformers are similar to each other, the overall effect is to broaden the observe features and to shift their relative intensities, depending on the structural flexibility of the metal complex. The H^+(Asn) spectrum has the broadest bands, since it is the most flexible. The Cs^+(Asn) and K^+(Asn) spectra contain more contribution from the bidentate [CO, CO] conformer, and a feature around 1100 cm^{-1} (C-N stretching) that may be related to their structural flexibility. But even for the more rigid Na^+(Asn) and Li^+(Asn), significant broadening is observed due to contributions from vibrational profiles at high temperature.

2.5 Reference

[1] Ogawa, H.; Toyoshima, C. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 15977.

[2) Jorgensen, P. L.; Hakansson, K. O.; Karlish, S. J. D. *Annu. Rev. Physiol.* **2003**, *65*, 817.

[3] Robinson, A. B.; Rudd, C. *Current topics in cellular regulation* **1974**, *8*, 247.

[4] Robinson, A. B.; Robinson, N. E. *Molecular clocks: deamidation of asparaginyl and glutaminyl residues in peptides and proteins*; Althouse: Cave Junction, 2004.

[5] Heaton, A. L.; Armentrout, P. B. *Journal of Physical Chemistry B* **2008**, *112*, 12056.

[6] Heaton, A. L.; Moision, R. M.; Armentrout, P. B. *J. Phys. Chem. A* **2008**, *112*, 3319.

[7] Heaton, A. L.; Armentrout, P. B. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 852.

[8] Heaton, A. L.; Armentrout, P. B. *Journal of the American Chemical Society* **2008**, *130*, 10227.

[9] Heaton, A. L.; Ye, S. J.; Armentrout, P. B. *J. Phys. Chem. A* **2008**, *112*, 3328.

[10] Heaton, A. L.; Bowman, V. N.; Oomens, J.; Steill, J. D.; Armentrout, P. B. *J. Phys. Chem. A* **2009**, *113*, 5519.

[11] McLean, A. D.; Chandler, G. S. *J. Chem. Phys.* **1980**, *72*, 5639.

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

[13] Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai,

H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; 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. *Gaussian03*, revision D.01; Gaussian, Inc.: Pittsburgh, PA, **2005**.
[14] Hay, P. J.; Wadt, W. R. *J. Chem. Phys.* **1985**, *82*, 299.
[15] VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. *Comput. Phys. Comm.* **2005**, *167*, 103.
[16] *The CP2K developers group, http://cp2k.berlios.de/.*
[17] Lippert, G.; Hutter, J.; Parrinello, M. *Theor. Chem. Acc.* **1999**, *103*, 124.
[18] Krack, M. *Theor. Chem. Acc.* **2005**, *114*, 145.
[19] Goedecker, S.; Teter, M.; Hutter, J. *Phys. Rev. B.* **1996**, *54*, 1703.
[20] Hartwigsen, C.; Goedecker, S.; Hutter, J. *Phys. Rev. B.* **1998**, *58*, 3641.
[21] Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865.
[22] Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010**, *132*, 154104.
[23]Becke, A. *Phys. Rev. A.* **1988**, *38*, 3098.
[24] Lee, C.; Yang, W.; Parr, R. *Phys. Rev. B.* **1988**, *37*, 785.
[25] Martyna, G.; Tuckerman, M. *J. Chem. Phys.* **1999**, *110*, 2810.
[26] McQuarrie, D. A., Statistic Mechanics; Harper-Collins Publishers; New York. **1976**.
[27] Gaigeot, M. P.; Vuilleumier, R.; Sprik, M.; Borgis, D. *J. Chem. Theory Comput.* **2005**, *1*, 772.
[28] Rodgers, M. T.; Armentrout, P. B.; Oomens, J.; Steill, J. D. *J. Phys. Chem. A* **2008**, *112*, 2258.
[29] Armentrout, P. B.; Rodgers, M. T.; Oomens, J.; Steill, J. D. *J. Phys. Chem. A* **2008**, *112*, 2248.
[30] Citir, M.; Hinton, C. S.; Oomens, J.; Steill, J. D.; Armentrout, P. B. *J. Phys. Chem. A* **2012**, *116*, 1532.
[31] Citir, M.; Stennett, E. M. S.; Oomens, J.; Steill, J. D.; Rodgers, M. T.; Armentrout, P. B. *Int. J. Mass Spectrom.* **2010**, *297*, 9.

2.6 Appendix

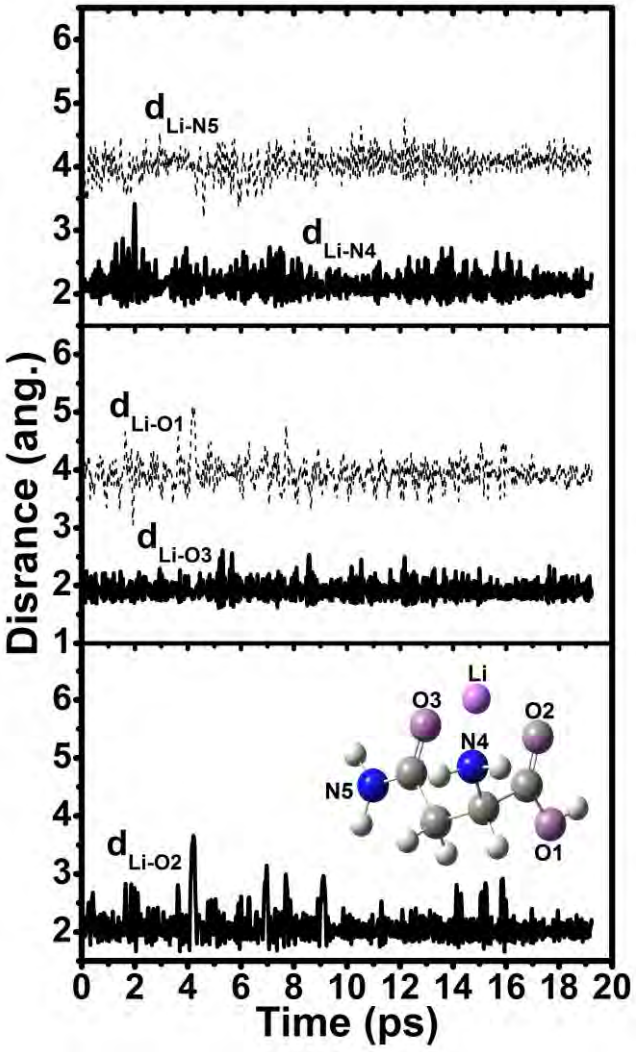

Figure 2.S1. Fluctuation of the distance from Li^+ to the N and O atoms in the [N,CO,CO]tggt conformer of Li^+(Asn) at 500 K.

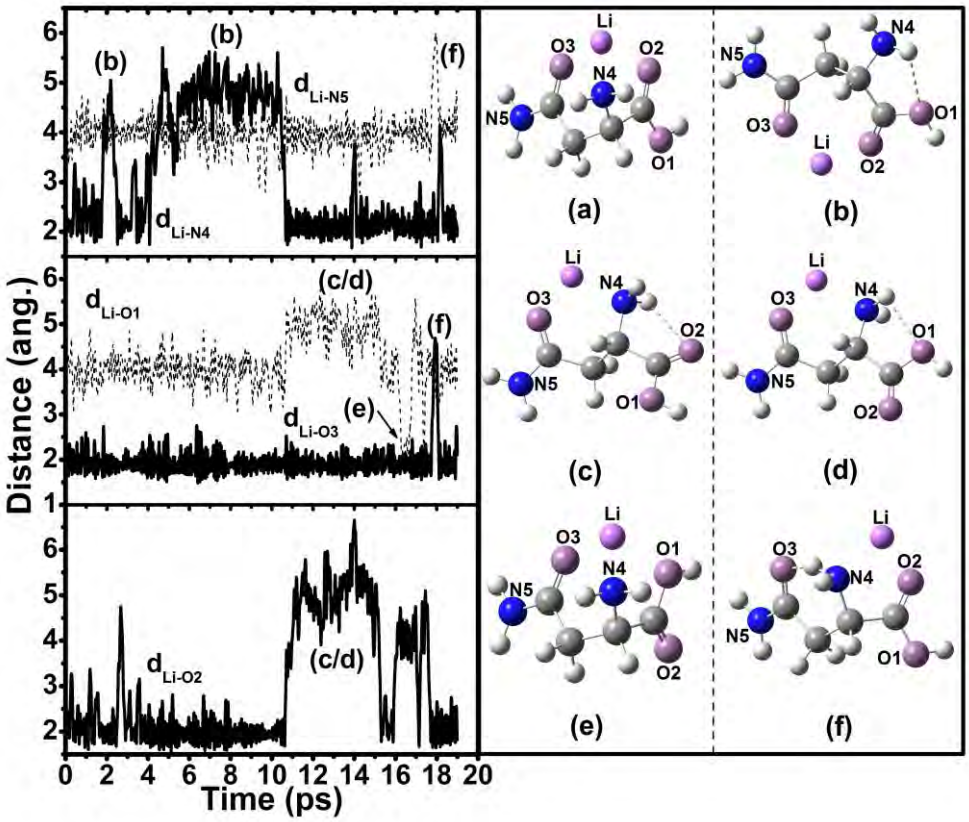

Figure 2.S2. Fluctuation of the distance from Li^+ to the N and O atoms in the [N,CO,CO]tggt conformer of Li^+(Asn) at 1000 K.

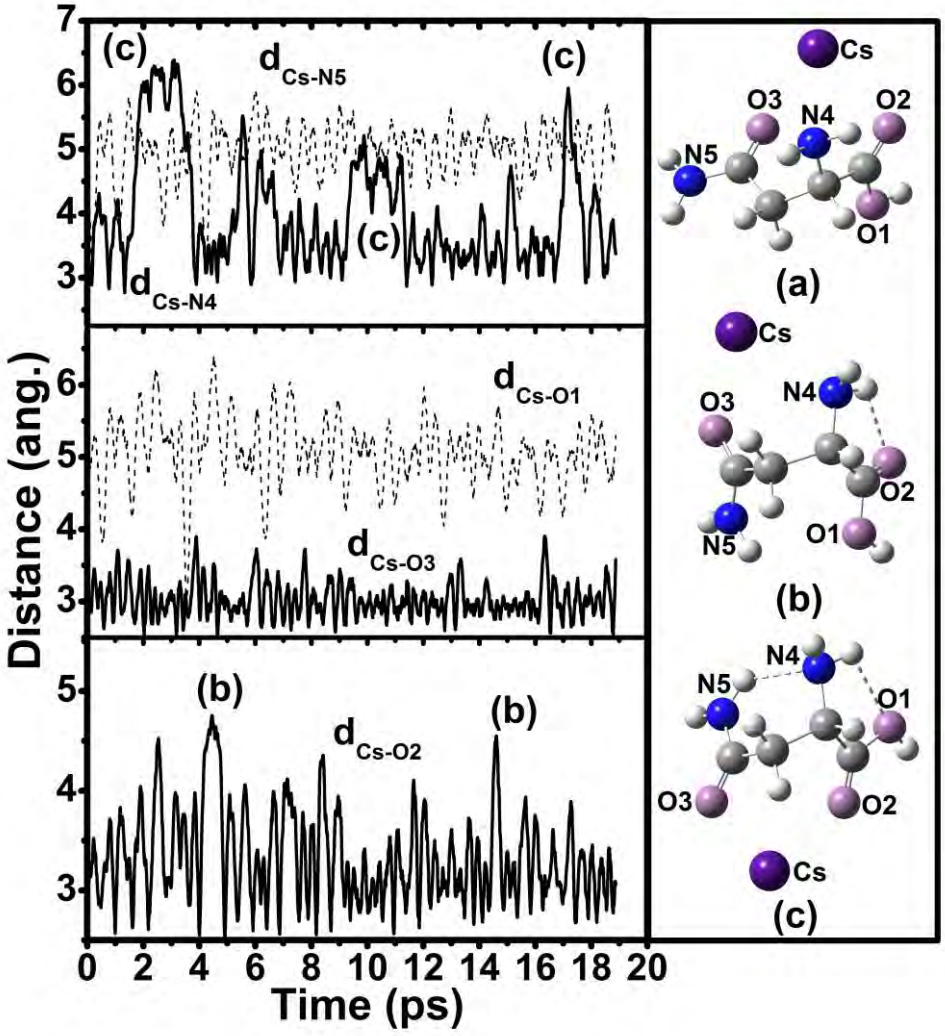

Figure 2.S3. Fluctuation of the distance from Cs^+ to the N and O atoms in the [N,CO,CO]tggt conformer of Cs^+(Asn) at 500 K.

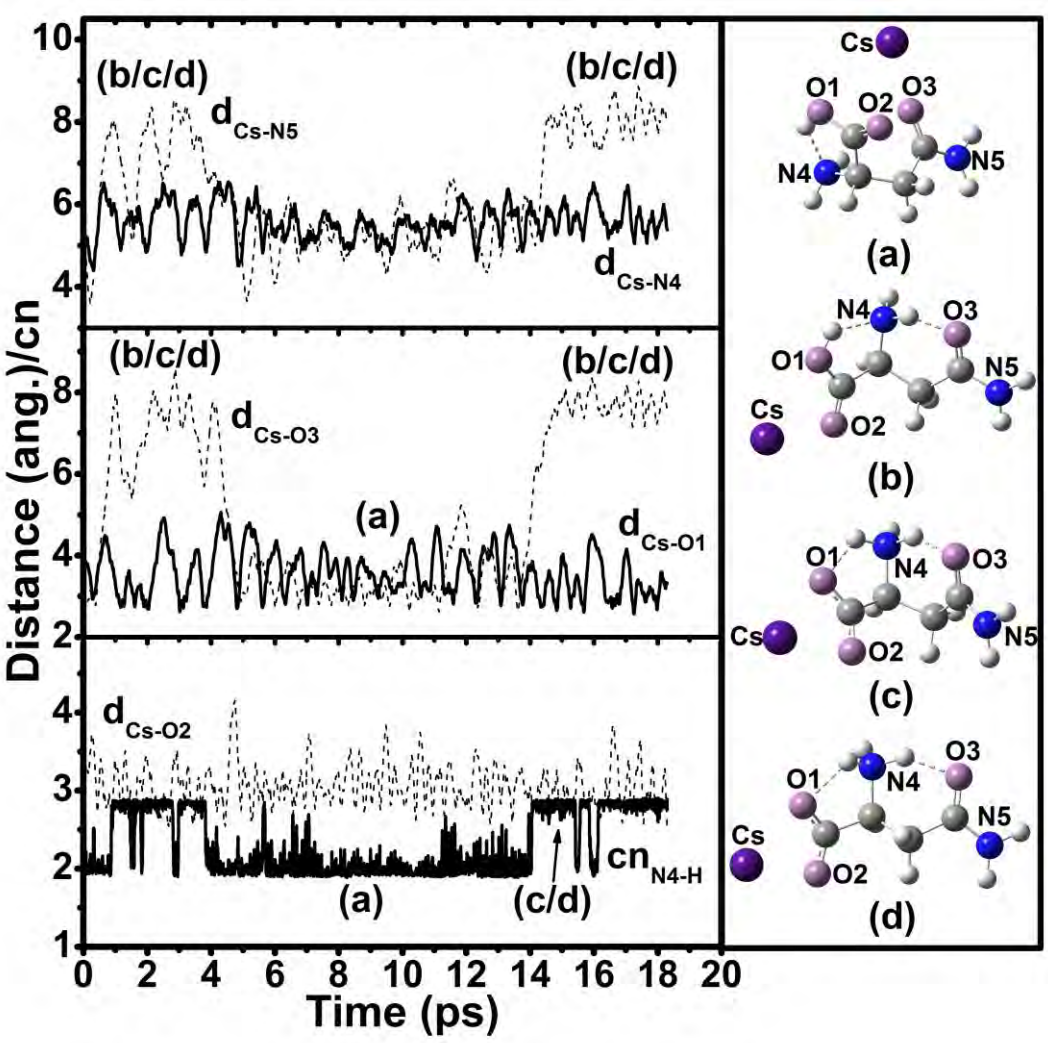

Figure 2.S4. Fluctuation of the distance from Cs^+ to the N and O atoms in the [COOH,CO]tggt conformer of Cs^+(Asn) at 500 K.

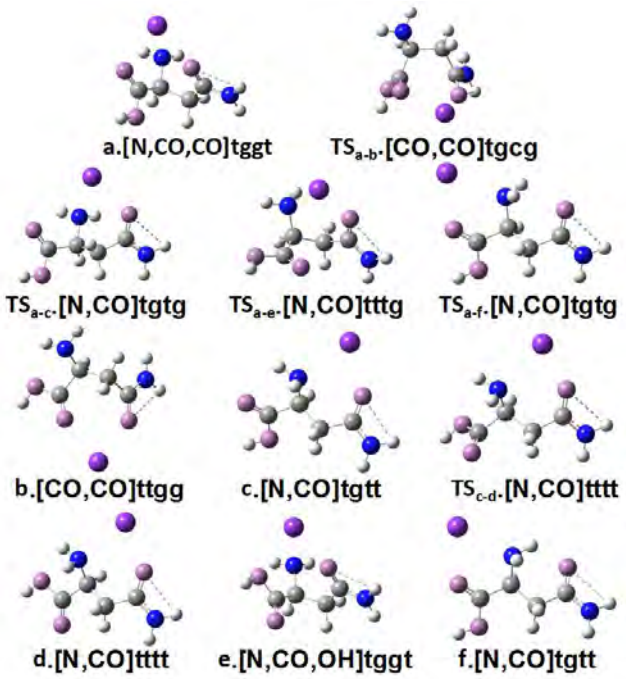

Figure 2.S5. Optimized conformers and transition states for M^+,[N,CO,CO] (M=Li, Na, K, and Cs) and the corresponding relative energy levels are shown in Figure 3.S6.

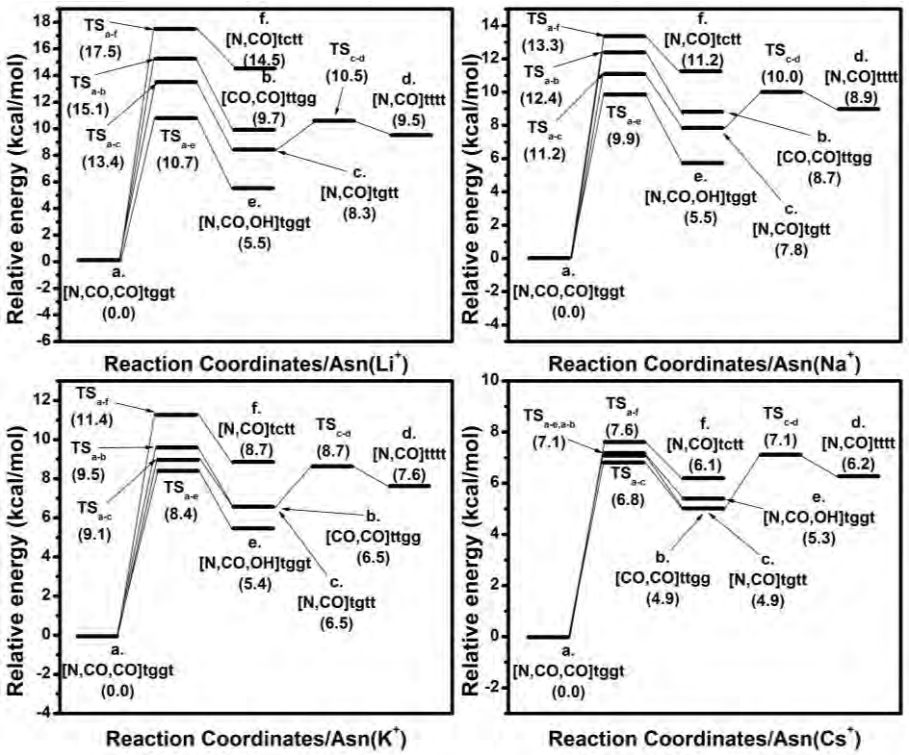

Figure 2.S6. Relative energies between various isomers observed in the finite temperature molecular dynamics for M^+,[N,CO,CO]tggt (M=Li, Na, K, and Cs). Relative energies in kcal/mol are listed in the brackets. The values are calculated at B3LYP/6-311+G(d,p) level (LANL2DZ-ECP basis set for Cs^+) with zero point energy corrections. The molecular conformers listed in each panel are shown in Figure 3.S5.

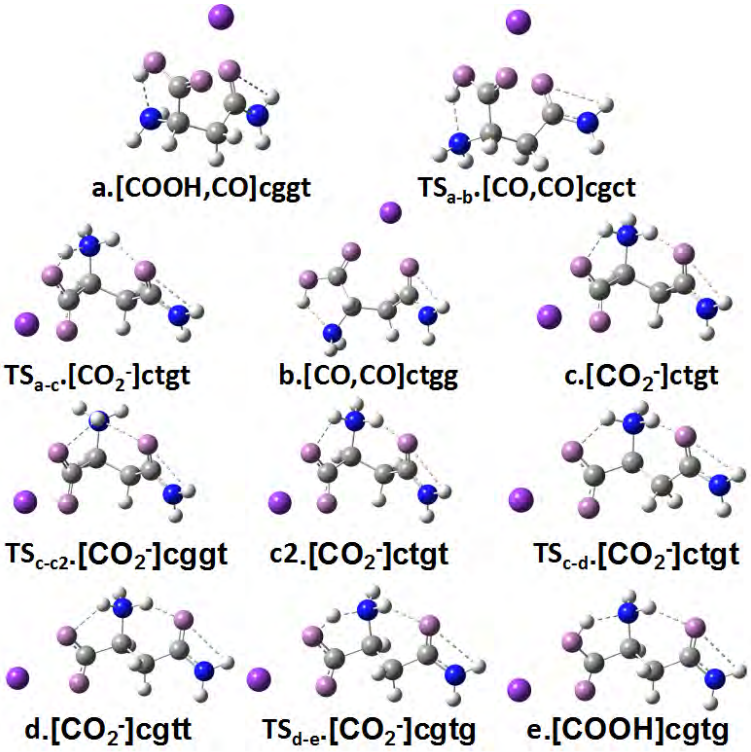

Figure 2.S7. Optimized conformers and transition states for M^+,[COOH,CO] (M= K and Cs) and the corresponding relative energy levels are shown in Figure 2.S8.

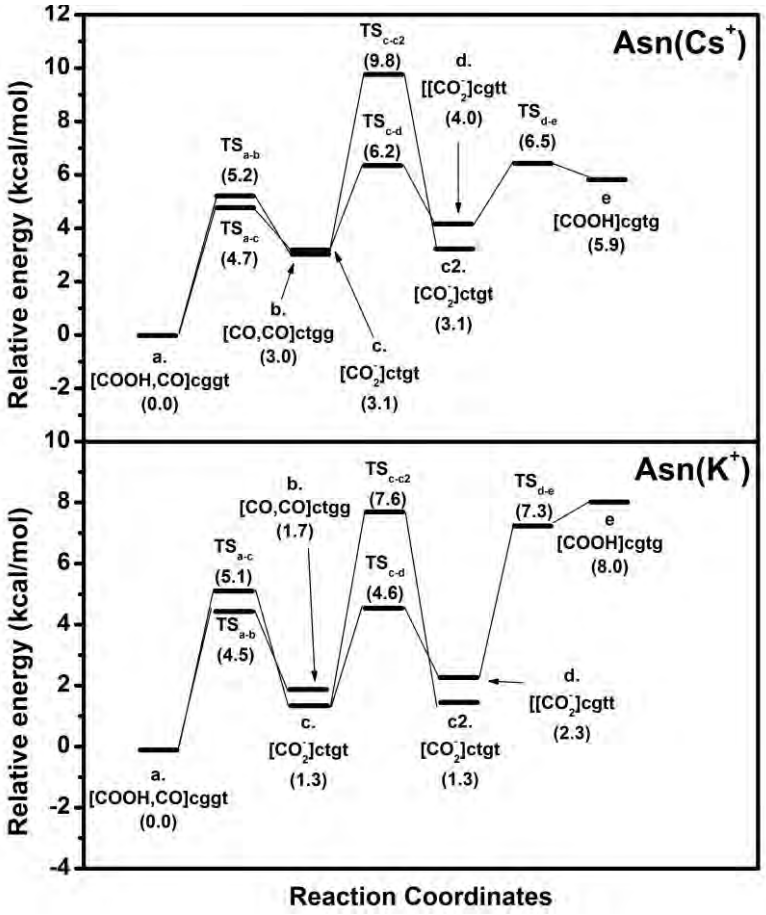

Figure 2.S8. Relative energies between various isomers observed in the finite temperature molecular dynamics for M^+,[COOH,CO]tggt (M=K and Cs). Relative energies in kcal/mol are listed in the brackets. The values are calculated at B3LYP/6-311+G(d,p) level (LANL2DZ-ECP basis set for Cs^+) with zero point energy corrections. The molecular conformers listed in each panel are shown in Figure 2.S7.

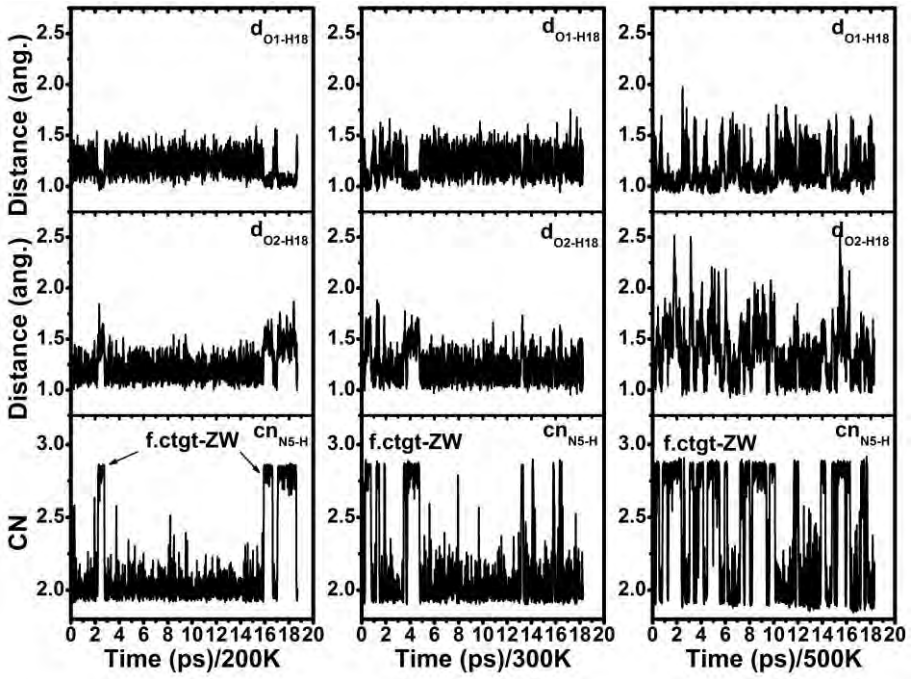

Figure 2.S9. Fluctuation of the atomic distances and coordination numbers for the [CO,CO]ctgt conformer of H^+(Asn) at 200, 300, and 500 K. For the labeling of atoms, see Figure 3.6a.

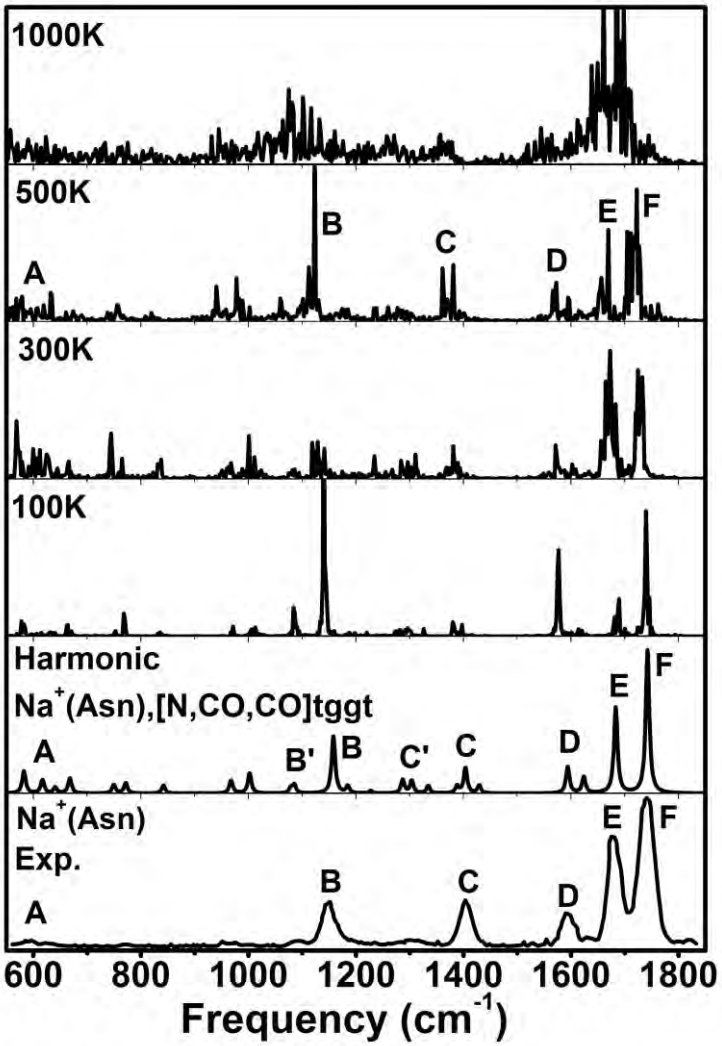

Figure 2.S10. Vibrational profiles for Na^+(Asn),[N,CO,CO]tggt at finite temperatures obtained by AIMD simulations, together with the harmonic and experimental spectrum.

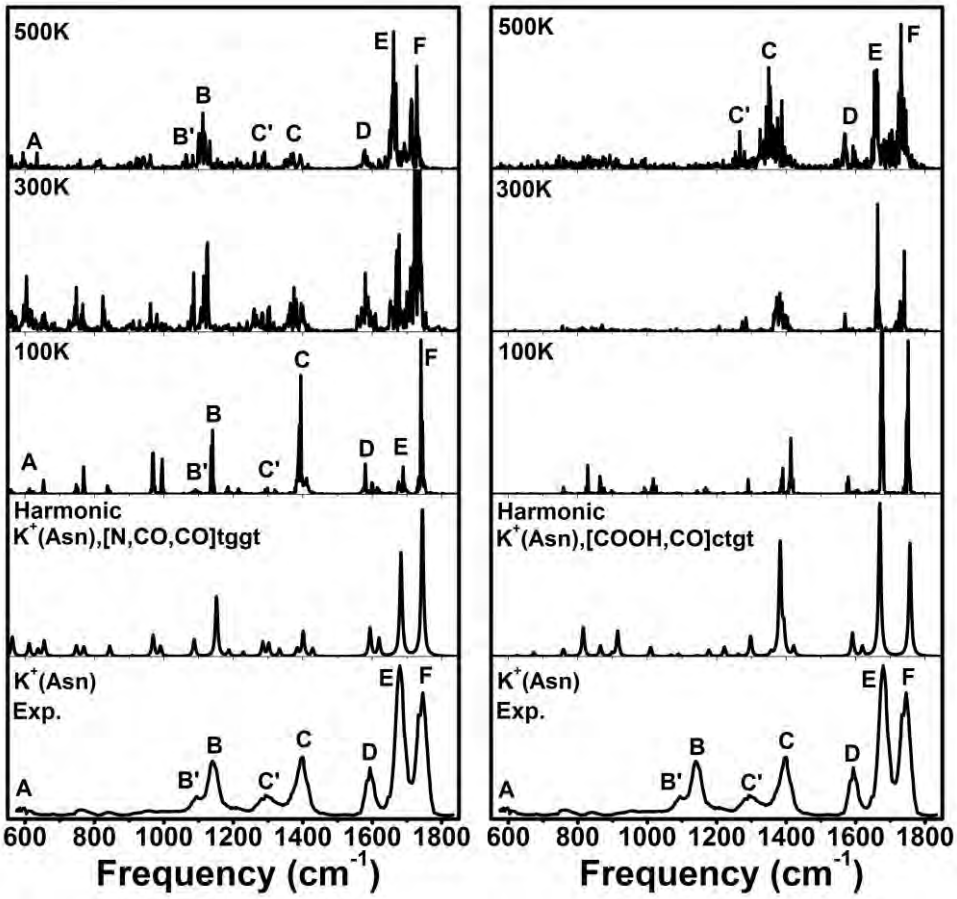

Figure 2.S11. Vibrational profiles for K^+(Asn) at finite temperatures obtained by AIMD simulations. Left: [N,CO,CO]tggt conformer; Right: [COOH,CO]ctgt conformer. Both are in comparison with the harmonic and experimental spectrum.

Chapter Three

AIMD simulated Infrared spectroscopy of Hydrated Bicarbonate Anion Clusters: $HCO_3^-(H_2O)_{1-7}$

3.1 Introduction

Bicarbonate anion is a ubiquitous species in aqueous chemistry and plays a crucial role in the acid-base equilibrium when CO_2 is dissolved in water. Thus this species is important in processes such as pH buffering system[1] of the human body, formation of external skeletons by calcifying organisms,[2] and pH homeostasis in oceans.[3] The solvent interactions of HCO_3^- in aqueous solution[4,5] and infrared spectra of gas-phase $HCO_3^-.(H_2O)_{1-10}$ clusters have been studied.[6] The charged hydrated bicarbonate anion clusters were investigated by IRMPD experimental technique with free electron laser for Infrared experiments facility which was established to probe the vibrational properties of floppy systems in gas phase. The spectroscopic signatures from IRMPD could provide the structural and solvation dynamics information of hydrated gas-phase HCO_3^-.

Density functional theory (DFT) calculations are often used to identify the structures and to assign the observed IR features. The matching of spectral positions and relative intensities between experimental and simulated theoretical results provides the identification of the dominant isomer(s). Normally, the harmonic approximation is used. The drawback is that the vibrational anharmonicities and temperature effects are neglected in the static simulation. This problem can now be fixed by using molecular dynamics. The Berry phase method in the modern theory of polarization and the maximally localized Wannier function made it possible to calculate the dipole moment,[7] and the infrared spectrum can be calculated from the Fourier transform of the dipole time correlation functions[8,9] which are collected along the MD trajectories.

In the previous work, harmonic approximation combined with density functional and MP2 methods are used to study the IR spectra of the gas-phase $HCO_3^-.(H_2O)_{1-10}$ clusters.[6] The hydrated bicarbonate anion clusters are flexible at finite temperature which indicates the dynamic and anharmonic effects cannot be overlooked in this system. As mentioned in the previous work,[6] for some clusters, the comparing the experimental and calculated spectra does not always identity a single isomer as being responsible for the spectrum. In this project, we will study IR spectra of the gas-phase $HCO_3^-.(H_2O)_{1-7}$ clusters at various temperatures. The advantages of AIMD are that with this method, temperature, conformational dynamics and anharmonicities effects are naturally taken into account.[10,11]

3.2 Methods

The conformers are rebuilt previously reported results in Ref. [6] and optimized at B3LYP/6-311++G(d, p) level. Harmonic IR spectra calculations were performed by the Gaussian03[12] package at B3LYP/6-311+G(d,p) level with a scaling factor 0.975.

Finite temperature DFT-based Born-Oppenheimer molecular dynamics simulations (BOMD) were performed with Quickstep in CP2K[13,14] package. The density functional theory implemented in Quickstep is based on the mixed Gaussian and plane wave scheme (GPW),[15] where a dual representation of the electron density is implemented. The wave functions are represented by Gaussians, and the electron density is represented using Gaussians and auxiliary plane waves basis set, for which the advantage is the use of fast Fourier transforms to compute the electrostatic energy. Goedecker-Teter-Hutter (GTH)[16-18] type pseudopotentials and PBE[19] exchange and correlation functional have been used in throughout the present MD simulations. The potentials included Grimme's dispersion correction at D3 level.[20] Our pre-calculations shows that PBE+D3 provided better active bands position in comparison to experimental values than BLYP[21,22]+D3 for the system studied in this work. A double-ζ (DZVP) Gaussian basis set and a 320 Ryd. plane wave density cutoff have been used. The gas-phase simulations were carried out with the decoupling technique of Martyna and Tuckerman[23] in order to eliminate the effect of

the periodical charge density images. All the cubic boxes were selected with box lengths of 18 Å. The convergence criterion for the SCF procedure was set to 10^{-7} a.u. at each step to guarantee the total energy conservation. The dynamics were performed with an equilibration period of 5~6 ps with a control of temperature though temperature rescaling within 20 K and followed by data collection in the NVE ensemble (without temperature rescaling) with a time step of 0.5 fs.

The IR absorption coefficient was calculated by means of the relation derived from linear response theory involving the Fourier transform of the dipole time correlation function (DCFT, eq. 1.53).[24] The assignment of each active band into individual atomic displacements or vibrational modes is done by using the vibrational density of states (VDOS/PDOS, eq. 1.54).[24,25]

3.3 Results and Discussion

There are mainly six peaks in the experimental spectra for $HCO_3^-(H_2O)_{1-7}$ marked from **A** to **F** as shown in Fig. 3.1.[6] According to harmonic analysis in the former analysis, peak **A** which located below 800 cm^{-1} was assigned to the weak CO_2 bending and various water rocking and wagging modes. This band can only be observed as n=6 and above. **B** and **C** located between 800-1000 cm^{-1} were assigned to CO_3^- out-of-plane bending and C-OH stretching modes, respectively. Peak **D** was assigned to the C-O-H bending and peak **E** was assigned to the CO_2 symmetric stretching motion. **F** came from the CO_2 asymmetric stretching which overlapped with the water bending mode. In this work, we will focus on the anharmonic and dynamic effects of these clusters and try to understand the absent of peak **A** in $HCO_3^-(H_2O)_{1-5}$ and the absorption bands change of **D** as the cluster size increases.

3.3.1 Results for $HCO_3^-(H_2O)$

3.3.1.1 Dynamics and stability for $HCO_3^-(H_2O)$ at finite temperatures

For n=1, three isomers were identified in previous work, as shown in Fig. 3.2, with energy differences less than 2 kcal/mol. There are three binding positions (α, β, γ)

for water as marked in **1-1**. The first site ***α*** includes two C=O bonds and each could accept more than one hydrogen bond. The second binding site is ***β***, which consists one C=O and C-O. The third one (***γ***) composes COH and one C=O which act as a hydrogen donor and acceptor, respectively.

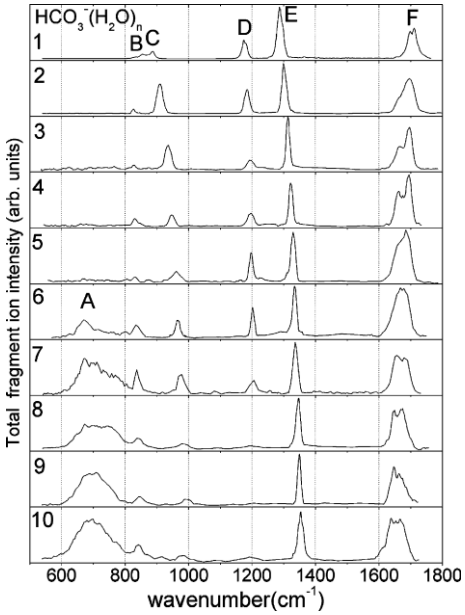

Figure 3.1 Experimental IRMPD spectra of $HCO_3^-(H_2O)_n$ ions with n=1-10 from Ref. [6]. The total fragment ion yield is plotted as a function of irradiation wavenumber (cm^{-1}).

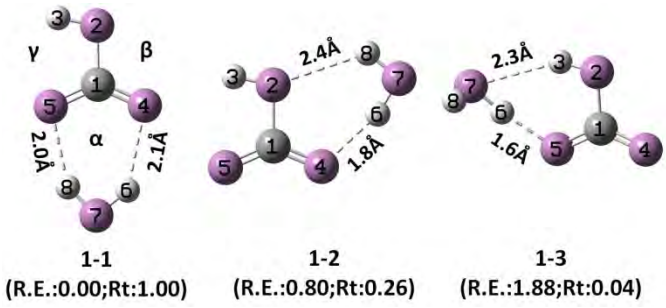

Figure 3.2 Optimized structures for n=1. The relative energies (R.E.) are listed in the parenthesis and the unit is kcal/mol. The right value (Rt.) is the population ratio calculated with Boltzmann distribution at 298.15 K with respected to **1-1**.

The hydrogen bond (Hbond) distance is an indicator of its strength. For **1-1**, two hydrogen bonds at ***α*** with typical bond distances of 2.0 and 2.1 Å, respectively. Hbond O5…H8 is shorter by about 0.1 Å than O6…H4, due to the internal Hbond O5…H3. Mulliken population for O5 is about -0.67 which is slightly bigger than O4(-0.62) at B3LYP/6-311+G(d,p) level. As a result, Hbond O5…H8 is a little

stronger than O6…H4. For **1-2**, the water molecule binds at ***β***. In **1-2**, the hydrogen bond formed between water and C=O acts as hydrogen bond acceptor with a bond length of 1.8 Å which is shorter by 0.3 Å than that in **1-1**. The other Hbond is relatively weak which is formed between water hydrogen and C-O with a typical length 2.4 Å. The solvation type of **1-3** is quite different from **1-1** and **1-2**. Hydroxyl group COH3 in **1-3** forms an Hbond as hydrogen donor with a typical bond length at 2.3 Å. Water hydrogen H6 bonded to C=O5 in **1-3** is the strongest in all the three isomers, with the bond length of 1.6 Å. And the energy barrier for the water flipping (O7H8) is around 1.3 kcal/mol and this process can happen even at 0 K, due to the zero point vibration and tunneling effects of the light atoms. For the transitions from **1-2**, **1-3** to **1-1**, the barriers are also small and the calculated values are around 2.8 kcal/mol.

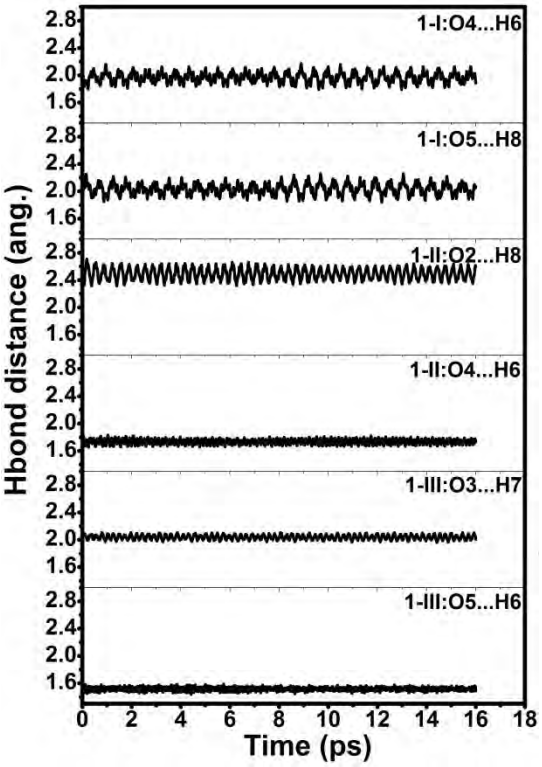

Figure 3.3 Time evolution of hydrogen bond distances: O4…H6, O5…H8 for **1-1**, O2…H8, O4…H6 for **1-2**, and O3…H7, O5…H6 for **1-3**, all at 20K.

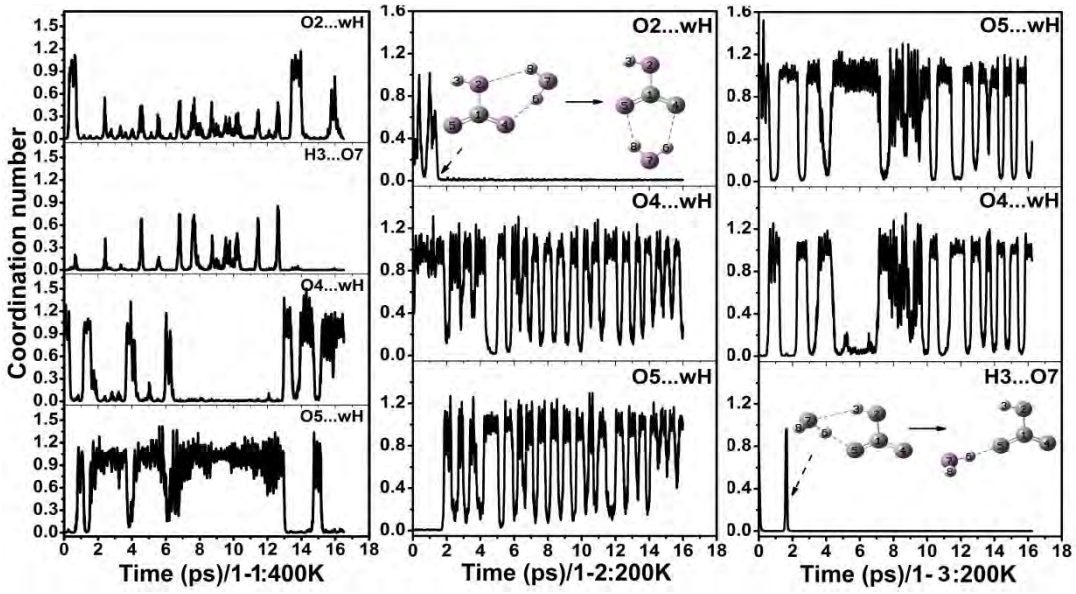

Figure 3.4 Structural transformation observed for **1-1** at 400 K, and **1-2** and **1-3** at at 200 K in the AIMD simulations.

The relative strength of the hydrogen bonds determines the dynamics features of the clusters as temperature arises. Both **1-2** and **1-3** are stable in the AIMD simulations at 20, 50, 100, and 150 K, as demonstrated in Fig. 3.3 (20 K are shown). However, by 200 K, both **1-2** and **1-3** are transformed into **1-1**. And in most cases, the hydrogen bonds involving COH group could be broken with only one hydrogen bond formed between C=O and water (CO5…H6), as shown in Fig. 3.4(right). **1-1** is very stable at 20, 50, 100, 150, 200, and 300 K. And by 400 K, **1-1** can be transformed to **1-2** or **1-3** in part of the simulation, as the calculated hydrogen bonds shown in the left of Fig. 3.4.

3.3.1.2 Harmonic spectra for $HCO_3^-(H_2O)$

The experimental IRMPD spectra were measured in the finger print region, 550-1800 cm^{-1}.[6] The harmonic spectra for **1-1**, **1-2** and **1-3** are calculated at B3LYP/6-311+G(d,p) level and shown in Fig. 3.5, together with experimental spectrum. The simulated harmonic IR in this work are very similar to the previous.[6] For **1-1**, the positions of water scissoring and C=O asymmetric stretching mode are very close to each other, at ~1700 cm^{-1}. And both of the vibrational modes should contribute to experimental peak **F**. However, the discrepancy is that this band, for **1-1**,

is too sharp and the height is bigger than the symmetric C=O stretching mode (s-C=O). The positions of peaks **C**, **D**, and **E** are well reproduced by **1-1**, despite of some discrepancies on the peak intensities. **B** is a broad band with a small tail on the left. And there is no corresponding absorption band for **1-1** in this position. The intensity of CO_3 out-of-plane motion (o-CO_3) in **1-1** is quite small and may offer its contribution to the tail of **B**. And water rocking mode in 700-800 cm^{-1} for **1-1** has no counterpart in experimental spectrum. This is another discrepancy between theoretical and experimental spectrum. For **1-2**, asymmetric C=O stretching and water scissoring/bending are well resolved, but the gap between them is too large. The position for symmetric C=O stretching mode in **1-2** is similar to **1-1** (red-shift ~16 cm^{-1}). It seems that the water at ***β*** has little effect on this mode. This is also true for COH bending mode in **1-2** (red-shifts ~23 cm^{-1}). C-OH stretching mode red-shifts about 50 cm^{-1} relevant to **1-1**, due to the hydrogen bond between O2 and water hydrogen H8, and this peak should be responsible for experimental peak **B**. o-CO_3 mode of **1-2** is nearly in the same position with **1-1** (red-shift ~11 cm^{-1}). And water rocking in **1-2** is also absent in experimental spectrum. For **1-3**, a water binds at ***γ*** and forms a hydrogen with a typical length and a strong hydrogen bond, which makes the water scissoring and C=O asymmetric stretching modes well resolved. There are a little components of water scissoring and C-O-H bending in the intense peak a-C=O. s-C=O position is also stable for **1-3** by comparing to **1-1** and **1-2**. Interestingly, COH bending mode is much affected by hydrogen bond H3…O7, and this mode shifts to blue by 57 cm^{-1} in comparison with the naked COH in **1-1**. However, this peak has no experimental counterpart. One additional peak around ~1000 cm^{-1} comes from water wagging motion (OH8) in **1-3** and this peak shifts to red for **1-1** and **1-2**. C-OH stretching mode in **1-2** is less affected by hydrogen bond H3…O7 and it should correspond to peak **C** in the experimental spectrum. o-CO_3 vibrational mode in **1-3** is also less sensitive to the water binding position and can be assigned to the tail of peak **B**. Again, water rocking and wagging modes (~600 cm^{-1}) also have no counterpart in experimental spectrum for **1-3**.

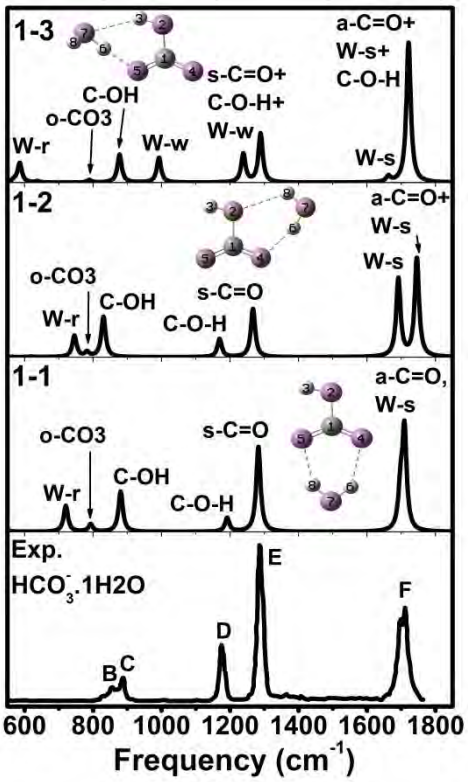

Figure 3.5 Harmonic vibrational spectra for **1-1**, **1-2** and **1-3** in the finger print region, and the corresponding experimental spectrum for n=1.

3.3.1.3 Dynamic effect for each peak of $HCO_3^-(H_2O)$

The dynamic effect on the vibration modes can be simulated through the dipole time correlation function (DTCF), as shown in Fig. 3.6-3.8 for **1-1**, **1-2** and **1-3**, respectively. At 12 K, the simulated DTCF spectra are very similar to the harmonic spectra, despite discrepancies of peak positions which is due to different functionals used. From the harmonic spectra analysis, experimental peak **F** is assigned to the asymmetric stretching of C=O and water bending modes in **1-1**. And the sharpness of this peak does not change much as temperatures rise to 100 K for the three isomers. And this peak is broadened at 150 K and above. The broaden band **F** in **1-2** and **1-3** at 200 K is similar to **1-1**, due to the transformation from **1-2** to **1-1**, and **1-3** to **1-1** in the AIMD simulation.

Peak **E** comes from the symmetric C=O stretching. And this mode is less sensitive to the water binding positions. For **1-1**, this peak is gradually broadened as temperature arises, as shown in Fig. 3.6 which is similar to **1-2** in Fig. 3.7. For **1-3**,

the dynamic effect for some peaks is unique even at 12 K as shown in Fig. 3.8. Peak **E** in **1-3** shifts to red slightly when temperature rises to 100 K. As discussed above, the water molecule in **1-3** binds both the CO-H and the C=O, with later hydrogen bond (O5…H6) being much stronger. The red-shift of symmetric C=O stretching mode indicates that the bonding strength between C=O and water is somewhat weakened as temperature arises. At 150 K, this peak is dramatically broadened due to the large fluctuation of hydrogen bond distance between COH and water oxygen. When temperature goes up to 200 K, the intensity of peak **E** increases dramatically and is well separated.

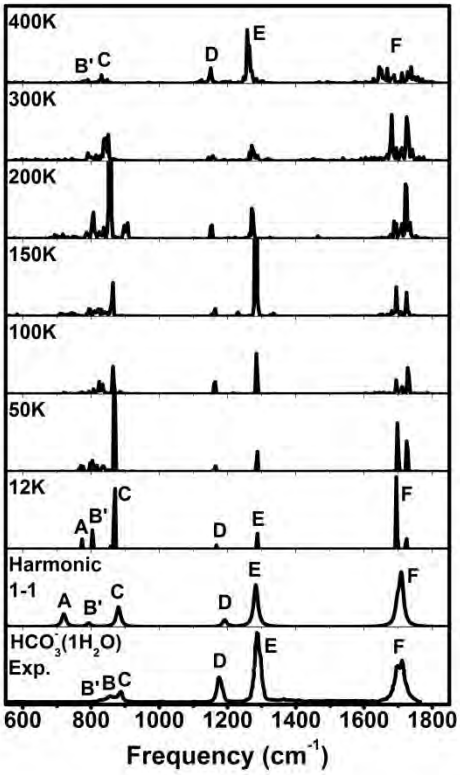

Figure 3.6 Comparison between DTCF, harmonic and experimental spectra for **1-1**.

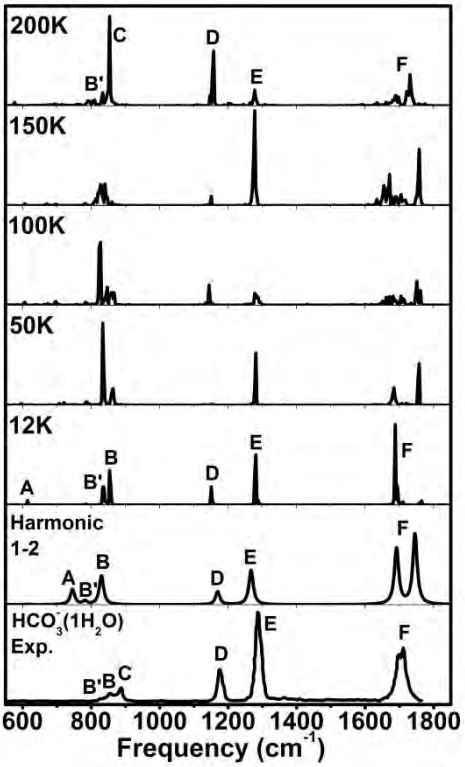

Figure 3.7 Comparison between DTCF, harmonic and experimental spectra for **1-2**.

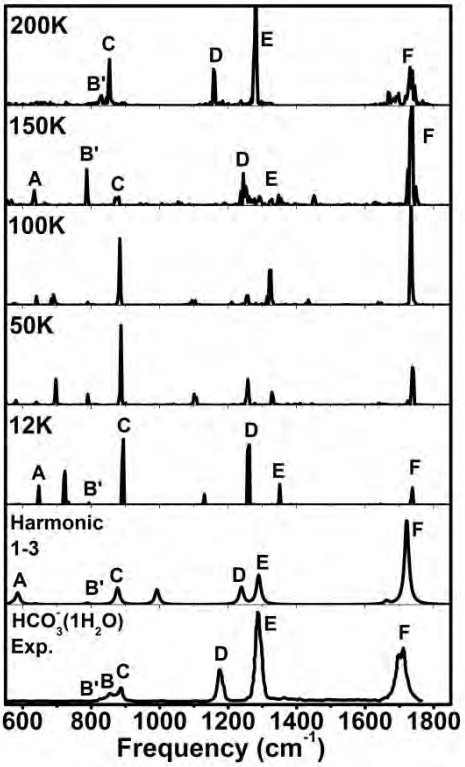

Figure 3.8 Comparison between DTCF, harmonic and experimental spectra for **1-3**.

Peak **D** is assigned the bending mode of naked COH group. The peak positions for **1-1** and **1-2** are very stable at various temperatures, and it is slightly broadened as temperature rise to 150 K and above. However, this mode shifts to blue for **1-3**, due to the solvation of COH. The water molecule in this site in **1-3** will strengthen the COH bending motion. As temperature rises to 150 K, this peak is broadened, similar to **E**, and again is due to the temperature dependent weakened hydrogen bond. At 200 K, peak **D** is well resolved since COH becomes naked in most of the simulation time.

Peak **B** and **C** are very closed in position, which come from C-OH stretching. For **1-1** and **1-2**, the O in C-OH is naked and should be assigned to experimental peak **C** and this peak is slightly broadened as temperature arises. Peak **B** is assigned to the solvated C-OH group in **1-2**, in which the strength of C-OH is weakened by water, as discussed in harmonic analysis. The long tail (**B'**) of peak **B** is assigned to the CO_3 out-of-plane bending mode of all the three isomers. At various temperatures, this mode is only slightly broadened, but the position is nearly unchanged (even for n=2-7) and locates around 790 cm^{-1}.

The water rocking mode (**A**) is absent in the experimental spectrum, although it is well resolved in both harmonic and DTCF spectra at low temperatures. This peak is gradually broadened due to the large motion of water molecule as temperature arises. At 200 K and above, the conformer with one H atom in water dangling to one of the C=O groups is observed during the most simulation time, which makes the water rocking and wagging modes out of resonance with laser. An important consideration is that the experimental spectrum is measured by the dissociation induced by multi-photon absorption. The energy of photon with a frequency of 800 cm^{-1} is around 2.3 kcal/mol, and at least two photons should be absorbed to induce one of the hydrogen bonds breakage. And the experimental spectrum is dependent on both the low and high temperature vibration profiles. After absorbing two photons, one of the hydrogen bonds will be broken and off resonance with laser, which will prevent absorption of additional photons and then no signal is observed for the water rocking motion in experiment. Similar situation is also valid for water wagging mode (harmonic: at 993 cm^{-1}) in **1-3**.

3.3.2 Results for $HCO_3^-(H_2O)_{2-7}$

3.3.2.1 Dynamics and stability for $HCO_3^-(H_2O)_{2-7}$ at finite temperatures

For n=2, three isomers were identified in the previous work.[6] As shown in Fig. 3.9, the energy differences between these isomers are within 2 kcal/mol as listed in the parenthesis. For **2-1**, two water molecules locate at ***α*** and ***β*** sites, respectively. The hydrogen bond distances between water hydrogen and C=O are similar to that in **1-1** and **1-2**, only with the H9…O4 distance increasing ~0.1 Å in comparison with **1-1** (H6…O4). Two water molecules bind at ***α*** site in **2-2** and **2-3** with the COH group being naked. The hydrogen bond distances (H(water)…O4 and H(water)…O5) are ~1.9 and 1.7 Å, respectively, which are shorten by 0.2 and 0.3 Å by comparing to **1-1** due to less angle constraint in water.

In order to discuss the dynamics of each isomer at finite temperatures conveniently, we introduce a labeling method: (a***α***,b***β***,c***γ***) in which a, b, c, are number of water molecules binding at the ***α***, ***β***, and ***γ*** sites, respectively. Following this approach, **2-1** can be labeled as (1***α***,1***β***). Since there is no water binding at ***γ***, we do not label it. For n=2, **2-1**, **2-2**, and **2-4** are selected as the initial structures for AIMD simulations at 150 and 250 K. **2-1** is quite stable at 150 K and it will be transformed into **2-5**(1***β***,1***γ***) at 250 K. **2-2** cannot be maintained even at 150 K, and it can be changed into **2-1** with one watering moving to ***β***. **2-4** is also unstable and it will convert into **2-5**(1***β***,1***γ***) in 1~8 ps at 150 and 250 K by kicking the water molecule (O6) from ***α*** to ***β***. **2-6** belongs to (2***β***) type whose structure is quite different from the other isomers, however this type is rarely observed in the dynamics at 150 and 250 K.

For n=3, **3-1**, **3-5**, and **3-7** are selected as the initial isomers for the AIMD simulations. Using the same scheme, **3-1** in Fig. 3.10 is labeled as (3***α***), with the COH group being naked. And the three water molecules in **3-1** form a ring with each water donating one hydrogen bond to solvate the two C=O groups. **3-3** is of this type, but with different of water linkage. **3-1** is stable at 150 K, but large motion amplitude for hydrogen bonds between water molecules is observed. More isomers appear at 250 K,

such as **3-2**, **3-4**, **3-6**, **3-7**, **3-8**, and **3-10**. **3-5(**1***α***,**1***β***,**1***γ***)** is not stable at 150 and 250 K, it will change into **(**2***α***,**1***β***)** and **3-2** and **3-7** belong to this type. The dynamics for **3-7** at 150 and 250 K are quite similar to **3-5**. One water molecule could be bumped into the second solvation shell occasionally at 250 K in all the dynamics. The left isomers, such as **3-9**, **3-11**, and **3-12**, are rarely observed in the dynamics too, due to the large motion amplitude of water at high temperatures.

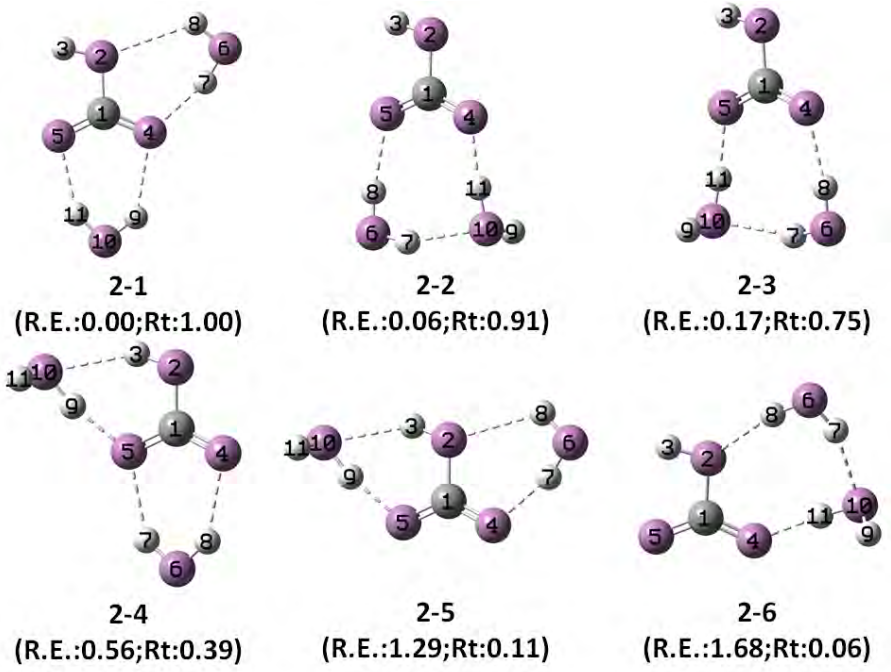

Figure 3.9 Optimized structures for n=2. The relative energies (R.E.) are listed in the parenthesis and the unit is kcal/mol. The right value (Rt.) is the population ratio calculated with Boltzmann distribution at 298.15 K with respected to the ground state.

Similar to n=4, as shown in Fig. 3.11, water molecules can bind to all the three sites of the anion core with the **4-1**(4***α***) being the lowest-energy isomer. Since its IRMPD similarity to n=5, this cluster will not be discussed in detail.

For n=5, **5-1**, **5-3**, and **5-6** are selected to perform molecular dynamics simulations. For **5-1**, there are five water molecules in ***α*** with four hydrogen bonds connected to the two C=O groups and the left water stays in the second solvation shell. This solvation type is stable at 150 K, and position of the water molecule in the second shell is less rigid, by arranging the hydrogen bonds with the water in the first shell, it can be transformed into **5-2**. At 250 K, one water molecule will move into ***α***/***β*** to form (4***α***,1***β***)/(4***α***,1***γ***) type occasionally, such as **5-3** to **5-7** in Fig. 3.12. The

dynamics of **5-3** and **5-6** are quite similar to **5-1**. In general, for n=5, the (5α) type is still dominant in each dynamics, although one or two water molecules can get into β or γ sites occasionally.

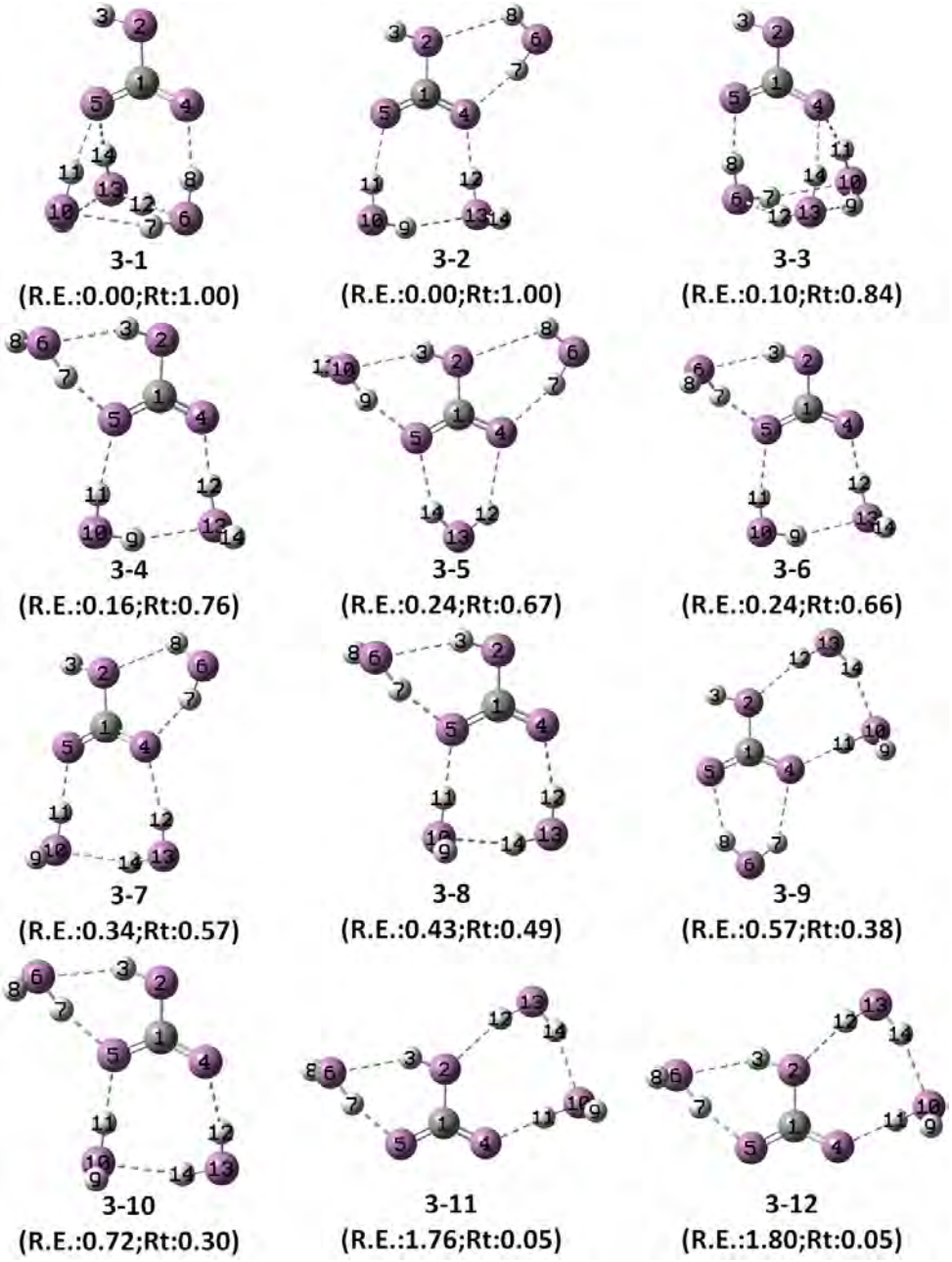

Figure 3.10 Optimized structures for n=3. The relative energies (R.E.) are listed in the parenthesis and the unit is kcal/mol. The right value (Rt.) is the population ratio calculated with Boltzmann distribution at 298.15 K with respected to the ground state.

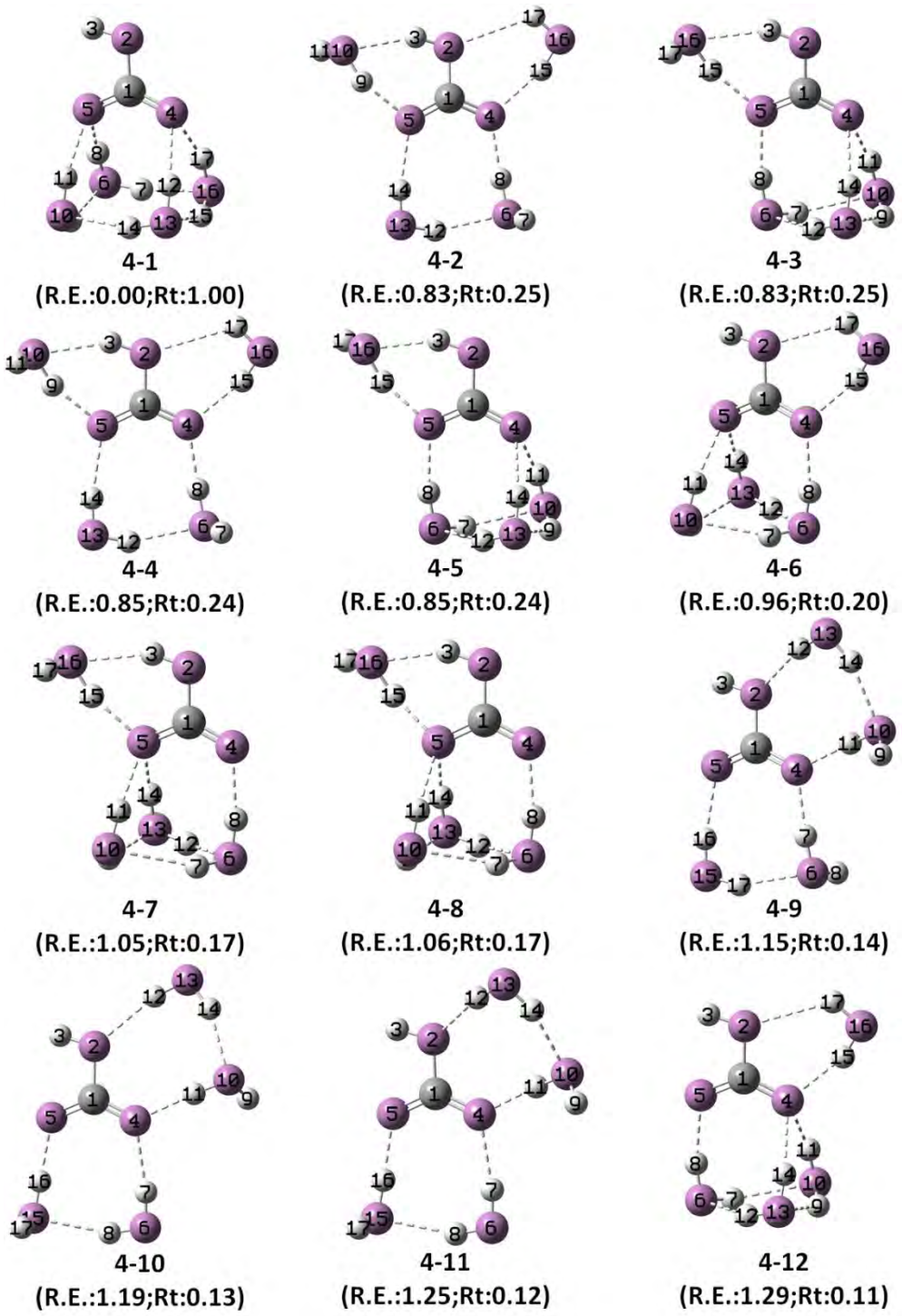

Figure 3.11 Optimized structures for n=4. The relative energies (R.E.) are listed in the parenthesis and the unit is kcal/mol. The right value (Rt.) is the population ratio calculated with Boltzmann distribution at 298.15 K with respected to the ground state.

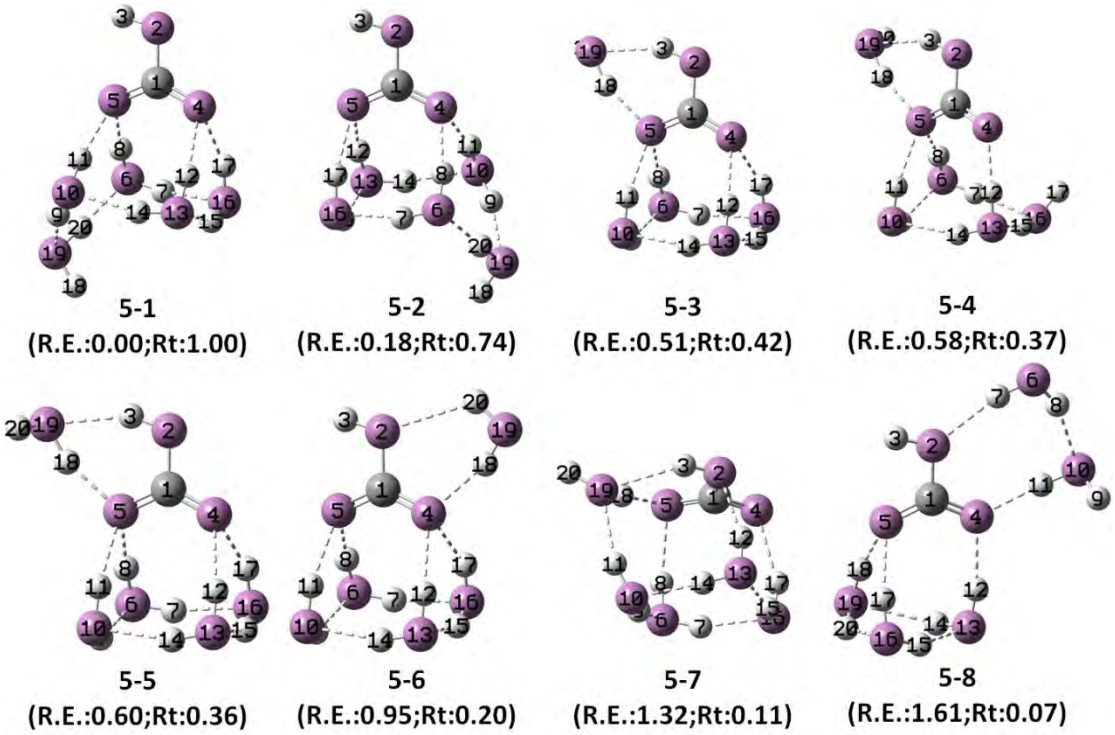

Figure 3.12 Optimized structures for n=5. The relative energies (R.E.) are listed in the parenthesis and the unit is kcal/mol. The right value (Rt.) is the population ratio calculated with Boltzmann distribution at 298.15 K with respected to the ground state.

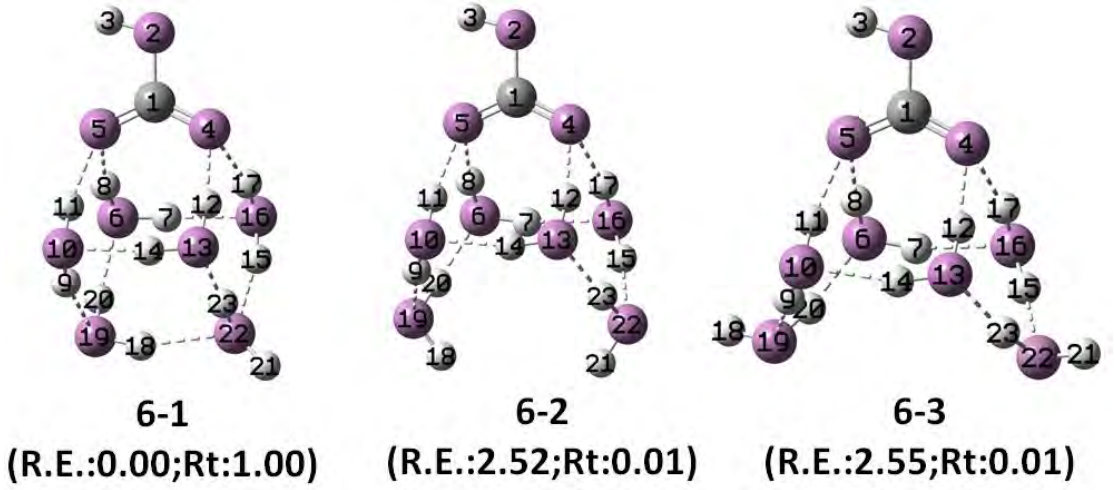

Figure 3.13 Optimized structures for n=6. The relative energies (R.E.) are listed in the parenthesis and the unit is kcal/mol. The right value (Rt.) is the population ratio calculated with Boltzmann distribution at 298.15 K with respected to the ground state.

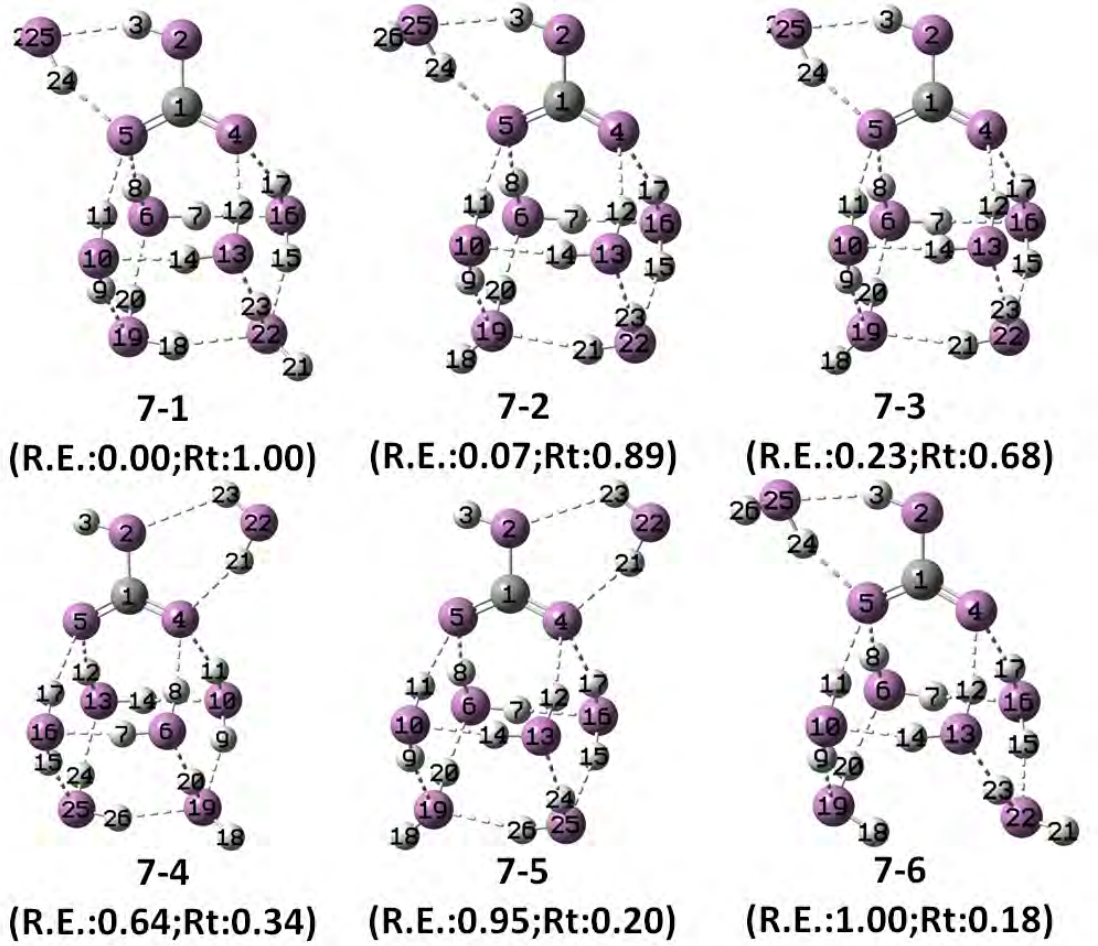

Figure 3.14 Optimized structures for n=7. The relative energies (R.E.) are listed in the parenthesis and the unit is kcal/mol. The right value (Rt.) is the population ratio calculated with Boltzmann distribution at 298.15 K with respected to the ground state.

The relative energy for **6-1**(6α) is much lower than the other isomers (~2 kcal/mol), as shown in Fig. 3.13. Due to its large population, only this conformer is selected as the initial structure in the AIMD simulation. The water molecules in **6-1** form two layers. The lowest layer (second solvation shell) includes two water molecules (O19 and O22) which are linked to the upper layer by four hydrogen bonds. The each four water molecules in the upper layer (first solvation shell) donates one hydrogen bond to one of the C=O groups of the anion core. This solvation type is very stable at 150 K. However, at 250 K, the double layers type cannot be maintained, and more water molecules can go into β and γ sites with the a popular type of **6-**(4α,1β,1γ).

The structure for the lowest-energy conformer (**7-1**) in n=7 is different from the smaller size clusters, as shown in Fig. 3.14. Two water layers are found in the ***α*** site of **7-1**, similar to **6-1**, with one additional water molecule bonded at ***γ***. Whereas for n=1-6, the ground states are either (a***α***) (n=1, 3-6) or (a***α***,1***β***) (n=2) type. **7-1** is very stable at 150 K. As temperature arises to 250 K, water molecules can move into ***β*** and ***γ*** sites too, with the observed dominant type of (5***α***,1***β***,1***γ***). **7-II**, in the previous paper[6] is selected as the initial structure for the AIMD simulations, which is (5***α***,1***β***,1***γ***) type, similar to **5-1**, with two additional water molecules located at ***β*** and ***γ***, respectively. And the dynamics for this isomer is similar to **7-1**. **7-4** belongs to (6***α***,1***β***) and it is stable at 150 K. Whereas at 250 K, this conformer cannot be maintained, and it will also change into the (5***α***,1***β***,1***γ***) type.

In general, for n=1-5, the dominant type is (a***α***) in the AIMD simulations. As cluster size increases, more water molecules can be transferred into ***β*** and ***γ*** sites, and also into the second solvation shell from n=6. This solvation type change will be helpful in the following assignment of IRMPD spectra.

3.3.2.2 DTCF spectra for $HCO_3^-(H_2O)_{2-7}$

To disentangle the discrepancies between the harmonic and experimental IRMPD spectra, the DTCF spectra were simulated, which accounts for anharmonic as well as the dynamic effects. Two sets of long AIMD simulations for n=2,3,5-7 were performed at 150 and 250 K, respectively, for more extensive sampling of the phase space and each lasting for around 100 ps. And each trajectory was then cut into 10 ps interval for Fourier transform, and then added them up to produce the DTCF spectra for a specific temperature. The DTCF spectra from each trajectory differ slightly from each other in the clusters as shown in Fig. 3.15-3.17, respectively. There is considerably better agreement in peak positions between the experimental spectra and the DTCF spectra at 150 and 250 K, especially for bands **C**-**F**. And the broadened bands **A** for n>5, it is also reproduced qualitatively in the DTCF spectra.

For the simulated DTCF spectra of n=2 and 3 in 1000-1400 cm^{-1}, as shown in Fig. 3.15, the experimental peaks **D**, **E**, and **F** are well reproduced in this region. The

intensity decreasing for the solvated POH bending mode of (aα,bγ) between **D** and **E** in each dynamics is due to the broadening. While the bands involving the heavier atoms are not much affected at higher temperature (250 K) in the simulations. Consequently, experimental peaks **D** and **E** correspond to free COH bending and symmetric C=O stretching motion. Similar phenomenon is also observed for the water rocking and wagging modes (PDOS analysis) below 1000 cm^{-1}, since these modes also involve light H atoms too, they will be broadened at a temperature above 150 K. Only the o-CO_3 bending and C-OH stretching modes are left, which is in good agreement with the experimental observations of peaks **B** and **C**.

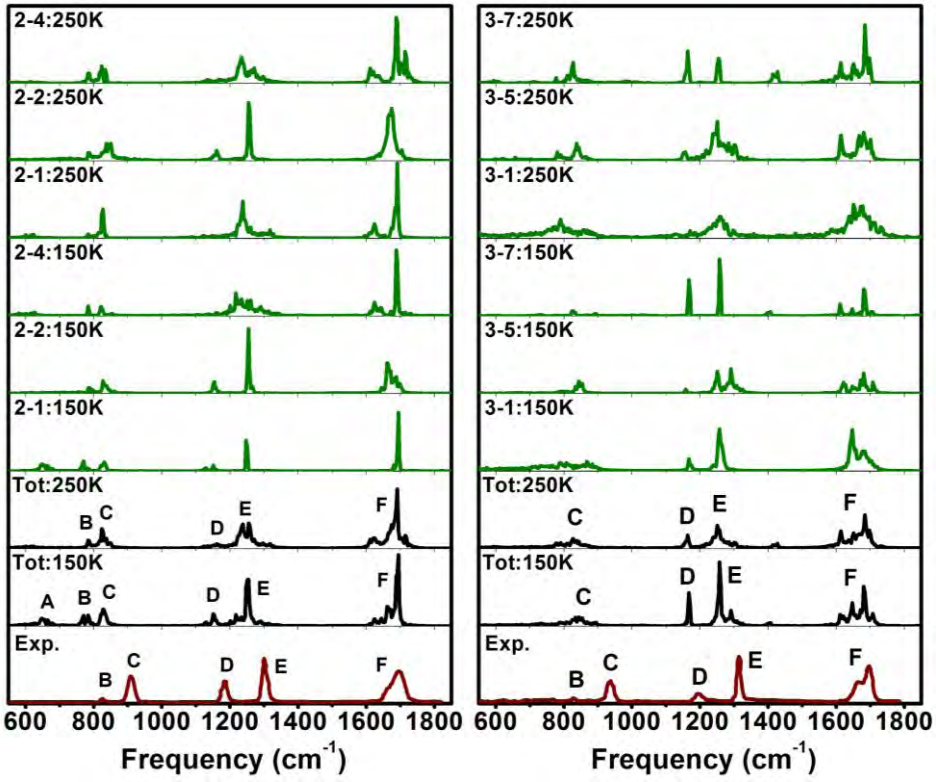

Figure 3.15 Comparison of simulated DTCF spectra with IRMPD spectrum for n=2 and 3, respectively.

Similar broadening phenomenon was also observed for n=5. As discussed above, when more water molecules are available, one water can move into the second solvation shell in the dynamics, which will produce the broadened band **A** in Fig. 3.16. And, this is also true for n=6 and 7, as shown in Fig. 3.16 and 3.17. Since the double C=O groups can only accept four hydrogen bonds at α site, and the left two or three water molecules can either bind at β and γ or in the second solvation shell, which are

observed in the AIMD simulations, especially from n=6. And the peak intensity increasing for **A** from n=6 can be assigned to the water rocking and wagging motion in the two solvation shells.

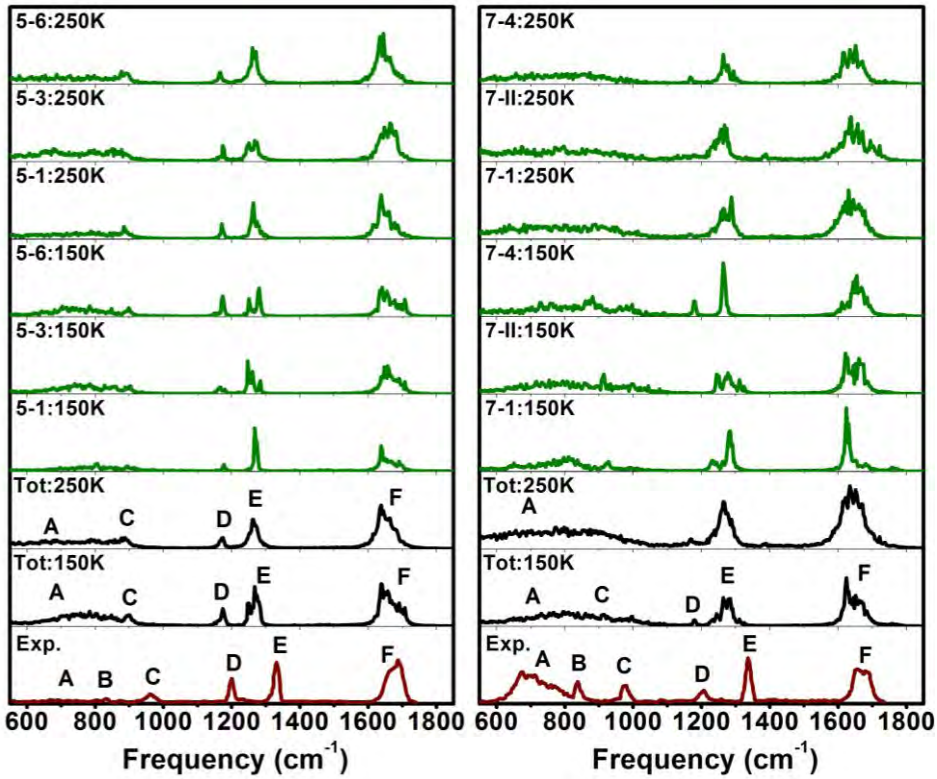

Figure 3.16 Comparison of simulated DTCF spectra with IRMPD spectra for n=5 and 7, respectively.

The bands position changes for different clusters are also well reproduced in the DTCF spectra for peaks **C**, **D**, and **E**, as listed in table 3.1. For band **C**, as cluster size increases, this peak shifts to blue gradually and this change is captured in the simulated DTCF spectra. Similar change for the blue-shifted band **D** in DTCF spectra is also in good agreement with experiment. And the blue-shift in peak D is mainly due to the electrostatic interactions between increasing number of water and the anion core.

Only the band **B** is worse reproduced in the DTCF spectra for n>2. And, this peak indeed shows its intensity for n=2 around 780 cm^{-1}. One reason may explain this discrepancy. As shown in the harmonic analysis, the peak intensity of this mode so quite small that it may merges into the rocking and wagging modes of the water, especially in the higher temperature dynamics. On the other hand, the intensity

changes of peaks **C** and **D**, as cluster size increases, are also not captured by the DTCF spectra, partially due to the sampling efficiency of the AIMD simulations. Trying to explain this phenomenon is a challenge. The possible explanation will be presented in the next part.

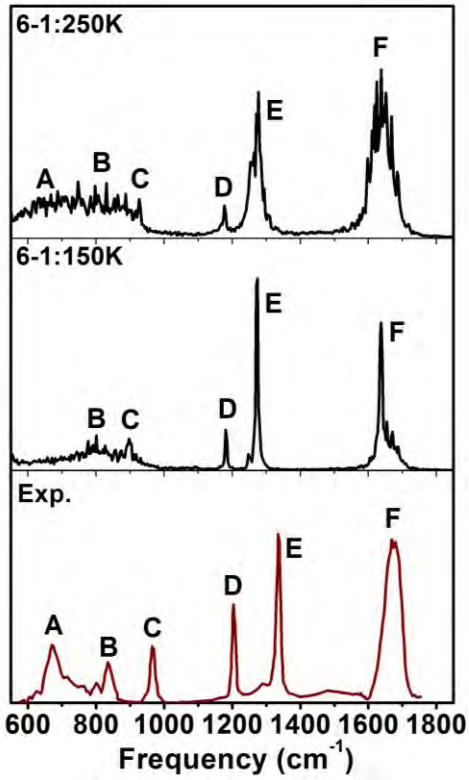

Figure 3.17 Comparison of simulated DTCF spectra with IRMPD spectrum for n=6.

Table 3.1 Simulated peak positions of **C**, **D**, and **E** in DTCF spectra for n=2, 3, 5-7 at 250 K and comparing with experimental peak positions in the parenthesis (unit: cm^{-1}).

	C	**D**	**E**
n=2	827.2 (808.3)	1150.7 (1184.4)	1247.5 (1299.5)
n=3	843.9 (837.2)	1164.1 (1198.1)	1250.1 (1314.0)
n=5	897.2 (961.0)	1174.1 (1199.5)	1264.8 (1331.5)
n=6	897.3 (964.5)	1177.4 (1203.8)	1277.5 (1334.7)
n=7	913.9 (980.2)	1180.6 (1207.1)	1264.2 (1336.0)

3.3.2.3 Peak intensities change in IRMPD spectra for $HCO_3^-(H_2O)_{1-7}$

It is interesting to notice that each absorption band in the experimental spectra of $HCO_3^-(H_2O)_{1-7}$ shows a different intensity change as the cluster size increases, as shown in Fig. 3.18 (left). For peak **A**, this peak is absent for n=1 and 2 and it shows its hint in a very wide distribution for n=3-5. From n=6, peak **A** becomes intense. For experimental peak **B**, its intensity also increases from n=6. The experimental absorption bands **C** and **D** display unique intensity change, as the cluster sizes increase. For **C**, its intensity increases from n=1 to 2 and then decreases from n=2 to 5, and again increases from n=5 to 6 and then decreases for n=7. And band **D** shows a similar change The positions and relative band intensities for **E** for each cluster are stable and give their big intensity around 1250 cm^{-1}. The intensity for peak **F** increases as cluster size growing, apparently due to more water molecules available.

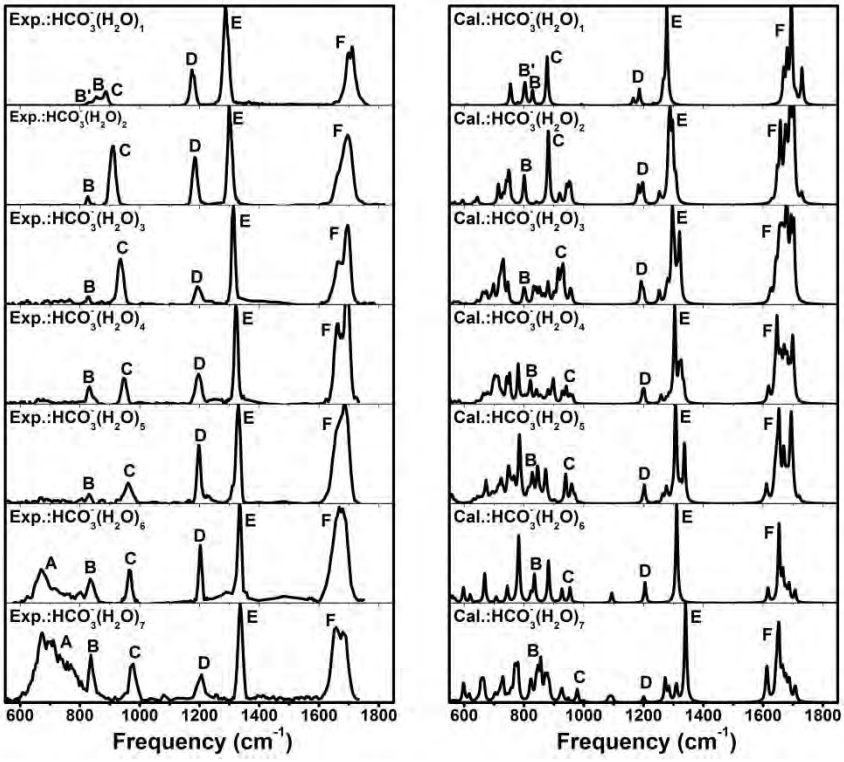

Figure 3.18 Comparison between experimental and calculated IR spectra for n=1-7. The IR spectra are calculated at B3LYP/6-311+G(d,p) level with a scale factor of 0.975. All the isomers in each cluster are added together according to its weight at 298.15 K and then the intensities are scaled respect to peak **E**.

The harmonic IR spectra can be assigned to the low temperature (isomerization still can happen due to the zero point vibration and tunneling effects, especially for light atoms) IRMPD absorption profiles, hence the excitation by the first IR photon is the main factor limiting the dissociation yield of the entire process. Based on the harmonic analysis, for some clusters, comparing the experimental with the calculated spectra does not always identify a single isomer as being responsible for the spectrum.[6] For the AIMD simulation, sampling efficiency is a key factor to guarantee the accuracy of the DTCF spectrum. Additionally, IRMPD is a multiple absorption process, in which the IR photon is absorbed gradually to the dissociation threshold, which means the simulations should be performed at various temperatures. On the other hand, both of the harmonic and AIMD spectra neglect one important factor that the absorption cross sections of each vibrational mode are different, even in the same mode at different temperatures. This will cause discrepancy between experimental and calculated spectrum in the relative band intensity. Here, we suppose that, for the same vibrational mode, the absorption cross section remains the same after absorptions of several photons. Another assumption is that the populations of each isomer are frozen in the cooling process in the experiments, since the experimental pressure is very low and collisions between isomers are rare. By calculating harmonic spectra of various isomers in each cluster and adding the spectra according to their weights according to Boltzmann distribution will offer a different view point to the intensity change of experimental **C** and **D** as the cluster size increases.

The comparison of experimental and density scaled harmonic spectra for n=1-7 is show in Fig. 3.18. For the calculated harmonic IR spectra, the spectra for several isomers in each cluster are summed up by multiplying the Boltzmann population weights at 298.15 K and then scaled the intensities respect to peak **E**. For n=1, **1-1** is the dominant conformer and **1-2** also shows its contribution to peak **B**. **1-3** is the minority and calculated ratio is only 0.04 with respect to **1-1**. For n=2, the populations for **2-2** and **2-3** are quite large, at 0.91 and 0.75, respectively. The large populations of these three conformers will give rise to the intensity of both peak **C** and **D**. And, this

could explain the intensity change from n=1 to n=2 in experimental spectra. For n=3, **3-1** and **3-2** should be the dominant structures. The calculated thermal energies for these two isomers are nearly the same. **3-3**, **3-4**, **3-5**, **3-6**, **3-7**, and **3-8** also show a large population with each ratio above 0.50. For **3-4**, **3-5**, **3-6**, and **3-8**, one water binds at γ and it will decrease the intensity of free COH bending around 1150 cm^{-1} and C-OH stretching modes, due to the broadening as discussed for n=1. And this should be the reason of the intensity decreasing for **C** and **D** from n=2 to n=3. **4-1** is the dominant conformer for n=4 and the population ratios of other isomers are basically less than 0.25. However, these minorities, especially 4-(...,1γ) type **4-2**, **4-3**, **4-4**, **4-5**, **4-7**, and **4-8** will again decrease the intensities of **C** and **D**, as shown in the right panel of Fig. 3.18. For n=5, **5-2** gives a population ratio of 0.74 and the rest are below 0.5. The naked COH groups in both **5-1** and **5-2** will give rise to the intensity of peak **D**, and similar situation is also valid for n=6. This may explain the intensity of peak **D** increasing again for n=5 and 6. For n=7, the (aα,bγ) type, such as **7-1**, **7-2**, and **7-3**, are large in populations and the calculated ratios are above 0.60. On the other hand, the populations of the other solvation types, such as **7-4** and **7-5**, are quite small, about 0.34 and 0.2, respectively. As discussed above, the dominant (aα,bγ) type will decrease the relative intensity of **D**. As the cluster size increases, the water wagging and rocking modes are broadened in a very wide range from ~600 to 1000 cm^{-1} and overlap with peaks **B** and **C**. This should be the reason of the discrepancies between experimental and calculated harmonic spectra.

For n=1-5, the water molecules distribute in the first solvation shell, and at higher temperature ~250 K, the forming and breaking of the hydrogen bonds will cause the normal modes off resonance of laser frequency and furthermore smear out these modes in a wide range in DCTF spectra. For n=6-7, at least two water molecules locate in the second solvation shell. Obviously, the wagging and rocking modes from water in these structures will give rise to the intensities of experimental bands **A**, **B**, and **C**. Hence, the intense peak **A** should be the signature of the existence of the second solvation shell.

3.4 Conclusion

In this work, the experimental IRMPD spectra for $HCO_3^-.(H_2O)_{1-7}$ are studied with both harmonic and AIMD simulations. As discussed in the previous work, comparing the experimental and calculated spectra does not always identify a single isomer as being responsible for the spectrum since it is difficult to distinguish between the isomers with water binding in the ***β*** or ***α*** positions, because they have similar spectra in the energy range and similar peak positions. In this work, we suppose that many isomers coexist with different populations. And the experimental peak intensity changes of **C** and **D** can be understood as the contributions of the less dominant isomers. Interestingly, for n=6, the dominant structure **6-1** with a cubic cage formed by the six water molecules and two oxygen atoms (O4 and O5) in HCO_3^- is very robust at 150 K. For n=7, the large population of ***γ*-**site solvated type isomers decrease the intensity of experimental band **D** dramatically. The calculated weighted harmonic spectra is roughly in agreement with the experiment for peak intensity change in each cluster.

Anharmonic and dynamic effects play a very important role in hydrated HCO_3^- clusters, the hydrogen bonds forming and breaking in the first solvation shell for n=1-5 will smear out the water rocking and wagging motions. For n=6-7, water cages are formed in the ***α*** position and the stabilities of these structures will give rise to the band intensities of **A**, **B**, and **C**. However, the simulated DCTF spectra do not give the accurate band intensity change between each cluster and it is mainly believed to be the sampling efficiency.

3.5 Reference

[1]. http://www.biology.arizona.edu/biochemistry/problem_sets/medph/intro.html
[2]. Orr, J. C.; et al. Nature **2005**, 437, 681.
[3]. Caldeira, K.; Wickett, M. E. Nature **2003**, 425, 365.
[4]. Peng, Z.; merz, K. M. J. Am. Chem. Soc. **1993**, 115, 9640.
[5]. Davidson, M. M.; Hillier, I. H.; Hall, R. J.; Burton, N. A. Mol. Phys. **1994**, 83, 327.
[6]. Erienne, G.; et al. J. Am. Chem. Soc. **2010**, 132, 849.
[7]. A. L. L. East, Z. F. Liu, C. McCague, K. Cheng, and J. S. Tse, J. Phys. Chem. A **1998** 102,10903.

[8]. R. Resta, Rev. Mod. Phys. **1994**, 66, 899.
[9]. D. W. Boo, Z. F. Liu, A. G. Suits, J. S. Tse, and Y. T. Lee, Science **1995**, 269, 57.
[10].M. S. Lee, F. Baletto, D. G. Kanhere, and S. Scandolo, J. Chem. Phys. **2008**, 128, 214506.
[11].P. L. Silvestrelli, M. Bernasconi, and M. Parrinello, Chemical Physics Letters **1997**, 277, 478.
[12]. (1) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb,M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K.N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.;Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li,X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; 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. *Gaussian03*, revision D.01; Gaussian, Inc.: Pittsburgh, PA, **2005**.
[13. VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. Comput. Phys. Comm. **2005**, 167, 103.
[14]. The CP2K developers group, http://cp2k.berlios.de/.
[15]. Lippert, G.; Hutter, J.; Parrinello, M. Theor. Chem. Acc. **1999**, 103, 124.
[16]. Krack, M. Theor. Chem. Acc. **2005,** 114, 145.
[17]. Goedecker, S.; Teter, M.; Hutter, J. Phys. Rev. B. **1996**, 54, 1703.
[18]. Hartwigsen, C.; Goedecker, S.; Hutter, J. Phys. Rev. B. **1998**, 58, 3641
[19]. Perdew, J. P.; Burke, K.; Ernzerhof, M. Phys. Rev. Lett. **1996**, 77, 3865–3868.
[20]. Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. J. Chem. Phys. **2010**, 132, 154104.
[21]. Becke, A. Phys. Rev. A. **1988**, 38, 3098.
[22]. Lee, C.; Yang,W.; Parr, R. Phys. Rev. B. **1988**, 37, 785.
[23]. Martyna, G.; Tuckerman, M. *J. Chem. Phys.* **1999**, *110*, 2810.
[24]. McQuarrie, D. A., Statistic Mechanics; Harper-Collins Publishers; New York, **1976**.
[25]. Cimas, A.; et al. J. Chem. Theory Comput. **2009**, 5, 1068.

Chapter Four
Large Amplitude Motion in Cold $H_2PO_4^-(H_2O)$: AIMD Simulations

4.1 Introduction

A molecular level understanding of the solvation of anions in size selected cluster can lend insight into the behavior in bulk solution as they can probe the precise arrangement of water molecules comprising the solvation shell. Hydrated sulfate dianion,[1-4] bicarbonate,[5] bisulfate,[6] and nitrate,[7] are studied experimentally and theoretically. Dihydrogen phosphate is another important anion which plays a significant role in the environment and in life processes. At physiological pH, $H_2PO_4^-$ and HPO_4^{2-} are abundant and important in acid-base equilibria involved in metabolic pathways. Loss of water from dihydrogen phosphate ($H_2PO_4^-$) leads to metaphosphate (PO_3^-), which is proposed as a key intermediate in the aqueous hydrolysis of phosphate monoesters. However, its identification in solution remains elusive. How phosphate ions are hydrated or dehydrated at the molecular level is thus crucial for a mechanistic understanding of hydrolysis reactions, but difficult to extract from condensed phase measurements. Here, we use gas phase vibrational spectra of isolated monohydrated dihydrogen phosphate in combination with molecular dynamics simulations to extract structural and dynamical information of the $H_2PO_4^-(H_2O)$ complex.

The experimental IRMPD spectrum of $H_2PO_4^-(H_2O)$ was measured from 550-1800 and 2600-4000 cm^{-1} and followed via H_2O loss channel by our collaborators Ling Jiang (State Key Laboratory of Molecular Reaction Dynamics Dalian Institute of Chemical Physics, Chinese Academy of Sciences) and Knut R. Asmis (Fritz-Haber-Institut der Max-Planck-Gesellschaft, Berlin, Germany).

4.2 Methods

Optimized structures and harmonic frequencies are obtained from standard density functional theory (DFT) calculations, using the Gaussian 03 package.[8] The dynamic motion of the clusters are simulated by the Ab Initio Molecular Dynamics (AIMD) method, in which the atoms are treated as classical particles and the potential energy and forces on the atoms are calculated within the framework of DFT at each time step.

For 0 K structure, energy optimization is performed at the level of B3LYP/6-311++G(d,p)[9,10] by Gaussian 03 package.[8] Initial structures are generated by running molecular dynamic simulations over tens of thousands time steps at 200 K and taking random configurations along the trajectories. Harmonic frequencies are calculated by using a larger basis set, at the level of MP2=full/aug-cc-pVDZ. A scaling factor of 0.9993 is used for the low frequency region, and 0.9646 for the high frequency region. The resulting stick spectra are convoluted using a Gaussian line shape function with an fwhm width of 4 cm^{-1} to account for the laser bandwidth, as well as broadening due to rotational excitation.

CP2K package[11,12] is employed for the AIMD simulations. The wave functions are expanded in a double zeta Gaussian basis set,[13] while the electron density is expanded in Gaussians and auxiliary plane waves with an electronic density cutoff at 320 Rydberg. The atomic cores are modeled by the Goedecker-Teter-Hutter (GTH) type pseudopotentials.[14-16] The exchange and correlation energy is calculated by PBE functional, with additional Grimme's dispersion correction at D3 level,[17] which produces harmonic frequencies in better agreement with experiment in our test calculations. A scaling factor of 1.0100 for is used for the low frequency region with respect to experimental peak **G**, and 0.9830 for the high frequency region with respect to experimental peak **A**.

A cluster ion is put at the center of a periodic cubic box, and the effects of the periodic charge density images are corrected by the decoupling technique developed by Martyna and Tuckerman.[18] The convergence criterion for the SCF electronic

procedure is set to 10^{-7} a.u. at each time step. For molecular dynamics at a specific temperature, the temperature is controlled by a Nose-Hoover thermostat, with a time step of 0.5 fs. An equilibration period of up to 10 ps (10ps trajectory) is performed first, with the temperature scaled to an interval of 20 K around the intended value. A data collection run is then followed in the NVE ensemble. Two sets of long AIMD simulations (140 K and 180 K) were performed, for a more extensive sampling of the phase space. At each temperature, two trajectories were simulated, one starting with **1-1** and the other with **1-2**, each lasting 200 ps. Each trajectory was then cut into 10 ps interval for Fourier transform, and all 40 frequency profiles were then added up to produce the DTCF spectrum for a specific temperature.

Hydrated clusters are bound by hydrogen bonds, which are relatively weak and therefore fairly flexible at finite temperature. Dynamic simulations are essential for sampling the solvation structures and for examining the thermal stability of a particular structure. More importantly, the hydrogen bonds could have strong effects on the vibrations, which could be captured by the AIMD simulations. A vibrational spectrum can be directly simulated by the Fourier transformation of the dipole time-correlation function (DTCF, eq. 1.53),[19] with the dipole moment again calculated directly from first principles within the Berry phase scheme.

The vibrational density of state (VDOS, eq. 1.54) could also be easily obtained by the Fourier transform of the velocity time correlation function,[19]with the summation typically going over all atoms in a cluster. VDOS is not directly proportional to the signal intensity measured in a vibrational spectrum, which is determined by the dipole selection rule. But when the summation is restricted to one single atom, the total VDOS could then be decomposed into the contribution from this particular atom, which is helpful for the assignment of a vibrational spectrum.

4.3 Results

4.3.1 Experimental band positions of $H_2PO_4^-$ and $PO_3^-(H_2O)$

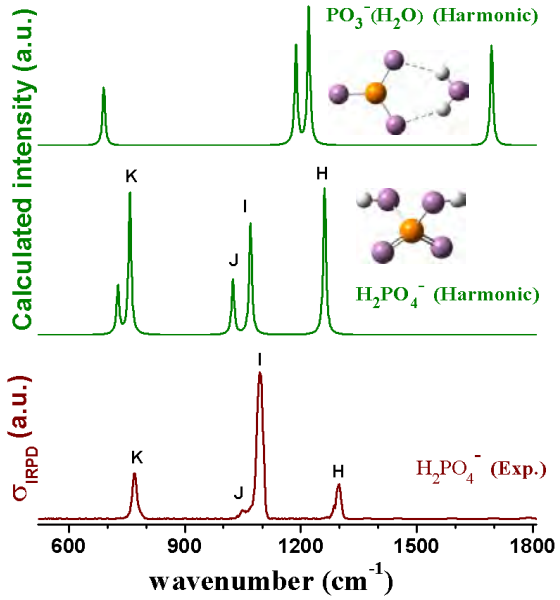

Fig. 4.1. Comparison of the experimental IRMPD spectrum of $H_2PO_4^-$ (bottom) to simulated MP2=full/aug-cc-pVDZ harmonic vibrational spectra of the global minimum energy structures of $H_2PO_4^-$ (middle) and $PO_3^-(H_2O)$ (top).

In order to check that the dihydrogen phosphate anions was actually formed by electrospray in experiments, the IRMPD spectrum of $H_2PO_4^-$ in the fingerprint region was measured by monitoring the H_2O loss channel as shown in Fig. 4.1. Comparison of the experimental band positions to those in the simulated MP2/aug-cc-pVDZ harmonic vibrational spectra of $H_2PO_4^-$ and $PO_3^-(H_2O)$ yields satisfactory agreement only with the spectrum of $H_2PO_4^-$, allowing an assignment of the four IR-active features to the asymmetric (**H**, 1299 cm^{-1}) and symmetric (**I**, 1094 cm^{-1}) P=O stretching, POH bending (**J**, 1049 cm^{-1}) and asymmetric P-OH stretching (**K**, 770 cm^{-1}) modes. Poorer agreement between 600-1400 cm^{-1} as well as the lack of any signal in the water bending region (~1700 cm^{-1}) rules out the contribution of the monohydrated metaphosphate anion, which is predicted to lie 37.4 kJ/mol higher in energy than dihydrogen phosphate anions. Discrepancies between IRMPD and the linear harmonic intensities are attributed to the high predicted dissociation energy of

$H_2PO_4^-$ (122.1 kJ/mol) for water loss channel (at MP2/aug-cc-pVDZ level with zero point energy and BSSE corrections). Thus, absorption of many IR photons is required for IRMPD.[20]

4.3.2 Relative strength of hydrogen bonds of $H_2PO_4^-(H_2O)$

The two most stable predicted structures for $H_2PO_4^-(H_2O)$ are the complexes containing a double donor (**DD**) configuration bound to the two P=O bonds and an acceptor/donor (**AD**) motif, labeled **1-1** and **1-2**, respectively, as shown in Fig. 4.2. In **1-1,** the H_2O molecule donates two HBs (hydrogen bonds) to the PO_2^- moiety, yielding a symmetric structure (C_2), in which the negative charge in $H_2PO_4^-$ is stabilized on the PO_2^- moiety, with a calculated Mulliken population number around −0.9. The HB distances of 2.07 Å in **1-1** are only slightly longer than the distance of 2.00 Å in the water dimer, implying that the interactions are of comparable strength. The structure of **1-2** is asymmetric. The H_2O molecule donates a strong HB (1.64 Å) to one P=O group, and accepts a weaker one (2.13 Å) from one of the hydroxyl groups. The hydrogen bond distance can be taken as an indicator its strength. As a result, the acceptor hydrogen bond (O3...H9) in **1-2** is the strongest. The naked P=O and P-OH groups do not interact with the water molecule, but form a weak internal hydrogen bond (2.43 Å).

The energetic ordering of these two isomers depends on the model used. B3LYP predicts **1-2** as the global minimum energy structure and **1-1** 0.8 kJ/mol higher in energy, including zero point energies (ZPE). In contrast, MP2 places **1-2** 0.1 kJ/mol above **1-1**. These structures are separated by a small barrier (B3LYP: 3.5 kJ/mol; MP2: 6.5 kJ/mol) at the first order transition state (TS) **1-3** (Fig. 4.2), indicating a fairly flat potential surface. In structure **1-3**, the H_2O molecule forms only a single HB with a phosphoryl group. The MP2/aug-cc-pVDZ binding energies between $H_2PO_4^-$ and water in **1-1** and **1-2** are 52.8 and 50.0 kJ/mol including zero point energy and BSSE corrections, respectively, which are close to the experimental value of 58.6 kJ/mol.[21]

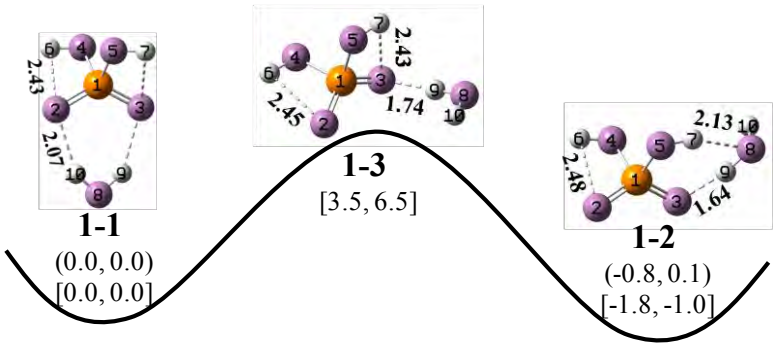

Fig. 4.2 Minimum energy (**1-1** and **1-2**) and first order transition state (**1-3**) structures for the $H_2PO_4^-(H_2O)$ complex. B3LYP/aug-cc-pVDZ and MP2=full/aug-cc-pVDZ relative energies (in kJ/mol) are listed with (inside round brackets) and without (inside square brackets) ZPE corrections. The MP2 distances of hydrogen bonds are given in angstrom.

4.3.3 Calculated harmonic spectra of $H_2PO_4^-(H_2O)$

Simulated MP2 harmonic spectra of **1-1** and **1-2** in the low frequency regions are shown in Fig. 4.3 (see Table 4.S1 for band positions and assignments in the appendix). As seen the vibrational modes in this figure, the scissoring motion of water contributes a sharp peak just above 1600 cm^{-1}, well separated from all the other peaks. But in the experimental spectrum, only a broad band is observed around 1700 cm^{-1}, which presents a significant discrepancy.

In **1-1**, which has a C_2 axis, each vibration is further divided into an asymmetric and a symmetric mode, which in total should produce four peaks, labeled as ***a1***, ***s1***, ***a2***, and ***s2*** in Fig. 4.3, in the order of decreasing frequency. A large component of P=O stretching is found in ***a1*** and ***s2***, while the two middle peak ***s1*** and ***a2*** are very close to each other, producing three peaks between 1000-1300 cm^{-1}. Below 800 cm^{-1}, there are further bands due to the symmetric and asymmetric stretching of the two P-OH bonds and the rocking of the hydrogen bonded water molecule.

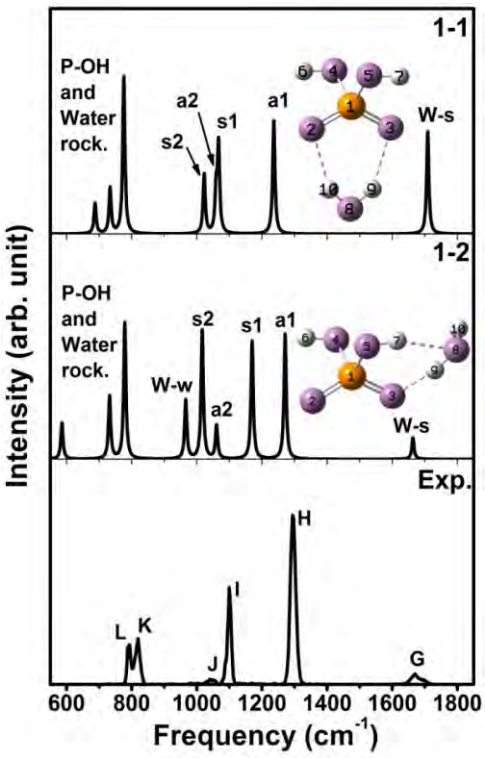

Fig. 4.3 Comparison of experimental spectrum of $H_2PO_4^-(H_2O)$ (bottom) to simulated MP2 harmonic spectra of the isomers **1-1** and **1-2** (upper) in the low frequency region.

The asymmetric **1-2** differs significantly from **1-1** in that one of the P-OH groups is now solvated by the water with the **AD** feature. However, there are again four peaks for the P-O-H bending and P=O stretching, and the same sequence of labels, ***a1,*** ***s1, a2,*** and ***s2***, can be borrowed to indicate the dominant component in each mode. Peak ***a1*** again has a large component of P=O asymmetric stretching, and its position is blue shifted, as the strong hydrogen bond (H9…O3) makes the structure more rigid. Peak ***s2*** also has a large component of P=O symmetric stretching, and there is little movement in its position. Symmetric P-O-H bending is mixed with a small amount of asymmetric P=O stretching in ***s1***, which is also significantly blue-shifted. The ***a2*** peak decreases in intensity, and is largely due to the bending of the solvation free POH group. In both ***s1*** and ***a2***, there is a large component of the H9 wagging. A new feature, also due to water wagging and labeled as "**W-w**", is found just below 1000 cm^{-1} due to the wagging motion of H10. Below 800 cm^{-1}, the stretching of the two P-OH groups is not much affected by the structural change.

Agreement between the harmonic and the experimental spectra is found only for the P-OH stretching features around 800 cm^{-1}. Based on the calculated relative energy and thermal stability, both **1-1** and **1-2** should make a contribution to the experimental spectrum. The discrepancy between harmonic and experimental spectra is therefore quite pronounced, and harmonic spectra alone cannot account for the experimental observations.

In the high frequency region, the features are due to the O-H stretching in the H_2O molecule and in the PO-H groups. The spectra for the two isomers are quite different from each other. **1-1** has C_2 symmetry, and the O-H stretching modes for the two O-H bonds of water are coupled, producing a symmetric and an asymmetric mode. The same applies to the two PO-H bonds. The frequency difference between the symmetric and asymmetric stretching modes is 47 cm^{-1} for H_2O, and only around 1 cm^{-1} for the two PO-H bonds.

1-2 is a low symmetry structure. In the water molecule, H9 forms a strong HB with O3 in one of P=O groups, while H10 is free. Similarly for the P-OH bonds, H7 forms a hydrogen bond with O8 of the water molecule, while H6 forms only a weak internal hydrogen bond with O2. Each of the four O-H stretching modes becomes local, as the stretching of just one O-H bond. PO4-H6 is naked, and its stretching occurs at almost the same position as the free PO-H peaks for **1-1**. Peak **A** in the experimental spectrum is therefore assigned to the free PO-H group stretching mode.

The PO5-H7 group in **1-2** forms a hydrogen bond of typical length and its stretching frequency is red-shifted by 140 cm^{-1} from the stretching of a free PO-H. It corresponds to peak **B** in the experimental spectrum, which is the signature of a solvated PO-H group. More dramatic changes are observed for the stretching of H_2O in **1-2**. The H9-O8 stretching is red-shifted by around 600 cm^{-1}, as the bond strength of H9-O8 is significantly affected by the strong hydrogen bond H9…O3. For the other bond H10-O8 in the same water molecule, its stretching is blue-shifted by 180 cm^{-1}, indicating a bonding strength of the free O8H10 bond is weakened. As for the peak heights, it is interesting to notice that the stretching of O-H bonds involved in hydrogen bonding atoms (H9-O8 in water, and PO5-H7) produce larger intensities.

In comparison with experiment, three bands can be readily assigned based on the calculated harmonic spectra. Peak **A** comes from the O-H stretching of a free PO-H, which is present in both **1-1** and **1-2**. Peak **B** is due to the O-H stretching of a solvated PO-H, which is only present in **1-2**. Peak **C** originates from the O-H stretching in water, contributed by the O-H stretching of water in **1-1**. It should be noticed that the hydrogen bonds in **1-1** is of typical strength, as indicated by its distance around 2.1 Å.

The discrepancy between experiment and harmonic spectra is also very obvious in the high frequency region. On the one hand, there is no feature in the harmonic spectra, corresponding to peak **D** and the very broad features **E** and **F** in the experimental spectrum. On the other hand, the H9-O8 stretching in **1-2** produces a strong peak in the harmonic spectrum, but is not observed in the experimental spectrum. Please notice that H9 forms a very strong hydrogen bond with a P=O group (O3), for which the dynamic effect could be very important.

In summary, satisfactory agreement between experiment and harmonic theory is observed for isomer **1-1**, in particular below 2000 cm^{-1}. The harmonic spectra for **1-1** and **1-2** account for all the experimentally observed peaks (**G-L**), while the spectrum of **1-2** predicts additional intense bands at 1170 cm^{-1} and 966 cm^{-1}, which are not observed in the experiment. Above 2000 cm^{-1}, the spectrum of **1-1** also accounts for peak **A** and the doublet **C** at ~3450 cm^{-1}. However, the harmonic spectrum of **1-1** leaves bands **B** and **D-F** unassigned. On the other hand, the simulated harmonic spectrum of **1-2** yields reasonable assignments for peaks **A**, **B** and **F**, but leaving **C-E** unassigned. In particular, the strongest hydrogen bond in **1-2** (1.64 Å) agrees well with the characteristically red-shifted band **F** (2700-3100 cm^{-1}), even though its width cannot be rationalized at the harmonic level. In sum, neither harmonic IR spectra of the two isomers, nor a linear combination of the two, can satisfactorily explain the experimental IRMPD spectrum.

To disentangle the discrepancies between the harmonic and experimental IRMPD spectra, ab initio molecular dynamics (AIMD) simulations were performed. Briefly, vibrational profiles at finite temperature are obtained by the Fourier transform of the

dipole time correlation function (DTCF), which accounts for anharmonic as well as dynamic effects.

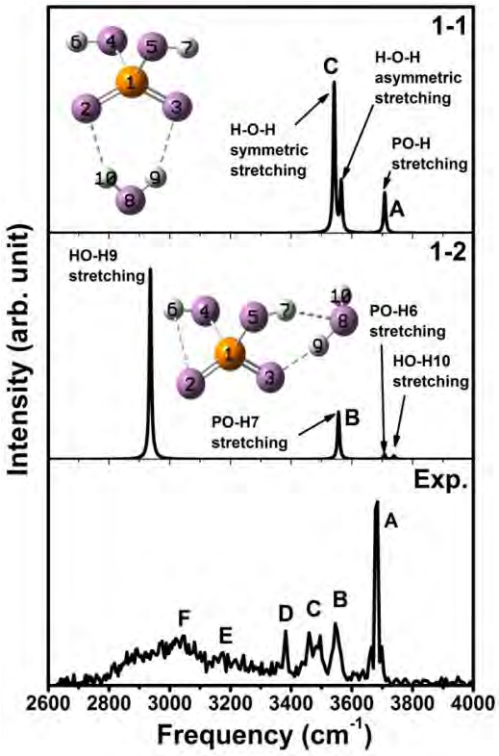

Fig. 4.4 Comparison of experimental spectrum of $H_2PO_4^-(H_2O)$ (bottom) to simulated MP2 harmonic spectra of the isomers **1-1** and **1-2** (upper) in the high frequency region.

4.3.4 Dynamics and stability at finite temperature

The relative strength of the hydrogen bonds determines the dynamic features of the cluster at raised temperature. Both **1-1** and **1-2** are stable during the AIMD simulations at 20, 60, and 100 K, as demonstrated in Fig. 4.5 (only dynamic at 60 K is shown). However, by 140 K, both **1-1** and **1-2** could be transformed into each other during part of the simulation as shown in Fig. 4.6. At 180 K, the PO-H donor hydrogen bond could be broken, and there is only one hydrogen bond left, between a P=O group and the water molecule, like a single thread holding the cluster together (**1-3**). These simulations indicate that both **1-1, 1-2** and **1-3** could all contribute to the experimentally observed spectra.

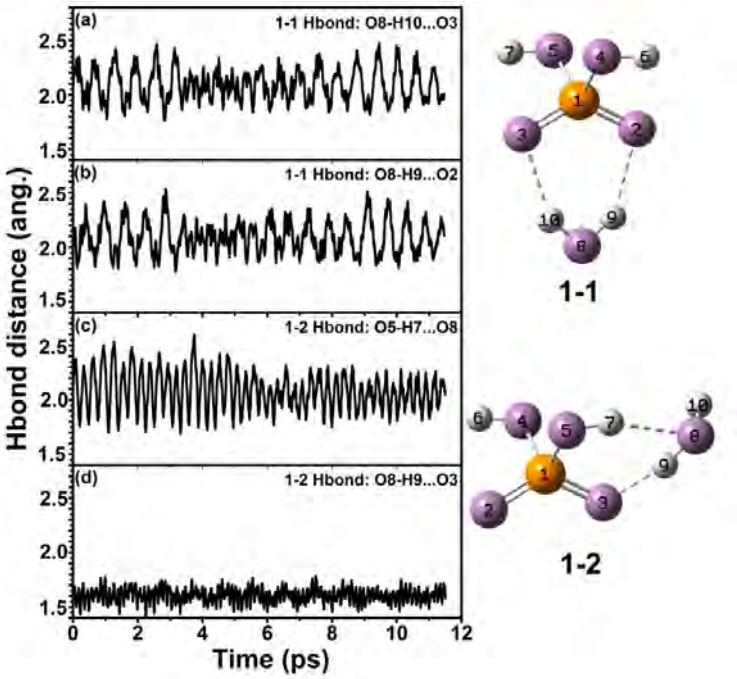

Figure 4.5 Time evolution of (a) H10…O3 and (b) H9…O2 distances for **1-1**, and (c) H7…O8 and (d) H9…O3 distances for **1-2**, both at 60 K.

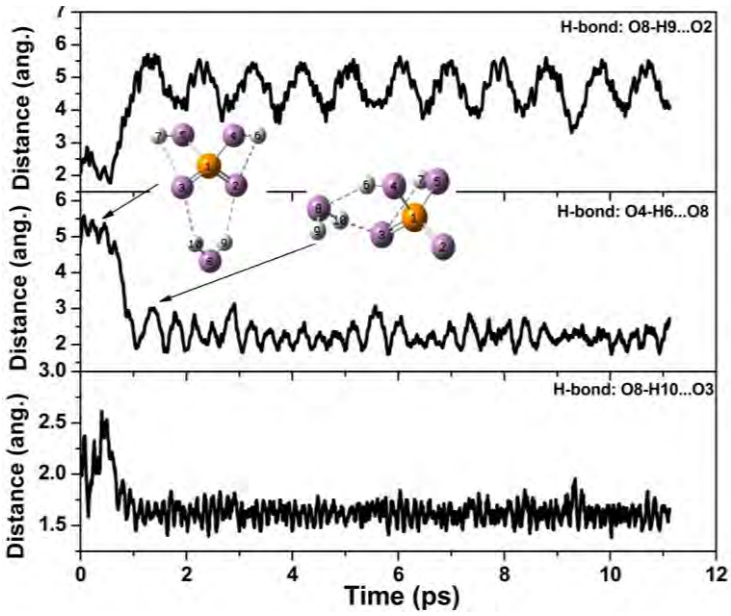

Figure 4.6 Structural transformation observed for **1-1** in an AIMD simulation at 140 K. Within 1 ps, the O8-H9…O2 hydrogen bond is broken, while at the same time, the PO4-H6…O8 bond is formed, and the O8-H10…O3 bond is strengthened, forming an **AD** structure which is maintained to the end of the 10 ps simulation.

4.3.5 Simulated DTCF spectra for $H_2PO_4^-(H_2O)$

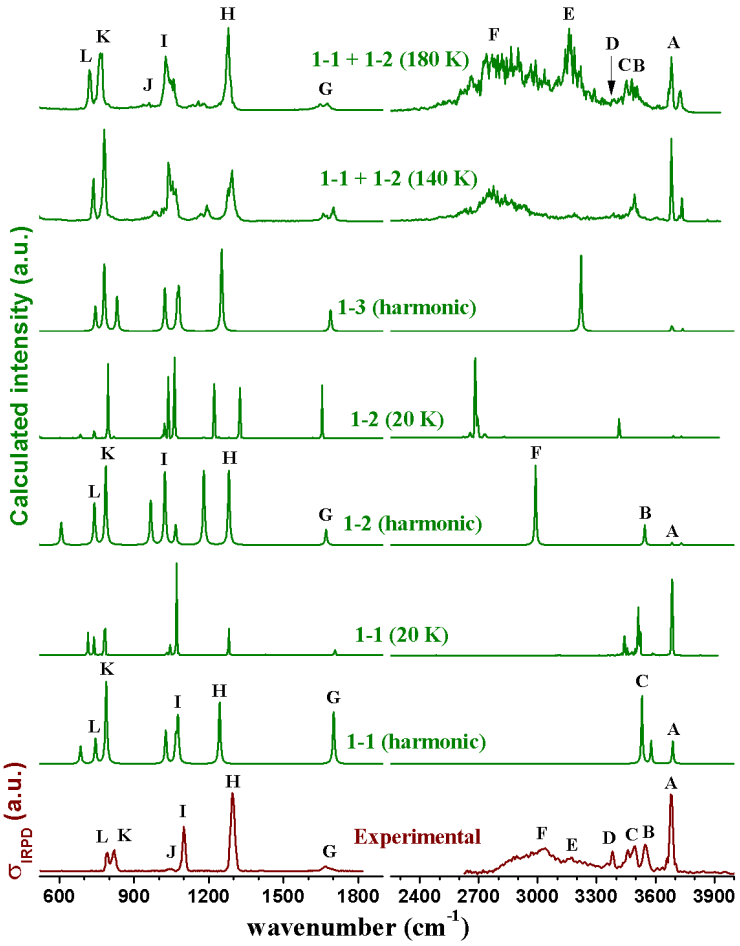

Figure 4.7 Comparison of the experimental IRMPD spectrum of $H_2PO_4^-(H_2O)$ (bottom) to simulated MP2/aug-cc-pVDZ harmonic and DTCF spectra from 200 ps trajectories of isomers **1-1** and **1-2**, and transition state **1-3**.

The AIMD simulations at 20 K are helpful to test the quality of the potential energy surface, but do not correspond to a physically achievable temperature, since zero-point energies are not considered. The general appearance of DTCF spectra at 20 K is indeed similar to the previously discussed harmonic spectra, with the hydrogen-bonded O-H stretching modes showing the largest shifts due to the use of different methods (PBE vs. MP2). Interestingly, the relative band intensities of experimental features **A-C** and **G-L**, with the exception of band **H**, are reproduced well by the simulation of **1-1** at 20 K comparing to the harmonic spectrum (Fig. 4.7). The **1-2** spectrum at 20 K, on the other hand, still mainly reflects the harmonic intensities, but does capture the pronounced red-shift to 2700 cm^{-1} of band **F**. To

consider ZPE and finite temperature effects, two sets of long AIMD simulations were performed at 140 K and 180 K, for a more extensive sampling of the phase space. At each temperature, two trajectories were simulated, one starting with **1-1** and the other with **1-2**, each lasting 200 ps. Each trajectory was then cut into 10 ps interval for Fourier transform, and all 40 frequency profiles were then added up to produce the DTCF spectrum for a specific temperature. The DTCF spectra from both trajectories differ only slightly from each other, indicating that **1-1** and **1-2** interconvert readily at these temperatures.

There is considerably better agreement between the experimental spectra and the DTCF spectra at higher simulation temperatures (140 K or 180 K) throughout the spectral range for band positions and relative intensities. At these temperatures, isomers **1-1** and **1-2** are interconverting and this is important to rationalize the experimental IRMPD spectra. The complexity of the features in the O-H stretching region as well as the number and relative intensities of the bands in the fingerprint region are qualitatively reproduced. The increased broadening of the hydrogen-bonded O-H stretching bands **B**-**F** with the strength of HB (the red shift), is also captured.

In detail, peak **A** remains sharp at 140 K and 180 K, indicative of a O-H stretching mode of a free PO-H group, and it rises in relative peak intensity at higher temperature. Both **1-1** and **1-2** contribute to peak **A**. Peak **B**, which is due to the O-H7 stretching in **1-2** (Fig. 4.2), is similar in width (around 50 cm^{-1}) to the double peaks in feature **C**, which are due to the O8-H9 and O8-H10 stretching motions in **1-1**, in agreement with similar hydrogen bond strengths (~2.10 Å). Band **F** is the broadest predicted and observed band which involves the strongest hydrogen bond (O3…H9) in **1-2**. Consequently, the integrated intensity of the sharp peak **F** in the harmonic spectrum (and also in the 20 K spectrum) of **1-2** is distributed over a much larger energy range and loses relative peak intensity. Similar broadening has been observed for other cluster ions.[22-28] The reduced relative intensity and broadening of the water bending mode (band **G**), relative to the harmonic spectra, is also reproduced by the DTCF spectra.

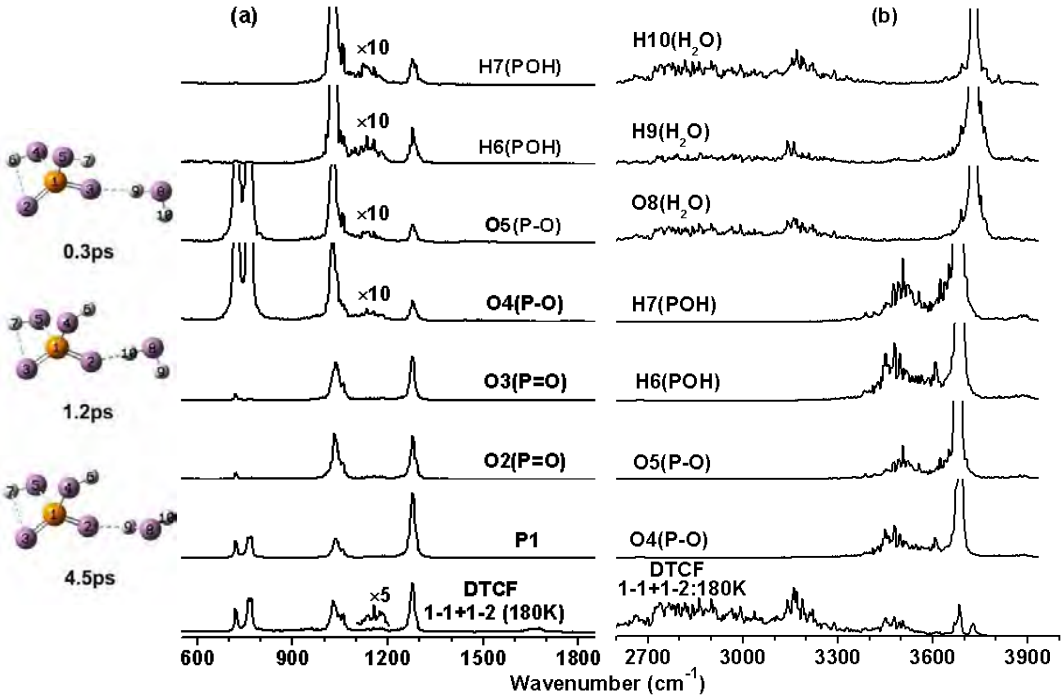

Figure 4.8 DTCF spectra (bottom) and PDOS plots (top) obtained from a 200 ps AIMD trajectory at 180 K for the specified atoms (Fig. 4.2) in $H_2PO_4^-(H_2O)$. Representative structures depicted in the left column are captured at 0.3, 1.2, and 4.5 ps during the AIMD simulation.

Between 1000 and 1300 cm^{-1}, eight IR active P=O stretching and the P-O-H bending modes, three for **1-1** and five for **1-2**, are predicted by the harmonic analysis, while in the experimental spectrum only three bands, two intense (**H** and **I** in Fig. 4.7) and one weaker (**J**) are observed. The DTCF spectra better reproduce this region too. The reason becomes apparent in the partial density of state (PDOS) plots shown in Fig. 4.8. The PDOS plots are based on the Fourier transform of the velocity correlation function for a specific atom. They indicate that there is loss of relative peak intensity due to broadening in the 900-1800 cm^{-1} region too. Interestingly, the broadening is limited to the POH bending bands (H6 and H7 contributions in Fig. 4.8(a)), while the widths of each band involving only the heavier atoms are not much affected in the simulations. Consequently, the two intense bands correspond to the asymmetric (**H**) and symmetric (**I**) P=O stretching modes, while the weaker feature **J** is attributed to the water wagging band.

Below 900 cm^{-1}, there are a number of bands related to wagging and rocking

modes (in harmonic analysis). Since these modes involve H atoms too, they are broadened and smeared out in the simulation above 140 K. Only the asymmetric (**K**) and symmetric (**L**) stretching modes of P-OH are left, which is in very good agreement with the experimental observation of peaks **K** and **L**.

Band **E** is still unassigned, which cannot be attributed to normal modes of either structure **1-1** or **1-2**. The DTCF spectrum at 180 K indeed reproduces a broad feature centered at 3161 cm^{-1}, which is also observed in the 140 K spectrum, but with much low intensity. Inter-conversion between **1-1** and **1-2** involves considerable displacement of the water molecule (as the structures shown in the left column of Fig. 4.8) across a nearly flat potential energy surface with a small barrier less than 7 kJ/mol. At 140 K, the cluster is mainly trapped in the potential well of **1-1** or **1-2** and does not visit **1-3** much. By increasing the simulation temperature to 180 K, it leads to a different situation. The cluster spends considerable more time around **1-3.** The transition region is loosely bound and favored by entropy, which can lead to temporary trapping in the transition state region. As shown in Fig. 4.8(b), the DTCF spectrum shows a prominent peak around 3161 cm^{-1}. PDOS plots show that this peak is indeed due to the O-H stretching of water in **1-3**. This assignment is further supported by MP2 harmonic analysis on **1-3**, which predicts this mode at 3220 cm^{-1} (Fig. 4.7). These results indicate that peak **E** is due to the hydrogen-bonded O-H stretching mode of water in the transient **1-3**.

The experimental peak **D** is left unassigned, too. Is there another conformer responsible for this band? And, is this band reproduced in the DTCF spectra? As discussed before, bond length is the indicator of the hydrogen bond strength, such as the strong HBs in both **1-2** (1.64 Å, 2989 cm^{-1}) and **1-3** (1.74 Å, 3221 cm^{-1}). Consequently, the HB in the contributors of peak **D** should be slightly longer than 1.74 Å in **1-3** and shorter than 2.07 Å in **1-1**. As the vibrational frequencies of hydrogen bonded water O8-H9 at each constraint PO3...H9 distances for **1-2** and **1-3** shown in Fig. 4.9, when the HB distances gradually increase to 1.9 Å, the frequencies approach to experimental value 3382 cm^{-1}. In sum, experimental peak **D** may originate from the strong HB with a new bonding type.

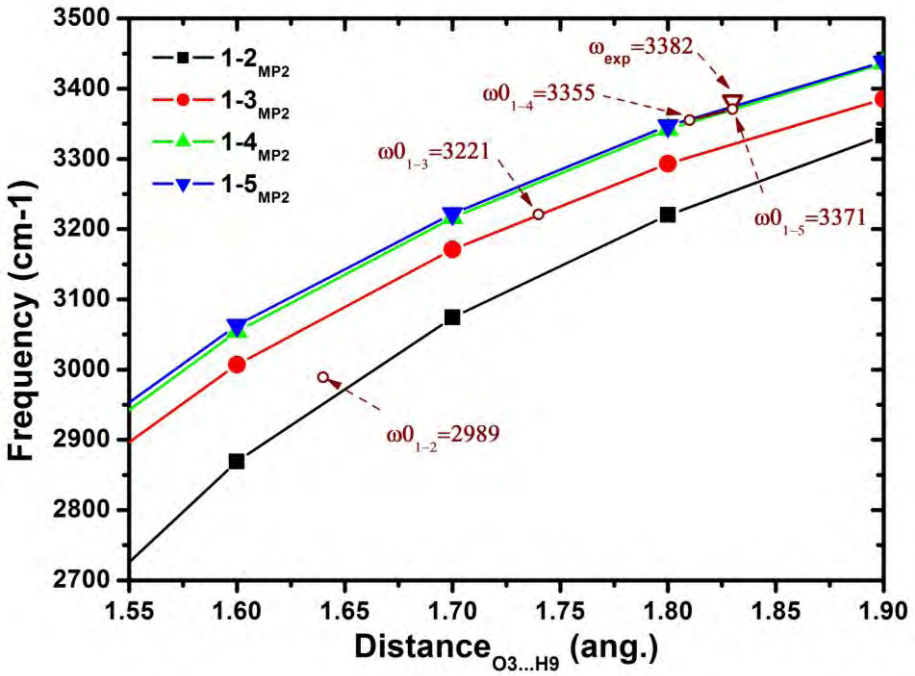

Figure 4.9 Simulated harmonic frequency of HB PO3...H9 for **1-2**, **1-3**, **1-4** and **1-5**, with constraint bond distances, 1.5-1.9 Å at MP2=full/aug-cc-pVDZ level. The wine round circles are the frequencies (ω0) of PO3...H9 for **1-2**, **1-3**, **1-4** and **1-5** in equilibrium structures. A scaling factor of 0.9646 is used.

Interestingly, we indeed found a new bonding type existing only with quite a small population in each trajectory, especially at 180 K. The new isomers were labeled as **1-4** and **1-5**, respectively. As shown in Fig. 4.10, both **1-4** and **1-5** contain a double donor structure (**DD**) water molecule, similar to **1-1**, but with one of the P-OH groups serving as an HB acceptor (PO4...H10). **1-4** is stable at B3LYP/aug-cc-pVDZ in which van der Waals interactions are absent, whereas it changes into **1-2** by breaking the relatively weak HB (P-O4...H10: **1-4**(~2.37 Å)) at MP2 with same basis set, since **1-2** is more energetically favorable (~2.4 kJ/mol lower than **1-4**). The MP2 relative energies and harmonic frequencies for **1-4** were calculated with constraint dihedral O2-P1-O8-O3. **1-5** is stable at both theoretical levels and MP2 places this isomer ~5.68 kJ/mol above **1-1**. The main difference between **1-4** and **1-5** is the orientation of the naked POH. Interestingly, when one H in water acts as an HB donor to P-OH, the strength of the strong HB (O3...H9) is slightly weakened (~1.8 Å) due to the water symmetry constraint (99º).

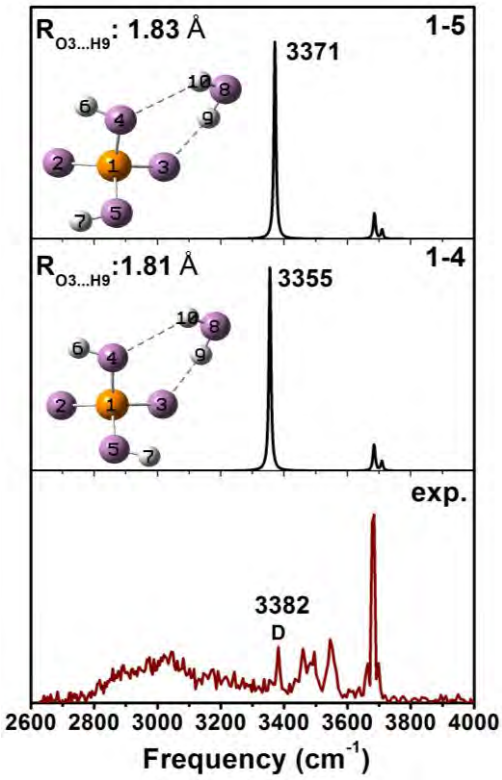

Figure 4.10 Simulated harmonic spectra for **1-4** and **1-5**, together with experimental spectrum for comparison. The relative energies and harmonic spectra are calculated at MP2=full/aug-cc-pVDZ level. A scaling factor 0.9646 is used for theoretical bands.

The MP2 harmonic spectra for **1-4** and **1-5** are shown in Fig. 4.10, together with experimental spectrum for comparison. The strong HB distances ($R_{O3...H9}$) listed in each panel are 1.81 and 1.83 Å, respectively, which are slightly longer than that in **1-3** (1.74 Å). And the corresponding vibrational frequencies of O-H are 3355 and 3371 cm^{-1} (shown both in Fig. 4.9 and 4.10) which are quite close to experimental band **D** (3382 cm^{-1}), indicating that the former hypothesis is correct. Comparison of distributions for P=O...HOH in the dynamics of each isomers to the equilibrium HB distances in MP2 yields satisfactory agreement, as shown in Fig. 4.11, the large distributions of H atoms (in water) are around ~1.6 Å for **1-2**, ~1.7 Å for **1-3**, and ~1.8 Å for **1-4**+**1-5**.

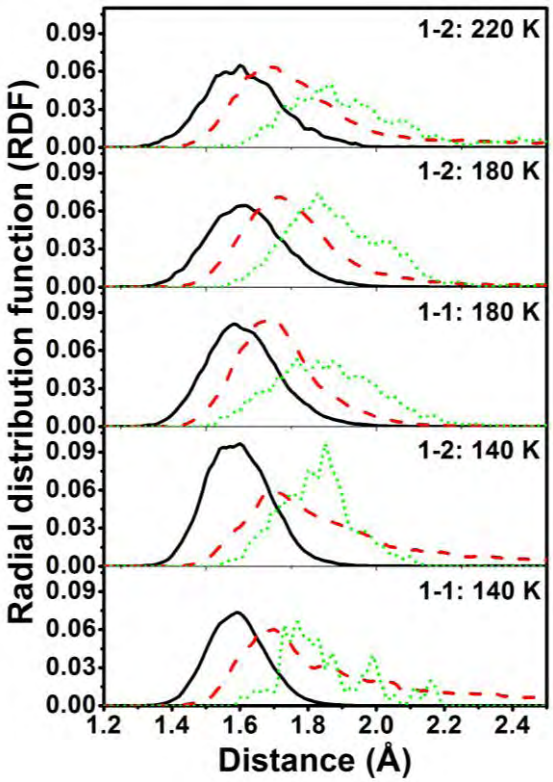

Figure 4.11 Radical distribution function of P=O and H atoms in water for initial conformers **1-1** and **1-2** at 140 and 180 K. The colored lines are the RDF for **1-2**(solid line), **1-3**(dash line), and **1-4**+**1-5**(dot line). The bumpy dotted lines are due to the low isomers populations.

For the second question, is experimental peak **D** reproduced in the DTCF spectra? Before answering this question, we should make clear the relationship between band intensities and isomer populations in DTCF spectrum. As known, in the harmonic approximation, the IR peak intensities of the vibrational modes are proportional to the first order derivative of dipole moments. For DTCF spectrum, the band intensity of each peak is no only related to the bonding strength, but also to the population of the corresponding isomer. Taking peak **E** in DTCF spectrum for example, the population for **1-3** is about 27% at 140 K, whereas at 180 K, this value increases up to 45%, which shows a big band intensity than that at 140 K. However, the population for **1-4**+**1-5** is extremely small in each dynamics, which is only ~1.5% at 140 K and 4.0% at 180 K. Interestingly, as temperature rises, the populations of these two isomers increase slightly, may due to the smoother free energy surfaces at higher temperature. However, the peak intensity produced by **1-4** and **1-5** is so small that it nearly merges

into the numerical nose. But still, experimental peak **D** is reproduced in DTCF spectra. One technique trick should be mentioned again, AIMD at finite temperatures is used to increase the sampling efficiency for simulating the large motion amplitude at cryogenic temperature, which is largely due to the zero point energy and atomic quantum tunneling effects, especially for light H atoms. These atoms are hotter than the relatively heavy atoms, such as O and P. However, all the atoms are treated equally as the hot atoms in AIMD, and the large motion amplitude of heavier atoms may easily destroy the weak hydrogen bond (e.g. O4...H10 in **1-5**) and gives low populations for **1-4** and **1-5** at higher temperatures simulations. This should be the reason of the low peak intensity for **D** in DTCF spectra. Similar situation also happens to the weak HBs in **1-1** (the population is similar to **1-4**/**1-5**, and even smaller at 180 K).

4.4 Conclusion

In summary, the IRMPD spectra of $H_2PO_4^-(H_2O)$ show evidence for isomerization at cryogenic temperatures, which leads to temporary trapping of the clusters in the transition state region and thus a complete breakdown of the harmonic approximation. AIMD simulations allow understanding these effects and qualitatively reproduce the experimental IRMPD spectra. The remaining differences can be attributed to approximations in the simulations, including the limited sampling time, the usage of pseudopotentials and the neglect of nuclear quantum effects. Experimentally, the measured IRMPD intensities deviate from the linear absorption cross sections. The isomerization at low temperatures observed here may be indicative of a highly functional water network around dihydrogen phosphate and therefore it will prove important to also study the larger hydrated clusters.

4.5 Reference

[1] B. Gao and Z.-F. Liu, J. Chem. Phys., **2005**, 123, 224302.
[2] Y. Miller, G.M. Chaban, J. Zhou, K.R. Asmis, and D.M. Neumark, J. Chem. Phys., **2007**, 127, 094305.
[3] J. Zhou, G. Santambrogio, M. Brümmer, D.T. Moore, L. Woste,G. Meijer, D.M. Neumark,and K.R. Asmis, J. Chem. Phys., **2006**, 125, 111102.
[4] M.F. Bush, R.J. Saykally, and E.R. Williams, J. AM. CHEM. SOC. **2007**, 129, 2220.
[5] E. Garand, T. Wende, D.J. Goebbert, R. Bergmann, G. Meijer, D.M. Neumark, and K.R. Asmis, J. AM. CHEM. SOC., **2010**, 132, 849.
[6] T.I. Yacovitch, T. Wende, L. Jiang, N. Heine, G. Meijer, D.M. Neumark, and K.R. Asmis, J. Phys. Chem. Lett., **2011**, 2, 2135.
[7] D.J. Goebbert, E. Garand, T. Wende, R. Bergmann, G. Meijer, K.R. Asmis, and D.M. Neumark, J. Phys. Chem. A, **2009**, 113, 7584.
[8] Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb,M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K.N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.;Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li,X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; 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. Gaussian03, revision D.01; Gaussian, Inc.: Pittsburgh, PA, **2005**.
[9] McLean, A. D.; Chandler, G. S. J. Chem. Phys., **1980**, 72, 5639.
[10] Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. J. Chem. Phys., **1980**, 72, 650.
[11] VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. Comput. Phys. Comm., **2005**, 167, 103.
[12] The CP2K developers group, http://cp2k.berlios.de/.
[13] VandeVondele J. and Hutter J. J. Chem. Phys., **2007**, 127, 114105.
[14] Krack, M. Theor. Chem. Acc., **2005**, 114, 145.
[15] Goedecker, S.; Teter, M.; Hutter, J. Phys. Rev. B., **1996**, 54, 1703.
[16] Hartwigsen, C.; Goedecker, S.; Hutter, J. Phys. Rev. B., **1998**, 58, 3641
[17] Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. J. Chem. Phys., **2010**, 132, 154104.
[18] Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. J. Chem. Phys., **2010**, 132, 154104.

[19] McQuarrie, D. A., Statistic Mechanics; Harper-Collins Publishers; New York, **1976**.
[20] T. I. Yacovitch, T. Wende, L. Jiang, N. Heine, G. Meijer, D. M. Neumark and K. R. Asmis, J. Phys. Chem. Lett., **2011**, 2, 2135-2140.
[21] A. T. Blades, Y. H. Ho and P. Kebarle, J. Am. Chem. Soc., **1996**, 118, 196-201.
[22] R. S. Walters, E. D. Pillai and M. A. Duncan, J. Am. Chem. Soc., **2005**, 127, 16599-16610.
[23] M. F. Bush, R. J. Saykally and E. R. Williams, J. Am. Chem. Soc., **2008**, 130, 15482-15489.
[24] J. D. Rodriguez and J. M. Lisy, J. Phys. Chem. A, **2009**, 113, 6462-6467.
[25] J. D. Rodriguez, T. D. Vaden and J. M. Lisy, J. Am. Chem. Soc., **2009**, 131, 17277-17285.
[26] J. P. Beck and J. M. Lisy, J. Chem. Phys., **2011**, 135, 044302.
[27] J. D. Rodriguez and J. M. Lisy, J. Am. Chem. Soc., **2011**, 133, 11136-11146.
[28] J. T. O'Brien and E. R. Williams, J. Am. Chem. Soc., **2012**, 134, 10228-10236.

4.6 Appendix

Table 4.S1. Assignment for the IR spectra of $H_2PO_4^-(H_2O)_{0,1}$ shown in Fig. 4.1, 4.2 and 4.8.

Label	$H_2PO_4^-$		$H_2PO_4^-(H_2O)$			Assignment
	Exp.	MP2 Harmonic	Exp.	MP2 Harmonic (isomer)	1-1+1-2 (180K)	
A	-	-	3684	3688(**1-1**) 3684(**1-2**)	3684	Free PO-H stretching
B	-	-	3546	3545(**1-2**)	3504	Solvated PO-H stretching
C	-	-	3410-3520	3578(**1-1**) 3531(**1-1**)	3479 3453	O-H stretching for water
D	-	-	3382	3355(**1-4**) 3371(**1-5**)	3380	Solvated O-H stretching of water in **1-4** and **1-5**
E	-	-	3100-3300	-	3080-3280	Solvated O-H stretching of water in transient **1-3**
F	-	-	2700-3100	2989(**1-2**)	2600-3100	Broadened O-H stretching for the strongly hydrogen bonded
G	-	-	1671	1701(**1-1**) 1671(**1-2**)	1674 1644	Water bending
H	1299	1290	1294	1243(**1-1**) 1280(**1-2**)	1277	Asymmetric P=O stretching
I	1094	1070	1099	1079(**1-1**) 1023(**1-2**)	1027	Symmetric P=O stretching
J	1049	1024	1020-1070	966(**1-2**)	~961	P-OH bending
K	770	758	820	787(**1-1**) 786(**1-2**)	760	Asymmetric P-OH stretching
L	-	727	793	744(**1-1**) 740(**1-2**)	720	Symmetric P-OH stretching

Chapter Five
AIMD simulated Infrared spectroscopy of $(HO)_2PO_2^-(H_2O)_n$, with n=2-4, 6, and 8

5.1 Introduction

The IRMPD spectra of $H_2PO_4^-(H_2O)$ studied in the former chapter show evidence for isomerization already at cryogenic temperatures, which leads to temporary trapping of the clusters in the transition state region and consequently a complete breakdown of the harmonic approximation. AIMD simulations qualitatively reproduce the experimental IRMPD spectrum. The remaining differences can be attributed to approximations in the simulations, including the limited sampling time, the use of pseudopotentials and the neglect of nuclear quantum effects. The isomerization at low temperatures observed for $H_2PO_4^-(H_2O)$ may be indicative of a highly functional water network around dihydrogen phosphate and therefore it will be important to larger hydrated clusters. In this chapter both harmonic and AIMD methods are employed to study the anharmonic features of $H_2PO_4^-(H_2O)_{n=2\text{-}4,6,\text{ and }8}$. Again, the experiments were performed by our collaborators Ling Jiang (State Key Laboratory of Molecular Reaction Dynamics Dalian Institute of Chemical Physics, Chinese Academy of Sciences) and Knut R. Asmis (Fritz-Haber-Institut der Max-Planck-Gesellschaft, Berlin, Germany).

5.2 Methods

Optimized structures and harmonic frequencies are obtained from standard density functional theory (DFT) calculations, using the Gaussian 03 package.[1] The dynamic motion of the clusters are simulated by the Ab Initio Molecular Dynamics (AIMD) method, in which the atoms are treated as classical particles and the potential energy and forces on the atoms are calculated within the framework of DFT at each time step.

For 0 K structure, energy optimization is performed at the level of B3LYP/6-311++G(d,p)[2,3] by Gaussian 03 package.[7] Initial structures are generated by running molecular dynamic simulations over tens of thousands time steps at 200 K and taking random configurations along the trajectories. Harmonic frequencies are calculated by using a larger basis set, at the level of B3LYP/aug-cc-pVDZ, and a scaling factor of 1.0022 for the low frequency region, and 0.9698 for the high frequency region.[4]

CP2K package [5,6] is employed for the AIMD simulations. The wave functions are expanded in a double zeta Gaussian basis set,[7] while the electron density is expanded in Gaussians and auxiliary plane waves with an electronic density cutoff at 320 Rydberg for the electron density. The atomic cores are modeled by the Goedecker-Teter-Hutter (GTH) type pseudopotentials.[8-10] The exchange and correlation energy is calculated by BLYP functional, with additional Grimme's dispersion correction at D3 level.[11] And the dispersion correction (from external code [11]) is also used for the relative energies corrections listed in following figures.

A cluster ion is put at the center of a periodic cubic box, and the effects of the periodic charge density images are corrected by the decoupling technique developed by Martyna and Tuckerman.[12] The box length is 16 Å for $(HO)_2PO_2^-(H_2O)_{2\text{-}4}$ and 20 Å for the larger $(HO)_2PO_2^-(H_2O)_{6,8}$. The convergence criterion for the SCF electronic procedure is set to 10^{-7} a.u. at each time step. For molecular dynamics at a specific temperature, the temperature is controlled by a Nose-Hoover thermostat, with a time step of 0.5 fs. An equilibration period of up to 6 ps is performed first, with the temperature scaled to an interval of 20 K around the intended value. A data collection run is then followed 100 ps in the NVE ensemble.

Hydrated clusters are bound by hydrogen bonds, which are relatively weak and therefore fairly flexible at finite temperature. Dynamic simulations are essential for sampling the solvation structures and for examining the thermal stability of a particular structure. More importantly, the hydrogen bonds could have strong effects on the vibrations, which could be captured by the AIMD simulations. A vibrational spectrum can be directly simulated by the Fourier transformation of the dipole

time-correlation function (DTCF),[13] with the dipole moment again calculated directly from first principles within the Berry phase scheme.

The vibrational density of state (VDOS) could also be easily obtained by the Fourier transform of the velocity time correlation function.[13] *VDOS* is not directly proportional to the signal intensity measured in a vibrational spectrum, which is determined by the dipole selection rule. But when the summation is restricted to one single atom, the total *VDOS* could then be decomposed into the contribution from this particular atom, which is helpful for the assignment of a vibrational spectrum.

5.3 Results

5.3.1 $(HO)_2PO_2^-(H_2O)_2$: structure, dynamics and labeling

The isomers for n>1 clusters differ from each other in how the P=O and PO-H groups are solvated, given that the number of water molecules is limited. The solvation of the two PO-H groups provides a convenient way to characterize the structures of the larger clusters with more than one water molecule. In the following discussion, the symbol ***(L1, L2)*** will be used for this purpose, with ***L1*** and ***L2*** as two labels indicating the interaction between each of the two PO-H groups and water molecules, respectively.

For n=2, four isomers are identified, as shown in Fig. 5.1, and the energy differences between these isomers are within 2.1 kcal/mol. **2-1** is the lowest-energy structure for n=2 with only one PO-H solvated, similar to **2-3**. However, the water bounded to one PO-H in **2-1** forms an additional hydrogen bond with the other water molecule (H12…O8). To characterize this feature, we shall label **2-1** as **(0, 1D+mW)**, to indicate that this hydrogen bond is further linked to other water molecules. As the cluster size grows and more water molecules are around, such "**1D+mW**" type would also become more common. In **2-2**, each of the two water molecule solvates one PO-H group, and the structure should be labeled as **(1D+1W, 1D+1W)**.

2-3 is labeled as **(0, 1D+1W)**. The label "**0**" indicates that one of the PO-H groups is free of solvation. The label "**1D+1W**" indicates that the other PO-H group is solvated by only one water molecule by forming a donor (**D**) hydrogen bond

(POH7…O11). This same water molecule also forms another hydrogen bond with a P=O group (P=O3...H12), which is remarkably short at a distance of only 1.69 Å, indicating strong interaction. It's a common structural motif which also presents in **2-1**, **2-2** and other structures discussed below. The POH group could also form an acceptor hydrogen bond on the oxygen atom, which will be labeled as "**A**", but it's weaker and less common than the donor bonds. By such a scheme, **2-4** is labeled as **(0, 0)**, with both PO-H groups naked. The two water molecules form a chain to solvate the two P=O groups. In AIMD simulation, **2-4(0, 0)** is unstable even at 20 K and it transforms into **2-1**.

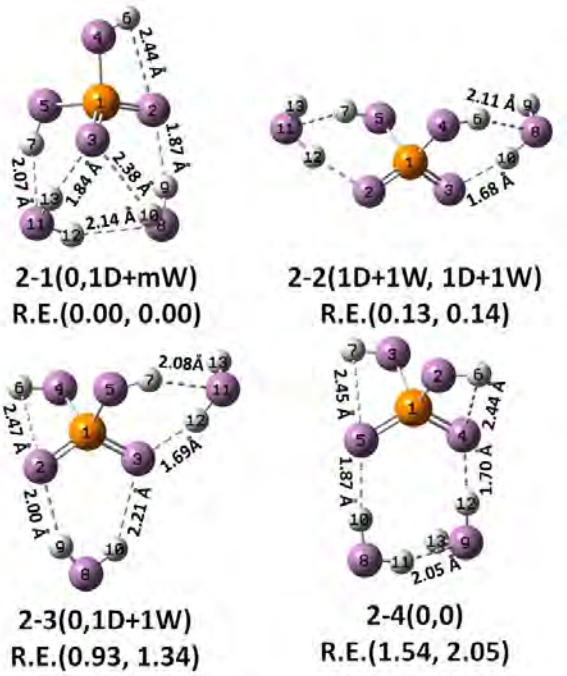

Figure 5.1 Optimized structure for n=2. The relative energies are listed in the brackets in kcal/mol. The left value is calculated at B3LYP/aug-cc-pVDZ level with zero point energy and Grimme's D3 dispersion corrections, while the right value is calculated enthalpies at 298.15 K at the same level.

In AIMD simulations, **2-1(0, 1D+mW)** could be easily transformed into other structures: breaking of H12…O8 leads it to **2-3(0, 1D+1W)**, breaking of H7…O11 and O3...H10 leads it to **2-4(0, 0)**, and forming H6...O8 to **2-2**. It's stable at 20 and 60 K (which cannot be achieved in experiments due to the atomic zero-point vibrations as discussed for n=1 in the former chapter), but at 140 K and above, the structures flip

back and forth between the four isomers. Consequently, all conformers should make contributions to the observed experimental spectrum.

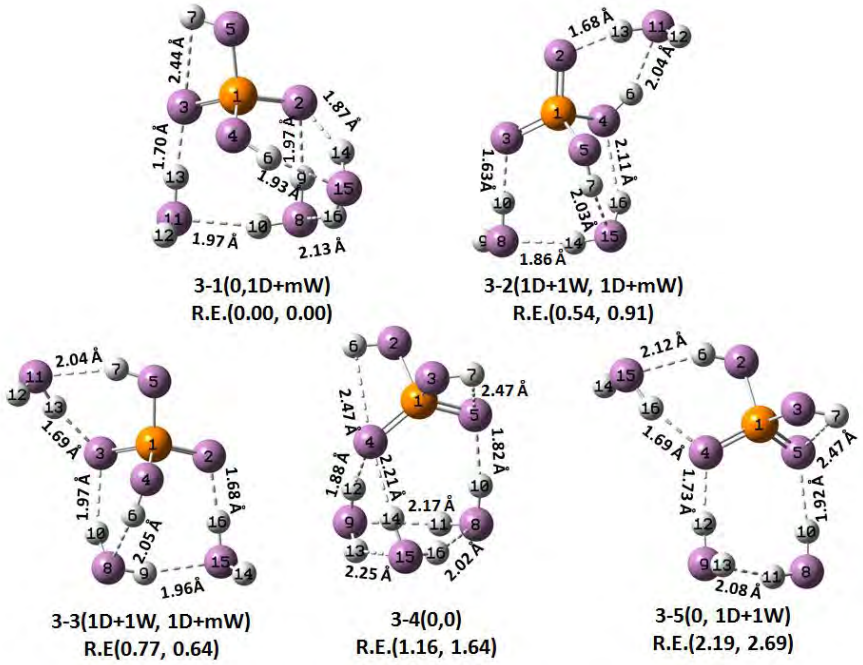

Figure 5.2 Optimized structure for n=3. The relative energies are listed in the brackets in kcal/mol. The left value is calculated at B3LYP/aug-cc-pVDZ level with zero point energy and Grimme's D3 dispersion corrections, while the right value is calculated enthalpies at 298.15 K at the same level.

5.3.2 $(HO)_2PO_2^-(H_2O)_3$: structure and dynamics

By the same scheme, **3-1** is labeled as **(0, 1D+mW)** and **(1D+1W,1D+mW)** for **3-2** in Fig. 5.2. **3-4** is labeled as **(0, 0)**, with both PO-H groups free from solvation. All the isomers except **3-3** are selected as the initial structures in the molecular dynamic simulations, since its solvation type is similar to **3-2**. The three water molecules form a ring structure in **3-4**, to solvate the two P=O groups. Similar to **2-4(0, 0)**, **3-4(0, 0)** is unstable upon heating and transformed into other structures in AIMD simulations at 140 K and above with one of the PO-H groups became naked. In other words, the $(H_2O)_3$ ring in **3-4(0, 0)** is not stable upon heating. The last structure **3-5** is labeled as **(0, 1D+1W)**. In AIMD simulations, **3-2(1D+1W, 1D+mW)** can be maintained at 140 K, whereas at 180 K, it is transformed into **3-1** and **3-4** with one or two POH groups naked. Both **3-1(0, 1D+mW)** and **3-5(0, 1D+1W)** are stable

at 140 K. But by 180 K, they can change into **3-(0, 1D+mW)**. Despite of large populations of **(0, 1D+1W)** and **(0, 1D+mW)** types at 180 K in the four dynamics, large motion amplitude of water molecules are also observed.

5.3.3 $(HO)_2PO_2^-(H_2O)_4$: structure and dynamics

A **(0, 1D+mW)** structure **4-1** is also found by energy minimization, as shown in Fig. 5.3, which is stable up to 140 K. Even at 180 K, **4-1(0, 1D+mW)** type can still be maintained with one of PO-H groups remained free, although there are reshufflings in the water molecules of the **1D+mW** feature. **4-2** in Fig. 5.3 is again a structure in which both PO-H groups are naked, and the four water molecules form a ring to solvate the two P=O groups. With more water molecules available, interactions between solvent molecules become more important. By forming a four water ring, with each water provided one HB and each P=O accepted two HBs, **4-2(0, 0)**, can be maintained at 140 K. At 180 K, **4-2(0, 0)** is transformed into **4-1(0, 1D+mW)**, when one PO-H group flips its orientation and is solvated by the water molecules.

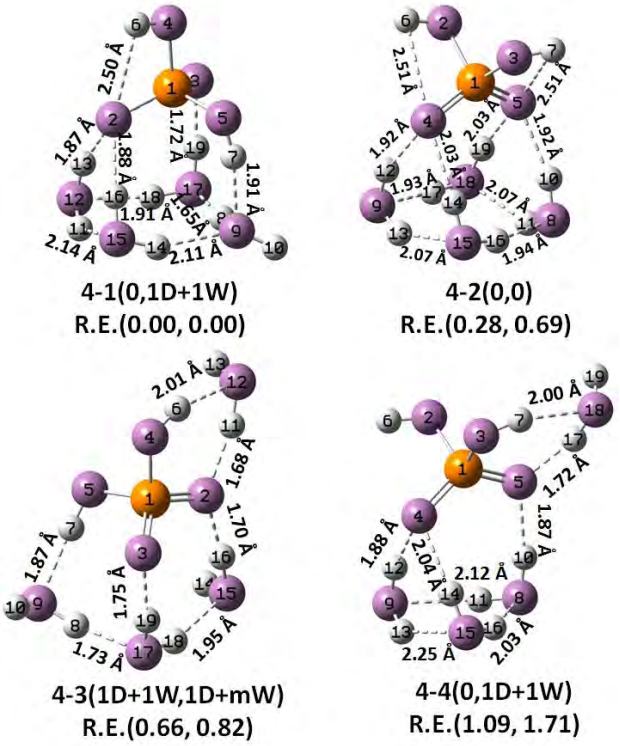

Figure 5.3 Optimized structure for n=4. The relative energies are listed in the brackets in kcal/mol. The left value is calculated at B3LYP/aug-cc-pVDZ level with zero point energy and Grimme's D3 dispersion corrections, while the right value is calculated enthalpies at 298.15 K at the same level.

4-3 belongs to **(1D+1W, 1D+mW)** type, with both PO-H groups being solvated. Albeit some differences in water arrangement, similar to **3-3(1D+1W, 1D+mW)**, **4-3(1D+1W, 1D+mW)** is less stable, with one of the PO-H group being free from solvation **(0, 1D+mW)** in most time of the AIMD simulations at 140 and180 K. **4-4** is also a structure with only one PO-H being solvated, but it is a **(0, 1D+1W)** structure. This structure is not robust either and it will be converted into **(0, 1D+mW)** with an HB being formed between O18 and the left water molecules. For both n=3 and 4, the dominant features observed in MD are **(0, 1D+mW)**, although more water molecules are now available.

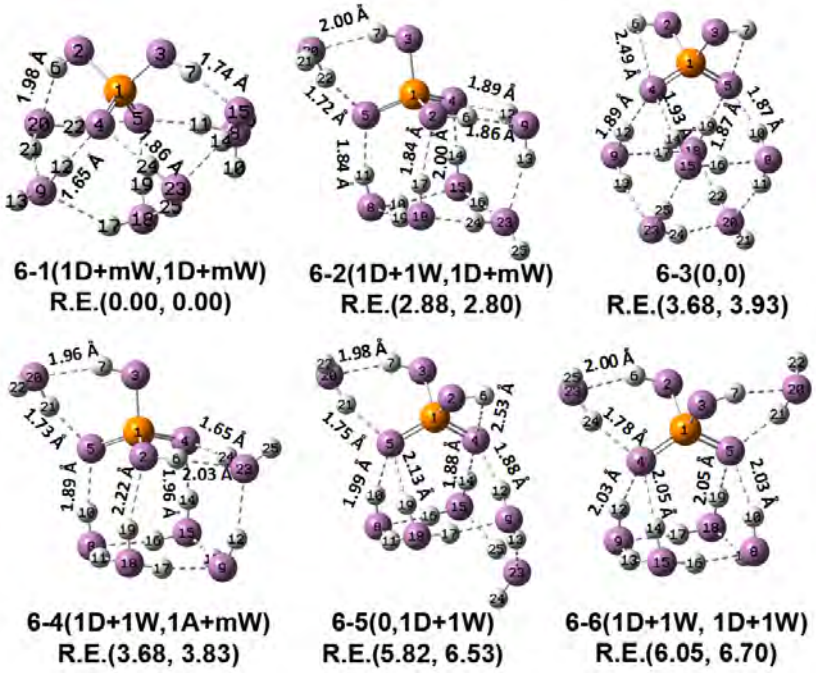

Figure 5.4 Optimized structure for n=6. The relative energies are listed in the brackets in kcal/mol. The left value is calculated at B3LYP/aug-cc-pVDZ level with zero point energy and Grimme's D3 dispersion corrections, while the right value is calculated enthalpies at 298.15 K at the same level.

5.3.4 $(HO)_2PO_2^-(H_2O)_6$: structure and dynamics

More isomers are found at n=6, as shown in Fig. 5.4, due to the increasing number of water molecules. The dynamics were performed by selecting **6-2**, **6-3**, **6-5**, and **6-6** as the initial structures. **6-1** (~2.0 kcal/mol lower in energy than **6-2**) is extracted from the dynamics of **6-3** at 180 K. **6-2** and **6-4** belong to **6-(1D+1W,1D+mW)**, with one of PO-H groups solvated by one water molecule, while the other is solvated together

with the two P=O groups by the left five water molecules. With increasing number of water molecules, there are now considerable variations in the **1D+mW** feature, which is double layered. In **6-2**, the lower layer is a $(H_2O)_4$ ring, with each water provides a hydrogen atom, pointing up as hydrogen bond donor, similar to the ring observed in **4-2(0,0)**. In the upper layer, one water molecule (H_2O20) actually forms a **1D+1W** feature, and together with the two P=O groups, these four oxygen atoms (O2, O4, O5, O9) form the upper layer, with each accepting a hydrogen bond from the $(H_2O)_4$ ring. This solvation type is stable at both 140 and 180 K in AIMD simulations, although there are rearrangements of the HB in water at high temperature.

6-3(0, 0) is stable at 140 K. In this symmetric structure, the six water molecules are stacked into two layers, with two at the bottom and four at the top to bond the two P=O groups. Upon heating, the two PO-H groups could bend to one side to form a donor hydrogen bond with one of the water molecules for a short period. But the bond distance is larger than 2 Å, and is not enough to disrupt the **6-3(0, 0)** structure. However, upon heating to 180 K, **6-3(0, 0)** will change into **(1D+mW, 1D+mW)** by destroying the water network. Interestingly, for **6-3** at high temperature, the six water molecules tend to aggregating on both PO-H groups separately, such as the water in **6-1** observed at 180 K.

In **6-4**, the upper water molecule (O23) flips its orientation and donates a hydrogen bond to the lower layer. Accordingly, the $(H_2O)_4$ ring also has one water molecule (H_2O9) with its orientation flipped and now donates three and accepts one hydrogen bonds. **6-5** belongs to **6-(0,1D+1W)** type, but with one water molecule in the second solvation shell (O23). **6-6** is **(1D+1W,1D+1W)** type, similar to **2-2(1D+1W,1D+1W)**, with each POH solvated by one water molecule and the left four water molecules form a ring around the two P=O groups, similar to **4-2(0,0)**. **6-5** is unstable even at 140 K, with the naked PO-H group solvated by the other water molecules and the structure transforms into **6-(1D+1W, 1D+mW)** and **6-(1D+mW, 1D+mW)** at 180 K. Similar to **6-5**, **6-6** also is not stable upon heating, and it changes into **6-(1D+1W, 1D+mW)** but with one PO-H occasionally became naked.

Overall, three solvation types are identified for n=6. The first is **6-(1D+1W, 1D+mW)**, this type appears frequently in most of the dynamics for n=6. **6-2**, **6-4**, and **6-6** belong to this type. The second is **6-(0, 1D+mW)**, with one free PO-H group, the other PO-H group is solvated together with the two P=O groups, which is only observed occasionally in the dynamics of **6-6**. In the final type, **6-(1D+mW, 1D+mW)**, two PO-H groups are solvated by more than one water (such as **6-1**), which is the dominant structure observed at 180 K in the molecular dynamics of **6-3**.

Except for **6-1** and **6-6**, in the other structures, at least one water is observed in the second solvation shell (O23 in **6-2**, O20 and O23 in **6-3**, O9 in **6-4**, and O23 in **6-5**). For n=6, structures with both of the PO-H groups solvated have a significant presence.

5.3.5 $(HO)_2PO_2^-(H_2O)_8$: structure and dynamics

The structure **6-1(1D+mW, 1D+mW)** for n=6 provides a good basis for understanding two isomers **8-1(1D+mW, 1D+mW)** and **8-2(1D+mW, 1D+mW)** for n=8, as shown in Fig. 5.5. Both of the PO-H groups are solvated in **8-1** and **8-2**, similar to **6-1**, the eight water molecules solvate the two PO-H groups equally. In **8-2** one of the PO-H groups can also act as an HB acceptor. These two conformers are extracted from the dynamics with initial isomers of **8-3** and **8-4** at 180 K, respectively.

The two layers structure in **8-3** consists a $(H_2O)_4$+$(H_2O)_2$ stack to solvate the two P=O groups, is actually the same as that in **6-3(0, 0)**. Each of the two additional water molecules solvates one POH group as a **1D+1W** feature. Similar to **6-6(1D+1W, 1D+1W)**, **8-1(1D+1W, 1D+1W)** is unstable in AIMD simulations at 140 K, and it is transformed into **8-(0, 1D+mW)** or **8-(1D+1W, 1D+mW)**. At a higher temperature 180 K, it will change into **8-1(1D+mW, 1D+mW)**.

8-5 belongs to **(1D+1W, 1D+mW)** type, similar to **8-4**, this conformer is unstable at 140 and 180 K, and it will be converted into **8-(1D+mW, 1D+mW)**, which differs from **8-1** and **8-2** in water molecules arrangement. In **8-6(0, 0)**, the two more water molecules are added to the bottom layer to produced two stacked $(H_2O)_4$ rings. The two P=O groups are solvated, but the two P-OH groups are free. Such an **8-6(0, 0)**

structure is stable in AIMD simulations at 140 K, with the two P-OH groups bending side-wise and forming a weak hydrogen bond with water molecule for a short period of time. However, upon heating to 180 K, **8-6(0, 0)** will change into **8-(1D+mW, 1D+mW)** by destroying the water stack.

In sum, for n=2 and 3, the dominant solvation type for the two PO-H groups observed in MD, especially at 180 K, are **(0, 0)** and **(0, 1D+mW)**, although there are instantaneous formation of **(1D+1W, 1D+1W)** or **(1D+1W, 1D+mW)** for n=3. With only two or three water molecules binding to one PO-H and two P=O groups, large motion amplitude of water molecules still can be observed. For n=4, none of the initial structures can be maintained in MD at high temperature. However, with one water added, the solvation type for n=4 appears like **(0, 1D+mW)**, which is very dominant solvation in all the four dynamics. In **4-1(0, 1D+mW)**, with an HB network being formed between water molecules, double P=O and one PO-H, the motion of this PO-H can be stabilized at high temperature, despite of HBs arrangement in water. For n=6 and 8, two dominant solvation type **(1D+1W, 1D+mW)** and **(1D+mW, 1D+mW)** are observed in the MD simulations, and the last type is more robust and popular. Interestingly, from n=6, the two PO-H groups are solvated evenly at high temperature, such as **6-1**, **8-1**, and **8-2**. As the cluster size increases from 2 to 8, PO-H groups show a different binding motif which will be helpful for the analyzing IRMPD spectra.

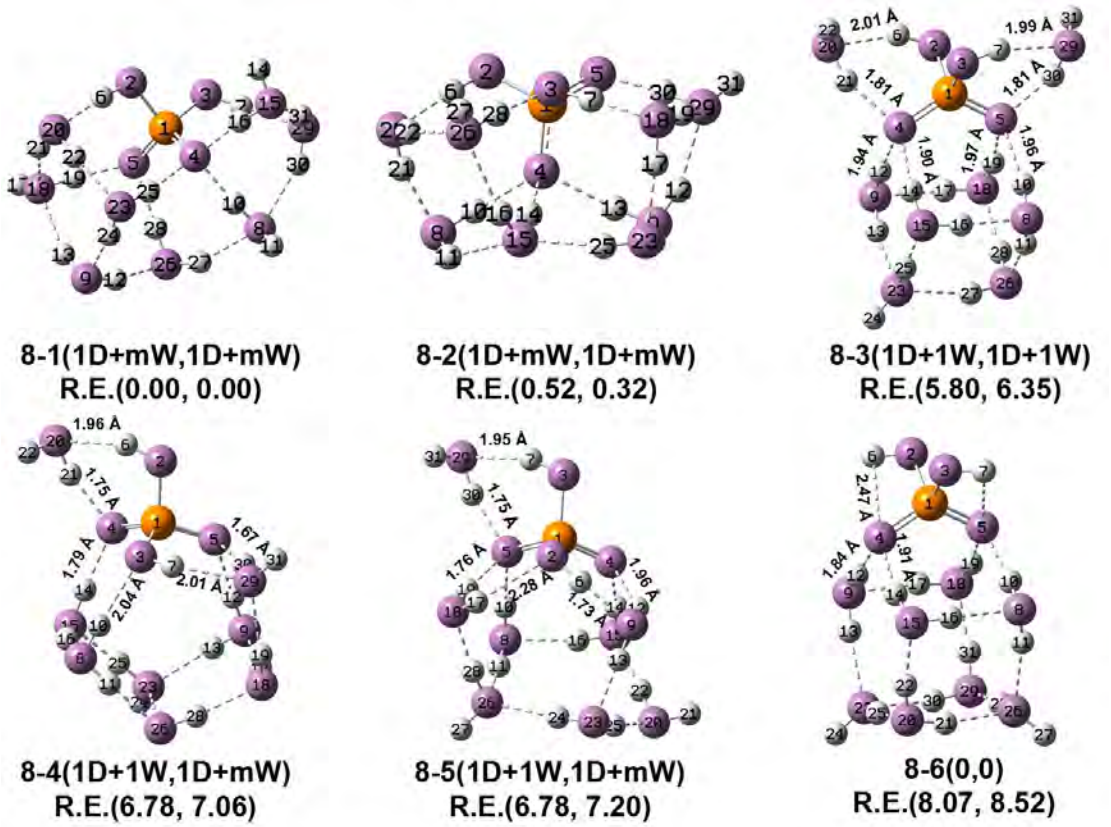

Figure 5.5 Optimized structure for n=8. The relative energies are listed in the brackets in kcal/mol. The left value is calculated at B3LYP/aug-cc-pVDZ level with zero point energy and Grimme's D3 dispersion corrections, while the right value is calculated enthalpies at 298.15 K at the same level.

5.3.6 Vibrational spectra in high frequency region for n=2-4

As the cluster size grows, there are more water molecules and more variety of the O-H stretching modes. For the case of n=2, it is still possible to categorize and arrange these O-H stretching modes, according to the strength of HBs, as shown in the Fig. 5.6, with the related peaks positions listed in Table 5.1.

There are two types of hydrogen atoms that do not participate in hydrogen bonding, either in a free PO-H group, such as H6 in **2-3**, and H6 and H7 in **2-4**, or in a water molecule, H13 in **2-4**, and H13 and H9 in **2-2**. Their O-H stretching frequencies are the highest, around 3700 cm^{-1}.The majority of the hydrogen atoms are involved in hydrogen bonds. For the PO-H group, the donor hydrogen bond is close to 2.1 Å in length, which shifts the O-H frequency to the red by almost 200 cm^{-1}. For this mode, there is no significant difference between the **1D+1W** and the **1D+mW** types, and it is

labeled as "**B**" in Fig. 5.6 (comparing peak **B** in **2-1** and **2-2**). There are many variations in the hydrogen bonds for the water molecules among the four identified isomers, with distances ranging from 2.1 Å of a typical strength to 1.7 Å for a very strong hydrogen bond. The **1D+1W** features in **2-2** produce the most shifted O-H peaks to the red. And all the other O-H peaks fall between it and the peaks for free O-H.

The finite temperature DTCF spectra are shown in Fig. 5.7 for n=2-4. Four trajectories at 180 K of the clusters were cut into 10 ps interval for Fourier transform, and totally 40 frequency profiles were added up to produce the DTCF spectra in the high frequency region. Although the DTCF spectra are too noisy, they do provide important clues for the assignment of the experimental spectra. The two free O-H peaks remain sharp at 180 K. The one highest frequency peak in the experimental spectrum, labeled as **A**, is indeed sharp. They must be due to the stretching mode of free PO-H. The O-H peaks involved in the strongest hydrogen bonds in the **1D+1W** feature produce the broad band **F**, just as in the case of n=1. The other OH peaks involved in hydrogen bonds are also broadened, although not as much as **F**. In fact, the O-H stretching for the solvated PO-H group in the **1D+1W** is still visible as peak **B**, although it certainly has overlaps with other broadened O-H features of water. And the other hydrogen bonds make their contributions to the broad band between **B** and **F**. However, the simulated DTCF spectra are too noisy and prevent additional assignments of these peaks in IRMPD spectra. But understandably, the relative intensity of **F** versus **A** increases with the cluster size.

Table 5.1 Calculated harmonic frequencies for O-H stretching in the four isomers of $H_2PO_4^-(H_2O)_2$.

Frequencies (cm^{-1})	Structure	O-H bond	Hbond	Hbond distance (Å)
3799	2-2	O8-H9	-	-
		O11-H13		
3737	2-1	O8-H10	-	-
3743	2-4	O9-H13	-	-
3740	2-3	O11-H13	-	-
3711	2-3	O4-H6	-	-
3709	2-1	O4-H6	-	-
3707	2-4	O2-H6	-	-
		O3-H7		
3636	2-3	O8-H10	OH10...O3	2.25
3582	2-4	O8-H11	OH11...O9	2.06
3536	2-2	O4-H6	OH6...O8	2.09
		O5-H7	OH7...O11	2.09
3528	2-1	O5-H7	OH7...O11	2.11
		O11-H12	OH12...O8	2.03
3519	2-3	O5-H7	OH7...O11	2.05
3512	2-1	O5-H7	OH7...O11	2.11
		O11-H12	OH12...O8	2.03
3507	2-3	O8-H9	OH9...O2	2.00
3459	2-4	O8-H10	OH10...O5	1.93
3363	2-1	O11-H13	OH13...O3	1.89
3139	2-4	O9-H12	OH12...O4	1.70
3043	2-1	O8-H9	OH9...O2	1.69
3012	2-3	O11-H12	OH12...O3	1.67
2997	2-2	O8-H10	OH10...O3	1.66
		O11-H12	OH12...O2	1.66
2975	2-2	O8-H10	OH10...O3	1.66
		O11-H12	OH12...O3	1.66

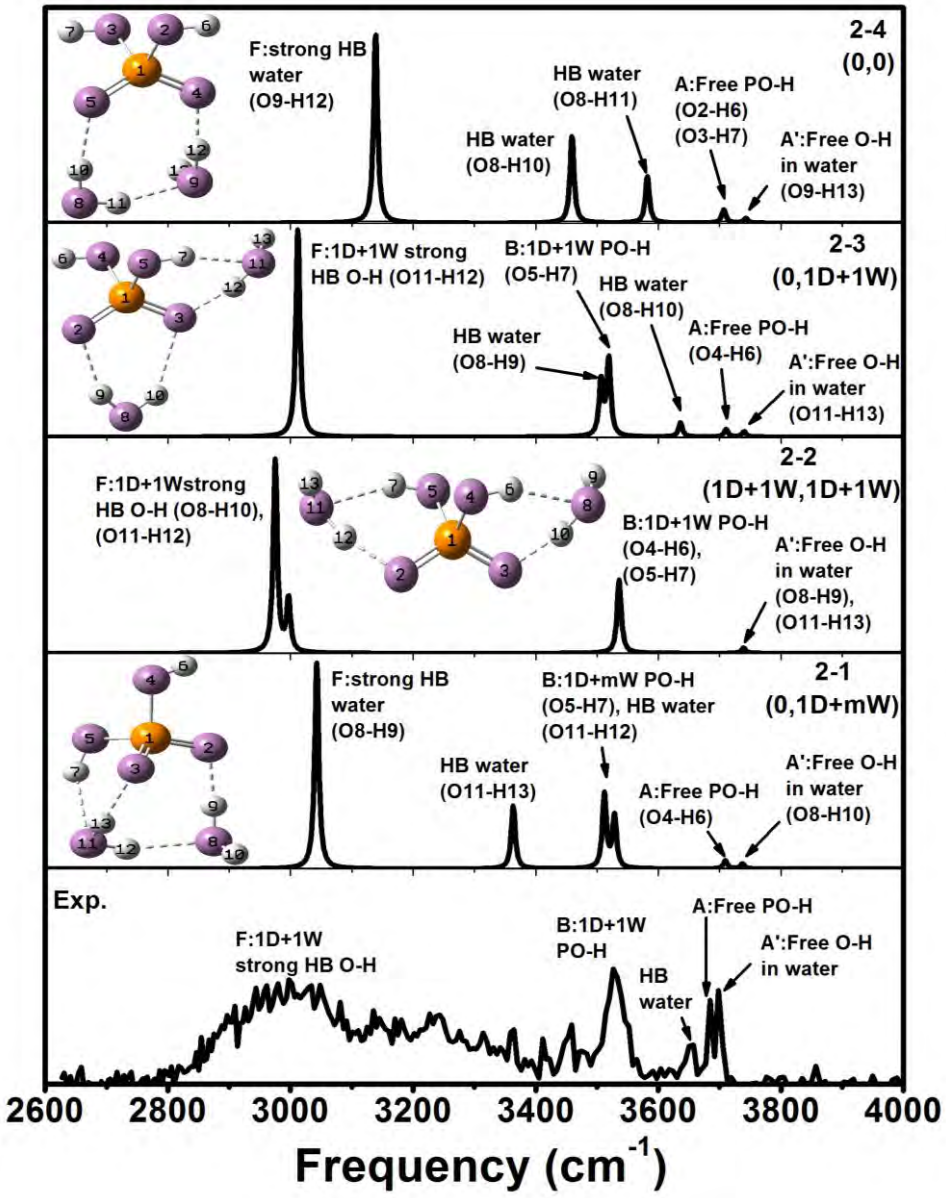

Figure 5.6 Calculated harmonic O-H stretching frequencies for n=2 at B3LYP/aug-cc-pVDZ level with a scaling factor of 1.0022 to experimental spectrum.

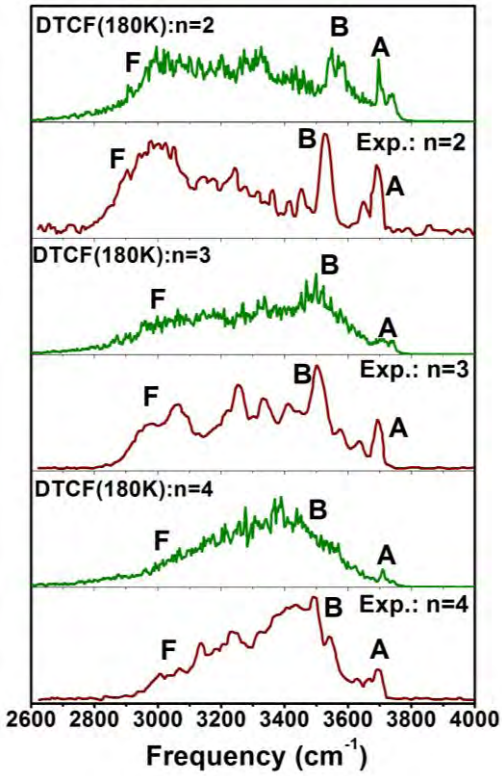

Figure 5.7 DTCF spectra for n=2−4 at 180 K together with experimental spectra. No scaling factor is used for DTCF spectra.

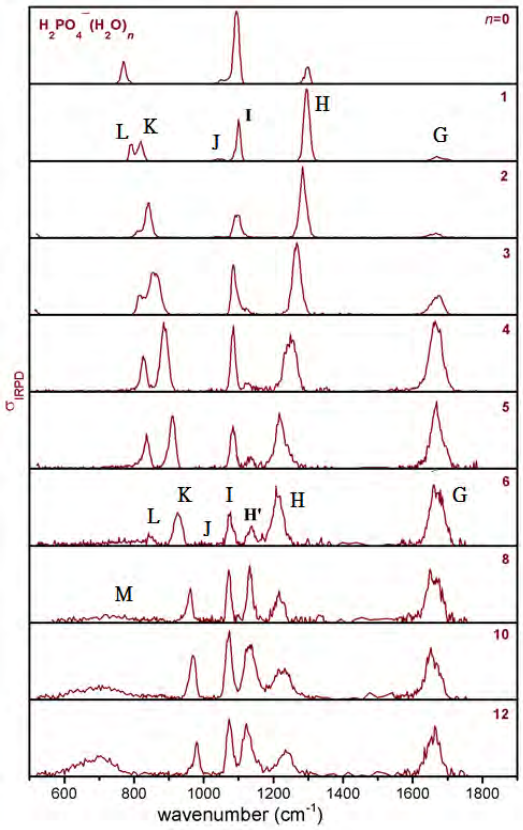

Figure 5.8 Experimental spectra in the low frequency region for n=0-6, 8, 10, and 12.

5.3.7 Harmonic vibrational spectra in low frequency region

The experimental spectra in Fig. 5.8 for the low frequency region show a number of well-defined features. Peak **G** is obviously due to the scissoring mode of water. It's barely visible at n=1, as it is flattened by the large amplitude motion water molecule. Its intensity increases with n, simply because more water molecules are contributing to this band.

The other features are due to P=O stretching coupled with P-O-H bending (**H**, **I**), and P-OH stretching mixed with wagging and rocking of water molecules (**L**, **M**), based on understanding of the spectrum for n=1. There are several clear trends as the cluster size grows. While peak **H** is gradually red shifted, peaks **L** and **K** are gradually blue shifted, and peak **I** more or less stay at the same position. Beyond n=6, these features are stabilized and do not change much with the cluster size. At each n, there are contributions from many different isomers, and their vibrational profiles are sensitive to temperature changes. In the IRMPD process, it is anharmonic profile due to the large amplitude motion of hydrogen bonded water molecule, rather than the harmonic spectrum.

In the following, we shall base the analysis on a systematic examination of the harmonic spectra. The starting point is to divide the isomer structures according to how the two PO-H groups are solvated, which has already been shown as a good indicator to understand the evolution of the cluster structure. At low n, the solvation of PO-H groups has lower priority to the solvation of P=O groups and to solvent-solvent interaction. As n increase, PO-H groups become more solvated. And by n=8, both PO-H groups are solvated by at least one water molecule in most of the optimized isomer and during the AIMD simulations. The attention will be first focused on the P=O stretching and P-O-H bending region, between 1000 and 1300 cm^{-1}.

The first type are the **(0, 0)** structures, with both PO-H groups free from solvation. Their harmonic spectra are shown in Fig. 5.10, and the P-O-H bending and P=O stretching related features actually do not change much with the cluster size. The four

vibration modes are arranged as ***a1, s1, a2***, ***s2***, in the order of decreasing frequency. A large component of P=O stretching is found in ***a1*** and ***s2***, while the two middle peak ***s1*** and ***a2*** are very close to each other, producing three peaks between 1000-1300 cm^{-1}. The same situation applies to the spectra of all the other five **(0, 0)** structures in Fig. 5.10. The only noticeable change in this frequency region is a few extra small peaks which are due to the wagging and rocking vibration in the water rings for n=6 and 8. These water related modes also show up in the region below 1000 cm^{-1}, and mix quite extensively with the P-OH stretching modes. Upon the first glance, satisfactory agreement between experiment and harmonic spectra is observed for **(0, 0)** with n=2, 3, and 4, in particular for 1000-1300 cm^{-1}. However, the **(0, 0)** binding motif is unstable upon heating.

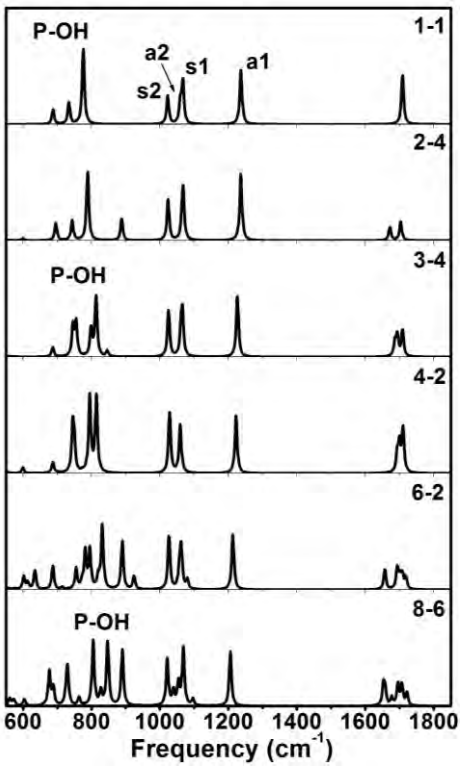

Figure 5.10 Calculated harmonic spectra for **(0, 0)** for n=1-4, 6, and 8.

The next types are the **(0, 1D+1W)**, with only one PO-H group solvated by a water molecule. For the smallest cluster, **1-2** in this series, its harmonic spectrum has been presented in the former chapter. There is the same "***a1-s1-a2-s2***" pattern, but the four peaks are well separated, as shown in Fig. 5.11. Peak ***s2*** remains the symmetric

P=O stretching, while the wagging of the water in the **1D+1W** feature is mixed into these vibration modes. As the number of water molecule, the basic pattern in this region again changes little. Significant changes are only seen in the region below 1000 cm^{-1}, which is due to the wagging and rocking of water mixed extensively with P-OH stretching modes.

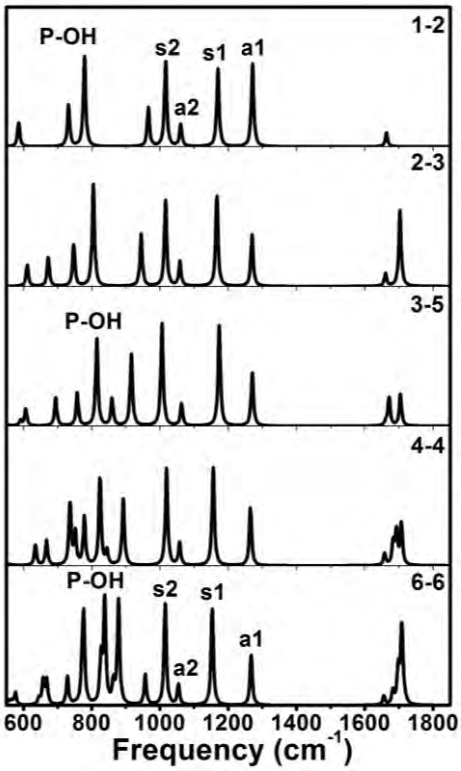

Figure 5.11 Calculated harmonic spectra for **(0, 1D+1W)** for n=1-4, 6, and 8.

The next level of solvation is **(1D+1W, 1D+1W)**, with both PO-H groups solvated by one water molecule, for which there are only three structures. Their harmonic spectra are shown in Fig. 5.12. The same "***a1-s1-a2-s2***" pattern is maintained, but there are actually important changes in the vibration modes. For the previous two types of structures, ***a1*** is due largely to asymmetric P=O stretching. For **(1D+1W, 1D+1W)**, ***a2*** replaces ***a1*** in that role. With both P-OH groups solvated, the P-O-H bending modes couple with the wagging of the solvent water, and are now dominating the two high frequency peaks, ***a1*** and ***s1***. Only peak ***s2*** remains the same symmetric P=O stretching.

Going further to **(1D+1W, 1D+mW)** shown in Fig. 5.13, when one of the **1D+1W** features is solvated by additional water molecules, the harmonic spectra changes a little, and the corresponding vibrational modes are the same. **(1D+mW, 1D+mW)** is another solvation type as shown in Fig. 5.14, **6-1**, **8-2** and **8-3** are of this type. Similar to **(1D+1W, 1D+mW)**, with the two PO-H groups solvated equally by water molecules, the "***a1-s1-a2-s2***" pattern still can be maintained, but with large component of P=O asymmetric stretching for intensive harmonic peak ***a2***. And the rest ***a1*** and ***s1*** in this types come from the solvated PO-H bending motion, and there is a small component of P=O asymmetric stretching motion coupled with PO-H bending in ***s1***. In other words, the evolution of the spectra in this region is stabilized and additional water molecules would not bring much change to the absorption mode from n=6.

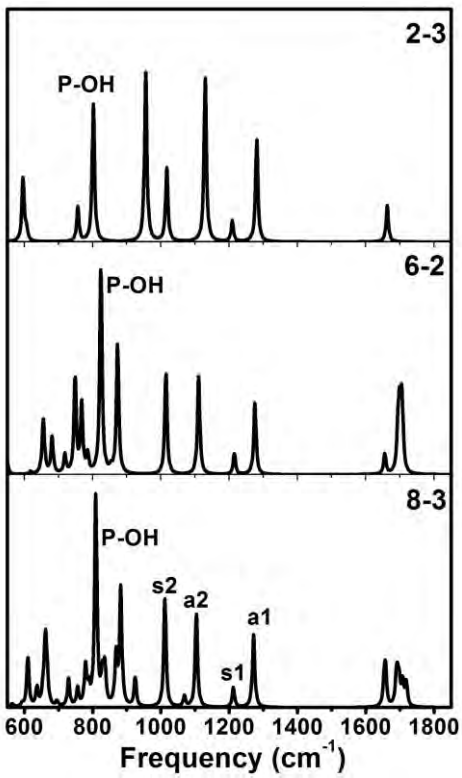

Figure 5.12 Calculated harmonic spectra for **(1D+1W, 1D+1W)** for n=2, 6, and 8.

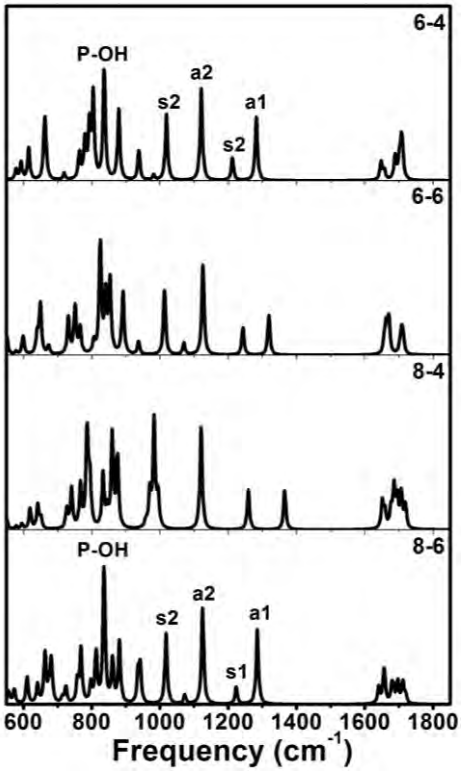

Figure 5.13 Calculated harmonic spectra for **(1D+1W, 1D+mW)** for n=6 and 8.

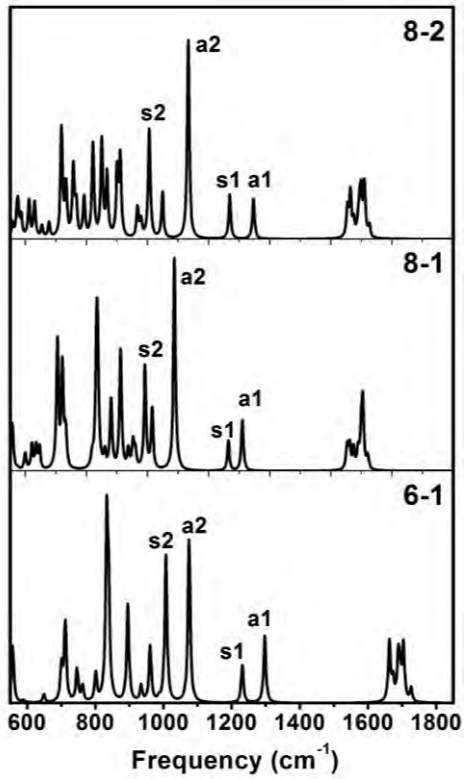

Figure 5.14 Calculated harmonic spectra for **(1D+mW, 1D+mW)** for n=6 and 8.

5.3.8 DTCF vibrational spectra in low frequency region

To disentangle the discrepancies between the harmonic and experimental IRMPD spectra, we perform AIMD simulations for n=2-4, 6, and 8. The vibrational profiles at finite temperatures are obtained by the Fourier transform of the dipole time correlation function (DTCF), which takes account of anharmonic as well as dynamic effects. Two sets of long MD simulations were performed at 140 and 180 K for extensive sampling of the phase space. At each temperature, four trajectories were simulated and each lasting 100 ps. Each trajectories was then cut into 10 ps interval for Fourier transform, and totally 40 frequency profiles were added up to produce the DTCF spectrum for a specific temperature.

There is an interesting peak change observed in the IRMPD spectra. Peak **I** is very unique in which its position is almost a constant, within a range of 20 cm^{-1} from n=1 to 12 as shown in Figure 5.8. This peak can be assigned to the symmetric P=O stretching. It is not only in the right position, as the peak with the lowest frequency in this region, but also shows small shifts in its frequency, in the calculated harmonic spectra for all the three types of structures. (**8-5** is the only exception, due to the mixing of water wagging.) In DTCF spectra obtained in AIMD simulations, this peak is also a fixed feature, showing little shift in position and always present at various temperatures.

For n=2 and 3, there are considerably better agreement between experimental and the DTCF spectra especially at 180 K, as shown in Fig. 5.15 and 5.16, respectively. In more detail, the water bending band **G** also is broaden like in n=1, but the relative intensity increases gradually due to the available of more water molecules, apparently. Between 1000 and 1300 cm^{-1}, more than four peaks (***a1***, ***s1***, ***a2***, and ***s2***) are predicted by harmonic analysis, while in the experimental spectra only two intensive bands (**I** and **H**) are observed. As discussed before, the dominant solvation types for these two clusters especially at 180 K are **(0, 0)** and **(0, 1D+mW)**, although there are instantaneous formations of **(1D+1W, 1D+1W)** or **(1D+1W, 1D+mW)**. Still, large motion amplitude is observed in water molecules, similar to n=1. And the coupling

mode (between **I** and **H**) between P-O-H bending and P=O stretching in Fig. 5.15 and 5.16 are broadened at high temperature in DTCF spectra. It also indicates that there is loss of relative peak intensity due to broadening in 1000-1300 cm^{-1}. Similar to n=1, the broadening is limited to the POH bending modes for n=2 and 3, and the bands intensities involving the heavier atoms are still not much affected. Consequently, the two intensive bands are assigned to the asymmetric (**H**) and symmetric (**I**) P=O stretching modes. In the region below 1000 cm^{-1}, there are a number of peaks related to wagging and rocking motion from water molecules in harmonic analysis. Since these modes all involve hydrogen atoms too, they are broadened and smeared out at 140 and 180 K. Only the asymmetric (**J**) and symmetric (**K**) P-OH stretching bands are left, which in good agreement with the experimentally observed peaks **K** and **L**.

For the cluster of n=4, the absorption band **H** redshifts dramatically according to n=2 and 3 by comparing the band gaps between **H** and **I**. For n=2 and 3 in experimental spectra, the band gaps are 188 and 186 cm^{-1}, respectively, whereas for n=4, this gap decreases to 167 cm^{-1}. This feature is also captured by the DTCF in AIMD simulations. As shown in Fig. 5.17, the intense peak **H** redshifts to 1164 cm^{-1}. Interestingly, the solvated PO-H bending mode also shows its absorption signature in the right-handed of **H**, but with quite small peak intensity. As discussed for the dynamics of n=4, the dominant binding motif is **4-(0, 1D+mW)**, in which the motion of the hydrated PO-H is restricted by the network formed by the four water molecules. Consequently, experimental band **H** in n=4 is assigned to asymmetric P=O stretching with a small component PO-H (solvated) bending from PDOS analysis. The broadened peak beside **H** in IRMPD is assigned to the pure bending motion of the solvated PO-H. Since this peak involves H atoms, too, it is also broadened and smeared out at high temperature in the AIMD simulations. The other bands in IRMPD spectrum of n=4 are similar to that of n=2 and 3. Broadened and more intensive peak **G** is also observed, which comes from water scissoring motion. **K** and **J** come from the asymmetric (**J**) and symmetric (**K**) P-OH stretching modes.

From n=6, pure bending modes of solvated PO-H groups (**H''**) are also broadened due large motion amplitude of light H atoms, similar to n<6. However,

there are some new absorption bands standing out, when more water molecules are available, such as **L** blow 900 cm^{-1} and intensive peak **H'** around 1200 cm^{-1}. These features are also captured in the simulated DTCF spectra at both 140 and 180 K, as shown in Figure 5.18. **H** comes from the asymmetric stretching of the double P=O group which is red-shifted for n=6 with reference to n=4. **H'** originates from P=O stretching of **6-1** since its gap between peak **I** is similar. The broadened band **G** in n=6 comes from water bending mode is also captured in DTCF spectra, as more water molecules are added, this peak will become more intensive and broadened. From n=6, symmetric (**K**) stretching of P-OH will be smeared out and merged into the water rocking and wagging modes and this feature is also captured by DTCF spectra started with **6-3** and **8-3**, as shown in Fig. 5.18 and 5.19.

The IRMPD spectra for n=8 shows a similar feature with n=6 in Fig. 5.19. And for n=10 and 12 shown in Fig. 5.8, the absorption features does not change too much with more water molecules available, and the band assignments for large clusters can be done according to n=6 and 8. With two more water molecules added, the intensities of absorption peaks (**L** and **G**) increase, which is also reproduced in the AIMD simulations for n=8. However, the large intensity increase for band **H** from 6 to 8 are not in agreement with the DTCF spectra (in which the peaks intensities of from each dynamics are un-biased), partially due to the sampling efficiency. As cluster size rises, the phase space cannot be completely sampled by AIMD simulations. But based on the simulated DTCF spectra, we can still assign each IRMPD band, qualitatively.

Overall, several the trends in peak positions re observed in the low frequency region of IRMPD spectra, such as intensity of **G**, blue-shift of **J** and **K**, red-shift of **H**, and appearance of **H'** from n=6. These trends are also captured qualitatively in the DTCF spectra. The transition is gradual and completed by n=8, when it becomes dominant structures with both POH groups solvated. The red shift of peak **H** could be due to P=O and PO-H groups solvated by more water molecules. The blue shifts of **J** and **K** can be attributed to the more favorable solvent-solvent interactions with increasing number of water molecules. And the broad band **L** from n=6 to n=12 is due

to wagging and rocking motions of the increasing number of water molecules. These trends can be rationalized by the changing relative importance of these types of structures as the cluster size increases.

For hydrated $(HO)_2PO_2^-$ anion, the solvation of the P=O groups is of higher priority than the solvation of the P-OH groups. For n=1-3, the clusters are often hold together just by the P=O…HOH hydrogen bonds at high temperature, and their spectra should be similar to the **(0, 0)** type spectrum. The main features are the asymmetric and symmetric P=O stretching bands in 1000-1300 cm^{-1}. As the cluster size increases, more water molecules are available, and PO-H groups are increasingly solvated. There is then a transition from **(0, 0)** to **(0, 1D+1w)**, and then to **(1D+1W, 1D+1W)**, **(1D+1W, 1D+mW)** and **(1D+mW, 1D+mW)**. And these trends are indeed in agreement with the stabilization of features in this frequency region in the IRMPD spectra, starting from n=8.

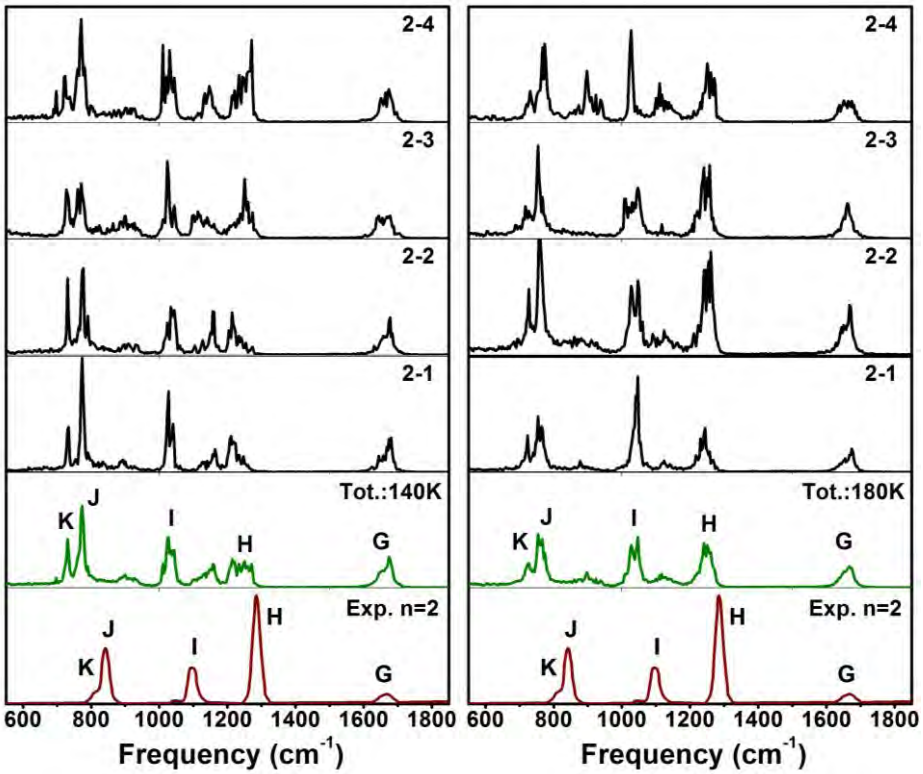

Figure 5.15 Comparison of the experimental IRMPD spectra of n=2 to the simulated DTCF spectra from 100 ps trajectories of four isomers.

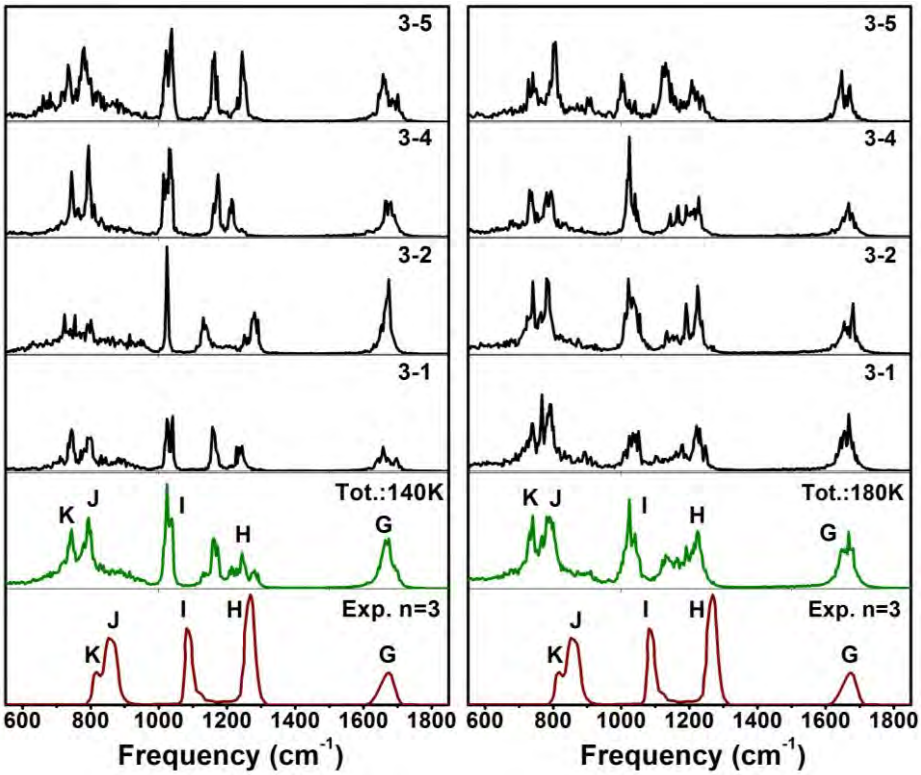

Figure 5.16 Comparison of the experimental IRMPD spectra of n=3 to the simulated DTCF spectra from 100 ps trajectories of four isomers **3-1**, **3-2**, **3-4**, and **3-5**.

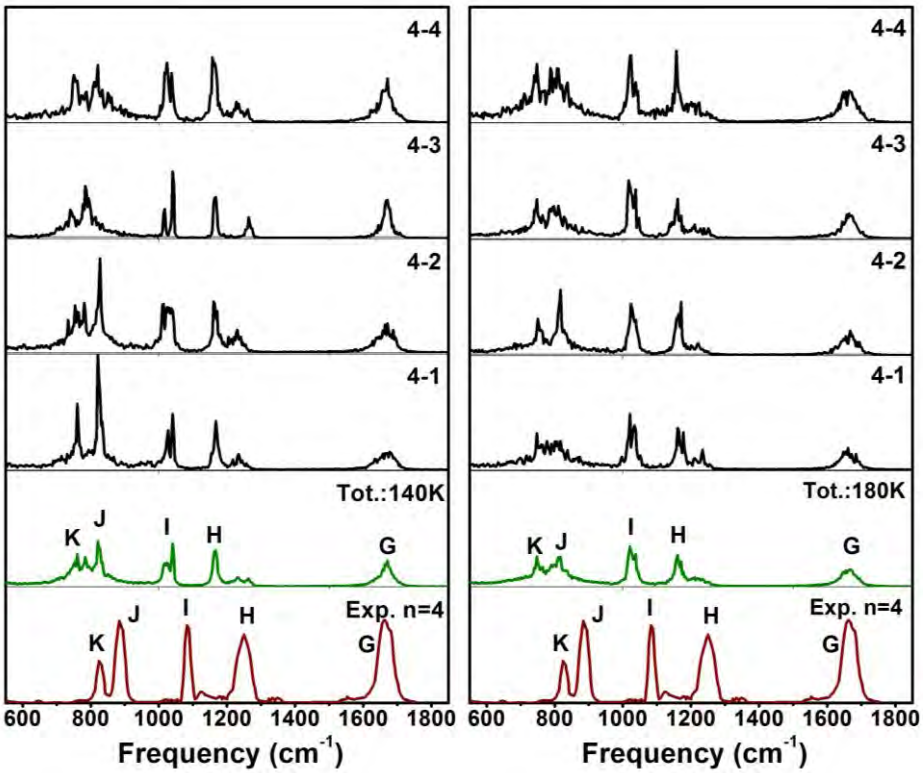

Figure 5.17 Comparison of the experimental IRMPD spectra of n=4 to the simulated DTCF spectra from 100 ps trajectories of four isomers.

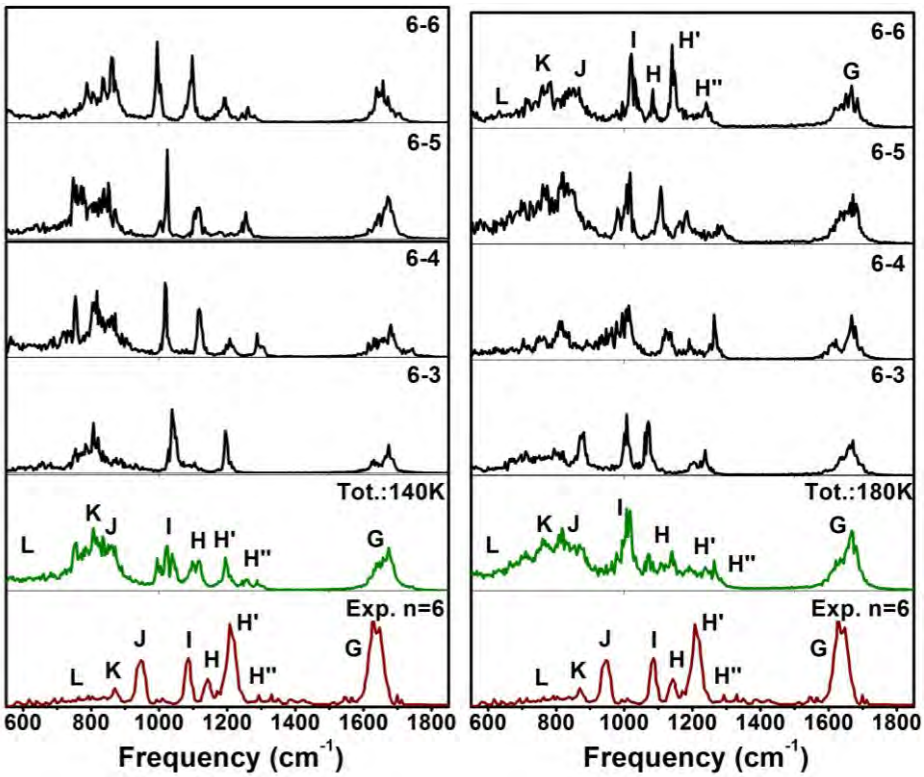

Figure 5.18 Comparison of the experimental IRMPD spectra of n=6 to the simulated DTCF spectra from 100 ps trajectories of four isomers **6-1**, **6-2**, **6-5**, and **6-6**.

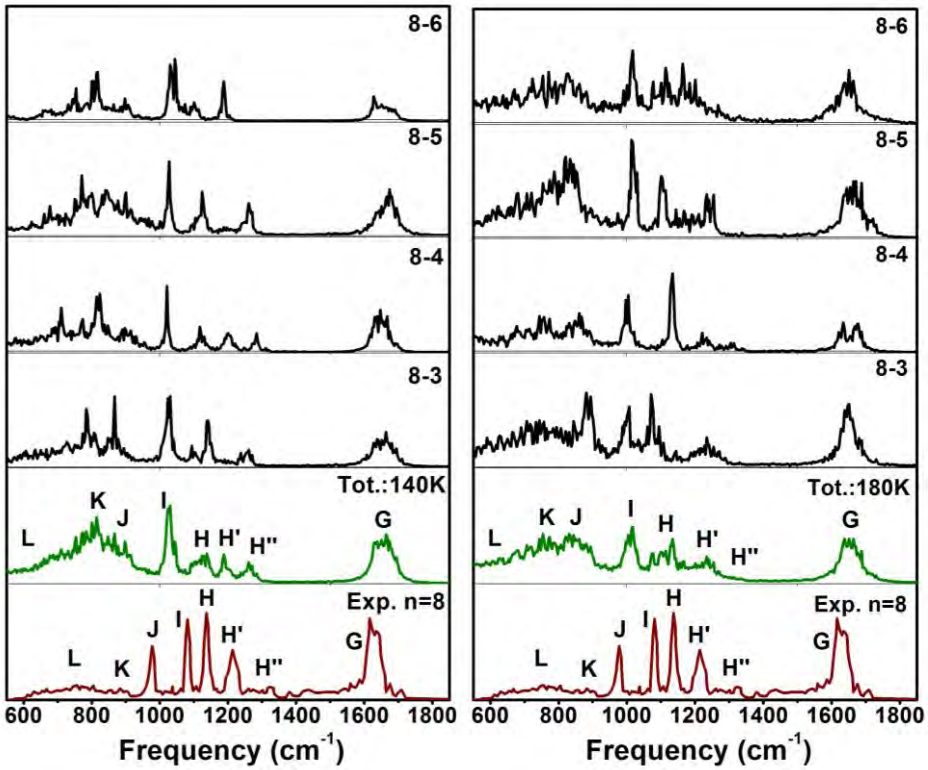

Figure 5.19 Comparison of the experimental IRMPD spectra of n=8 to the simulated DTCF spectra from 100 ps trajectories of four isomers **8-1**, **8-3**, **8-4**, and **8-5**.

5.4 Conclusion

In summary, the IRMPD spectra of $H_2PO_4^-(H_2O)_{n=2\text{-}4,6,8}$ also show evidence for isomerization at cryogenic temperatures and thus a complete breakdown of the harmonic approximation. There is an overall trend in the structural evolution for the hydrated $(HO)_2PO_2^-$ clusters. For n=2-4, there is one POH group left free. Whereas by n=6 and 8, both POH groups are solvated. This marks an interesting turning point in the structural evolution of the hydrated $(HO)_2PO_2^-$ clusters and is an important consideration in the assignment of their IRMPD spectra.

The trends in bands positions are observed in the low frequency IRMPD spectra. As the cluster size increases, PO-H groups are increasingly solvated from **(0, 0)** to **(0, 1D+1w)**, and then to **(1D+1W, 1D+1W)**, **(1D+1W, 1D+mW)** and **(1D+mW, 1D+mW)**. The red shift of peak **H** could be due to P=O and PO-H groups solvated by more water molecules. The blue shifts of **J** and **K** can be attributed to the more favorable solvent-solvent interactions with increasing number of water molecules. And the broad band **L** from n=8 to n=12 is due to wagging and rocking motions of the increasing number of water molecules. From n=4, the coupling mode between POH bending and asymmetric P=O stretching motion stands out, due to restricted POH orientation holding by the water network as more water molecules are available.

The finite temperature DTCF spectra in the high frequency region do provide important clues for the assignment of the experimental spectra. The two highest frequency peaks in the experimental spectrum, labeled as **A**, are due to the free PO-H stretching. The O-H peaks involved in the strongest hydrogen bonds in the **1D+1W/1D+mW** feature produce the broad band **F**, as in the case of n=1. The other OH peaks involved in hydrogen bonds are also broadened. And some of these features also present in the high frequency spectra for n=3 and 4.

Table 5.2 Assignment of the high frequency spectra for n=2, 3, and 4.

Label	Exp. freq.(cm^{-1})	n	Harmonic freq.(cm^{-1})	Assignment
A'	3698	n=2	3737(**2-1**)	Free H-OH stretching for n=2
			3739(**2-2**)	
			3740(**2-3**)	
			3743(**2-4**)	
A	3684	n=2	3709(**2-1**)	O-H stretching of the free PO-H
			3711(**2-3**)	
			3707(**2-4**)	
		n=3	3710(**3-1**)	
			3708(**3-4**)	
			3710(**3-4**)	
			3708(**3-5**)	
		n=4	3710(**4-1**)	
			3705(**4-2**)	
			3706(**4-2**)	
			3707(**4-4**)	
B		n=2	3512(**2-1**)	O-H stretching of the solvated PO-H
			3536(**2-2**)	
			3519(**2-3**)	
		n=3	3429(**3-1**)	
			3486(**3-2**)	
			3497(**3-2**)	
			3497(**3-3**)	
			3506(**3-3**)	
			3502(**3-5**)	
		n=4	3447(**4-1**)	
			3378(**4-3**)	
			3501(**4-3**)	
			3486(**4-4**)	
F	2800-3500	n=2	3043(**2-1**)	O-H stretching for hydrogen bonded H-O stretching in water.
			3363(**2-1**)	
			3528(**2-1**)	
			2975(**2-2**)	
			2997(**2-2**)	
			3012(**2-3**)	
			3507(**2-3**)	
			3636(**2-3**)	
			3139(**2-4**)	
			3459(**2-4**)	
			3582(**2-4**)	

Table 5.3 Assignment of the low frequency spectra for n=2, 3, 4, 6, and 8.

Label	Exp. freq.(cm^{-1})	n	Harmonic freq.(cm^{-1})	Assignment
G	1666	n=2	1661- 1704	Water scissoring mode
	1673	n=3	1659-1724	
	1664	n=4	1658-1717	
	1661	n=6	1648-1718	
	1649	n=8	1641-1728	
H'	1284	n=2	1236	Asymmetric P=O
	1268	n=3	1226	stretching from **(0,0)**
	1249	n=4	1222	structures
	1207	n=6	1231(**6-1**)	P-O-H bending from
			1243(**6-2**)	**(1D+1W, 1D+1W)**,
			1212(**6-4**)	**(1D+1w, 1D+mW)**, and
			1215(**6-5**)	**(1D+mW, 1D+mW)**
			1267(**6-6**)	structures.
	1216	n=8	1264(**8-1**)	
			1269(**8-2**)	
			1212(**8-3**)	
			1271(**8-3**)	
			1223(**8-4**)	
			1285(**8-4**)	
			1259(**8-5**)	
			1365(**8-5**)	
H	1121	n=3	1063	Bending of the free P-O-H
	1124	n=4	1057	from **(0, 1D+1w)**
				structures
	1138	n=6	1075(**6-1**)	Asymmetric P=O from
			1126(**6-2**)	**(1D+1W, 1D+1W)**,
			1121(**6-4**)	**(1D+1w, 1D+mW)**, and
			1153(**6-6**)	**(1D+mW, 1D+mW)**
	1132	n=8	1104(**8-3**)	structures.
			1125(**8-4**)	
			1120(**8-5**)	
			1111(**8-5**)	
I	1100	n=2	1017(**2-2**)	Symmetric P=O stretching
			1016(**2-3**)	
			1024(**2-4**)	
	1084	n=3	1025(**3-4**)	
			1006(**3-5**)	
	1086	n=4	1028(**4-2**)	
			1019(**4-4**)	
	1077	n=6	1015(**6-5**)	

	1071	n=8	1015(**6-6**) 1026(**6-3**) 1019(**6-4**) 1088(**8-1**) 1133(**8-2**) 1012(**8-3**) 1018(**8-4**) 1020(**8-5**) 1021(**8-6**)	
J	broad band		789-990	Water wagging.
K	840 853 885 924 962	n=2 n=3 n=4 n=6 n=8	799-805 795-843 814-844 824-838 827-885	Asymmetric P-OH stretching
L	810 816 827 842	n=2 n=3 n=4 n=6	747-795 743-766 749-778 759-785	Symmetric P-OH stretching
M	718	n=8	368-779	Water wagging and rocking

5.5 Reference

[1] Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb,M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K.N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.;Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li,X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; 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. Gaussian03, revision D.01; Gaussian, Inc.: Pittsburgh, PA, 2005.

[2] McLean, A. D.; Chandler, G. S. J. Chem. Phys. **1980**, 72, 5639.

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

[4] P. Sinha, S.E. Boesch, C. Gu, R.A. Wheeler, and A.K. Wilson, J. Phys. Chem. A, **2004**, 108 9213.

[5] VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. Comput. Phys. Comm. **2005**, 167, 103.

[6] The CP2K developers group, http://cp2k.berlios.de/.

[7] VandeVondele J. and Hutter J. J. Chem. Phys. **2007**, 127, 114105.

[8] Krack, M. Theor. Chem. Acc. **2005**, 114, 145.

[9] Goedecker, S.; Teter, M.; Hutter, J. Phys. Rev. B. **1996**, 54, 1703.

[10] Hartwigsen, C.; Goedecker, S.; Hutter, J. Phys. Rev. B. **1998**, 58, 3641

[11] Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. J. Chem. Phys. **2010**, 132, 154104.

[12] Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. J. Chem. Phys. **2010**, 132, 154104.

[13] McQuarrie, D. A., Statistic Mechanics; Harper-Collins Publishers; New York, **1976**.