Modelling the structure and IR spectra of micro-hydrated ions

Chandramohan Jana

École doctorale de l'École Polytechnique

THÈSE

pour obtenir le grade de

DOCTEUR DE L'ÉCOLE POLYTECHNIQUE

Spécialité : Chimie

Par

CHANDRAMOHAN JANA

Modelling the structure and IR spectra of micro-hydrated ions

Rapporteurs :
M. Jean-Philip PIQUEMAL Professeur - UPMC Sorbonne Universités

M. Philippe MAÎTRE Directeur de recherche au CNRS - Université Paris Sud

Examinateurs :
M. Jean-Pierre DOGNON Chercheur CEA - CEA/SACLAY

M. Gilles FRISON Chargé de recherche au CNRS - École Polytechnique

Co-directeur de thèse :
M. Gilles OHANESSIAN Directeur de recherche au CNRS - École Polytechnique

Directeur de thèse :
Mme Carine CLAVAGUÉRA Chargée de recherche au CNRS – École Polytechnique

Laboratoire de Chimie Moléculaire (LCM)
UMR N^{o} 9168 CNRS- École Polytechnique
91128 Palaiseau Cedex France

École doctorale de l'École Polytechnique

THÈSE

pour obtenir le grade de

DOCTEUR DE L'ÉCOLE POLYTECHNIQUE

Spécialité : Chimie

Par

CHANDRAMOHAN JANA

Modelling the structure and IR spectra of micro-hydrated ions

Rapporteurs :
M. Jean-Philip PIQUEMAL Professeur - UPMC Sorbonne Universités

M. Philippe MAÎTRE Directeur de recherche au CNRS - Université Paris Sud

Examinateurs :
M. Jean-Pierre DOGNON Chercheur CEA - CEA/SACLAY

M. Gilles FRISON Chargé de recherche au CNRS - École Polytechnique

Co-directeur de thèse :
M. Gilles OHANESSIAN Directeur de recherche au CNRS - École Polytechnique

Directeur de thèse :
Mme Carine CLAVAGUÉRA Chargée de recherche au CNRS – École Polytechnique

Laboratoire de Chimie Moléculaire (LCM)
UMR N^{o} 9168 CNRS- École Polytechnique
91128 Palaiseau Cedex France

To my parents
For sowing the seeds of my education

Acknowledgements

With this thesis, I happily conclude my research work which was rigorous and demanded patience. I take this opportunity to thank this institute, which is quite open and free and also thank all those who have contributed knowingly or unknowingly in this endeavor.

First and foremost my deepest heartily gratitude goes to my thesis supervisor ***Dr. Carine Clavaguera*** *and has no words to thanks her for help, guidance, support and encouragement throughout my Ph.D research work. This work would not be possible without her constant support. I successfully overcame many difficulties with her supports and learned a lot from her. I could not have imagined having better advisor and mentor for my Ph.D. study and I would try to inculcate the virtue of equanimity in my future endeavors. I appreciate all her contributions of time, ideas, and funding to make my Ph.D. experience productive and stimulating.*

I extremely thankful to my co-supervisor ***Dr. Gilles Ohanessian,*** *formerdirector of the laboratoire des MécanismesRéactionnels, for the opportunity to work in this laboratory and complete the study of thisthesis. More importantly, I am very thankful for his support throughout my research work. His interactions and discussions have always helpful to put valuable addition in my research time to time and also his interactions have always benefited me tremendously in broadening my understanding and encourage me to learn.*
Both ***Dr. Carine Clavaguera*** *And* ***Dr. Gilles Ohanessian*** *are really great inspiration for me.*

In addition, I have no words to say thanks to the people, with whom I shared my office: ***Ashwani, Ophélie, Manjeet, Jianqing and Saurabh.*** *It has always been a pleasure working and talking with you. Besides my special thanks to* ***Ashwani*** *for his always supports and care for in all situations.*

My grateful thanks go to all the staff, researchers (or former researchers) and students in LCM lab: Thérèse, Julien, Vanessa, Madan, Sophie, Stephane, Michel, Christophe, Guy,……..You have made this three years a memory that I will bring with me forever. I am truly grateful for your friendship.

*My gratitude also goes to **Prof. Jean-Philip Piquemal**and **Dr. Philippe Maitre**for reviewing my thesis and to **Dr. Jean-Pierre Dognon** and **Dr. Gilles Frison**for evaluating my work as my thesis committee member.*

*Dear **Mom** and **Dad**, despite my favor for languages, it is hard to express my thanks to you in words. Your understanding, the faith, advises, intellectual sprit and indescribable support to me throughout my whole life are invaluable. I appreciate your honesty, including complements and criticisms. I am very grateful about your care, your love and trust in me, and further thank you for your constant interest and positive stimulation. My special thanks also go to my family and elder brother **Shyamsundar** for their continuous supports and love.*

Also I thank my special friends: Giridhari, Dip, Pranabesh, Sourav, Rohit, Rajkumar, Krishna Gopal, Bangshilal.

My time at Ecole Polytechnique was made enjoyable in large part due to the many friends and groups that became a part of my life.

Table of Contents

Chapter 1

General introduction

The structure of water is usually described by a dynamical network of hydrogen bonded clusters whose properties are modified by temperature and pressure changes, as well as by the introduction of solutes. In the case of ionic solutes, the balance of intermolecular forces is disturbed by the water molecules that surround the ion due to strong ion-water interactions and modified water-water interactions[1,2]. Positively or negatively charged, ionic hydration is an exothermic process in which energy is released when a hydrated ion forms from the gaseous free ion. This energy is mainly coming from the ionic electrostatic forces in the ion-water interactions. Figure 1 provides a standard picture of the hydration spheres of a metal cation having a first solvation shell of six water molecules as usual for many 3d metal ions[2]. Three regions are defined for the organization of water around the ion: the first hydration shell in which the water molecules are strongly oriented by the presence of the ion, the second hydration shell in which the water molecules are more or less weakly oriented by the ion, depending on its charge and nature, and hydrogen bonded to the first shell, and finally the outermost region can be considered as a third hydration shell in some cases or as bulk water. Generally, cations are known to be more strongly hydrated than anions due to strong interactions between the positive charge and the negatively polarized oxygen atoms of first-shell water. However, halide ions are known to have a well-defined first hydration shell even though their interactions via the hydrogen atoms are weaker[3].

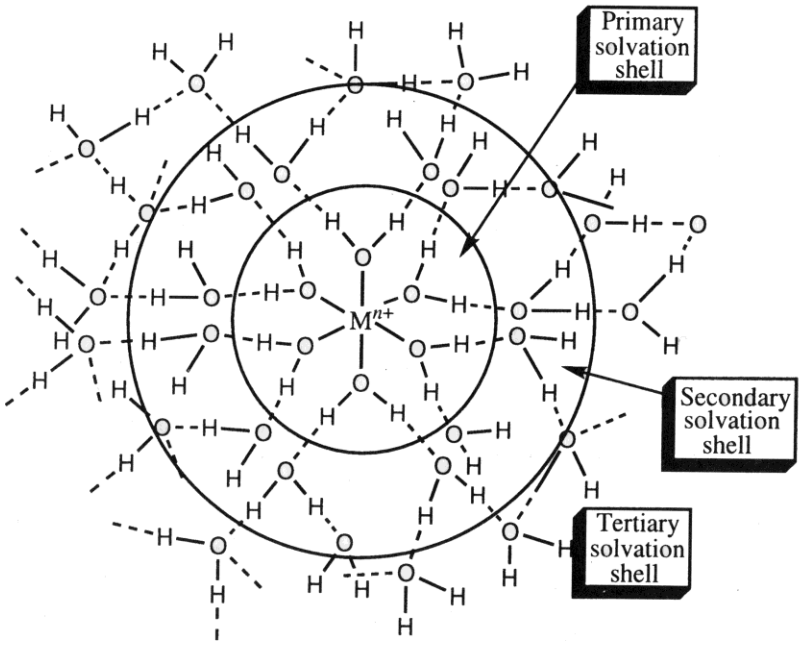


Figure 1 Structure of a hydrated metal cation in aqueous solution (a first hydration number of 6 is assumed). The Figure is taken from reference 2.

Understanding the changes occurring to water in the presence of an ion implies the use of various experimental and theoretical techniques with the aim to apprehend the behavior of ions in chemical and biological systems and processes. Consequently, several experimental methods, mainly X-ray and neutron diffraction and Extended X-Ray Absorption Fine Structure (EXAFS), have been applied to obtain structural information on the ion-water interactions, e.g. coordination numbers and ion-water distances[3]. In addition, ^{17}O NMR is generally used to get the coordination number of the first shell and the water exchange dynamics[4]. For 3d metal transition cations, it is established that the water molecules in the first hydration shell are organized in a six-coordination with an octahedral arrangement. However, the hydration shell organization of other ions is not fully determined due to limited information available on their structural and dynamics properties. In order to go further, the combination of experimental and theoretical methods is essential to extend the knowledge of these hydration properties[5].

At the other end of the size spectrum, hydration can be studied by stepwise construction of the hydration layers around isolated ions, one molecule at a time. Isolated ions are easily formed in mass spectrometers; their micro-hydrated clusters are most commonly formed by electrospray ionization (ESI)[6]. Mass spectrometric techniques allow for mass selection, enabling the study of size-selected cluster ions. Various activation techniques can then be implemented in the mass spectrometer to gain insight into structures and dynamics. Collision-induced dissociation (CID) and Blackbody InfraRed Dissociation (BIRD) allow for the determination of stepwise water binding energies. More recently, infrared-induced photodissociation, either by single (IRPD) or multiple (IRMPD) photon absorption, has provided a wealth of vibrational data on gaseous ions[7-9]. Most of those dealing with hydrated ions address only the very first steps of micro-hydration[10], however IR spectroscopy of ions surrounded by dozens of water molecules is beginning to emerge[11,12].

Most of the time, assigning IR bands to specific vibrational modes is done by quantum mechanics calculations in the static double harmonic approximation. However, these calculations are limited by the size of the systems. Furthermore, temperature and anharmonic effects are not taken into account, leading to large deviation from experiment. Usually, scaling factors applied to the computed frequencies are used to reduce the difference with experimental values[13]. Intense efforts were made to explicitly include anharmonic effects in vibrational frequency calculations[14,15]. These methods developed to solve the vibrational Schrödinger equation are very powerful, providing accuracy of ca.10 cm^{-1} or less. However, they remain limited to small-size systems and thus are not appropriate to the study of hydrated ions except for the smallest.

Among simulation techniques, molecular dynamics is a powerful tool for the analysis of both static and dynamical properties of solvated ions. It has been extensively used in the last

decades for the study of aqueous solutions[3,16-21]. The use of force fields in combination with molecular dynamics is justified by the need for statistical and dynamical properties, and by the occurrence of rather rare events taking place over hundreds of picoseconds or more. Consequently, *ab initio* molecular dynamics, which is limited to few picoseconds, cannot provide an appropriate picture of such properties. Classical force fields typically treat intermolecular interactions as sums of pairwise van der Waals and electrostatic terms. Many-body interactions are thus only taken into account implicitly. The inclusion of an explicit polarization term in the potential energy expression constitutes an improvement for classical force fields. This has led to a new generation of "polarizable force fields". Among them, the AMOEBA (Atomic Multipole Optimized energetics for Biomolecular Applications) force field has been developed to provide accurate structural and energetic description of the hydration of ions, organic molecules and biomolecules[21-24]. Molecular dynamics is also a method of choice for the computation of IR spectra which thus intrinsically include anharmonicity and temperature effects[25-27]. To this end, the AMOEBA force field has been recently extended to enable the calculation of accurate IR spectra of gaseous ions in our group[28].

Using this methodology, modeling the IR spectra of micro-hydrated ions is presented in this thesis. The combination of quantum chemistry and polarizable molecular dynamics was used to determine their structural and spectroscopic properties with close comparison with recent experimental data. The manuscript is organized as follows.

The methodological background is detailed in chapter 2.

Chapter 3 focuses on the hydration of Zn^{2+} with the ambition to bridge the gap between theoretical studies in the condensed as well as in the gas phase. Extensive simulations have been

run for clusters growing from 6 to 216 water molecules to derive the zinc coordination number as a function of cluster size and temperature. IR spectra of small clusters were computed to understand temperature and charge transfer effects on O-H stretching bands.

Chapter 4 presents a detailed study of sodiated tryptamine solvated by one or two water molecules. Emphasis is put on force field parameterization of the electrostatic term to improve the spectral features in comparison with available experimental IRPD spectra. Band attribution was done by a detailed analysis of the trajectories.

Chapter 5 is devoted to the hydration of the sulfate dianion in nanodroplets. This study is motivated by intriguing recent IRPD experiments which provide spectral signatures of the presence of dangling O-H bonds at the surface of sulfate-water droplets[29]. MD simulation tools were adapted to the study of isolated clusters containing 25 to 100 water molecules. Then, finite temperature IR spectra were computed and compared with experimental results.

References

(1) Conway, B. E. *Ionic hydration in chemistry and biophysics*; Elsevier: Amsterdam, **1981**.

(2) Richens, D. T. *The Chemistry of Aqua Ions*; John Wiley and Sons, Inc.: New York, **1997**.

(3) Ohtaki, H.; Radnai, T. *Chem. Rev.* **1993**, *93*, 1157.

(4) Helm, L. ; Nicolle, G. ; Merbach, A. E. *Adv. Inorg. Chem.* **2005**, *57*, 327.

(5) Kunz, W. *Specific ion effects*; World Scientific: Singapore, **2010**.

(6) Blades, A. T.; Kebarle, P. J. *Am. Chem. Soc.* **1994**, *116*, 10761.

(7) Polfer, N. C.; Dugourd. P. *Laser photodissociation and spectroscopy of mass-separated biomolecular ions*; Springer: Switzerland, **2013**.

(8) MacAleese, L ; Maître, P.*Mass Spectrom. Rev.* **2007**, *26*, 583.

(9) Polfer, N. C. ; Oomens J.*Mass Spectrom. Rev.* **2009**, *28*, 468.

(10) Nicely, A. L.; Lisy, J. M. *J. Phys. Chem. A* **2011**, *115*, 2669.

(11) O'Brien, J. T.; Williams, E. R *J. Am. Chem. Soc.* **2012**, *134*, 10228.

(12) Nagornova, N. S.; Rizzo, T. R.; Boyarkin, O. V. *Science* **2012**, *336*, 320.

(13) Bowman, J. M. *J. Chem. Phys.* **1978**, *68*, 608.

(14) Merrick, J. P.; Moran, D.; Radom, L. *J. Phys. Chem. A* **2007**, *111*, 11683.

(15) Gerber, R. B.; Ratner, M. A. *Chem. Phys. Lett.* **1979**, *68*, 195.

(16) Chang, T. M. ; Dang, L. X. *Chem. Rev.* **2006**, *106*, 1305.

(17) Brancato, G; Barone, V.*J. Phys. Chem. B* **2011**, *115*, 12875.

(18) Lamoureux G.; MacKerell A. D.; Roux B.*J. Chem. Phys.* **2003**, *119*, 5185.

(19) Lamoureux G.; Roux B. *J. Phys. Chem. B* **2006**, *110*, 3308.

(20) Dang, L. X. *J. Phys. Chem. C* **2014**, ASAP.

(21) Grossfield, A.; Ren, P.; Ponder, J. W. *J. Am. Chem. Soc.* **2003**, *125*, 15671.

(22) Ponder, J. W.; Wu, C.; Ren, P.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio, R. A.; Head-Gordon, M.; Clark, G. N. I.; Johnson, M. E.; Head-Gordon, T. *J. Phys. Chem. B* **2010**, *114*, 2549.

(23) Ren, P.; Wu, C.; Ponder, J. W. *J. Chem. Theory Comput.* **2011**, *7*, 3143.

(24) Shi, Y.; Xia, Z.; Zhang, J.; Best, R.; Wu, C.; Ponder, J. W.; Ren, P. *J. Chem. Theory Comput.* **2013**, *9*, 4046.

(25) Fanourgakis, G. S.; Xantheas, S. S. *J. Chem. Phys.* **2008**, *128*, 074506

(26) Schultheis, V.; Reichold, R.; Schropp, B.; Tavan, P. *J. Phys. Chem. B* **2008**, *112*, 12217.

(27) Schropp, B.; Wichmann, C.; Tavan, P. *J. Phys. Chem. B* **2010**, *114*, 6740.

(28) Semrouni, D.; Sharma, A.; Dognon, J-P.; Ohanessian, G.; Clavaguéra, C. *J Chem. Theory Comput.* **2014**, ASAP.

(28) Ramanathan, V.; Crutzen, P. J.; Kiehl, J. T.; Rosenfeld, D. *Science* **2001**, *294*, 2119.

Chapter 2

Computational methodology

1. Quantum mechanical models

Many quantum mechanical (QM) models have been developed over the years. For optimizing the structure of a molecule and calculating its properties, such as its vibrational spectrum, each method/basis set combination has its own figures of merit. It is thus of interest to recall the main features of the methods used in the following chapters. More detailed description are available in a number of textbooks, see e.g. references[1,2,3].

As an example, a standard method in computational chemistry is "Coupled Cluster including Single and Double excitations, plus a perturbative treatment of connected Triples [CCSD(T)]. It is often used as a reference for quantitative energetics, however it suffers from being very computationally costly for systems larger than about ten atoms. Density Functional Theory (DFT) is often the choice for calculations on larger systems however it has limitations, e.g. most early functionals fail to describe weakly interacting systems. DFT is still developing at high pace with e.g. the Minnesota series of functional[4] and Grimme's empirical corrections for dispersion[5]. A less intensive alternative is the Hartree-Fock (HF) formalism which provides molecular orbitals, and is a starting point for CCSD(T). Intermediate in complexity is MP2 theory in which second order perturbation theory is used to add to the single HF configuration, the effect of excited configurations in an approximate manner. Understanding the basic theory in each case helps identify the reasons for successes and drawbacks of each approach. There remains however some need for calibration on each new system studied, i.e. running calculations and then using the end result to determine whether a method/basis set combination is suitable.

In the work presented in this thesis, all QM results are obtained using DFT (B3LYP or M06) and MP2 calculations. In general, these QM methods are used for finding low energy structures and their energies by geometry optimizations and single point energy calculations. Vibrational

frequencies are also obtained at the optimized geometries. The following section provides some qualitative background about the computational procedure.

1.1. Hartree-Fock method

QM methods explicitly describe the motion of electrons in a calculation, and it is possible to derive a molecular property that depends upon the electronic distribution. In particular, computational quantum chemistry involves solving the Schrödinger equation to obtain properties originating from the electronic structure of molecule. To carry out QM calculations on a static structure, one needs to solve the time-independent electronic Schrödinger's equation: $\boldsymbol{H\Psi = E\Psi}$. This equation cannot be solved exactly in general. The Born-Oppenheimer approximation can be used, according to which the motion of the electrons is decoupled from that of the nuclei. Even within this approximation, an exact solution is not available except for simple single electron species such as H_2^+. However, the variational principle provides us with a mechanism for answering this question. The theorem states that any trial wave function has an energy value greater than or equal to that of the exact solution and that the best wavefunction is obtained when the energy is a minimum. This forms the basis of many QM methods, including HF as the simplest. At the latter level, the trial wavefunction is the Slater determinant (SD), which satisfies the antisymmetric property upon exchanging all coordinates of any pair of electrons[1,2,3].

For a single SD and non-relativistic Hamiltonian, applying the variational principle leads to a set of eigenvalue equations (known as the HF equations), with eigenfunctions named molecular orbitals. In canonical form, they read: $F_i\phi_i = \epsilon_i\,\phi_i$ where F_i is known as the Fock operator. Since the latter depends upon the ϕ_i, solutions must be searched for in an iterative, self-

consistent manner. The molecular orbitals ϕ_i are themselves expanded as linear combinations of basis functions, generally taken in Gaussian form.

The use of a single SD amounts to restricting electron-electron interactions to averaged exchange and Coulomb interactions. As a consequence HF calculations, even if carried out close to the infinite basis set limit, can show large deviations from experiment. They can be corrected by using methods that utilize the Hartree-Fock method as a starting point.

1.2. Møller–Plesset perturbation theory

Instantaneous electron correlation effects are neglected in HF theory, introducing significant error into energy calculations. It provides however a qualitatively correct picture of a number of molecular properties. In addition, it is convenient to use the HF wavefunction as a starting point toward more accurate wavefunctions and energies. One can use of the perturbation theory to improve upon the HF results. A popular improvement is the Møller–Plesset (MP) version of Rayleigh-Schrodinger perturbation theory[1,2,3]. It assumes that if the exact Hamiltonian H is expanded into the sum of the HF Hamiltonian H_0 and a complementary term λVs[2]: $H = H_0 + \lambda \mathrm{V}$, the latter is small enough that the expressions below for energy and wavefunction can be truncated at small order (truncation at order n is termed MPn).

$$E = E_0 + \lambda E_1 + \lambda^2 E_2 + \cdots + \lambda^n E_n$$

$$\Psi = \Psi_0 + \lambda \Psi_1 + \lambda^2 \Psi_2 + \cdots + \lambda^n \Psi_n$$

Substituting the above two equations into $\boldsymbol{H\Psi} = \boldsymbol{E\Psi}$, and equating terms for each power n, it appears that truncation at second order, MP2, is the first level of perturbation theory which improves upon the energy obtained from HF theory. The MP2 energy is given by:

$$E_{MP2} = \sum_{i<j}^{occupied} \sum_{a<b}^{virtual} \frac{| < \Psi_0 | V | \Psi_{ij}^{ab} > |^2}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b}$$

where Ψ_{ij}^{ab} involves exciting electrons from occupied orbitals *i* and *j* to virtual orbitals *a* and *b* and ε_i, ε_j, ε_a, and ε_b are the corresponding orbital energies. This shows that only double electron excitations are included at second order (in an approximate way). Increasing the value of n, e.g. to n=4 (MP4) allows inclusion of triply and quadruply excited configurations, making the calculation much more time consuming but generally more accurate[1,2,3].

1.3. Coupled cluster method

The coupled cluster (CC) method is among the most sophisticated post-HF theories as it includes higher order excited configurations. Instead of calculating exactly the coefficients of highly excited configuration, they are deduced from those of lower excitation degrees, e.g. the coefficients of quadruples are products of coefficients for doubles. When only products of single and double excitations are included, the method is termed CCSD. An efficient approximation to CCSD is known as CC2[6]. The best known coupled cluster method includes products of all singles, doubles, and a perturbative estimate of connected triples, thus its CCSD(T) acronym. This method is often considered as the "gold standard" in computational chemistry, because of its high accuracy, which of course is achieved at high computational cost. A more detailed description can be found in reference 1.

1.4. Density functional theory

Density functional theory (DFT) has become the most popular quantum chemical method over the last two decades[1,2,3]. It solves for the one-electron density of a molecular system, which

is known to contain all information about the electrons of this system. While the electron-nucleus attraction and the Coulombic part of electron-electron interaction operators can be readily expressed in terms of the one-electron density, this is true neither for the kinetic energy nor for the exchange part of electron-electron interaction. Thus the very simple and appealing formal basis of DFT has led to rather complicated developments in order to bring it to its present status. Those terms for which there is no simple expression known are globally named the exchange correlation energy. They include the effects of quantum mechanical exchange, correlation, correction for self-interaction, and the difference of the kinetic energy between a fictitious non-interacting system and the real one. There have been several general levels of increasing complexity to approximately describe the exchange correlation correction. In the local density approximation (LDA), exchange correlation functional is determined solely from the density at that location, while in the generalized gradient approximation (GGA), the gradient of the density is included in addition to the local density. Hybrid functionals also include some fraction of Hartree-Fock exchange[7], while the meta-GGA approximation amounts to including terms that depend on the kinetic energy density[1,2,3]. In the work presented in this thesis, all DFT results are obtained with either the B3LYP (hybrid GGA) or the M06 (hybrid meta-GGA) functional. M06 is a member of the Minnesota density functional family[4]. This series involves functionals with various levels of approximation. They are all heavily parameterized, with no attempt to model specific types of interactions (polarization, dispersion…). Rather, parameter fitting relies on several databases of experimental or highly accurate computational results. As a result, each Minnesota functional is more or less object-oriented with variable performances for thermochemistry, kinetics, non-covalent interactions, etc. Each can be recommended for specific types of applications[4] .

An alternative approach to address the problem of accurately describing non-covalent interactions is to add specific empirical correction terms, either within the functional expression, or else as a correction to the DFT calculation. The latter approach has become largely used following Grimme's prescriptions[5].

1.5. Basis sets

All of the methods described above rely on a representation of the one-electron wavefunctions (molecular orbitals) or density. Development as a linear combination of pre-defined basis functions is used nearly universally, and Gaussian basis functions are by far the most common. A molecular orbital is thus described as a linear combination of Gaussian basis functions whose coefficients are determined using the iterative self-consistent field method (SCF)[1,2,3]. Since individual Gaussians are not good models of atomic orbitals, linear combination of Gaussians are used instead, often with predefined coefficients for part of them (called contracted Gaussians). Several series of basis sets of varying sophistication have been developed. For light main group elements, the most prominent series are those of Pople et al.[8] and of Dunning et al.[9]. We have used both series, at the valence double- or triple-ζ level with various numbers and types of polarization functions and also diffuse functions in the case of anions. The Dunning series has been designed to converge systematically to the complete basis set limit (CBS) at the MP2 level, using empirical extrapolation. This CBS limit is very useful for method calibration when high accuracy is required.

1.6. Calibration of functionals on hydrated sulfate

A typical example of calibration is provided by the work of Head-Gordon et al. on hydrated sulfate dianions $SO_4^{2-}(H_2O)_n$[10,11]. The existence of a variety of different isomeric structures due

to the relative populations at different temperatures, even in small hydrated-sulfate clusters, offers complexity in assigning their spectral bands. Since the energetic differences between these isomers are of few tenths of kcal.mol^{-1}, it is very important to estimate the accurate energy ordering of these isomeric structures. In a recent report, Head-Gordon and co-workers have performed an extensive search for energetic ordering of sulfate–water (n = 3–6) clusters using high-level electronic structure calculations at CCSD(T) level of theory in combination with an estimate of basis set effects at the CBS limit, using MP2 theory[10]. CCSD(T) calculations provide energetic ordering as a reference, compared to other methods, but these calculations are very demanding for clusters including up to 7 water molecules. In order to find a substitute for CCSD(T) calculations, the same group has performed a detailed calibration study of density functional calculations for relative and binding energy analysis of sulfate-water clusters[11]. They have included 24 density functionals including GGA, meta-GGA, hybrid GGA, hybrid meta-GGA and double hybrid density functionals, as well as dispersion-corrected density functionals, along with Hartree-Fock and the MP2 methods with respect to CCSD(T)/CBS reference energies. The MP2 method was found to perform best with respect to the CCSD(T)/CBS reference. The Truhlar M11 was found to be the best density functional. Furthermore, other density functionals such as XYG3, ωB97X-2 and XYGJ-OS were also found to provide good estimates of the relative and binding energies compared to the reference CCSD(T)/CBS method.

All QM results presented in this thesis have been obtained using the Gaussian09 software package[12]. Structure display and manipulation were done using the Molden software[13].

2. Empirical force field models: molecular mechanics

2.1. General description

Molecular mechanics (MM) methods describe the potential energy of a molecule in terms of a set of equations physically meaningful in classical mechanics. Therefore, these methods ignore the electronic motions and calculate the energy of the system as a function of the nuclear positions only. In an all-atom description, atoms are the smallest fundamental units of the system. The set of energy terms and associated parameters is called a force field. Many force fields are widely available for a variety of problems in chemistry, biochemistry and materials science.

Some of the most common and standard force fields are AMBER, OPLS, CHARMM, GROMOS, MM2, etc. AMBER[14] (Assisted Model Building and Energy Refinement) is widely used for proteins, nucleic acids and DNA. OPLS[15,16] (Optimized Potential for Liquid Simulations) is developed using a part of the formalism of the AMBER force field. A distinctive feature of the OPLS force field is that it reproduces densities and extended other properties of liquid organic molecules to fit experimental data. CHARMM[17,18] (Chemistry at HARvard Molecular Mechanics) is widely used for both small molecules and macromolecules and its formalism is similar to that of AMBER. GROMOS[19] (GROningen MOlecular Simulation package) is developed for molecular dynamics simulation on biomolecules and macromolecules such as proteins, nucleotides and sugars in aqueous or apolar solutions. There are other force fields for a broad range of applications such as UFF[20] (Universal Force Field) for which the parameters are estimated using general rules based only on the element, its hybridization and its connectivity. Allinger *et al.* have developed a series of force fields called MM2[21], MM3[22,23] and MM4[24] for organic molecules. This MM family is often regarded as the gold standard, as these

force fields have been derived from a large set of *ab initio* and experimental data. The covalent potential energy includes additional terms of anharmonicity (see below) and some terms are coupled. The parameterization allows distinguishing the following types of carbon atoms: sp^3, sp^2, sp, carbonyl, cyclopropane, radical, cyclopropene, and carbonium ion. The MM3 force field is considered as a good compromise between accuracy and computational cost for modeling organic molecules.

Force fields mentioned up to this point are called first or second-generation force fields. The most important difference between the various models is the approach to describing electrostatic interactions. Force fields such as CHARMM, OPLS, AMBER and GROMOS use partial charges to represent the electron distribution. These charges are often, but not necessarily, atom-centered and other charge sites can also be included at bond centers or other off-atom sites to improve the model. Another important issue is that polarization effects are not explicitly addressed by these force fields.

The potential energy of the system $U_{potential}$ is expressed by typical components:

$$U_{potential}=[U_{bond}+ U_{angle}+ U_{torsion}+ U_{cross\text{-}term} + U_{oop}]+ [U_{vdW} + U_{elec}]$$

The first five terms describe covalent interactions, and the last two terms represent the noncovalent contributions. They are described in sections 2.2 and 2.3.

2.2. Covalent interactions

U_{bond} represents the bond stretching between covalently bonded atoms *i* and *j* and takes the form of a Taylor expansion around a reference bond distance (r_{ij}^0):

$$U_{ij}^{bond} = k_2(r_{ij} - r_{ij}^0)^2 + k_3(r_{ij} - r_{ij}^0)^3 + k_4(r_{ij} - r_{ij}^0)^4 \ldots$$

k_2, k_3, k_4… are called the force constants. This term is frequently restricted to harmonic oscillators (i.e. only k_2 is non-zero) but higher-order terms can be added to model anharmonicity as in the MMn force fields.

As this formalism does not have the correct behavior to dissociation, a Morse potential[25] is an alternative function that satisfies these limiting conditions:

$$U_{ij}^{Morse} = D\left\{1 - \exp[-a\,(r_{ij} - r_{ij}^{0})]\right\}^2$$

Where, D is the dissociation energy and $a = \omega\sqrt{(\mu/2D)}$ where μ is the reduced mass and ω is the frequency of the bond vibration. The Morse potential exhibits correct dissociation behavior, however, this potential gives very small restoring forces for large values of r_{ij} and therefore, causes slow convergence in geometry optimization. A number of molecular force fields adopt the Morse potential function such as UFF, Extensible systematic force field (ESFF)[26] and DREIDING[27]. However, the Taylor expansion formulation is more widely used for its low computational cost and flexibility in fitting bond vibrational frequencies.

Similarly, the angle bend potential U_{angle} usually employs a Taylor expansion around a reference angle (θ^0) formed between two bonds and often truncated to the first term:

$$U_{\theta}^{angle} = k_2(\theta - \theta^0)^2 + k_3(\theta - \theta^0)^3 + k_4(\theta - \theta^0)^4 \ldots$$

The torsional energy U_{torsion} is a four-body term describing the rotational barrier along a bond. This is usually implemented as a Fourier series:

$$U_{\omega}^{torsion} = \sum_{n} V_n\,[1 + \cos(n\omega - \gamma)]$$

V_n gives a qualitative indication of relative rotation barriers and ω is the torsion angle. n is the bond multiplicity; its value gives the number of minimum points in the function as the bond is rotated through 360^0 with typically range from 1 to 3. The phase factor γ determines where the

torsion energy passes through its minimum value and is often assigned 0 and π for odd and even *n*, respectively.

U_{oop} represents out-of-plane bending motions to enforce planar geometry at sp^2-hybridized centers. There is significant energetic penalty associated with moving the central atom away from the trigonal plane.

$$U_{\chi}^{oop} = a\ \chi^2$$

$$U_{d}^{oop} = a\ d^2$$

The out-of-plane term can be a function of either an out-of-plane bending angle (χ) or a distance (d) to the trigonal plane formed by the three attached atoms.

Finally, cross-term energy ($U_{\text{cross-term}}$) function represents the coupling between bonds, angles and torsions, which has been found significant to reproduce vibrational frequencies. The most commonly implemented term is stretch-bend coupling, but other cross-terms may also be used, depending on the force field model.

2.3. Noncovalent interactions

Noncovalent interactions may be decomposed into the global electrostatic and the van der Waals terms. The van der Waals (vdW) energy is repulsive at short distance due to the explicit overlap of electron clouds named Pauli repulsion, and it is attractive surrounding the ideal vdW contact distance to account for dispersion attractions between atoms. One of the oldest functions for vdW interaction is the Lennard-Jones potential between atom *i* and *j*[28]:

$$U_{ij}^{LJ} = \varepsilon \left[\left(\frac{r_{ij}^0}{r_{ij}} \right)^{12} - \left(\frac{r_{ij}^0}{r_{ij}} \right)^{6} \right]$$

ε is the energy well depth and r the interatomic distance. Alternative vdW functions can be also used such as the "buffered 14-7"[29] and the Buckingham potentials[30]. The Buckingham potential is represented by:

$$U_{ij}^{Bucking\ ham} = Ae^{-Br} - \frac{C}{r^6}$$

A, *B* and *C* are parameters. Even if this potential is a better description of strong repulsion, it goes toward -∞ for very small distance and it may become difficult to compute. For these reasons, this potential is not usually employed in current force fields.

The electrostatic interaction E^{elec} between two partial charges (q_i and q_j) is generally described by a simple Coulomb potential:

$$E_{ij}^{elec} = \frac{q_i q_j}{\varepsilon\, r_{ij}}$$

with ε the effective dielectric constant. This term is easy to compute but the reproduction of the electrostatic potential around a molecule can be improved by a more sophisticated description. The main idea is to include contributions from higher order electric moments, such as dipoles, quadrupoles, octupoles, etc. Several force fields use such a multipolar description, e. g. AMOEBA[31-33], SIBFA[34], NEMO (Non Empirical Molecular Orbital)[35], EFP (Effective Fragment Potential)[36], and the PE (Polarizable Embedding) QM/MM method[37]. One way to obtain these higher order moments is the distributed multipole analysis (DMA) of quantum electron densities as developed by Stone[38].

Furthermore, an important part of electrostatic interactions comes from the presence of a third body, named the polarization effects. This coupling between the local environment and the multipoles model is neglected in standard force fields[39]. The Drude oscillator methods, also known as "shell models", have been developed to account for polarization effects. Polarization is

introduced by attaching massless charged particles to atoms via harmonic springs, which are allowed to move in response to the electric field[40]. Alternatively, one can compute the induced dipole moment $\mu_{i,\alpha}^{ind}$ explicitly in response to the external field $E_{i,\alpha}$:

$$\mu_{i,\alpha}^{ind} = \alpha_i E_{i,\alpha}$$

where, α_i is the polarizability of atom i. This is the model implemented in the AMOEBA force field[41]. Furthermore, the SIBFA force field[34] for organic molecules and proteins, developed by Gresh and collaborators includes both polarization effects and charge transfer. This force field is formulated and calibrated on the basis of *ab initio* super molecule computations. It has shown very successful performances, especially on biomolecules attached to a metal cation. Unfortunately, long molecular dynamics simulations with the aim of extracting vibrational spectra are out of the scope of this accurate force field.

AMOEBA has already been extensively tested for the simulation of various systems and it has proved particularly suited to the computation of dynamical properties of metal cations of biological interest[42-44]. Hence, we choose AMOEBA force field for this thesis as it was previously extended to the structure and the vibrational spectra of gas-phase biomolecules in the laboratory[45,46]. A brief description of the force field is presented in the next section.

2.4. AMOEBA, a polarizable force field based on distributed multipoles

The AMOEBA (Atomic Multipole Optimized Energetics for Biomolecular Applications) force field is an effort to achieve chemical accuracy of conformational and interaction energies similar to that of reliable quantum mechanical models; it is developed by Ponder et al.[31,47,48]. This force field provides a computationally efficient potential energy function relative to QM calculations and is able to capture the polarization effects that fixed charged force fields lack.

Accordingly, AMOEBA demonstrated a series of successes with modeling the hydration of the monovalent ions K^+, Na^+, and Cl^-[42], as well as divalent ions such as Ca^{2+} and Mg^{2+}[43]. Additional studies of ion hydration with a first molecular mechanics step followed by QM calculations have concluded that AMOEBA is efficient in generating ensemble configurations for post-analysis, and that fixed charge force fields have difficulties with mimicking local charge rearrangements[49]. It has been shown also that, when used in conformational searches, AMOEBA is able to reproduce ca. 80% of the conformations of flexible organic molecules and peptides[33,45]. Recently, AMOEBA vs. DFT study indicates that the physically based energy terms of the force field allow for a sound description of the interactions in packed conformations, in which accurate electrostatic, polarization and dispersion terms are required[50]. In addition, the force field showed a better performance in comparison with OPLSAA, AMBER and CHARMM for the conformations and stabilities of a β-hairpin model thanks to an accurate treatment of essential aromatic-aromatic and cation−π interactions[51]. A valence bond (VB) model parameterized on QM calculations has been incorporated in AMOEBA for aqueous Cu^{2+} and Zn^{2+} ions with a significant improvement of specific geometries and relative energies in comparison with the original AMOEBA[52].

We chose it for its ability to reproduce various physicochemical properties both in the gas condensed phases, and also for its complete intramolecular flexibility which is necessary for the computation of vibrational spectra. The AMOEBA force field was used in this work as implemented in the TINKER software package[53].

2.4.1. Energetic terms in AMOEBA

The bonded interactions are represented by bond stretching, angle bending, bond-angle cross term, out-of-plane bending and torsional rotation terms. Non-bonded interactions are described by a van der Waals term and permanent and induced electrostatic contributions. The components of the potential energy of the force field are:

$$U_{potential} = [U_{bond} + U_{angle} + U_{torsion} + U_{stretch\text{-}bend} + U_{oop}] + [U_{vdW} + U_{ele}^{perm} + U_{ele}^{ind}]$$

The valence functional terms are transferred from the MM3 model[22,54]. The out-of-plane bending is controlled by a Wilson-Decius-Cross function at sp^2-hybridized trigonal centers[55]. The different energy terms are:

$$\boldsymbol{U_{bond}} = K_b\,(b - b_0)^2[1 - 2.55(b - b_0) + \frac{7}{12} * 2.55^2(b - b_0)^2]$$

$$\boldsymbol{U_{angle}} = K_\theta(\theta - \theta_0)^2\,[1 - 0.014(\theta - \theta_0) + 5.6 \times 10^{-5}(\theta - \theta_0)^2 - 7.0 \times 10^{-7}(\theta - \theta_0)^3 + 2.2 \times 10^{-8}(\theta - \theta_0)^4]$$

$$\boldsymbol{U_{stretch-bend}} = K_{b\theta}\,[(b - b_0) + (b' - b_\theta')](\theta - \theta_0)$$

$$\boldsymbol{U_{oop}} = a\,\chi^2$$

$$\boldsymbol{U_{torsion}} = \sum_n K_{n\omega}\,[1 + \cos(n\omega \pm \gamma)]$$

Bond lengths b and b', bond angles θ, and energies are in units of Å, degrees and kcal.mol^{-1}, respectively, with the force constants K in corresponding units.

Van der Waals interactions

The repulsion-dispersion interactions are represented by a buffered 14-7 potential[29].

$$U_{vdW}^{ij} = \varepsilon_{ij}\left(\frac{1.07}{\rho_{ij} + 0.07}\right)^7\left(\frac{1.12}{\rho_{ij}^7 + 0.12} - 2\right)^{14}$$

Here, the potential is a function of the separation distance, R_{ij}, between atoms i and j and $\rho_{ij} = \frac{R_{ij}}{R_{ij}^0}$where R_{ij}^0 is the minimum energy distance. In addition, ε_{ij} is the potential well depth. The combination rules for atom pairs are defined as:

$$\varepsilon_{ij} = \frac{4\varepsilon_{ii}\varepsilon_{jj}}{(\varepsilon_{ii}^{1/2} - \varepsilon_{jj}^{1/2})^2}$$

$$R_{ij}^0 = \frac{(R_{ii}^0)^3 + (R_{jj}^0)^3}{(R_{ii}^0)^2 + (R_{jj}^0)^2}$$

Electrostatic interactions

The electrostatic term is usually the largest contribution to the energy of a polar system. The AMOEBA force field contains both permanent and induced atomic multipoles to take into account environmental effects in comparison with simple fixed-charge force fields[56]. The permanent atomic multipoles at each atom include monopole, dipole and quadrupole moments. Permanent electrostatic interactions are computed with higher order moments.

$$M_i = [q_i, d_{ix}, d_{iy}, d_{iz}, Q_{ixx}, Q_{ixy}, Q_{ixz}, \dots, Q_{izz}]^T$$

where the permanent atomic multipoles, M_i, are included for each atom, *i*, *q* is its point charge, *d* its dipole and *Q* its quadrupole components, respectively.

Partial charges as well as dipoles and quadrupoles are extracted from high-level QM calculations on small molecules. The multipoles are localized at atomic centers in a local frame defined by bonded or adjacent atoms. In principle, the choice of the frame should respect local symmetry such that axes are placed along major chemical determinants. The multipole sites are defined using a "z-then-x" local frame.

Polarization term

Polarization effects are explicitly included in the electrostatic component via atomic dipole induction. Electronic polarization describes the redistribution of electron density due to an external field; it represents a major contribution to the overall many-body energetic description of molecules, clusters and the condensed phase. Polarizable point dipoles are utilized at atomic centers to reproduce these effects. The induced dipole, $\mu_{i,\alpha}^{ind}$, at site *i* is:

$$\mu_{i,\alpha}^{ind} = \alpha_i \left(\sum_j T_\alpha^{i,j} M_j + \sum_{j'} T_{\alpha\beta}^{i,j'} \mu_{j'\beta}^{ind} \right)$$

where α_i is the atomic dipole polarizability, *T* is the usual interaction matrix for sites *i* and *j*, and M_j the permanent multipole components. The first term in the equation represents the dipole on site *i* induced by the permanent multipoles at other sites while the second term corresponds to the dipole on site *i* induced by the induced dipoles produced at the other atoms. To avoid polarization catastrophe at very short distance, a polarization-damping scheme is used via a smeared charge distribution as proposed by Thole[57,58]:

$$\rho = \frac{3a}{4\pi} \exp(-au^3)$$

with $u = R_{ij} / \left(\alpha_i \alpha_j\right)^{1/6}$. u is the effective distance between atoms *i* and *j*, α_i and a_j are the respective atomic dipole polarizabilities, and a is a dimensionless width parameter of the smeared charge distribution which controls the strength of the damping.

2.4.2. AMOEBA parameter adjustment and validation

There are two necessary steps before running a calculation with AMOEBA for a new system. The first is to determine the parameter set. AMOEBA parameters are available for various systems including the most common organic molecules and almost all amino acids, different metal ions, solvent, etc., however new parameters can be extracted if necessary. The second step is the validation of this parameter set by comparing results to those of benchmark electronic structure calculations for e.g. relative conformational energies. The force field should be able to reproduce or predict relative conformational energies of model systems, which are complex enough to contain realistic features, but simple enough to be treated by electronic structure methods that can yield reliable benchmark results. Results of validation are presented in chapter 4, by comparison of AMOEBA and MP2 relative energies and structures.

Extraction of atomic multipoles

Distributed multipole analysis (DMA) developed by Stone[59,38,60] is a simple, fast procedure to assign multipole moments to each atom or other specified sites in a molecule. In this model, point charges, dipoles and higher multipoles are derived from QM electron density and are distributed at the different sites of the molecule. A key feature of DMA is the use of multiple sites for the multipole expansion, typically at atom centers, to provide accurate multipole expansion at short intermolecular distances. It was shown that the use of higher-order moments yielded better agreement with the quantum mechanical reference potential than point charges representation[61]. Fowler and Buckingham[62,63] showed that a multipole representation for electrostatics is essential to obtain correct geometries of complexes dominated by hydrogen bonds or nonpolar interactions.

For the AMOEBA force field, only point charges, dipoles and quadrupoles are used. They are calculated using the GDMA (Gaussian Distributed Multipole Analysis) program[59]. A general scheme with the different steps of the procedure is presented in Figure 1. For this purpose, we used the "original" DMA formulation, and not the newer spatial partitioning scheme recently proposed[59]. During the procedure, special care was taken to use an appropriate hydrogen atom radius, as recommended by Stone. Indeed, when explicit multipoles are introduced for hydrogen atoms, they have been found to be much more sensitive to the atomic radius than the ones of the other atoms. A radius of 0.325 Å for hydrogen atoms has been found to give better results than the default of 0.65 Å, especially for hydrogen-bonded systems. The "Poledit" routine was then used to finalize the extraction of the multipoles and to rotate them in the "z-than-x" local frame for each atom. Finally to ensure an accurate reproduction of the electrostatic potential, the resulting multipoles are refined by fitting the QM electrostatic potential using the "Potential" routine available in Tinker. Usually, the electrostatic potential is obtained at the MP2/cc-pVTZ level. The application of this procedure to the tryptamine system and the associated results on chemical properties will be presented in chapter 4.

An alternative method is to calculate the distributed multipoles from the Gaussian Electrostatic Model (GEM)[64-66]. Recently, Cisneros reported a successful application of GEM distributed multipoles in AMOEBA[67] and an example of its use for IR spectroscopy is presented in chapter 4.

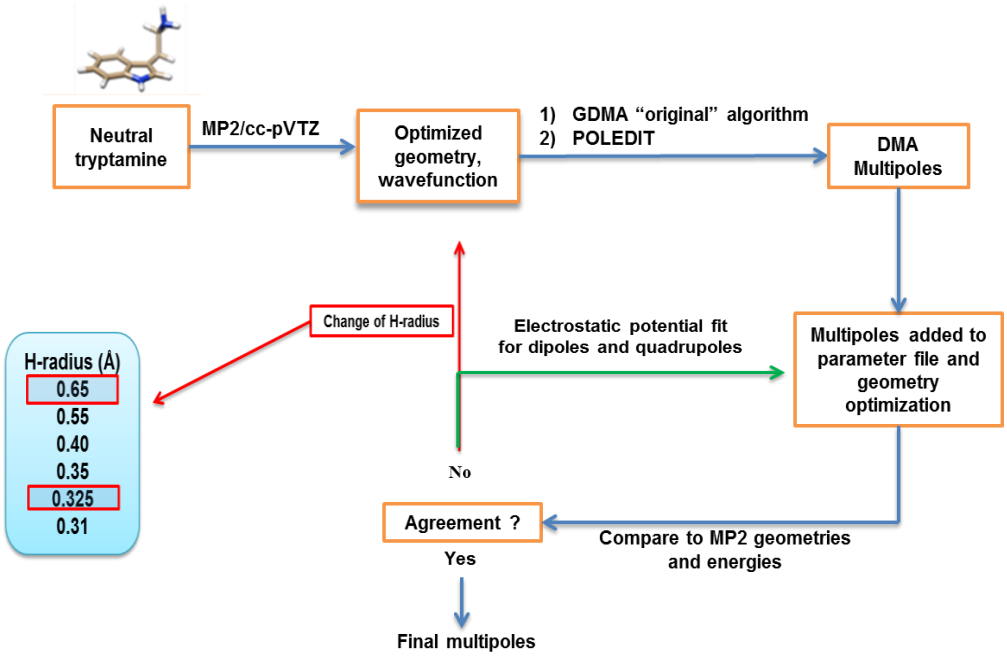


Figure 1 Scheme of the multipole extraction procedure on neutral tryptamine as an example.

3. Molecular dynamics simulations

Two main types of simulations are used in computational chemistry to study macroscopic and microscopic properties of chemical and biochemical systems: Molecular Dynamics (MD) and Monte Carlo (MC). These techniques are described in detail in the textbooks "Computer Simulations of Liquids" (Allen and Tildesley 1987) and "Understanding molecular simulation" (Frenkel and Smit 2002)[68,69].

All simulations in this work were done using MD, which provides explicit access to time-dependent properties, an advantage over MC for our purposes. The method is based on calculating the forces between atoms in a system and allowing them to move according to Newton's law of motion:

$$F_i = m_i a_i$$

with F_i is the force applied on particle i, m_i is the mass of particle i and a_i is the acceleration of i. The particles are moved over a small time-step, forces and velocities are recalculated and the system is moved forward a new time-step. For each time-step, the properties of the system such

as energy and temperature are monitored. A simulation is carried out for a number of time-steps necessary to obtain reliable averages. The result of a simulation is a trajectory that specifies how the positions and velocities of the particles vary with time.

3.1. Integration algorithms

In order to obtain the positions and the forces of the system to generate the MD trajectory, many algorithms have been developed over the years to solve numerically the equations of motion. The widely used ones are Verlet, "leap-frog", velocity Verlet and Beeman algorithms. In these methods, the positions *r* are approximated as Taylor series expansions. In this work we have used the Beeman[70] algorithm :

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

$$v(t+\delta t) = v(t) + v(t)\delta t + \frac{1}{3}\,a(t)\delta t + \frac{5}{6}\,a(t)\delta t - \frac{1}{6}\,a(t-\delta t)\delta t$$

since it provides a more accurate expression of the positions and velocities and a better energy conservation thanks to a direct calculation of the kinetic energy from the velocities, as compared to simpler algorithms such as Verlet.

3.2. Thermostats

MD simulations are often performed in the microcanonical ensemble (NVE) in which the total energy is conserved. However, simulations can be also performed in the canonical ensemble (NVT) to modulate accurately the temperature of the system, for example to match some

experimental conditions. Then, a thermostat is required and different methods are available such as Berendsen[71], Andersen[72] or Nosé-Hoover thermostats[73,74], all implemented in TINKER. The value of the temperature is related to the kinetic energy as follows:

$$\sum_{i=1}^{N} \frac{|P_i|^2}{2m_i} = \frac{k_B T}{2}(3N - N_c)$$

where, N_c is the number of constraints and so $3N-N_c = N_{df}$ is the total number of degrees of freedom. The average temperature T corresponds to the macroscopic temperature. P_i is the total momentum of particle i of mass m_i.

One way to control the temperature is the rescaling of velocities[75] according to :

$$P'_i = \sqrt{\frac{T_0}{T}} P_i$$

where, T_0 is the desired temperature, and T the actual temperature calculated from the kinetic energy. However, this method leads to discontinuities in the momentum part of the phase space trajectory due to the rescaling procedure at each time step. Furthermore, the method does not allow temperature fluctuations as in the canonical ensemble.

The Berendsen thermostat is similar to the velocity rescaling approach but gives a time scale for the updating of the velocities, rather than assuming that they are completely scaled to the target temperature at each time step. In this approach, the system is weakly coupled to an external heat bath with the coupling constant τ and the scaling factor λ is:

$$\lambda^2 = 1 + \frac{\delta t}{\tau}\left(\frac{T_0}{T(t)} - 1\right)$$

Here, δt is the time step in the MD simulation. Typically, $\tau = 0.1 - 0.4$ ps for a time step around 1 fs. However, this method also suffers from the same problems as velocity rescaling because the energy fluctuations are not captured correctly. The Andersen thermostat introduces a stochastic

element to the temperature by having random collisions of molecules with an imaginary heat bath at the desired temperature. The velocity of the particle is reassigned randomly from a Maxwell-Boltzmann distribution at the desired temperature. However, the presence of random collisions causes the velocities of particles to decorrelate. As a result, this thermostat does not produce smooth trajectories.

The Nosé-Hoover thermostat was introduced by Nosé and subsequently developed by Hoover[73,74]. It is nowadays broadly used in the simulation community. In this method, a canonical ensemble is rigorously generated during the MD simulation. Some methods have been developed to increase the robustness of the approach such as Nosé-Hoover chains in which multiple heat baths are linked to enhance temperature equilibration.

Earlier in our group[46], the three thermostats were tested on the N-methyl acetamide molecule for MD spectra and Nosé-Hoover thermostat was found to be the most accurate where both Berendsen and Andersen ones showed artifacts during the simulations. Furthermore, these thermostats were tested on zinc-water complexes (see chapter 3) and reliable results were obtained with the Nosé-Hoover thermostat. Consequently, all the MD simulations in this thesis were carried out using only this thermostat.

3.3. IR spectra using MD simulations

The experimental techniques for IR spectroscopy are well developed for chemical structure determination and bioanalytical applications. Theoretical calculations are a useful and complementary tool to link vibrational spectra to molecular structures. In the most basic approach, the computations are carried out within the harmonic approximation. This means that the potential energy surface is assumed to behave like a harmonic potential around the minimum

obtained by geometry optimization. Hence vibrational frequencies are directly derived from the second derivatives of the potential energy with respect to molecular coordinates. One or several scaling factors are then applied to take into account temperature and anharmonic effects which are not obtained from the harmonic calculations. A more direct way to include both anharmonicity and temperature effects in the computed spectrum is to perform molecular dynamics simulations. Classical, Born-Oppenheimer or Car-Parrinello MD associated to the Fourier transform of the Dipole Auto Correlation Function (DACF) was shown to be an effective approach to compute IR spectra for small molecules of biological interest both in gas-phase and in solution.

In 1977, Noid *et al.* first used DACF to calculate spectra from MD simulations in the direct analysis of system dynamics through time series analysis[76]. The Wiener-Khinchin theorem is a fundamental part of signal analysis[77]. For a stationary random process $f(t)$, it relates the power spectrum $S_f(\omega)$ to the Fourier transform of the autocorrelation function $R_f(\tau)$.

$$S_f(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} R_f(\tau) e^{-j\,2\pi\omega\tau}\, d\tau$$

For applications such as simulated IR spectra from MD trajectories only a finite set of sampling points is available. For a signal known over an infinitely long interval $[-T, T]$, the normalized autocorrelation function is defined as:

$$R_f(\tau) = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} f(t) f(t+\tau) dt$$

In this framework, by choosing the dipole moment as the observable, the power spectrum obtained from the Fourier transform of its autocorrelation function is an infrared vibrational spectrum. Due to the discrete nature of molecular dynamics trajectories, the autocorrelation function applied to the dipole moment $\vec{\mu}$ is defined as:

$$C(m) = \frac{1}{n}\sum_{i=1}^{n} \vec{\mu}(m+i).\vec{\mu}(i)\, with \;\; m \leq n$$

where, m is the time-lag of the autocorrelation function, n is the total number of time steps and $\vec{\mu}(i)$ the dipole at time step i. The spectral density function $I(\omega)$ is then obtained by computing this discretized autocorrelation function $C(m)$ followed by its Fourier transform.

Previously in our group, the Tinker program was modified in order to calculate and store the instantaneous dipole moment of the system resulting in an increase of the global CPU time as a function of the number of atoms[46]. The Fourier program was used to calculate the Fourier transform of the dipole auto-correlation function in order to obtain the power spectrum[78]. Classical time evolution of the nuclei induces an approximation in the calculation of the spectral density. The formula for the coefficient of infrared absorption by unit of length in the approximation of the electric dipole is given as a function of the spectral density by:

$$\alpha(\omega) = \frac{4\pi^2\omega}{3V\hbar cn(\omega)}\left(1 - e^{-\beta\hbar\omega}\right)I(\omega)$$

with V the volume of the sample, $n(\omega)$ the index of refraction at temperature $T = \frac{1}{k_B\beta}$. For gas phase spectroscopy, n is close to 1. In the harmonic approximation (HA), the final absorption coefficient has the expression: $\alpha_{HA} = \frac{4\pi^2\omega^2}{3Vcn(\omega)}I_{CL}(\omega)$. More details about the quantum corrections can be found in reference[79] and [80].

In this study, MD trajectories aimed at computing IR spectra are propagated by a Beeman integrator with a time step of 0.1 fs. At each time step, structures are saved and the corresponding dipole moment is recorded. This time step of 0.1 fs is small enough to be well below the X-H (X=O, N, C) full stretch motion cycle, the shortest in the systems considered herein. As discussed in the previous section, the Nosé-Hoover thermostat has been found to be

the best to maintain the temperature during the simulation. Usually, an equilibration period at the desired temperature is done over a period of 50 ps, then simulations of 200 ps are performed from different initial conditions to record the dipole moment. During this period of 200 ps, the system is thermally isolated except potentially for the starting point. In general, three to five trajectories have been started from the same initial geometry and the kinetic energy was distributed randomly over the atomic velocities, yielding different initial conditions. Then, statistical average of the dipole autocorrelation function is performed to limit artifacts due to the choice of starting point.

The resulting spectra are plotted using Origin 9.1 software. The Savitzky–Golay smoothing algorithm was used as a noise reduction technique[81]. This procedure is more efficient than simple point average. The idea is to perform a least squares fit of a small set of consecutive data points to a polynomial and take the calculated central point of the fitted polynomial curve as the new smoothed data point. It was shown that a set of integers could be derived as weighting coefficients to carry out the smoothing operation. The use of these weighting coefficients, known as convolution integers, turns out to be exactly equivalent to fitting the data to a polynomial, but it is computationally much more effective.

4. References

(1) Ostlund, N. S.; Szabo, A. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure and Theory*; Dover Publications: Mineola NY, **1996**.

(2) Cramer, C. J. *Essentials of Computational Chemistry: Theories and Models, 2nd ed.*; John Wiley & Sons: Chichester, **2008**.

(3) Jensen, F. *Introduction to Computational Chemistry, 1st ed.*; John Wiley & Sons: Chichester, **1999**.

(4) Zhao, Y.; Truhlar, D. G. *Acc. Chem. Res.* **2008**, *41*, 157.

(5) Grimme, S. *WIREs Comput. Mol. Sci.* **2011**, *1*, 211.

(6) Christiansen, O.; Koch, H.; Jorgensen, P. *Chem. Phys. Lett.* **1995**, *243*, 409.

(7) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

(8) Frisch, M. J.; Pople, J. A.; Binkley, J. S. *J. Chem. Phys.* **1984**, *80*, 3265.

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

(10) Lambrecht, D. S.; McCaslin, L.; Xantheas, S. S.; Epifanovsky, E.; Head-Gordon, M. *Mol. Phys.* **2012**, *110*, 2513.

(11) Mardirossian, N.; Lambrecht, D. S.; McCaslin, L.; Xantheas, S. S.; Head-Gordon, M. *J. Chem. Theory Comput.* **2013**, *9*, 1368.

(12) Frisch, M. J. T., G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; 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.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09, Gaussian, Inc., Wallingford CT,* **2009**.

(13) Schaftenaar, G.; Noordik, J. H. *J. Comput. Aid. Mol. Des.* **2000**, *14*, 123.

(14) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.

(15) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. B* **2001**, *105*, 6474.

(16) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225.

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

(18) MacKerell, A. D.; Jr.; Brooks, B.; Brooks, C. L.; III; Nilsson, L.; Roux, B.; Won, Y.; Karplus, M. *CHARMM: The Energy Function and Its Parameterization with an Overview of the Program*; In Schleyer, P.v.R.; et al. The Encyclopedia of Computational Chemistry 1: John Wiley & Sons: Chichester, **1998**.

(19) Oostenbrink, C.; Villa, A.; Mark, A. E.; Van Gunsteren, W. F. *J. Comput. Chem.* **2004**, *25*, 1656.

(20) Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. *J. Am. Chem. Soc.* **1992**, *114*, 10024.

(21) Allinger, N. L. *J. Am. Chem. Soc.* **1977**, *99*, 8127.

(22) Allinger, N. L.; Yuh, Y. H.; Lii, J. H. *J. Am. Chem. Soc.* **1989**, *111*, 8551.

(23) Lii, J. H.; Allinger, N. L. *J. Am. Chem. Soc.* **1989**, *111*, 8566.

(24) Allinger, N. L.; Chen, K. S.; Lii, J. H. *J. Comput. Chem.* **1996**, *17*, 642.

(25) Morse, P. M. *Phys. Rev.* **1929**, *34*, 57.

(26) Shi, S. H.; Yan, L.; Yang, Y.; Fisher-Shaulsky, J.; Thacher, T. *J. Comput. Chem.* **2003**, *24*, 1059.

(27) Mayo, S. L.; Olafson, B. D.; Goddard, W. A. *J. Phys. Chem.***1990**, *94*, 8897.

(28) Jones, J. E. *Proc. R. Soc. Lond. A* **1924**, *106*, 463.

(29) Halgren, T. A. *J. Am. Chem. Soc.* **1992**, *114*, 7827.

(30) Buckingham, R. A. *Proc. R. Soc. Lond. A* **1938**, *168*, 264.

(31) Ren, P.; Ponder, J. W. *J. Phys. Chem. B* **2003**, *107*, 5933.

(32) Ponder, J. W.; Wu, C.; Ren, P.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio, R. A.; Head-Gordon, M.; Clark, G. N. I.; Johnson, M. E.; Head-Gordon, T. *J. Phys. Chem. B* **2010**, *114*, 2549.

(33) Rasmussen, T. D.; Ren, P.; Ponder, J. W.; Jensen, F. *Int. J. Quant. Chem.* **2007**, *107*, 1390.

(34) Gresh, N.; Cisneros, G. A.; Darden, T. A.; Piquemal, J-P. *J. Chem. Theory Comput.* **2007**, *3*, 1960.

(35) Engkvist, O.; Åstrand, P. O.; Karlström, G. *Chem. Rev.* **2000**, *100*, 4087.

(36) Gordon, M. S.; Fedorov, D. G.; Pruitt, S. R.; Slipchenko, L. V. *Chem. Rev.* **2012**, *112*, 632.

(37) Sneskov, K.; Schwabe, T.; Kongsted, J.; Christiansen, O. *J. Chem. Phys.* **2011**, *134*, 104108.

(38) Stone, A. J. *Chem. Phys. Lett.* **1981**, *83*, 233.

(39) Patel, S.; Mackerell, A. D.; Brooks, C. L. *J. Comput. Chem.* **2004**, *25*, 1504.

(40) Anisimov, V. M.; Lamoureux, G.; Vorobyov, I. V.; Huang, N.; Roux, B.; MacKerell, A. D. *J. Chem. Theory Comput.* **2005**, *1*, 153.

(41) Ren, P.; Ponder, J. W. *J. Comput. Chem.* **2002**, *23*, 1497.

(42) Grossfield, A.; Ren, P.; Ponder, J. W. *J. Am. Chem. Soc.* **2003**, *125*, 15671.

(43) Jiao, D.; King, C.; Grossfield, A.; Darden, T. A.; Ren, P. *J. Phys. Chem. B* **2006**, *110*, 18553.

(44) Clavaguéra, C.; Sansot, E.; Calvo, F.; Dognon, J. P. *J. Phys. Chem. B* **2006**, *110*, 12848.

(45) Semrouni, D.; Ohanessian, G.; Clavaguéra, C. *Phys. Chem. Chem. Phys.* **2010**, *12*, 3450.

(46) Semrouni, D. Énergétique et spectroscopie de peptides attachés à un cation métallique, champ de force de seconde génération et chimie quantique. Ph.D. Thesis. , École Polytechnique, **2010**.

(47) Ren, P.; Wu, C.; Ponder, J. W. *J. Chem. Theory Comput.* **2011**, *7*, 3143.

(48) Shi, Y.; Xia, Z.; Zhang, J.; Best, R.; Wu, C.; Ponder, J. W.; Ren, P. *J. Chem. Theory Comput.* **2013**, *9*, 4046.

(49) Beck, T. L. *J. Stat. Phys.* **2011**, *145*, 335.

(50) Semrouni, D.; Clavaguéra, C.; Ohanessian, G.; Parks, J. H. *J. Phys. Chem. B* **2013**, *117*, 1746.

(51) Zheng, X.; Wu, C.; Ponder, J. W.; Marshall, G. R. *J. Am. Chem. Soc.* **2012**, *134*, 15970.

(52) Xiang, J. Y.; Ponder, J. W. *J. Comput. Chem.* **2013**, *34*, 739.

(53) Ponder, J. W., TINKER: Software Tools for Molecular Design, Tinker 6; Washington University School of Medicine: Saint Louis, MO, **2009**.

(54) Lii, J. H.; Allinger, N. L. *J. Am. Chem. Soc.* **1989**, *111*, 8576.

(55) Wilson, E. B.; Decius, J. C.; Cross, P. C., Molecular Vibrations: The Theory of Infrared and Raman Vibrational Spectra. *New York: McGraw-Hill* **1955**.

(56) Kaminsky, J.; Jensen, F. *J. Chem. Theory Comput.* **2007**, *3*, 1774.

(57) van Duijnen, P. T.; Swart, M. *J. Phys. Chem. A* **1998**, *102*, 2399.

(58) Thole, B. T. *Chem. Phys.* **1981**, *59*, 341.

(59) Stone, A. J. *J. Chem. Theory Comput.* **2005**, *1*, 1128.

(60) Stone, A. J.; Alderton, M. *Mol. Phys.* **1985**, *56*, 1047.

(61) Williams, D. E. *J. Comput. Chem.* **1988**, *9*, 745.

(62) Fowler, P. W.; Buckingham, A. D. *Mol. Phys.* **1983**, *50*, 1349.

(63) Fowler, P. W.; Buckingham, A. D. *Chem. Phys. Lett.* **1991**, *176*, 11.

(64) Piquemal, J-P.; Cisneros, G. A.; Reinhardt, P.; Gresh, N.; Darden, T. A. *J. Chem. Phys.* **2006**, *124*, 104101.

(65) Cisneros, G. A.; Piquemal, J-P.; Darden, T. A. *J. Chem. Phys.* **2005**, *123*, 044109.

(66) Cisneros, G. A.; Piquemal, J-P.; Darden, T. A. *J. Chem. Phys.* **2006**, *125*, 184101.

(67) Cisneros, G. A. *J. Chem. Theory Comput.* **2012**, *8*, 5072.

(68) Allen, M. P.; Tildesley, D. J. In *Computer Simulations of Liquids*; Oxford University Press, New York: United States, **1987**.

(69) Frenkel, D.; Smit, B. *Understanding molecular simulation*; Academic Press, **2002**.

(70) Beeman, D. *J. Comput. Phys.* **1976**, *20*, 130.

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

(72) Andersen, H. C. *J. Chem. Phys.* **1980**, *72*, 2384.

(73) Nosé, S. *J. Chem. Phys.* **1984**, *81*, 511.

(74) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695.

(75) Woodcock, L. V. *Chem. Phys. Lett.* **1971**, *10*, 257.

(76) Noid, D. W.; Koszykowski; Marcus, R. A. *J. Chem. Phys.* **1977**, *67*, 404.

(77) McQuarrie, D. A. *Statistical Mechanics*; Harper Collins Publishers: New York, **1976**.

(78) Forbert, H.; Kleinschmidt, V. F. *Ruhr-Universität Bochum, **2002-2005** modified by A. Kohlmeyer, Version 21.02.2006.*

(79) Borysow, J.; Moraldi, M.; Frommhold, L. *Mol. Phys.* **1985**, *56*, 913.

(80) Marx, D.; Hutter, J. *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods*; Cambridge University Press, Cambridge, **2009**.

(81) Savitzky, A.; Golay, M. J. E. *Anal. Chem.* **1964**, *36*, 1627.

Chapter 3

Structure, dynamical coordination and IR spectra of the hydrated divalent metal ion Zn^{2+}

1. Introduction

Zinc solvation and its coordination behavior are of large interest in bioinorganic chemistry such as in the functioning of certain proteins and metalloenzymes[1-3]. Zinc is heavily used in industries and is thus abundant in waste waters[4]. Thus, zinc is a biologically important metal and to understand its biological transport and impact in aqueous environments, a complete understanding of the hydration and organization of the water molecules around zinc in its divalent cation form Zn^{2+} is essential. To this end, establishing the number of water molecules in the inner solvent shells of Zn^{2+} has been the subject of numerous computational and experimental studies.

The coordination behavior of divalent zinc ion has been investigated in aqueous phase using a variety of experimental techniques, such as X-ray absorption fine structure (EXAFS)[5,6], Raman spectroscopy of aqueous $Zn^{2+}(ClO_4)_2$[7] and near-IR spectroscopy studies of aqueous ZnI_2[8]. Briefly, all of these experimental results are consistent with a coordination number (CN) of six, which until recently has been the commonly accepted configuration.

More recently, the coordination behavior of micro-hydrated Zn^{2+} has also been investigated by using various mass spectrometry methods, such as threshold collision-induced dissociation (CID) of gaseous $Zn^{2+}(H_2O)_n$ complexes (n = 6-10)[9,10], infrared photodissociation (IRPD) spectroscopy and blackbody infrared radiative dissociation (BIRD) (n = 6-12)[11]. All of these experiments suggest that Zn^{2+} favors a coordination of five in the gas phase. These results are consistent with the lowest-energy structures calculated at the MP2 level of theory, but they contrast with the well-established coordination number of 6 in the aqueous phase.

Quantum chemistry (QM) calculations have shown that the lowest-energy complex can have an inner shell of four, five, or six water molecules depending on the level of theory (DFT or *ab initio*) and basis set used[9,12-16]. For example, DFT calculations at the B3LYP/6-311G(2d, 2p)

level suggest that Zn^{2+} prefers a coordination of four up to 12 water molecules[15]. However, Ali and coworkers[16] reported that Zn^{2+} is found to be preferably four coordinated for smaller sizes of hydrated clusters but attains six coordination for larger sizes ($n \geq 11$) at the B3LYP level using cc-pVDZ basis set for H and O and a split valence 6-31G (d, p) basis set for Zn. In the same context, Armentrout and coworkers[9] reported that the low-energy structures of $Zn^{2+}(H_2O)_n$ (n = 6-10) may have a first coordination shell with 4, 5 and 6 water molecules depending on the level of theory. Their findings show that, at both B3LYP and B3P86 levels, zinc favors a CN of 4 up to n = 8 where, it switches to 5. In contrast, M06 and MP2(full) calculations show a preference of an inner shell of 5 for n = 5 and 7-10, and of 6 for n = 6.

The CN and water residence time of Zn^{2+} in bulk water have also been studied by molecular dynamics simulations such as QM/MM-MD, *ab initio* molecular dynamics (AIMD) and combined *ab initio* and classical molecular dynamics methods (AIMD/MM)[16-21]. The results have shown that the first coordination shell of Zn^{2+} is 6, however these methods have not reached simulation times long enough to observe the exchange of water molecules between the first and second shells. The result showed that first hydration shell contains 6 water molecules. No exchanges were seen between the first and the second hydrations shells. However, many water transfers between the second hydration shell and the bulk were observed.

In recent years, as an alternative to QM, simulations with accurate force fields that include many-body effects such as SIBFA[22,23] (Sum of Interactions between Fragments *ab initio* computed) and AMOEBA[24] have been developed and gained interest to study Zn^{2+} complexes. Such techniques are computationally more efficient and provide reliable potential energy surfaces in comparison with QM. Particularly for Zn^{2+}, SIBFA, which treats both polarization and charge transfer contributions, has been shown to be accurate and has enabled the study of

binding modes and intermolecular interaction energies of Zn^{2+} with various ligands and of active sites of metalloproteins[25-28], however without a dynamic point of view. On the other hand, AMOEBA has been shown to be particularly suited to the computation of dynamical properties of metal cations[29-31]. Even without an explicit treatment of charge transfer, it has been proved to be very accurate in reproducing the dynamical behavior of divalent zinc in water. In this regard, Ren and coworkers[32] reported first-shell water CN and water residence time of the zinc divalent cation from MD simulations with the AMOEBA force field in the condensed phase (512 water molecules, 298 K). The results obtained from their study show that the first-shell CN of Zn^{2+} is 6 and that the water residence time in the first solvation shell around Zn^{2+} is at least 2 ns. This is consistent with an earlier experimental studies on Zn^{2+} to water binding time with incoherent quasi-elastic neutron scattering (IQENS) by Salmon et al.[33] reported that the lower and upper limits of water residence time in the first shell of zinc(II) in solution are 0.1 and 5 ns, respectively.

Furthermore, micro-hydration of several divalent transition metal ions including zinc with 5–8 water molecules has been reported by Williams et al. using the IRPD spectroscopy technique[34]. One of the most relevant spectra is that of $Zn^{2+}(H_2O)_6$, reported in Figure 1. The latter figure shows that at 215 K, spectral intensities exist in both the bonded-OH (2800-3500 cm^{-1}) and free-OH (3550-3800 cm^{-1}) stretch regions while a CN of six would not involve water-water interactions. The weak band that occurs in the H-bonding region centered near 3400 cm^{-1} is assigned to inner-shell water molecules that donate one or two H-bonds to double acceptor, second shell water molecules. This assignment of at least one water molecule to the second shell indicates that the average CN is less than 6. In addition, they have also reported the effect of temperature on O-H stretch bands by measuring spectra of $Mn^{2+}(H_2O)_6$ and $Ni^{2+}(H_2O)_6$ clusters

at 305 and 331 K, respectively. These spectra are compared to the lower temperature (215 K) experiments in Figure 2. The results show broadening of the free-OH bands at higher temperature. Particularly for $Mn^{2+}(H_2O)_6$, the antisymmetric stretch of free-OH bonds has shoulders on both sides. In addition, a peak in the H-bonding region at ~3400 cm^{-1} appears in the higher temperature spectrum of $Mn^{2+}(H_2O)_6$ but not in the Ni^{2+} one. This peak indicates that the population of water molecules in the second solvation shell increases for $Mn^{2+}(H_2O)_6$ at 305 K leading to an average CN smaller than 6. However, analogous experiments could not been carried out for Zn^{2+}, although this would have been particularly interesting in the present context. $M^{2+}(H_2O)_6$ ions being formed via collision dissociation of larger clusters, it turned out that they could not be formed in sufficient abundance for IRPD spectroscopy at elevated temperatures in the case of Zn^{2+} (E.R. Williams, private communication).

In the present work, we have aimed at providing the first reliable values of average coordination numbers of Zn^{2+} at three different temperatures ranging from 200 to 300 K for clusters ranging from 6 to 216 water molecules using the AMOEBA force field. The study has the aim to make the bridge between gas phase and condensed phases. Transitions between different coordination types at different temperatures are characterized. In a second part, finite temperature IR spectra of hydrated divalent zinc cations are calculated in an attempt to characterize charge and temperature effects on the spectra.

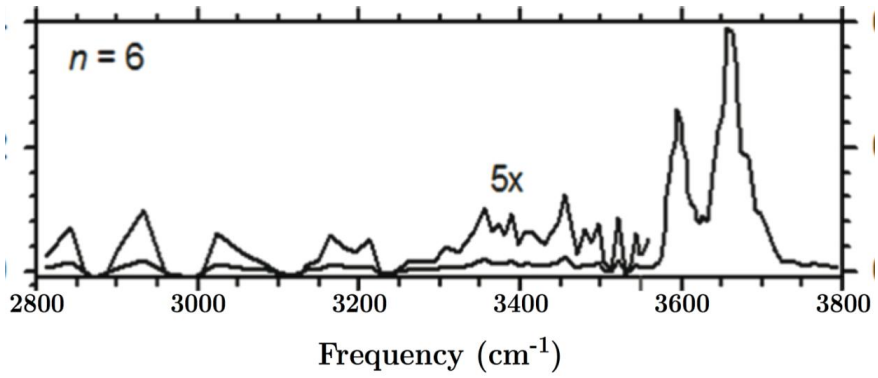


Figure 1 IRPD spectrum of $Zn^{2+}(H_2O)_6$ from 2800 – 3800 cm^{-1} at 215 K. The spectrum is taken from reference34.

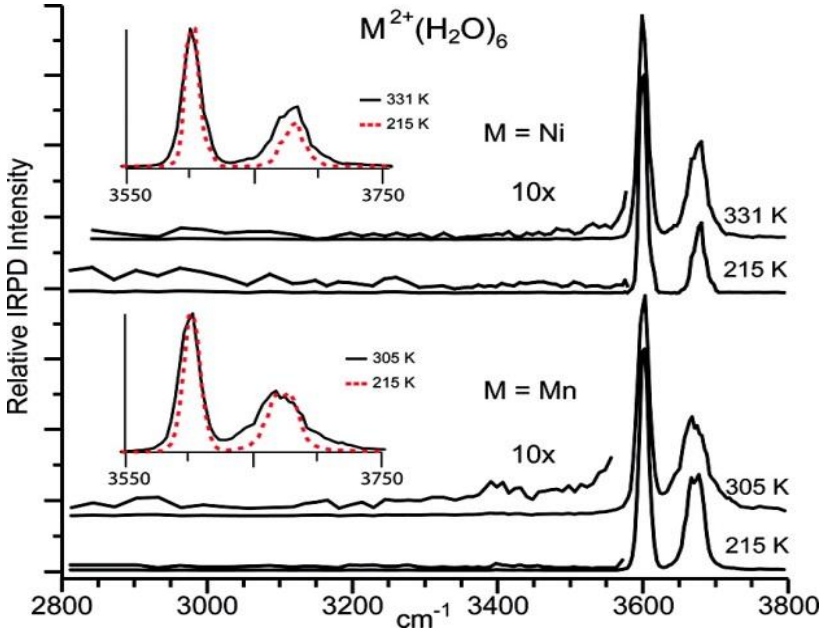


Figure 2 Comparison of higher temperature IRPD spectra of $Ni^{2+}(H_2O)_6$ and $Mn^{2+}(H_2O)_6$ (331 K and 305 K, respectively) to spectra acquired at 215 K. Inset shows an overlay of the spectra in the free-OH region. The spectra are taken from reference 34.

2. Computational details

2.1. Quantum chemistry calculations

In the present work, we have selected two small clusters of zinc-water complexes with six and seven water molecules. The low energy structures for these cluster sizes were taken from earlier reports[9,10]. Four starting structures were selected for each cluster size of $Zn^{2+}(H_2O)_6$, i.e.

6+0, 5+1_AA, 4+2_2AA, 4+2_AA_A, and of $Zn^{2+}(H_2O)_7$, i.e. 6+1_AA, 5+2_2AA, 4+3_2AA_A, 4+3_AA_2A. These different starting structures were chosen for diversity in the numbers of water molecules in first, second and third solvation shells and, more importantly, hydrogen-bonding motifs. The following nomenclature describes these different hydrogen bonding patterns. When two different inner shell water molecules are hydrogen bonded to a single second shell water molecule, the latter is a double acceptor (AA), whereas a second shell water molecule that forms only one hydrogen bond to an inner shell water molecule is a single acceptor (A). For example, a 4+2_AA_A structure contains 4 water molecules in the first and 2 water molecules in the second solvation shell with one "AA" and one "A" type H_2O. These bonding patterns are represented in Figure 3. Water molecules in the third and fourth shells hydrogen-bond to second and third shells, respectively, are also named accordingly. For each structure, the geometry optimization and frequency calculation were performed at DFT/B3LYP and DFT/M06 levels using 6-311+G(d,p) and aug-cc-pVTZ basis sets. Optimized geometries and vibrational frequencies were obtained at the MP2(full)/6-311+G(d,p) level. In addition, single point energy calculations were performed at MP2/aug-cc-pVTZ//MP2/6-311+G(d,p) level.

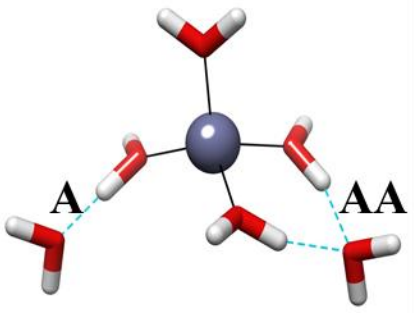


Figure 3 4+2_AA_A structure with "AA" and "A" type hydrogen bonding motifs. The number 4 stands for the number of water molecules in the first solvation shell and 2 stands for the number of water molecules in the second solvation shell.

2.2. AMOEBA force field parameters

The AMOEBA parameterization of Zn^{2+} was reported by Ren and coworkers[32] where parameters were derived from gas-phase *ab initio* calculations of the Zn^{2+}-H_2O complex. The Thole-based dipole polarization is fitted on energy decomposition analysis such as the Constrained Space Orbital Variations (CSOV). The damping factor "*a*" was adjusted so that the AMOEBA polarization energy matched the CSOV values as much as possible. Parameters for the van der Waals interactions, radius and well-depth, were derived by comparing the total ion-water binding energy computed by AMOEBA to the *ab initio* values at various metal-oxygen distances. The reported parameters for Zn^{2+} cation are 0.2600 $Å^3$, 0.2096, 2.6800 Å and 0.2220 $kcal.mol^{-1}$ for the polarizability, the damping coefficient, van der Waals radius and well-depth, respectively. Charge transfer is not taken into account explicitly; it is nevertheless implicitly included via the van der Waals parameters to a certain extent. It was shown to be limited for metal-water clusters with a complete solvation shell[32,35]. For H_2O, we have used available AMOEBA water parameters but for which Urey-Bradley, angle-bending and bond-stretching parameters were modified in order to reproduce symmetric and anti-symmetric O-H bond stretching frequencies for the isolated water molecule. The properties of bulk water, as reported in references[24,36,37] are not affected by these changes. These water parameters have been obtained previously in our group and are now available in the AMOEBA parameter set (see http://dasher.wustl.edu/ffe/distribution/params/water.prm).

2.3. Simulation protocol

For small clusters (n = 6-25), optimized structures obtained at the B3LYP level were selected as starting points for the simulations. To control the temperature, it is to use a suitable thermostat

to avoid bias in the simulations. Four different thermostats, Nosé-Hoover, Berendsen, Andersen and Bussi, have been tested. We found that the system is kept trapped artificially in one conformation by using Berendsen and Andersen thermostats. Contrastingly, both Nosé-Hoover and Bussi thermostats were able to keep equilibrium between different conformations while favoring the most stable. As the Bussi thermostat generates large standard deviations in the mean coordination number values, the Nosé-Hoover thermostat has been chosen to run simulations at 200, 250 and 300 K. For each starting point, 10 molecular dynamics simulations of 5 ns have been performed with a time step of 1 fs and structures were saved every 10 fs. Some simulations of 50 ns have been also performed to see the influence of the sampling on the results. The CN of Zn^{2+} is defined as the total number of O atoms of water molecules directly attached to the Zn^{2+}. This CN has been estimated by calculating the total number of O atoms present within 3 Å from Zn^{2+}. The distance of 3 Å was taken based on the Zn^{2+}-O radial distribution function (see for example Figure 4 for $Zn^{2+}(H_2O)_{10}$. Final average of CN is obtained over 10 trajectories for each size and each starting structure. The cluster with 216 water molecules around Zn^{2+} was confined at 300 K by spherical boundary conditions with a van der Waals wall which is defined by a 12-6 Lennard-Jones potential. Tests were carried out with different values of the wall radius of 14, 14.5, 15, and 15.5 Å. It was found that the exact value of the radius has little influence on the first coordination shell. Finally, a value of 15 Å was chosen.

We have also calculated IR spectra of hydrated Zn^{2+} using MD simulations with AMOEBA force field by taking the Fourier transform of the autocorrelation function over time of the total dipole moment following the protocol described in chapter 2.

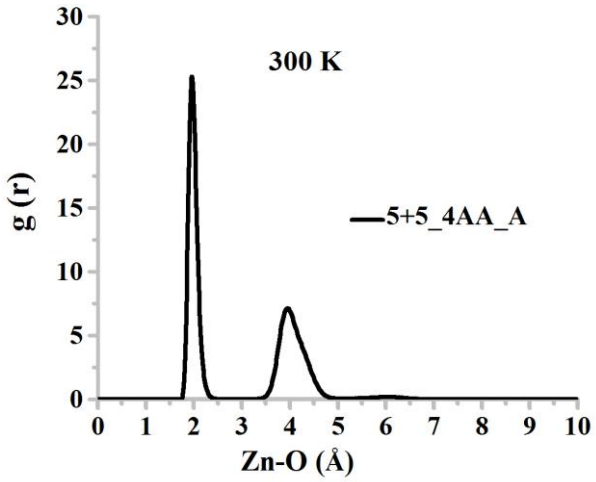


Figure 4 Zn-O radial distribution function for $Zn^{2+}(H_2O)_{10}$ obtained from the 5 ns MD simulation at 300 K (5+5_4AA_A structure as a starting point).

3. Structure and coordination number of hydrated Zn^{2+}

3.1. Clusters with 6 and 7 water molecules

3.1.1. Quantum chemistry results

The low energy structures for the clusters of $Zn^{2+}(H_2O)_6$ and $Zn^{2+}(H_2O)_7$ are presented in Figure 5. The relative electronic and free energies of these structures are provided in Table 1 at different QM levels. As previously highlighted[9,12-16], the results are very sensitive to the method (DFT vs. MP2) and to a lesser extent to the basis set used (6-311+G(d,p) and aug-cc-pVTZ). B3LYP relative energies are very different than with M06 or MP2 ones while M06 and MP2 relative energies are much more similar. For instance for the 4+2_2AA structure, the energy difference between B3LYP and M06 levels using the 6-311+G(d,p) basis set is 29.5 $kJ.mol^{-1}$, while it is only 4.9 $kJ.mol^{-1}$ between MP2 and M06. Furthermore, B3LYP optimized geometries provide larger Zn^{2+}-O distances (ca. 0.03-0.04 Å) than the ones obtained both at the M06 and MP2 levels which have been found to be very similar. The results show also that the basis set effect is smaller than the method effect, especially on relative electronic energies, e.g. 4 $kJ.mol^{-1}$ at the M06 level. However, the difference can reach 10 $kJ.mol^{-1}$ for free energies.

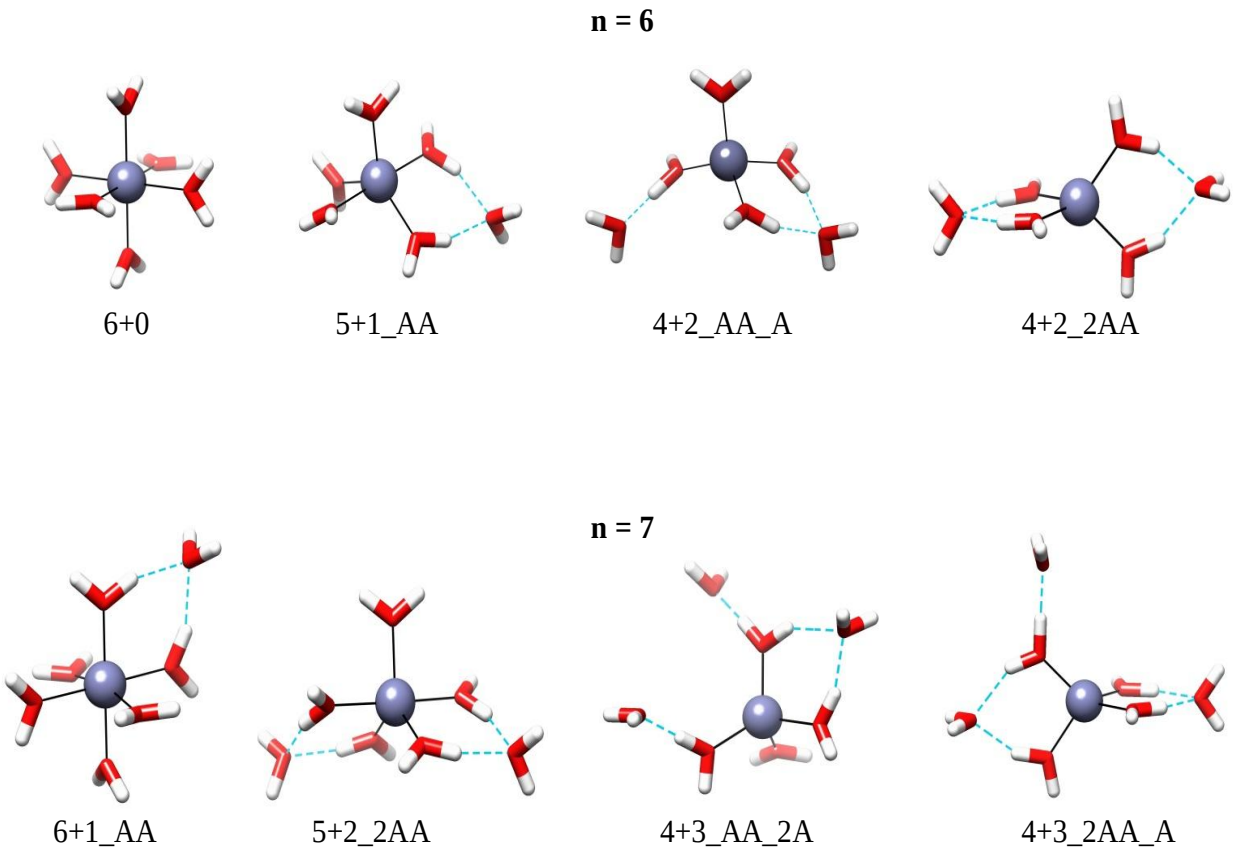


Figure 5 Starting structures for MD simulations of $Zn^{2+}(H_2O)_n$ system (n = 6 and 7).

For n=6, the coordination with 4 water molecules is the most stable at the B3LYP level. The 4+2_2AA and 4+2_AA_A structures are the most stable in terms of relative and free energy, respectively. Nevertheless, at both M06 and MP2 levels, the 6+0 structureis the most stable in favor of a CN of 6, except at the MP2/aug-cc-pVTZ level where 4+2_2AA becomes again the most stable. Still at the MP2 level, structures 6+0 and 5+1 are quasi degenerate in terms of electronic energy. For n=7, both the 4+3_2AA_A and the 5+2_2AA structures are found to be energetically close at the B3LYP level. However at the M06 and MP2 levels, the 5+2_2AA structure is the most stable using the largest basis set. Hence, for n = 7, CN of 4 and 5 are

preferred at the B3LYP level while a CN of 5 is favored at MP2 and M06 levels. The QM results are very sensitive to the different methods used, with again M06 results close to the MP2 ones.

Table 1 $Zn^{2+}(H_2O)_6$ and $Zn^{2+}(H_2O)_7$: relative electronic energies with relative free energies in parentheses at 298 K ($kJ.mol^{-1}$).

Complex	**B3LYP/TZ[a]**	**B3LYP/aTZ[b]**	**M06/TZ[a]**	**M06/aTZ[b]**	**MP2/TZ[a]**	**MP2/TZ[c]**
6+0	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0
5+1_AA	-9.5 (- 6.1)	-9.6 (-8.4)	9.0 (6.6)	6.2 (7.5)	2.8 (13.2)	0.1
4+2_2AA	-21.4 (-7.7)	-21.7 (-17.3)	8.1 (9.1)	3.2 (4.0)	4.0 (15.3)	-5.1
4+2_AA_A	-10.7 (-13.3)	-13.7 (-20.3)	24.1 (14.1)	17.3 (10.5)	14.2 (13.4)	6.4
6+1 _AA	22.4 (19.4)	24.2 (25.3)	-9.7 (-11.3)	-5.1 (-6.2)	-5.4 (-4.0)	4.6
5+2 _2AA	-1.6 (2.2)	0.3 (9.0)	-17.4 (-8.0)	-13.8 (-9.7)	-14.9 (-9.6)	-9.8
4+3 _2AA_A	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0
4+3 AA_2A	12.1 (0.1)	9.6 (-0. 7)	17.3 (9.2)	15.2 (3.4)	11.5 (3.0)	12.9

a: TZ = 6-311+G(d,p)
b: aTZ = aug-cc-pVTZ
c: TZ = aug-cc-pVTZ//MP2/6-311+G(d,p)

3.1.2. MD results

For n=6, the three optimized structures (6+0, 5+1_AA and 4+2_2AA) are selected as starting points (Figure 5) to see the effect of the starting conformation with the temperature. The values of mean CN are provided in Table 2 at 200, 250 and 300 K. The evolution of CN with time is also reported in order to capture the transitions between different types of coordination and the influence of the starting conformation. Figure 6 represents the transitions of the different CNs for each starting point of $Zn^{2+}(H_2O)_6$ at three different temperatures. This figure is obtained from one trajectory, which is fairly representative of the 10 trajectories.

n = 6

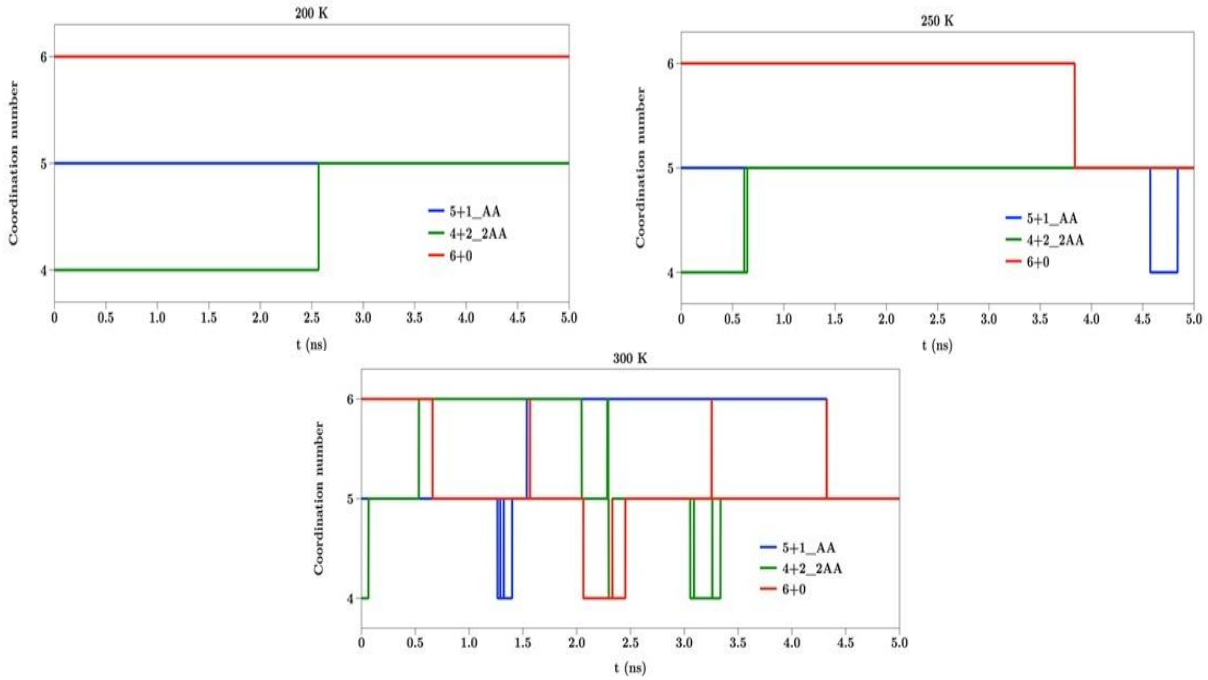


Figure 6 CN as a function of time for three different starting points for the $Zn^{2+}(H_2O)_6$ complex at 200, 250 and 300 K. Blue, green and red lines correspond to the 5+1_AA, 4+2_2AA and 6+0 structures, respectively.

Table 2 Mean CN of $Zn^{2+}(H_2O)_n$ from 10 trajectories of 5 ns each.

$Zn^{2+}(H_2O)_n$	**Starting structure[a]**	**200 K**	**250 K**	**300 K**
$Zn^{2+}(H_2O)_6$	4+2_2AA	4.27	4.83	5.15
	5+1_AA	5.00	5.04	5.21
	6+0	6.00	5.81	5.39
$Zn^{2+}(H_2O)_7$	4+3_2AA_A	4.74	5.01	5.20
	4+3_AA_2A	4.57	5.22	5.20
	5+2_2AA	5.00	5.21	5.25
	6+1_AA	6.00	5.58	5.38
$Zn^{2+}(H_2O)_{10}$	5+5_4AA_A	5.07	5.28	5.51
	5+3+2_3AA_2A	5.00	5.34	5.59
	6+4_4AA	5.86	5.89	5.62
$Zn^{2+}(H_2O)_{12}$	5+5+2_4AA_3A	5.20	5.72	5.75
	6+5+1_4AA_2A	6.00	5.95	5.77
$Zn^{2+}(H_2O)_{15}$	5+5+5_4AA_6A	5.88	5.89	5.94
	6+5+4_4AA_5A	6.00	5.99	5.94
$Zn^{2+}(H_2O)_{20}$	5+5+5+5_4AA_11A	5.96	5.98	5.99
$Zn^{2+}(H_2O)_{25}$	5+5+5+5+5_4AA_16A	5.99	5.99	5.99

a: x, y and z stand for the numbers of water molecules in the first, second and third solvation shell, respectively. This nomenclature can be extended to further shells, e.g. for n=20 and 25.

At 200 K, the mean coordination numbers (over 10 trajectories) from 6+0, 5+1_AA and 4+2_2AA starting points are found to be 6.0, 5.0 and 4.27, respectively. In Figure 6, it can be seen that starting from 6 or 5-coordination, the complex remains in this coordination. However, one transition occurs from 4+2 to 5+1 conformations. At 250 K, the mean CNs for 6+0, 5+1_AA and 4+2_2AA are5.81, 5.04 and 4.83, respectively. Figure 6 displays that the three types of conformations are explored during dynamics and few transitions occur. However, the 5+1_AA conformation seems to be preferred at 250 K. At 300 K, the different types of structures are highly explored with several transitions resulting in mean CNs of 5.39, 5.21 and 5.15 for 6+0, 5+1_AA and 4+2_2AA, respectively. As the time spent in 4-coordination is short, the mean CNs obtained at 300 K are all found to lie between 5 and 6. However the results are still dependent on the starting conformations.

These results suggest that the activation barriers are hardly surmountable at 200 K, easily crossed at 300 K, and that the internal energy at 250 K could be close to the barriers, leading to interconversions as rare events. In order to strengthen the link between the observed changes in dynamical exchanges as a function of temperature and the potential energy landscape, MP2 calculations were carried out on the 6+0, 5+1_AA and 4+2_2AA minima and two transition states connecting 5+1 to 4+2 and 6+0 to 5+1. The results are gathered in Table 3 and Figure 7. Based on frequency calculations, the thermal energies at 200 and 298 K were also obtained. It turns out that the two transition barriers lie 26 and 28 $kJ.mol^{-1}$ higher in energy than the lowest energy structure, while the internal energy at 200 K is only 24 $kJ.mol^{-1}$. Thus, interconversion is impossible at this temperature. At 300 K, the thermal energy amounts to 45 $kJ.mol^{-1}$, making interconversions easy and fast. When the energy supplement is smaller as at 250 K, it is enough to surmount barriers, however as a slow process of small probability. It thus appears that the

dynamical behavior obtained at the MD/AMOEBA level is in good agreement with the static picture derived from MP2 calculations.

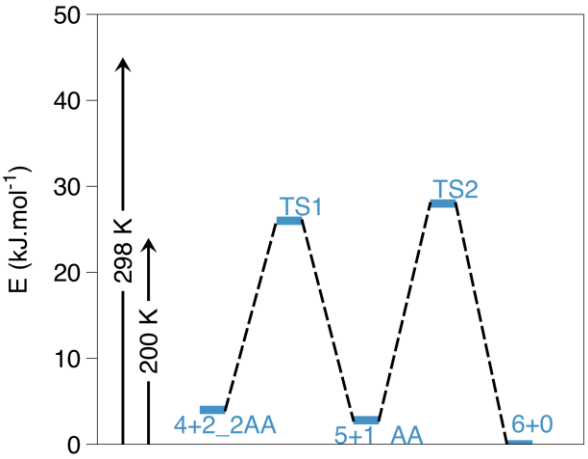


Figure 7 Activation barriers for interconversion between 4+2_2AA and 5+1_AA and between 5+1_AA and 6+0 at MP2 level.

Table 3 $Zn^{2+}(H_2O)_6$: AMOEBA and MP2 energies ($kJ.mol^{-1}$) of minima and transition states at AMOEBA and MP2 geometries, respectively.

$Zn^{2+}(H_2O)_6$	**MP2/6-311+G(d,p)**	**AMOEBA**
6+0	0.0	0.0
5+1_AA	2.8	3.0
4+2_2AA	4.0	8.5
TS 5+1→4+2	25.7	
TS 6+0→5+1	27.9	

For n=7, four structures, i.e. 6+1_AA, 5+2_2AA, 4+3_2AA_A and 4+3_AA_2A, were taken as starting points (Figure 5). The values of average CN and the evolution of CN with time are presented in Table 2 and Figure 8. At 200 K, the mean CNs of 6+1_AA, 5+2_2AA, 4+3_2AA_A and 4+3_AA_2A are found to be 6.0, 5.0, 4.74 and 4.57, respectively. When starting from 6 or 5-coordination, the average CN remains at 6 and 5, while transitions between 4+3 and 5+2 structures are observed during the dynamics using a starting point in a 4+3 conformation. At 250 K, the mean CNs of 6+1_AA, 5+2_2AA, 4+3_2AA_A and 4+3_AA_2A are found to be 5.58, 5.21, 5.003 and 5.22, respectively, with few transitions occurring whatever the starting structure. As for $Zn^{2+}(H_2O)_6$, the time spent in the 4-coordination is very short and the 5+2_2AA conformation is preferred at 250K with only few transitions. At 300 K, the mean

CNs are 5.38, 5.25, 5.20 and 5.20 for the 6+1_AA, 5+2_2AA, 4+3_2AA_A and 4+3_AA_2A structures, respectively. All types of structures are largely explored at 300 K compared to simulations at 200 and 250 K with numerous transitions. However, there remains a small dependence of mean CN with the starting point. These results indicate that at temperatures below 300 K, exchange may be thermally limited or impossible so that structures will mainly depend upon the experimental formation conditions, e.g. desolvation conditions in electrospray ionization.

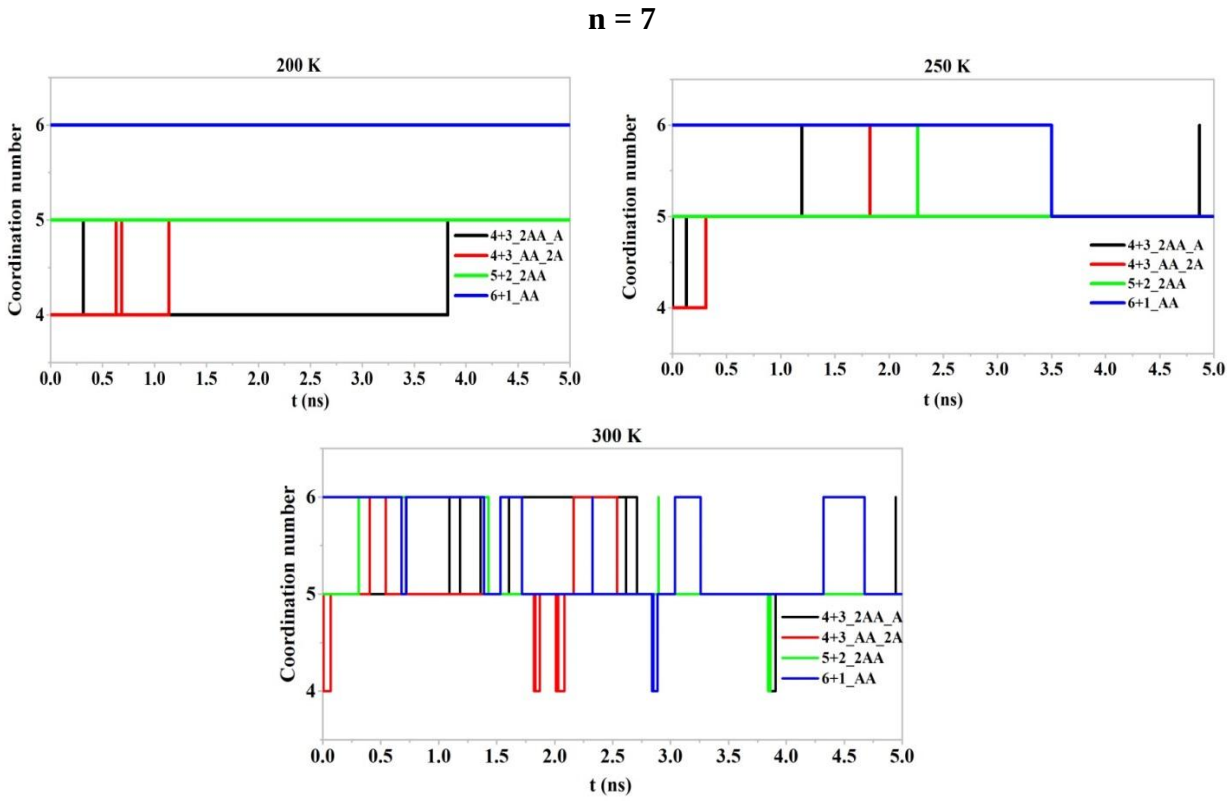


Figure 8 CN as a function of time for four different starting points for the $Zn^{2+}(H_2O)_7$ complex at 200, 250 and 300K. Black, red, green and blue lines correspond to the 4+3_2AA_A, 4+3_AA_2A, 5+2_2AA and 6+1_AA structures, respectively.

3.2. Zn^{2+} water clusters with n = 10, 12, 15, 20 and 25

In a similar manner, we have extended the study to larger complexes with 10 to 25 water molecules. The mean CNs are provided in Table 2 for each size and starting point. The latter are shown in Figure 9. The transitions between different coordination numbers are presented in Figures 11 and 12 for n = 10 and 12, while figures for n = 15, 20 and 25 are provided in Appendix. The starting structures were built by adding water molecules step by step from the 5+1_AA and 6+0 structures to keep five and six coordination, respectively. The Avogadro[38,39] software was used for this purpose and finally, these structures were optimized, using first the AMOEBA force field and then at the B3LYP/6-311+G(d,p) level.

Two structures have been selected from the MD trajectories for n = 15 at 300 K. They are shown in Figure 10 with the aim to show the change of hydrogen bonding patterns during the simulations. Structure "A" in Figure 10 is taken from the trajectory starting with structure 5+5+5_4AA_6A (see Figure 9). It has a CN of 6 with 2AA and 7A water molecules. Structure "B", taken in the trajectory from 6+5+4_4AA_5A, reflects that even if the CN remains at 6, there is a change in the hydrogen bonding pattern with 2A and 7AA water molecules. This illustrates that besides the CN which is the main focus here, numerous changes in the arrangement of the second and third shells occur along the trajectories.

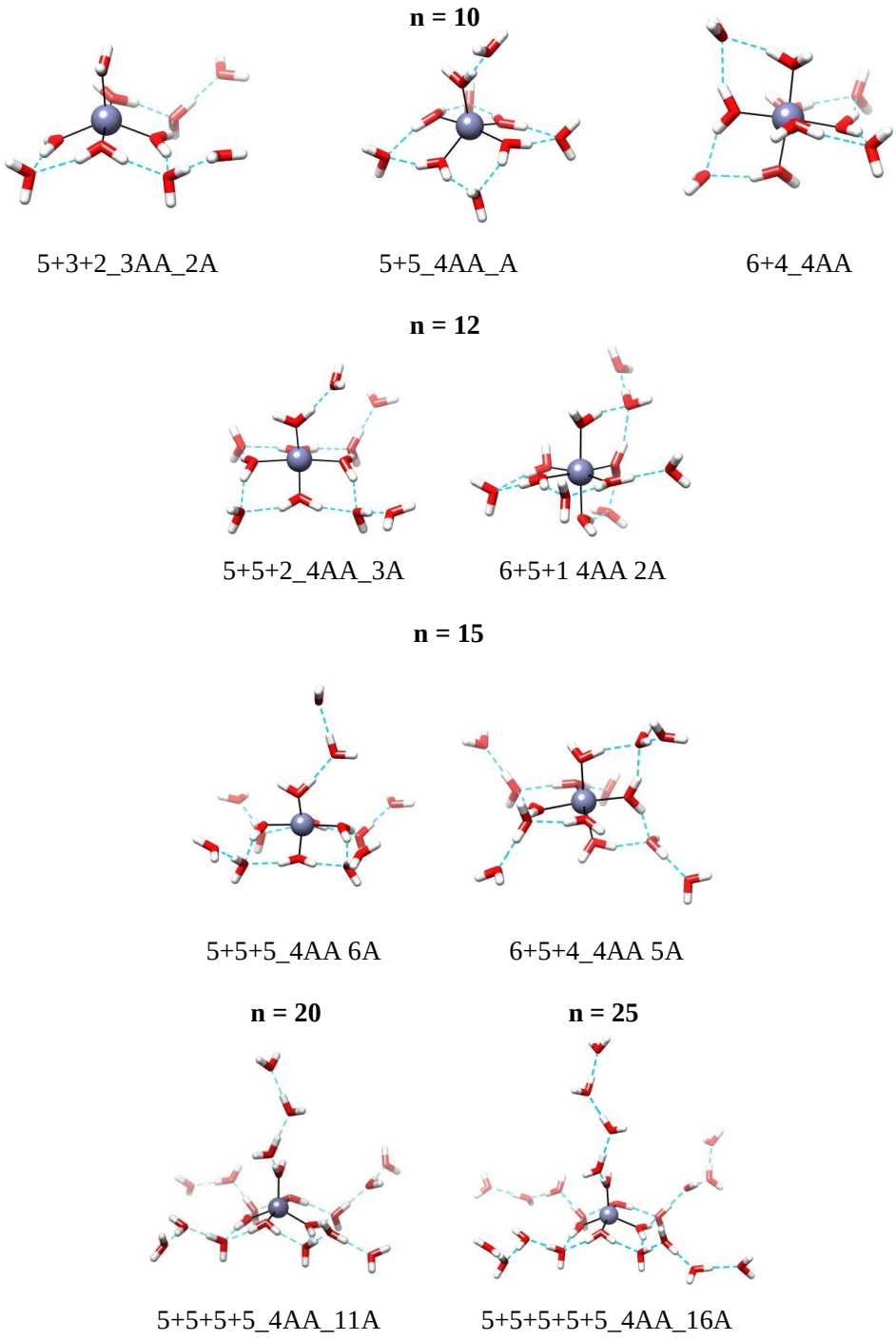
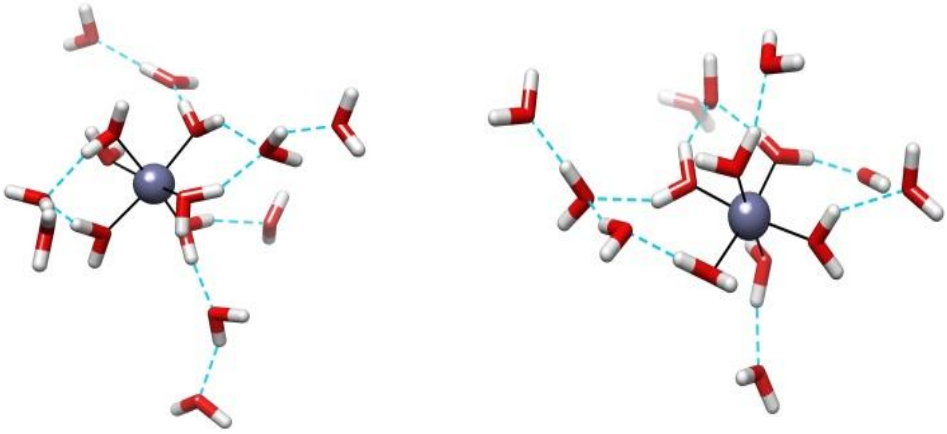


Figure 9 Starting structures for MD simulations of $Zn^{2+}(H_2O)_n$ system (n = 10 - 25).

For n = 10, three starting structures, i.e. 5+5_4AA_A, 5+3+2_3AA_2A and 6+4_4AA, were selected for MD simulations. At 200 K, the mean CN is close to 5 for a 5-coordination starting structure and 5.86 for the 6-coordination structure. This behavior is also displayed in Figure 11 with a single transition from 6 to 5 coordination for a 6+4 starting structure. Furthermore, 5+5 and 5+3+2 conformations stay in a 5-coordination. At 250 K, the mean CNs of 5+5_4AA_A, 5+3+2_3AA_2A and 6+4_4AA are found to be 5.28, 5.34 and 5.89, respectively. All structure types are explored during dynamics whatever the starting point, and various transitions are observed between 5 and 6-coordinations. At 300 K, the average CN is ca. 5.5 for the different simulations and the system becomes highly fluxional in comparison with 200 and 250 K. Many transitions are seen between the two coordinations, so the starting point has no more influence on the CN.

n = 15



A (6+5+4_2AA_7A) B (6+7+2_2AA_7A)

Figure 10 Structures obtained during the simulations for $Zn^{2+}(H_2O)_{15}$ at 300 K: "A" has been taken as the simulation start for the 5+5+5_4AA_6A conformation and "B" is the simulation start for 6+5+4_4AA_5A.

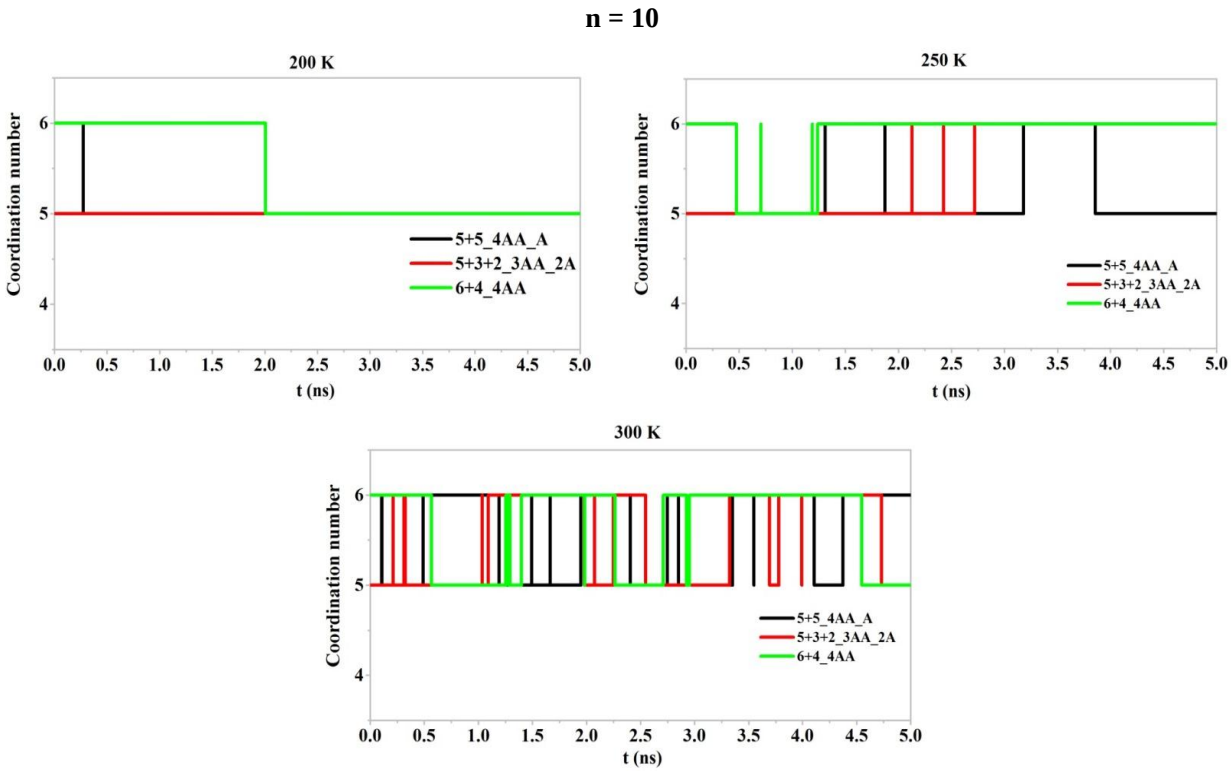


Figure 11 CN as a function of time for three different starting points for the $Zn^{2+}(H_2O)_{10}$ complex at 200, 250 and 300 K.

For n = 12, two starting structures, 5+5+2_4AA_3A and 6+5+1_4AA_2A, were used with 5 and 6 first shell coordinations. At 200 K, the mean CNs of 5+5+2_4AA_3A and 6+5+1_4AA_2A are 5.20 and 6.00, respectively. Figure 12 suggests that starting from 6-coordination, it is remaining at 6. From the mean CN value in Table 2, this is general for all 10 trajectories, with a mean CN of 6. A few transitions are observed for the 5+5+2 starting conformation. At 250 K, both coordinations are explored during the dynamics with a longer time spent in 6-coordination. The mean CNs are 5.72 and 5.95 for of 5+5+2_4AA_3A and 6+5+1_4AA_2A, respectively. However at 300 K, the average CN is 5.7 whatever the starting point with a strong dynamics between the 5 and 6-coordinations.

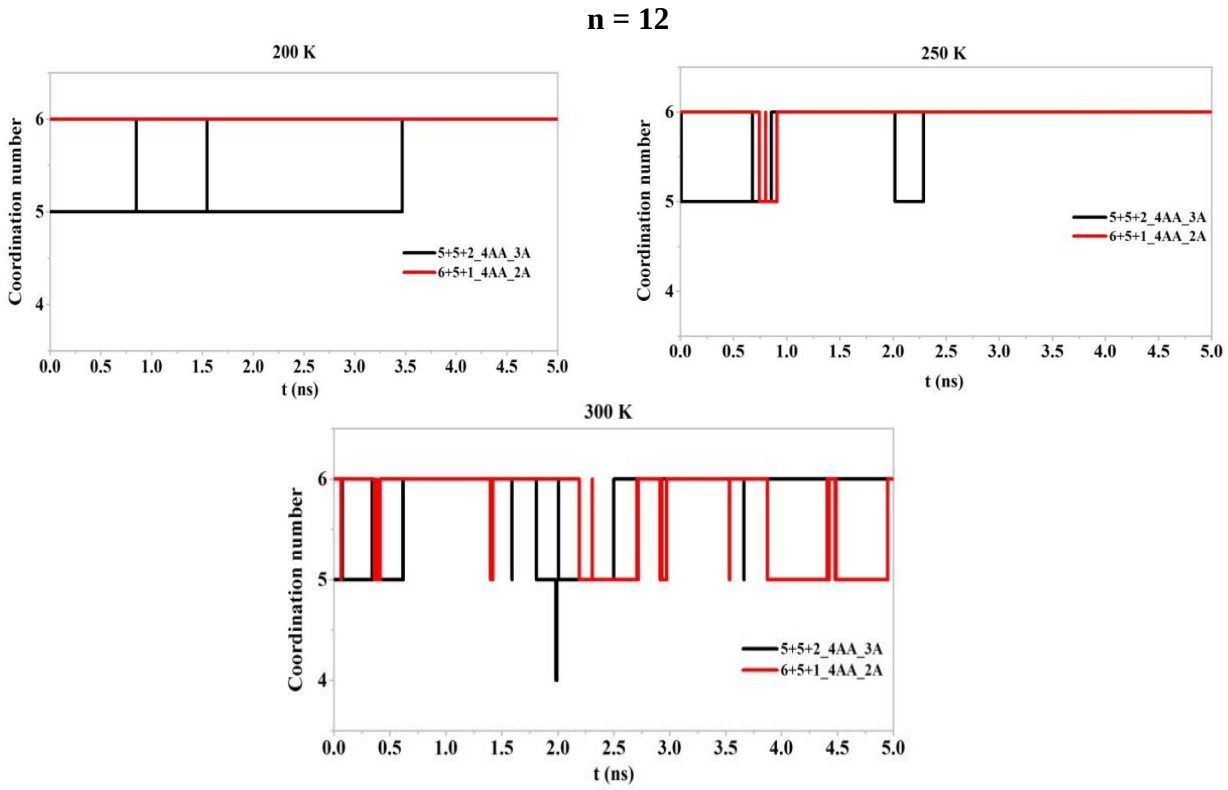


Figure 12 CN as a function of time for two different starting points for the $Zn^{2+}(H_2O)_{12}$ complex at 200, 250 and 300K.

For the cluster with n = 15, two starting structures, 5+5+5_4AA_6A and 6+5+4_4AA_5A were used as starting conformations. At 200 K, the mean CNs of 5+5+5_4AA_6A and 6+5+4_4AA_5A are found to be close to 6. Figure 20 shows that the time spent in 6-coordination is very long. When the simulation starts from 6-coordination, the CN remains at 6. Similarly at 250 and 300 K, the overall average CNs are 5.99 and 5.94, indicating limited but significant transitions between 5 and 6-coordinations during the simulations. The starting conformation has less influence on the CN.

MD simulations were also carried out for a cluster with 20 water molecules for a starting structure of type 5+5+5+5_4AA_11A. The mean CNs at 200, 250 and 300 K are found to be

5.96, 5.98 and 5.99, respectively. Whatever the temperature, only few transitions occur between 5 and 6-coordinations (see Figure 21). The average CN is very close to 6 with no influence of the starting structure. Finally, an increase of the cluster size to 25 water molecules with a 5+5+5+5+5_4AA_16A starting structure provides a mean CN of 6 with the three temperatures. Only few transitions occur between 5 and 6 coordinations (Figure 22) and the conformations remain in a 6-coordination most of the time.

3.3. Transition to the aqueous phase

To contrast with small clusters, one MD simulation of 5 ns was performed on a droplet of 216 water molecules confined in spherical boundary conditions. The results highlight a CN of 6 as showed in Figure 13. Only one water exchange was observed with a very short period in a 5-coordination. Therefore, our results reveal that under droplet conditions, Zn^{2+} favors the 6-coordination, which is consistent with the experimentally well-established CN of Zn^{2+} in the aqueous phase[5]. Previously, Ren and coworkers[32] considered the water organization around Zn^{2+} by MD simulations with the AMOEBA force field and using periodic boundary conditions (box with 512 water molecules, 298 K). In this context, they have shown that the first-shell water coordination number is 6 and the residence time in the first solvation shell around Zn^{2+} is at least 2 ns. Our results in droplet conditions are consistent with these earlier results.

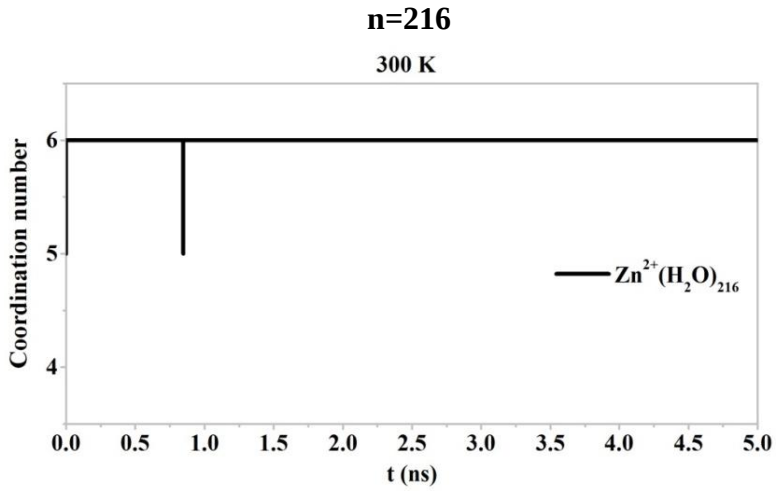


Figure 13 CN as a function of time for the $Zn^{2+}(H_2O)_{216}$ system at 300 K.

The first reliable values of the coordination number of the zinc divalent cation have been obtained for clusters growing from 6 to 216 water molecules at three different temperatures ranging from 200 to 300 K. The transitions between the coordinations have been described as a function of temperature. For smaller clusters, i.e. n = 6-12, the mean CNs depend on the starting points which is not the case for larger clusters (n = 15 - 216). The transitions between the different types of coordination suggest that 4, 5 and 6-coordinations are possible for Zn^{2+} for small cluster sizes (n = 6 - 7), but with a short time spent in the 4-coordination. Furthermore by increasing the cluster size (n = 10 - 216), Zn^{2+} prefers 5 and 6-coordinations. The time spent in the 5-coordination is becoming shorter for the largest sizes. In summary, Figure 14 shows the mean CN at 300 K for a cluster size growing from 6 to 25. Beyond n = 15, Zn^{2+} clearly favors a coordination of 6, similarly to the simulation with 216 water molecules and the reported studies in the condensed phase.

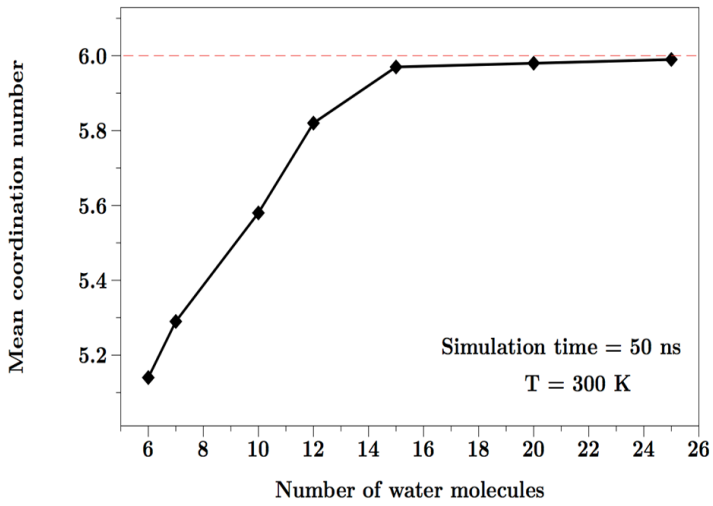


Figure 14 Change of the mean CNs for clusters growing from 6 to 25 water molecules at 300 K.

4. IR spectra of micro-hydrated Zn^{2+}

As mentioned previously, IRPD action spectroscopy was used to study the micro-hydration of several metal dications including zinc[11,34]. The experimental IRPD spectrum of $Zn^{2+}(H_2O)_6$ is reported in Figure 15 (bottom). The experimental temperature is estimated to be 215 K. The two intense peaks correspond to the symmetric and antisymmetric free O-H with maxima at ca. 3600 cm^{-1} and 3670 cm^{-1}, respectively. These bands occur at 3657 and 3756 $cm^{-1}$for an isolated water molecule[40]. The large red shift for these bands can be explained in part by a partial electron transfer from the water molecules to the metal cation. Furthermore, weak spectral intensity is also observed in the bonded O-H stretch region (3200 - 3500 cm^{-1}). The weak band centered near 3400 cm^{-1} is assigned to inner-shell water molecules that donate one or two hydrogen bonds to double acceptor, second shell water molecules. The presence of this band is characteristic of a second water shell and thus an average CN of Zn^{2+} smaller than 6. The clear distinction between the frequency ranges of bound and free O-H stretches makes these bands potentially diagnostic

of structure types. Thus, in complement to the previous results, we have computed IR spectra from MD trajectories as described in chapter 2.

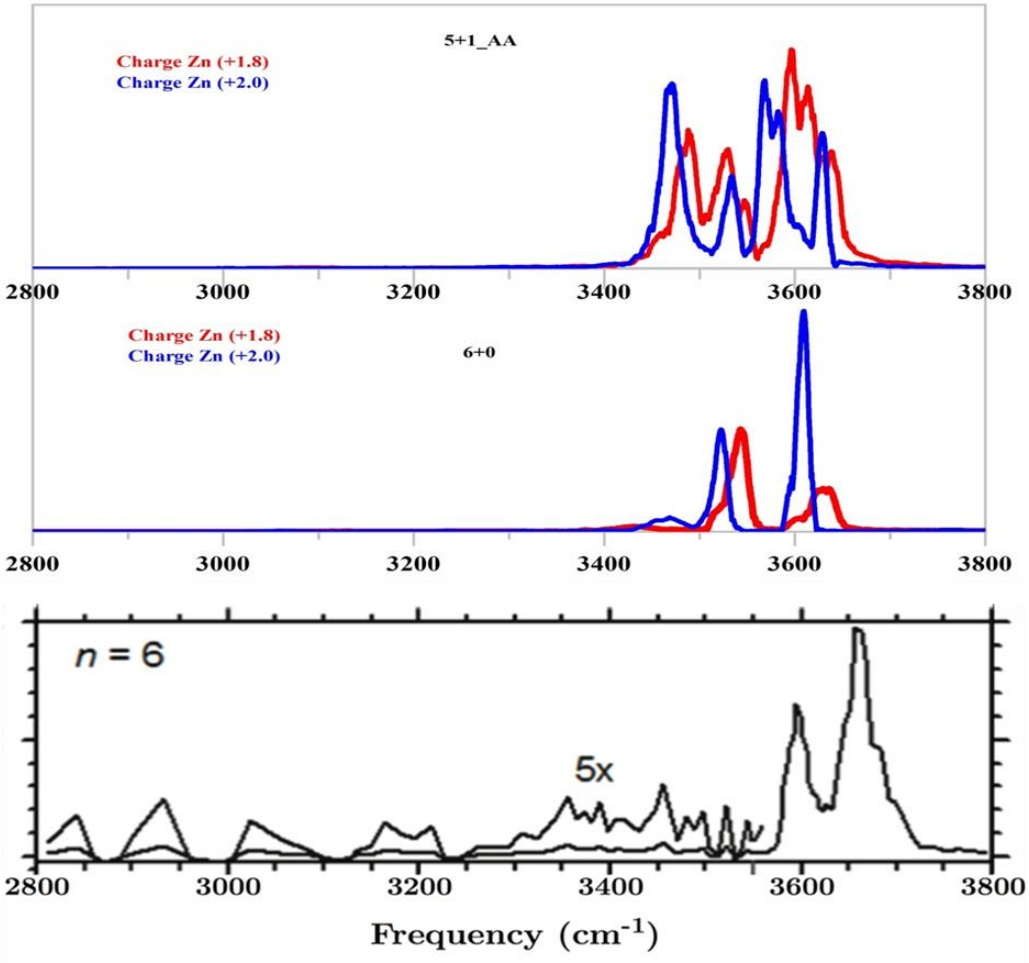


Figure 15 Comparison of the experimental IRPD spectrum for $Zn^{2+}(H_2O)_6$[34] at 215 K with calculated IR spectra for 5+1_AA and 6+0 starting structures from MD simulations at 200 K with a zinc charge of +2.0 and +1.8.

4.1. Charge transfer effect

The IR spectrum of $Zn^{2+}(H_2O)_6$ has been calculated at 200 K from the 6+0 and the 5+1_AA starting structures (Figure 5). The calculated spectrum starting from 5+1_AA presented in Figure 15 (top, blue) shows that two intense bands are present at ca. 3570 (as a doublet) and 3630 cm^{-1}

corresponding to symmetric and antisymmetric free O-H stretches, respectively. For the 6+0 starting point (middle, blue), these two bands are red-shifted to ca. 3520 and 3610 cm^{-1}, respectively. They are also red shifted in comparison with experimental frequencies. The doublet at ca. 3570 cm^{-1} in the 5+1_AA case is due to the mixing of two types of free O-H bond stretches for the water molecules present in the first and second solvation shells. There are also two bands at 3470 and 3535 cm^{-1} for symmetric and antisymmetric hydrogen bonded O-H stretches; they provide a signature of the formation of the second solvation shell. Consistently these bands are absent in the 6+0 case which is devoid of a second shell. The mean CN are found to be 5.0 and 6.0 during the simulation for 5+1_AA and 6+0 structures, respectively.

To unfold the reason behind the red shift and to obtain a better agreement with experiment, several changes have been attempted on Zn^{2+} parameters, i.e. vdW parameters and charge. Only a modification of the charge was found to have an impact on the band positions. In order to model the electron transfer from the water to the metal, the charge on zinc has been decreased gradually from +2.0 to +1.6. Using a +1.8 charge on the cation (Figure 15, top and middle, red), the symmetric and antisymmetric of the free O-H stretches are shifted to 3600 and 3640 cm^{-1}, respectively from 5+1_AA, while for 6+0 they are shifted to 3540 and 3630 cm^{-1}, respectively. These frequencies are in better agreement with experiment. Other bands are found to be blue-shifted as well. The doublet near 3540 cm^{-1} corresponds to the mixing of the two types of free O-H bond stretches as described above. The symmetric and antisymmetric hydrogen bonded O-H stretches are present at 3487 and 3530 cm^{-1}. Finally, the trajectory analysis shows that the mean CN is 5.0 during the 200 ps of the simulation, thus the charge adaptation does not have a strong effect on the mean CN.

The above approach can only provide a rough picture of charge transfer effects, since electron depletion in water is not taken into account and other parameters, which may be expected to be sensitive to electron transfer, are also left unchanged. Still the sensitivity of the computed spectrum to the metal charge illustrates the importance of implementing a model with an explicit charge transfer in the AMOEBA force field, which is currently under development (J.P. Piquemal, private communication).

4.2. Influence of the temperature

In the earlier experimental studies, temperature was found to have a significant impact on the O-H stretching bands of hydrated metal ions[11,34]. The key role of temperature on the IR spectra of hydrated metal ions has not been yet studied in details and to our best knowledge there is no detailed theoretical study. To obtain a better picture of temperature dependence of the IR spectrum, we have calculated IR spectra of $Zn^{2+}(H_2O)_{12}$ at 200 and 300 K with +1.8 charge on the zinc cation. The 5+5+2_4AA_3A and 6+5+1_4AA_2A starting points were used as in previous MD simulations. The computed spectrum for each starting point at 200 K and 300 K is shown in Figure 16 in the O-H stretching region. At 200 K, two distinct spectra are observed which are representative of the two starting structures. The average CNs during the simulations are 5 and 6 for 5+5+2_4AA_3A and 6+5+1_4AA_2A structures, respectively (Figure 16). At 300 K, the spectra are broader than at 200 K. Interestingly the band positions as well as the intensities are almost similar for the two simulations. The calculated mean CNs are also very similar and found to be 5.7 and 5.6 for 5+5+2_4AA_3A and 6+5+1_4AA_2A structures, respectively. Even during 200 ps, few transitions can be observed between 5 and 6-coordinations (Figure 17). The spectra obtained in Figure 16 are not in good agreement with experiments for

the O-H bonded region. Improvement of the position of these bands could be possible by modification of bond stretching parameters or by treating explicitly charge transfer in AMOEBA.

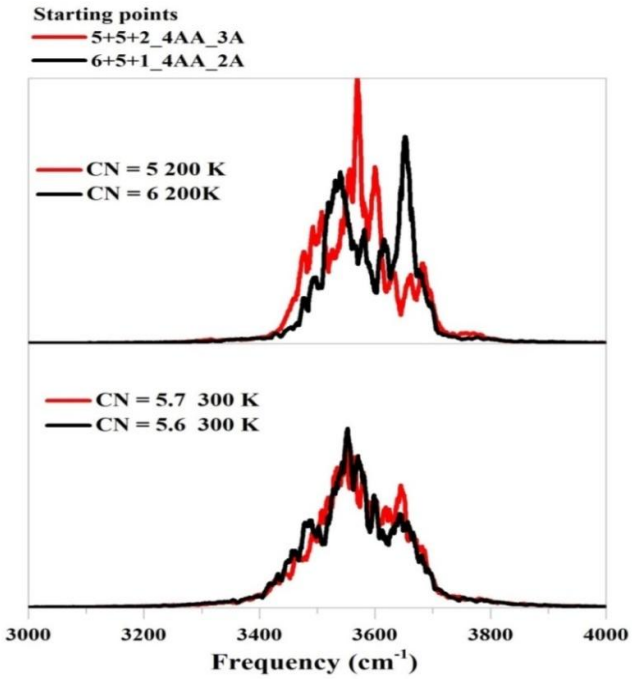


Figure 16 Calculated IR spectra of $Zn^{2+}(H_2O)_{12}$ at 200 and 300 K with average CN.

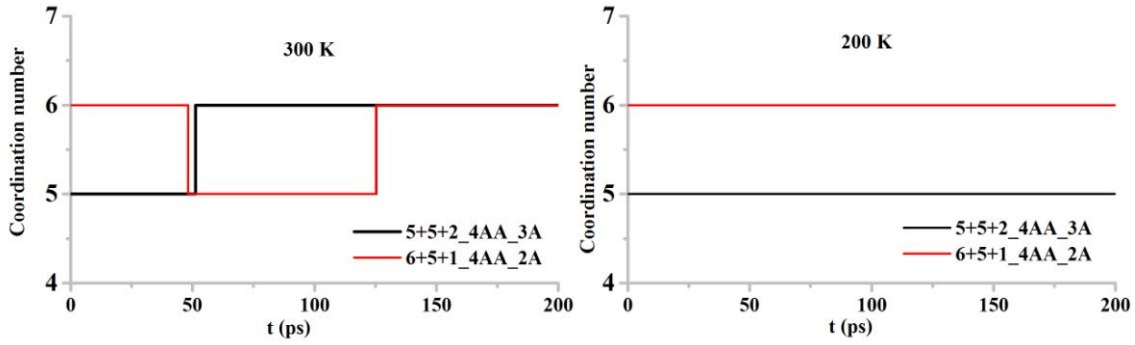


Figure 17 CN as a function of time for two different starting points for the $Zn^{2+}(H_2O)_{12}$ complex at 300 and 200 K.

Spiegelman and coworkers[41-44] have thoroughly addressed the role of temperature on the IR spectra of several π-complexes, aromatic hydrocarbons and molecular ions such as $[FePAH]^+$, $[SiPAH]^+$, $[PANH]^+$ (PAH = polycyclic aromatic hydrocarbon) and $(C_{24}H_{12})(H_2O)_n$ (n =3, 6, 8). IR spectra have been calculated from MD simulations on a self-consistent-charge density functional tight-binding (SCC-DFTB) potential energy surface at temperatures ranging from 0 to 850 K. The SCC-DFTB method is an approximated DFT based approach, which is able to provide much faster calculations than DFT. This method is very efficient for large systems to calculate on-the-fly energies and gradients in a quantum framework. They have performed Born-Oppenheimer MD simulations to obtain IR spectra by computing DACF as in the present work. The C-C and C-H bands of $[FeC_{24}H_{12}]^+$ are found to be broadened and red shifted linearly with increasing the temperature[41]. A similar analysis has been performed on $Zn^{2+}(H_2O)_7$, using the 5+2_2AA structure as starting conformation because we have shown that the 5-coordination is the most stable for this complex. Finite temperature IR spectra have been recorded between 50 and 300 K and are presented in Figure 18.We have also observed a red shift for both hydrogen bonded and free O-H bands, however with different magnitudes. Frequency as a function of temperature is represented in Figure 19 for band "a". The shifts are not strictly linear with increasing temperature but the general tendency is similar to the ones obtained by Spiegelman et al.

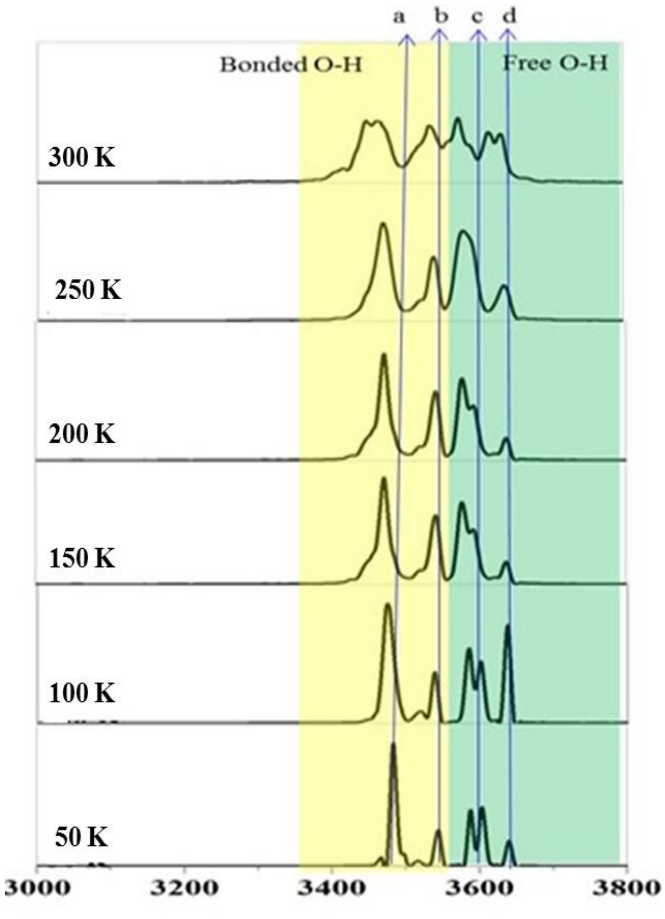


Figure 18 Calculated IR spectra for $Zn^{2+}(H_2O)_7$ within 50-300 K range. The charge on Zn is +1.8.

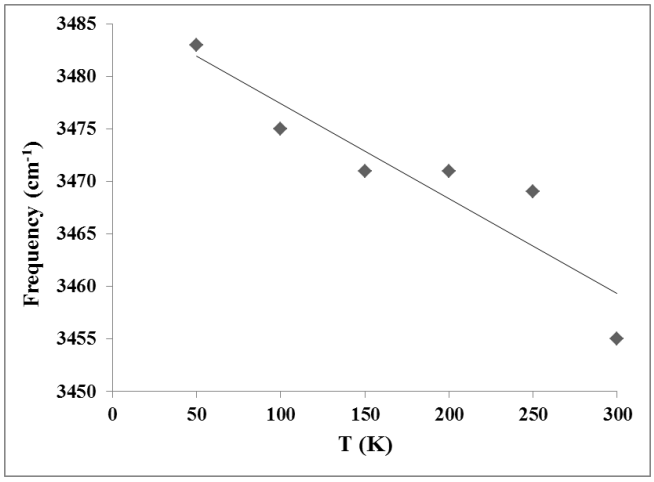


Figure 19 Change of frequencies as a function of temperature, obtained from Figure 18 for band "a".

5. Conclusions

This work provides insights into the structure, dynamics and IR spectroscopy of Zn^{2+}-water clusters. The AMOEBA force field associated to MD potential energy surface exploration is a useful tool to decipher the organization of the first coordination shell of water around the cation. The results are consistent with gas-phase experiments as well as simulations and experiments in condensed phase showing a progressive increase of the mean coordination number from 5 to 6 with increasing the cluster size. The IR spectra of some clusters have been computed at finite temperature in the O-H stretch region to study band shifts and broadenings with temperature. A realistic model of free O-H modes requires a decrease of zinc charge from 2.0 to 1.8 to mimic charge transfer from the water molecules to the metal cation. The force field limitations for the water model will deserve more work in the future to be able to reproduce the bonded O-H region. The model of zinc hydration with precise description of CN evolution with cluster size developed in this work represents a step toward future study understanding the detail of water structuration around Zn^{2+}, and thus of its transport and competitive attachment to other species. It may also improve our ability to understand zinc complexation, either functional or structural, to binding sites of zinc proteins and enzymes, in which changes in CN may be required.

6. References

(1) Spiro, T. G. *Zinc Enzymes*; J. Wiley: New York, 1983.

(2) Kimura, E. *Pure Appl. Chem.* **1993**, *65*, 355.

(3) Richens, D. T. *The Chemistry of Aqua Ions*; John Wiley and Sons, Inc.: New York, **1997**.

(4) Nriagu, J. O. *Zinc in the Environment*; Wiley: New York, **1980**.

(5) Chillemi, G.; D'Angelo, P.; Pavel, N. V.; Sanna, N.; Barone, V. *J. Am. Chem. Soc.* **2002**, *124*, **1968**.

(6) D'Angelo, P.; Barone, V.; Chillemi, G.; Sanna, N.; Meyer-Klaucke, W.; Pavel, N. V. *J. Am. Chem. Soc.* **2002**, *124*, 1958.

(7) W. Rudolph, W.; C. Pye, C. *Phys. Chem. Chem. Phys.* **1999**, *1*, 4583.

(8) Wakita, H.; Johansson, G.; Sandstrom, M.; Goggin, P. L.; Ohtaki, H. *J. Solution Chem.* **1991**, *20*, 643.

(9) Cooper, T. E.; Carl, D. R.; Armentrout, P. B. *J. Phys. Chem. A* **2009**, *113*, 13727.

(10) Cooper, T. E.; Armentrout, P. B. *J. Phys. Chem. A* **2009**, *113*, 13742.

(11) Cooper, T. E.; O'Brien, J. T.; Williams, E. R.; Armentrout, P. B. *J. Phys. Chem. A* **2010**, *114*, 12646.

(12) Bock, C. W.; Katz, A. K.; Glusker, J. P. *J. Am. Chem. Soc.* **1995**, *117*, 3754.

(13) Hartmann, M.; Clark, T.; vanEldik, R. *J. Mol. Model.* **1996**, *2*, 354.

(14) Hartmann, M.; Clark, T.; vanEldik, R. *J. Am. Chem. Soc.* **1997**, *119*, 7843.

(15) Pavlov, M.; Siegbahn, P. E. M.; Sandstrom, M. *J. Phys. Chem. A* **1998**, *102*, 219.

(16) De, S.; Ali, S. M.; Ali, A.; Gaikar, V. G. *Phys. Chem. Chem. Phys.* **2009**, *11*, 8285.

(17) Mohammed, A. M.; Loeffler, H. H.; Inada, Y.; Tanada, K.; Funahashi, S. *J. Mol. Liq.* **2005**, *119*, 55.

(18) Fatmi, M. Q.; Hofer, T. S.; Randolf, B. R.; Rode, B. M. *J. Chem. Phys.* **2005**, *123*.

(19) Fatmi, M. Q.; Hofer, T. S.; Randolf, B. R.; Rode, B. M. *J. Phys. Chem. B* **2006**, *110*, 616.

(20) Brancato, G.; Rega, N.; Barone, V. *Chem. Phys. Lett.* **2008**, *451*, 53.

(21) Cauët, E.; Bogatko, S.; Weare, J. H.; Fulton, J. L.; Schenter, G. K.; Bylaska, E. J. *J. Chem. Phys.* **2010**, *132*, 194502.

(22) Gresh, N.; Cisneros, G. A.; Darden, T. A.; Piquemal, J-P. *J. Chem. Theory Comput.* **2007**, *3*, 1960.

(23) Piquemal, J-P.; Williams-Hubbard, B.; Fey, N.; Deeth, R. J.; Gresh, N.; Giessner-Prettre, C. *J. Comput. Chem.* **2003**, *24*, 1963.

(24) Ponder, J. W.; Wu, C.; Ren, P.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio, R. A.; Head-Gordon, M.; Clark, G. N. I.; Johnson, M. E.; Head-Gordon, T. *J. Phys. Chem. B* **2010**, *114*, 2549.

(25) Tiraboschi, G.; Roques, B. P.; Gresh, N. *J. Comput. Chem.* **1999**, *20*, 1379.

(26) Tiraboschi, G.; Gresh, N.; Giessner-Prettre, C.; Pedersen, L. G.; Deerfield, D. W. *J. Comput. Chem.* **2000**, *21*, 1011.

(27) Gresh, N.; Piquemal, J-P.; Krauss, M. *J. Comput. Chem.* **2005**, *26*, 1113.

(28) Jenkins, L. M. M.; Hara, T.; Durell, S. R.; Hayashi, R.; Inman, J. K.; Piquemal, J-P.; Gresh, N.; Appella, E. *J. Am. Chem. Soc.* **2007**, *129*, 11067.

(29) Grossfield, A. *J. Chem. Phys.* **2005**, *122*, 024506.

(30) Jiao, D.; King, C.; Grossfield, A.; Darden, T. A.; Ren, P. *J. Phys. Chem. B* **2006**, *110*, 18553.

(31) Piquemal, J-P.; Perera, L.; Cisneros, G. A.; Ren, P.; Pedersen, L. G.; Darden, T. A. *J. Chem. Phys.* **2006**, *125*.

(32) Wu, C.; Piquemal, J-P.; Chaudret, R.; Reinhardt, P.; Ren, P. *J. Chem. Theory Comput.* **2010**, *6*, 2059.

(33) Salmon, P. S.; Bellissentfunel, M. C.; Herdman, G. J. *J. Phys.: Condens. Matter* **1990**, *2*, 4297.

(34) O'Brien, J. T.; Williams, E. R. *J. Phys. Chem. A* **2011**, *115*, 14612.

(35) Semrouni, D.; Isley III, W. C.; Clavaguéra, C.; Dognon, J. P.; Cramer, C. J.; Gagliardi, L. *J. Chem. Theory Comput.* **2013**, *9*, 3062.

(36) Ren, P.; Ponder, J. W. *J. Phys. Chem. B* **2003**, *107*, 5933.

(37) Ren, P.; Ponder, J. W. *J. Phys. Chem. B* **2004**, *108*, 13427.

(38) *Avogadro: an open-source molecular builder and visualization tool. Version 1.0.3.* *http://avogadro.openmolecules.net/*.

(39) Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeersch, T.; Zurek, E.; Hutchison, G. R. *J. Cheminform.* **2012**, *4,17*.

(40) Shimanouchi, T. *Molecular Vibrational Frequencies, 69th ed.; NIST Chemistry Webbook, National Institute of Standards and Technology: Gaithersburg, MD, 2010;* *http://webbook.nist.gov*.

(41) Simon, A.; Rapacioli, M.; Lanza, M.; Joalland, B.; Spiegelman, F. *Phys. Chem. Chem. Phys.* **2011**, *13*, 3359.

(42) Simon, A.; Rapacioli, M.; Mascetti, J.; Spiegelman, F. *Phys. Chem. Chem. Phys.* **2012**, *14*, 6771.

(43) Simon, A.; Spiegelman, F. *Comp. Theor. Chem.* **2013**, *1021*, 54.

(44) Simon, A.; Spiegelman, F. *J. Chem. Phys.* **2013**, *138*, 194309.

Appendix

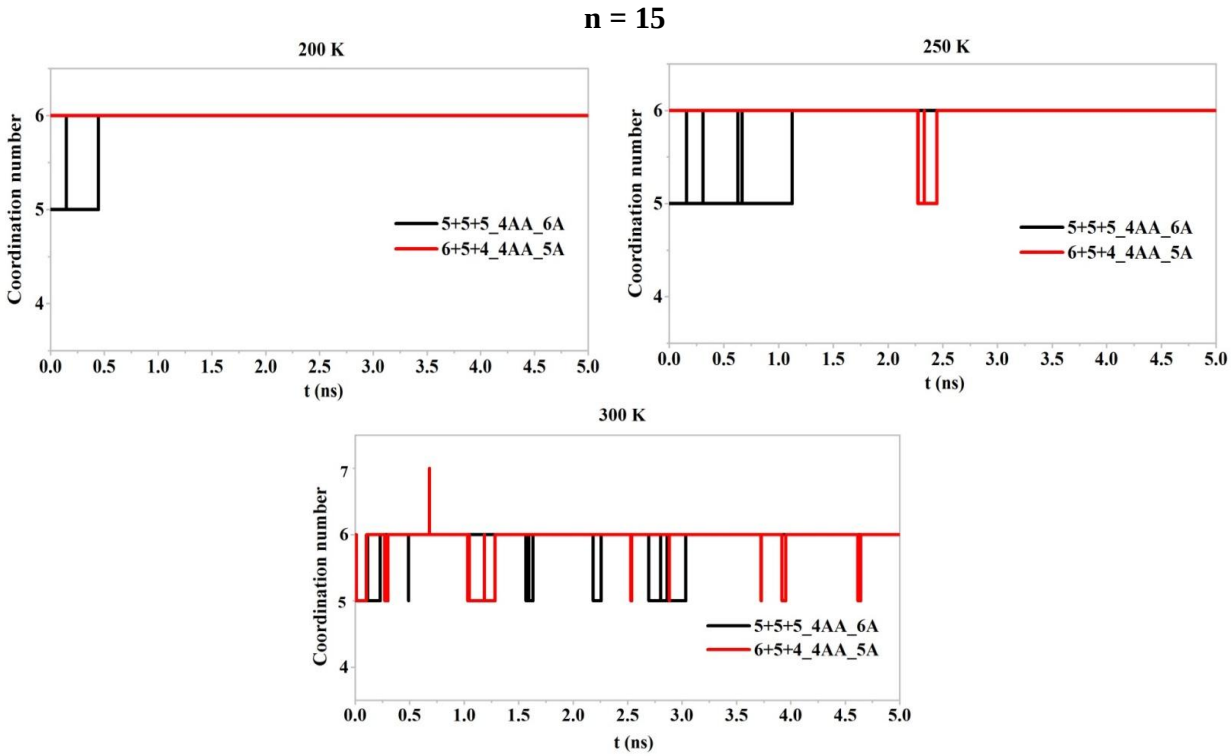


Figure 20 Change of coordination number at each time step for two different starting points of $Zn^{2+}(H_2O)_{15}$ system at 200, 250 and 300K.

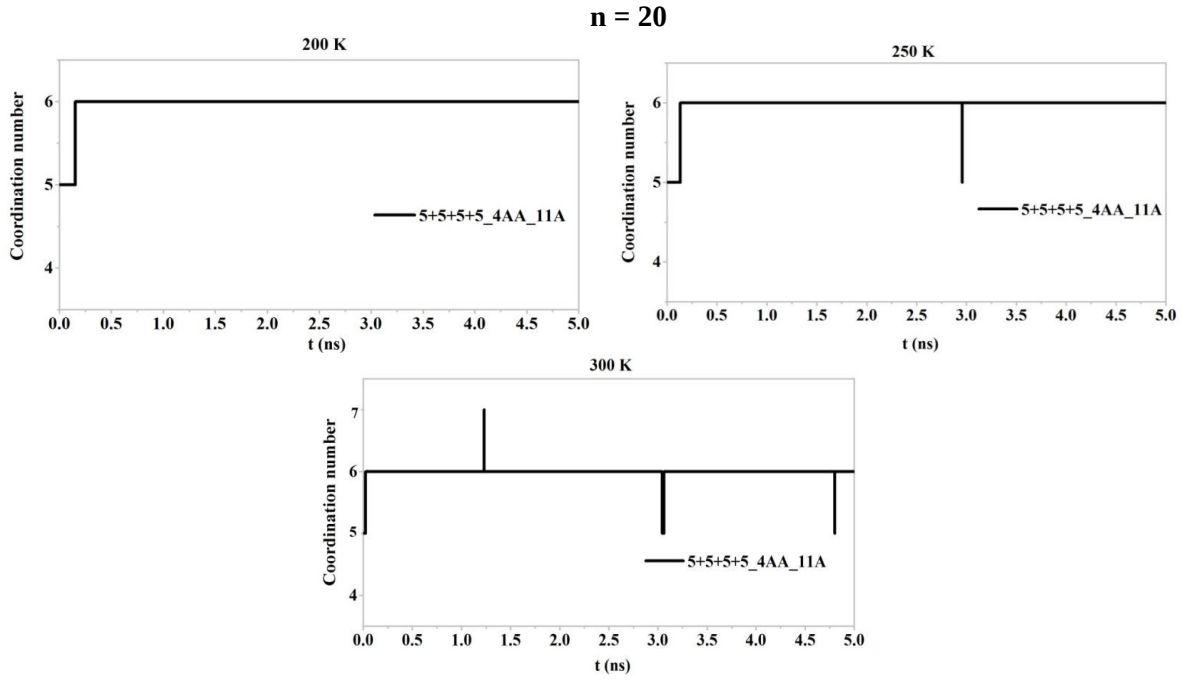


Figure 21 Change of coordination number at each time step for 5+5+5+5_4AA_11A starting point of $Zn^{2+}(H_2O)_{20}$ system at 200, 250 and 300K.

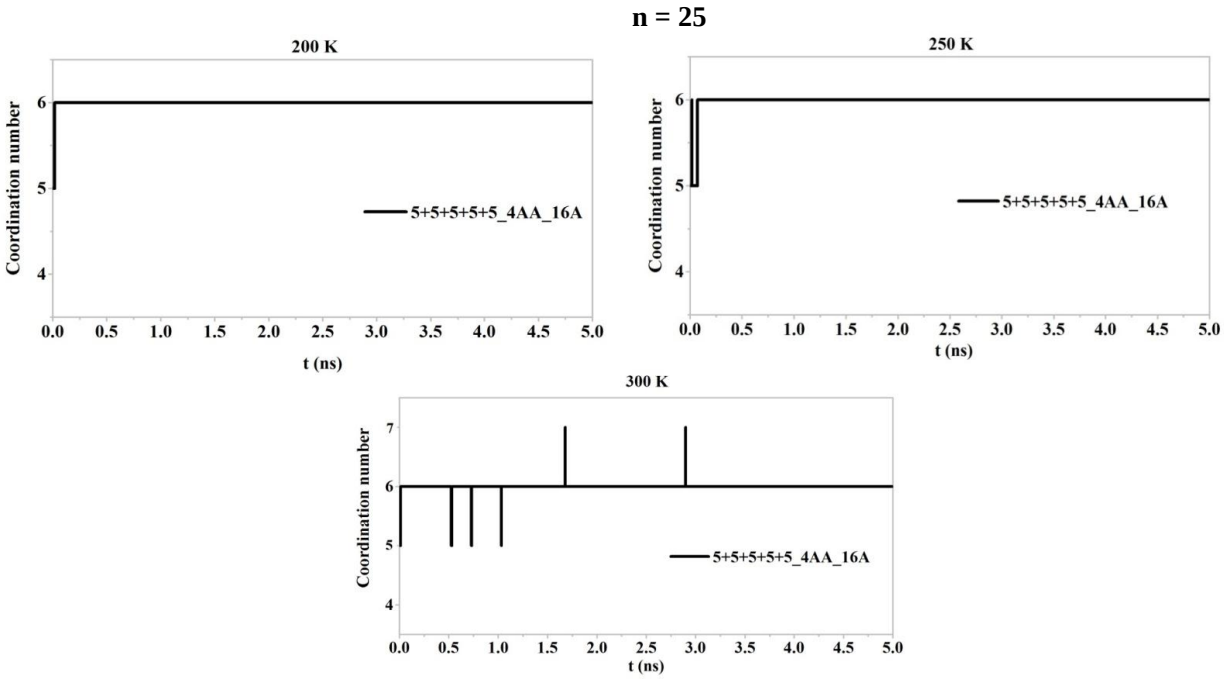


Figure 22 Change of coordination number at each time step for 5+5+5+5+5_4AA_16A starting point of $Zn^{2+}(H_2O)_{25}$ system at 200, 250 and 300K.

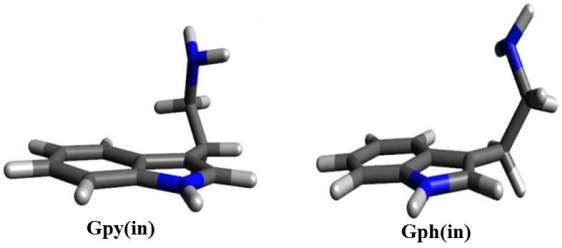
Chapter 4

Structures, dynamics and IR spectra of hydrated tryptamine ions

1. Introduction

Infrared spectroscopy has emerged in the last decade as a powerful technique in the structural elucidation of various gas-phase molecular and cluster ions. For example, this technique is widely used to study the structures of hydrated amino acids and their analogues[1-3]. In recent years, several studies have been carried out on M^+(amino acid) (M = metal ion) systems to resolve the structural discrepancies on metal ions from their bare amino acids. These studies highlight that the structures of M^+(amino acid) ions have strong dependency on the metal-ion interactions, leading to different structures of M^+(amino acid) than bare amino acids[4-8]. As an example, alkali metal ions such as sodium and potassium stabilize various organic molecules and amino acids by electrostatic interactions such as charge-charge, charge-dipole and charge-π interactions[9,10].

Tryptamine (Tryp) is a derivative of the amino acid tryptophan in which the carboxylic acid is replaced by a hydrogen atom. Along with several well-known substituted tryptamines, such as serotonin and melatonin, tryptamine is classified as a neurotransmitter and is believed to play a role in neuropsychiatric disorder[11]. Several studies have been performed to obtain a detailed picture of the conformations of neutral tryptamine[12,13]. Computational DFT studies suggest that neutral tryptamine has nine possible conformers of the ethylamine side chain. However, out of these nine conformers, only seven have been detected in molecular beam experiments. Two high energy conformers, named Gpy(in) and Gph(in), were not observed from experiments. In these conformers, the amino group of ethylamine side chain is in gauche orientation relative to the pyrrole side and phenyl side of indole, respectively. The complete set of tryptamine conformers and their nomenclature have been provided by Zwier and co-workers[13].



In recent reports, Lisy and co-workers have performed a detailed study of the complexes formed by Tryp and Na^+ or K^+, using infrared photodissociation (IRPD) spectroscopy[14,15]. Their experimental studies suggest that in the presence of sodium or potassium ions and zero to three water molecules, these two higher energy conformers are strongly stabilized and can then easily be detected by IRPD spectroscopy. The Na^+ or K^+ cation provides a positive charge that stabilizes efficiently these conformers by simultaneous interaction with the amino group and the indole π electrons. To demonstrate this, they have reported series of IRPD spectra of $M^+(Tryp)(H_2O)_{0-3}Ar_{0-1}$ (M = Na, K) systems. In addition, for these systems they have performed DFT/B3LYP calculations to identify stable structures and to aid the interpretation of the experimental spectra. All of these low energy structures were derived from high energy conformers Gpy(in) and Gph(in) of bare tryptamine. DFT calculations were used to show that several conformers are required to assign some of the experimental spectra, for example based on experimental O-H or aromatic C-H stretching frequencies.

The assignment of vibrational bands, needed to establish a relationship between a spectrum and a molecular structure, requires extensive molecular modeling. Apart from static quantum chemical calculations, such as those carried out by Lisy et al.[15], IR spectra can be calculated from molecular dynamics (MD) simulations. This alternative approach may provide considerable added value as it allows taking into account vibrational anharmonicities, dynamics and temperature effects. Note that the internal energy or effective temperature is an important issue for characterizing structures[16]. A dynamical approach may be particularly useful to reproduce experimental spectra when conformational changes are expected to take place. In the present

study, we aim at calculating spectra from MD simulations for the $Na^+(Tryp)(H_2O)$ and $Na^+(Tryp)(H_2O)_2$ cluster ions. This work turned out to be an opportunity to explore the importance of several "intermolecular" parameters of the AMOEBA force field, in particular the multipolar expansion used to calculate accurate electrostatics. Thus special emphasis is put in this chapter on the determination of these parameters.

2. Computational details

2.1. Quantum chemistry calculations

Lisy et al. identified different low energy conformers of $Na^+(Tryp)(H_2O)_n$ (n=1,2)[15] at the DFT/B3LYP level. They differ by the orientation of the amino group of the ethylamine side chain, water and sodium ion arrangements. These low energy structures are displayed in Figure 1, for n=1 in the upper part and n=2 in the lower. For n=1, in structures Na1a and Na1b, the water molecule is "sandwiched" between the sodium cation and the amino group of the ethylamine side chain, and the water molecule is hydrogen bonded to the amino nitrogen. These two conformers differ by the orientation of the ethylamine side chain. In Na1c and Na1d, the sodium cation is "sandwiched" between the water and the amino group, but the structures differ by the position of the water molecule. In Na1e the water molecule is acceptor of a hydrogen bond from the indole N-H. For $Na^+(Tryp)(H_2O)_2$, the seven conformers shown in Figure 1 were also derived depending upon the amino group, water and sodium ion arrangements. The Na2a and Na2b conformers are analogous to the Na1a and Na1b, respectively in which the additional water molecule is bound to Na^+ with some interaction with the indole ring. Na2c is the combination of Na1c and Na1d. Na2d is analogous to Na1d with the second water molecule only hydrogen bound to the first. Na2e and Na2f are the analogues of Na1a and Na1b with the

additional water molecule hydrogen bonded to the indole N-H. Finally, Na2g is analogous to Na1e with the additional water molecule interacting with the indole π electron cloud.

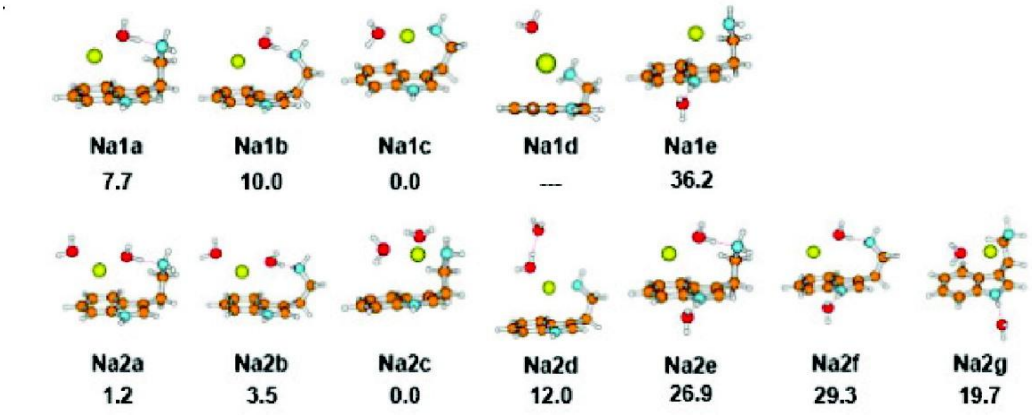


Figure 1 Calculated structures for $Na^+(Tryp)(H_2O)_{1-2}$. Relative energies at the B3LYP/6-31+G* level are in $kJ.mol^{-1}$. This figure was taken from reference 15.

In the present study, we have used these results as a starting point to perform geometry optimizations and frequency calculations at both M06 and MP2 levels with the 6-31G(d,p) basis set. Single point energy calculations have been performed at M06/cc-pVTZ//M06/6-31G(d,p) and MP2/cc-pVTZ//MP2/6-31G(d,p) levels. It has also been checked and found that there is no significant difference between energies computed with the aug-cc-pVTZ and cc-pVTZ basis sets. These calculations were deemed necessary to generate more reliable QM energetics for AMOEBA performance calibration.

2.2. AMOEBA parameter adjustment

AMOEBA parameters are available for the tryptophan residue, the sodium ion and the water molecule, but not specifically for tryptamine. Therefore, some parameter adjustments were necessary. Tryptophan parameters were used as much as possible for tryptamine and the proline parameters were used for the NH_2 terminal group. Standard parameters were used for sodium

and potassium cations. For water, the alternative AMOEBA parameters described in Chapter 2 were chosen to allow the right order of symmetric and anti-symmetric O-H bond stretching frequencies for the isolated molecule. At this stage, several problems remained to reproduce *ab initio* geometries and relative energies of Na^+(Tryp)(H_2O). A detailed analysis revealed that these problems arose from the atomic multipoles used for tryptamine that led us to extract new multipoles.

Multipoles were calculated for neutral tryptamine with the GDMA program, following the procedure described in chapter 2, section 2.4.2. The Gpy(in) conformer was taken to extract multipoles as it is found to be one of the most stable tryptamine conformers in the presence of sodium ion and water molecule[15]. The geometry was optimized atthe MP2/cc-pVTZ level of theory. GDMA sets atomic radii to calculate the atomic multipole expansion. It was found that the hydrogen radius is a sensitive parameter, thus different values were tested: 0.65, 0.55, 0.40, 0.35, 0.325 and 0.31 Å (vide infra). Finally, these DMA multipoles were refined by fitting the QM electrostatic potential at the MP2/cc-pVTZ level of theory to improve the description of electrostatic interactions.

2.3. Simulation protocol

Finite temperature IR spectra were obtained from MD simulations using the AMOEBA force field as described in chapter 2. MP2 optimized structures were used as starting geometries for 50 ps of equilibration followed by a 200 ps productive period to record the dipole moment. A final average spectrum is obtained over 5 such trajectories.

3. Static relative energies and geometries of the complexes

3.1. Quantum chemistry results

The results from this study are compared with the previously reported results[15] obtained at the B3LYP/6-31+G* level (Tables 1 and 2). MP2 optimized geometries for the various $Na^+(Tryp)(H_2O)_n$ (n=1,2) structures are shown in Figure 2. For $Na^+(Tryp)(H_2O)$ (Table 1), we can consider that the differences in relative energies (ΔE) are not significant between MP2 and M06 levels with both 6-31G(d,p) and cc-pVTZ basis sets, although basis set extension tends to reduce energy differences by a few $kJ.mol^{-1}$. Furthermore, the energetic order of the different structures is the same at both levels of theory. However, the difference is larger between B3LYP and M06 or MP2 levels, especially for Na1e. In terms of relative thermal enthalpies (ΔH) or thermal free energies (ΔG), differences between M06 and MP2 levels are always smaller than 3 $kJ.mol^{-1}$. Significant differences, up to 7 $kJ.mol^{-1}$ are obtained between ΔE and ΔG or ΔH values at a given level of theory in few cases. The computational level has a slightly stronger influence on the relative energies of the $Na^+(Tryp)(H_2O)_2$ structures but the trends remain unchanged (Table 2). For low energy conformers, significant energy differences of up to 10 $kJ.mol^{-1}$ are obtained between ΔE and ΔG values in a few cases.

Table 1 Relative energies (ΔE), thermal free energies (ΔG at 298 K) and thermal enthalpies (ΔH at 298 K) at B3LYP, M06 and MP2 levels of $Na^+(Tryp)(H_2O)$. All values are in $kJ.mol^{-1}$.

Structures	B3LYP/ 6-31+G*	M06/6-31G(d,p)			MP2/6-31G(d,p)			M06//M06	MP2//MP2
	ΔE[a]	ΔE	ΔG	ΔH	ΔE	ΔG	ΔH	ΔE[b]	ΔE[b]
Na1a	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Na1b	2.3	3.6	3.3	4.0	3.6	1.7	3.6	3.6	3.8
Na1c	-7.7	-8.2	-15.0	-8.5	-8.9	-16.8	-10.4	-4.9	-5.4
Na1d	-	-8.2	-15.0	-8.5	-8.9	-16.8	-10.4	-4.9	-5.4
Na1_new	-	-9.2	-14.2	-9.1	-9.8	-15.4	-10.9	-6.8	-6.8
Na1e	28.5	37.2	30.7	38.3	36.7	28.0	36.1	34.1	32.1

[a] Relative energies from ref. 15.

[b] Single point energies at the levels shown using cc-pVTZ basis set using geometries calculated with 6-31G(d,p) basis set.

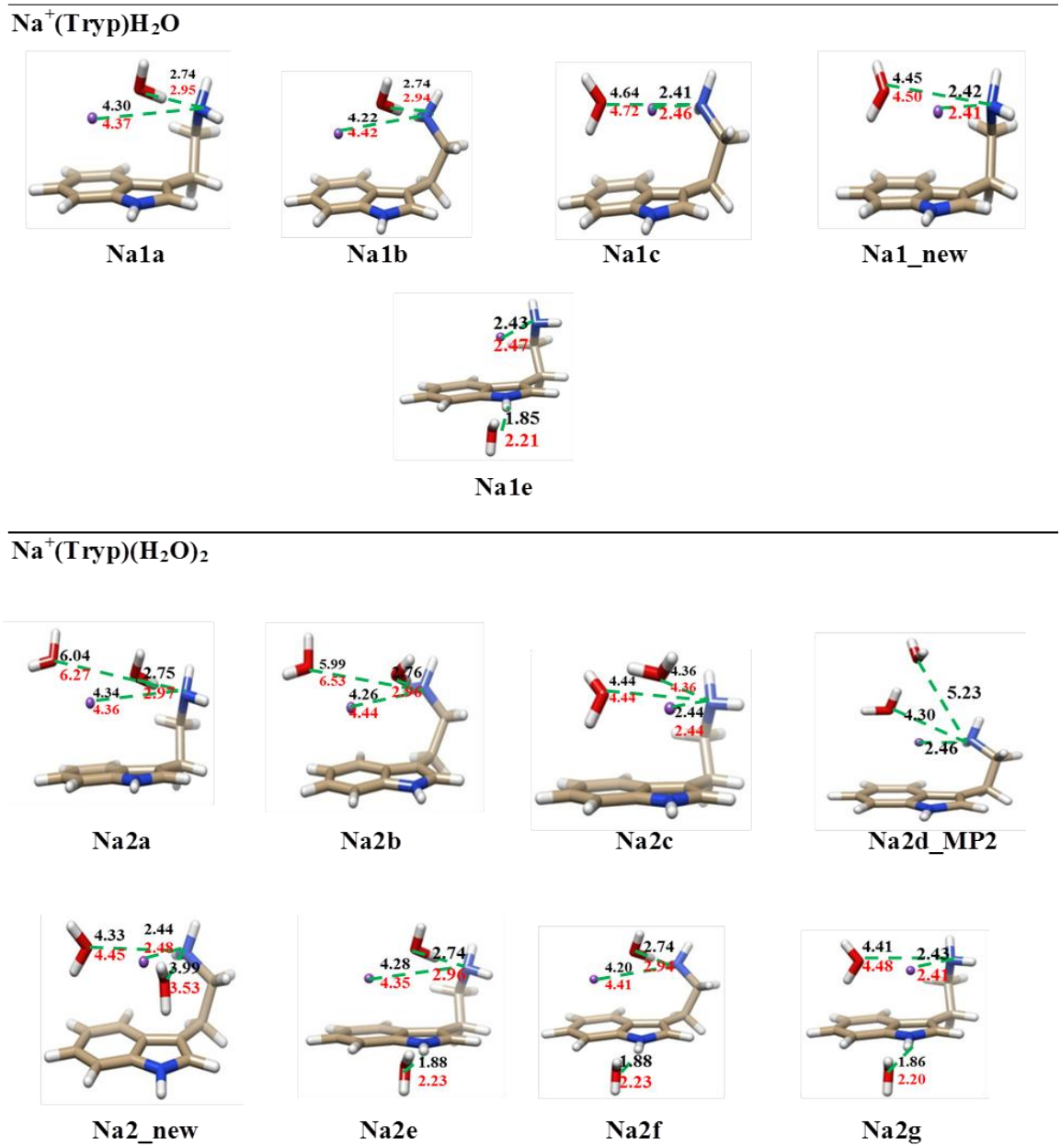


Figure 2 $Na^+(Tryp)(H_2O)$ and $Na^+(Tryp)(H_2O)_2$ structures with optimized distances (Å) at AMOEBA (in red, H-radius = 0.325 Å) and MP2 (in black) levels.

In the case of $Na^+(Tryp)(H_2O)$, at both M06 and MP2 levels, four geometries (Na1a, Na1b, Na1c and Na1e) are well reproduced within the five low energy structures identified previously[15]. However, Na1d switches to Na1c during geometry optimization, hence there is no minimum found for Na1d. The previous study identified that Na1c is the most stable conformer but we have found a new conformer "Na1_new" to be either the most stable energetically or close to Na1c depending on the level of calculation. The Na1c and Na1_new structures differ only by the orientation of the amino group in the ethylamine side chain of tryptamine. Furthermore, Na1e is found to be very high in energy at all levels.

For $Na^+(Tryp)(H_2O)_2$, all seven geometries Na2a-2g previously identified at the B3LYP level by Lisy et al. (see Figure 1) are well reproduced at both M06 and MP2 levels. While Na2a, Na2b, Na2c and Na2d correspond to low energy conformers, Na2e, Na2f and Na2g are found to lie high in energy. In addition, Na2c is found to be the most stable whatever the level of theory. Furthermore, we found a new conformer "Na2_new" as one of the lowest energy structures. These additional Na1_new and Na2_new conformers are discussed in the next section. All these QM calculations provide a database of relative energies and geometries to compare with the results obtained with the AMOEBA force field. Particularly, the MP2 level was chosen as the reference against which to validate the parameter set, to keep the standard procedure of AMOEBA parameterization[17].

Table 2 Relative energies (ΔE), thermal free energies and enthalpies (ΔG and ΔH at 298 K) at B3LYP, M06 and MP2 levels of $Na^+(Tryp)(H_2O)_2$. All values are in $kJ.mol^{-1}$.

Structures	B3LYP/ 6-31+G*	M06/6-31G (d,p)			MP2/6-31G (d,p)			M06//M06	MP2//MP2
	ΔE[a]	ΔE	ΔG	ΔH	ΔE	ΔG	ΔH	ΔE[b]	ΔE[b]
Na2a	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Na2b	2.3	3.9	3.8	4.6	4.0	1.0	3.9	4.3	4.3
Na2c	-1.2	-5.1	-4.8	-5.0	-5.9	-0.9	-6.1	-1.8	-2.8
Na2d	10.8	18.3	13.5	18.9	21.5	10.5	21.6	16.3	19.5
Na2_new	-	-2.9	-3.8	-3.5	-1.4	-0.4	-2.7	0.3	2.5
Na2e	25.7	29.7	33.6	30.1	31.5	37	32.4	26.4	24.1
Na2f	28.1	33.4	36.6	34.2	35.1	38.8	35.9	30.2	27.9
Na2g	18.5	20.4	19.8	20.5	20.8	21.5	20.4	19.4	16.4

[a] Relative energies from ref. 15.

[b] Single point energies at the levels shown using cc-pVTZ basis set using geometries calculated with 6-31G(d,p) basis set.

3.2. AMOEBA results and role of atomic multipoles

For each set of multipoles obtained for different hydrogen radii, i.e. 0.65, 0.55, 0.40, 0.35, 0.325 and 0.31 Å, we have calculated relative energies and geometries to validate the force field. The relative energies of $Na^+(Tryp)(H_2O)$ structures are given in Table 3 for AMOEBA optimized geometries starting from MP2 structures. These relative energies are compared with the MP2 relative energies and the global error is evaluated by RMSD (root mean square deviation). The errors on relative energies are found to be ca. 10% for hydrogen radii ranging from 0.65 - 0.35 Å. The error is at its minimum (5%) for a hydrogen radius of 0.325 Å. Moreover, the radius value has no significant influence on relative energies in the range of 0.65 - 0.35 Å, but large variations are found between 0.35 and 0.31 Å. The best compromise is thus obtained for a hydrogen radius of 0.325 Å. With this value, the AMOEBA and MP2 geometries are shown in Figure 2 with Na-$N(NH_2)$, O---N and OH---N distances. For $Na^+(Tryp)(H_2O)$, all five geometries are well reproduced by AMOEBA. The error on short $Na^+ \ldots N$ distances is usually small (ca.0.04 Å)while

ligand-ligand bond distances have somewhat larger variability for O---N (errors of ca. 0.21 Å) and more significantly, for O…H-N hydrogen bond to the indole N-H (errors of ca. 0.36 Å).

For the selected hydrogen radius of 0.325 Å, the relative energies of $Na^+(Tryp)(H_2O)_2$structures are provided in Table 4, with their RMSD. The error was also evaluated for "low energy" (Na2a, Na2b, Na2c and Na2_new) and "high energy" (Na2e, Na2f and Na2g) energy conformers. The agreement between AMOEBA and MP2 is slightly less good for $Na^+(Tryp)(H_2O)_2$ than for $Na^+(Tryp)(H_2O)$ relative energies. However, the error is very similar for both series of systems in terms of geometries. The force field is able to reproduce the MP2 minima except Na2d, for which the geometry optimization yielded a new minimum in the "low energy" class, named "Na2_new", which was found to exist at the MP2 level as well.

The role of the tryptamine conformation used as reference to extract the multipoles was also checked. The results are very similar to the ones described above if the multipoles are the average of those extracted for two tryptamine conformers, Gpy(in) and Gph(in), with the same hydrogen radius of 0.325 Å.

Table 3 Relative AMOEBA and MP2 energies of $Na^+(Tryp)(H_2O)$ conformers with global RMSD error ($kJ.mol^{-1}$).

Structures	**MP2/ccpVTZ// MP2/6-31G(d,p)**	**AMOEBA Hydrogen radius (Å)**					
		0.65	**0.55**	**0.40**	**0.35**	**0.325**	**0.31**
Na1a	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Na1b	3.8	9.5	10.4	10.3	10.5	7.3	6.0
Na1c	-5.4	-3.5	-3.4	-3.2	-2.6	-7.5	-11.5
Na1e	32.1	35.5	35.1	35.2	36.2	32.0	28.9
RMSD		3.45	3.76	3.76	4.17	2.04	3.62

Table 4 Relative AMOEBA and MP2 energies of $Na^+(Tryp)(H_2O)_2$ conformers with global error RMSD ($kJ.mol^{-1}$).

Structures	MP2/cc-pVTZ// MP2/6-31G(d,p)	AMOEBA H-radius= 0.325 Å
Na2a	0.0	0.0
Na2b	4.3	7.9
Na2c	-2.8	-6.8
Na2_new	2.5	-4.6
Na2e	24.1	33.6
Na2f	27.9	32.9
Na2g	16.4	29.9
RMSD		7.34
Low energy RMSD[a]		4.45
High energy RMSD[b]		9.97

a: Applies to low energy conformers Na2a, Na2b, Na2c and Na2_new
b: Applies to high energy conformers Na2e, Na2f and Na2g

4. IR spectra from MD simulations

4.1. Influence of the multipoles

4.1.1. IR frequencies and intensities

Based on static IR spectra computed at the B3LYP level, Lisy et al. assigned all IRPD bands for $Na^+(Tryp)(H_2O)$[15]. They had to invoke a mixture of two different types of structures, Na1c and Na1d, to assign all bands simultaneously. In the experimental spectrum (Figure 3, bottom), the prominent feature near 3500 cm^{-1} is due to the indole N-H stretch. The weaker features between 3550 and 3750 cm^{-1} are assigned to water O-H stretches. The fact that three bands are present in this region is indicative of at least two conformers, since a single one would generate only two bands. The proposed assignments are: a O-H…π hydrogen bond from Na1c (~3585 cm^{-1}), a symmetric OH stretch from Na1d (~3655 cm^{-1}), and a superposition of free O-H and asymmetric O-H stretches from Na1c and Na1d at indistinguishable frequencies (~3710 cm^{-}

[1]). The weaker amino asymmetric and symmetric stretches are found near 3350 and 3250 cm^{-1}. Finally, the C-H stretches were observed in the 2850-3100 cm^{-1} range (2850-2950 cm^{-1} for aliphatic C-H and 2950-3100 cm^{-1} for aromatic C-H stretches). The authors reported that Na1c accounts for most of the experimental features, however Na1d is required to complete the assignment. Indeed the band at 3655 cm^{-1} is only attributable to the symmetric O-H stretching a Na1d-typestructure. However, the static DFT/B3LYP calculations were not able to identify the Na1d isomer as a stable minimum on the potential energy surface, as it switches to Na1c during geometry optimization.

The computed IR spectrum at 100 K, displayed in Figure 3 (part a, red), is obtained after extraction of multipoles (AMOEBA_initial) with H-radius of 0.325 Å. In comparison with experiments, the indole N-H stretch is computed at 3504 vs 3495 $cm^{-1}$with very low intensity whereas this band is intense experimentally. On the contrary, the water asymmetric and symmetric stretches in Na1d are computed at 3713 and 3612 cm^{-1}, respectively, with much higher intensities than in the experimental spectrum. Thus the computed symmetric O-H stretch is ca. 40 cm^{-1} to the red of the experimental band, so that it is superposed with the O-H---π band of the Na1c structure, leading to only two apparent bands in the computed spectrum. The amino asymmetric stretch at 3365 $cm^{-1}$and symmetric stretches in the 3280-3350 cm^{-1}range also have much higher calculated than experimental intensity. The aromatic (3000-3060 cm^{-1}) and aliphatic (2850-2950 cm^{-1}) C-H stretch bands are in the right frequency ranges but with reverse relative intensities in comparison with experiment. Summarizing, in all cases but one, agreement is good for frequencies but not for intensities.

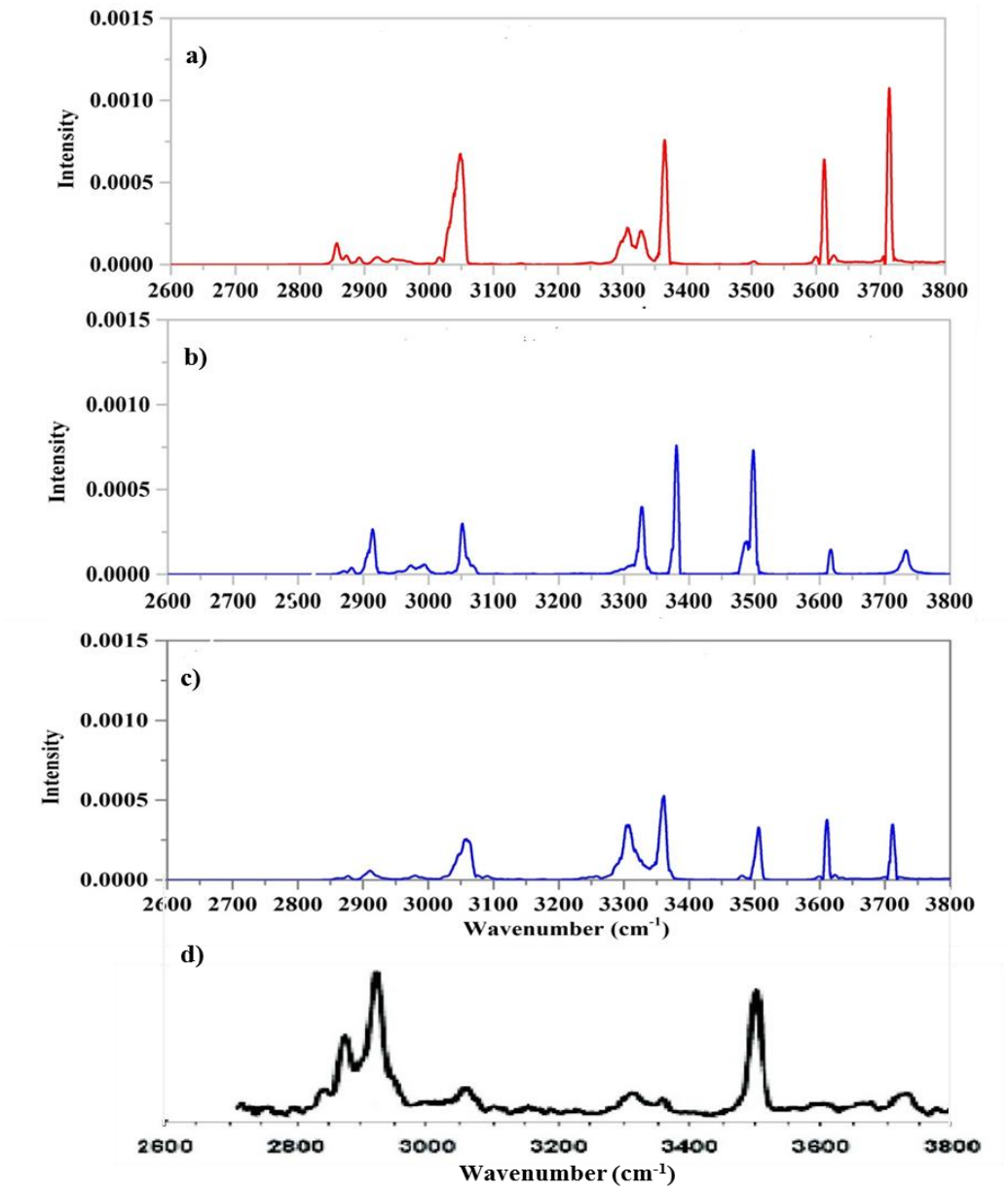


Figure 3 Calculated IR spectrum of Na^+(Tryp)(H_2O) from MD simulation at 100 K starting from Na1c conformer, with various multipole sets : (a) AMOEBA_initial; (b) with hydrogen charges decreased by 0.1 electron units except for water and NH_2 (AMOEBA_mod.1: global charge = 0.00); same as b plus increasing carbon charges by 0.1 electron units (AMOEBA_mod.2: global charge = 1.00). (d) Experimental IRPD spectrum from reference 15.

To overcome these intensity issues, various modifications of multipole parameters have been considered. Changing atomic charges strongly impacts bond dipoles and their derivatives. Since IR intensity of a mode is proportional to the squared derivative of the dipole moment with respect to the vibration coordinate of that mode, it is expected to be strongly dependent upon atomic charges. After various attempts, it has been found that the charges on H atoms are playing the main role for indole N-H and C-H band intensities. Our first approach was therefore to modify charges on hydrogen atoms of bare tryptamine. The charges on all hydrogen atoms were reduced by 0.1electronunits except for amino hydrogens. The resulting spectrum is shown in Figure 3b (AMOEBA_mod.1). It can be seen that frequency changes are relatively small with much improved agreement on intensities (typically 5, at most 20 cm^{-1}). However, this change leads to a global charge of 0.0 instead of +1.0. Therefore, to balance the global charge, our second attempt was to decrease the charges on all tryptamine H atoms by 0.1 electron as before, and simultaneously increase the charges on all carbon atoms by 0.1 electron. The corresponding spectrum is displayed in Figure 3c. Agreement with the experiment is good again for frequencies while it is improved, although still not perfect, for intensities while maintaining the correct global charge of +1.0. Intensity variations are explained in the next section from detailed trajectory analysis.

4.1.2. Analysis of the trajectories

The MD trajectory analysis for the spectrum shown in Figure 3a (AMOEBA_initial) is given in Figure 4. The latter consists in a histogram of number of occurrences as a function of the N_{indole}-H_{water} distance. This distance allows defining rough domains for various structural types. It can be seen that although the starting point is Na1c, it is Na1d which is mainly

populated. The system also explores regions where H_2O is far away from the indole, as well as, although marginally, a region in which H_2O is bound to the indole N-H. This region was subjected to AMOEBA and MP2 geometry optimizations, leading to the finding of a new low energy conformer (Na1_new1, see Table 5), not identified previously by Lisy et al. Note that Na1d in Figure 4 is taken directly from MD trajectory as Na1d always switches to Na1c during geometry optimizations as discussed above.

These diverse conformers lead to different signatures in the spectrum due to their different structural features. In the new conformer Na1_new1, one water O-H bond points toward the indole N through hydrogen bonding and the water molecule also directly interacts with the sodium ion. As a consequence it is also relatively close to the amino group of ethylamine side chain. While inNa1c, one water O-H bond interacts with the π-electron cloud of indole and the other O-H bond is free, in Na1d both O-H bonds are free. Thus, Na1d contributes only to asymmetric and symmetric free O-H stretches.

Figure 4 clearly indicates that Na1d is highly populated as compared to Na1c, with a very small population of Na1_new1. The histograms related to the spectra (AMOEBA_mod.1 and AMOEBA_mod.2) in Figures 3b and 3c are shown in Figure 5. For AMOEBA_mod.1, the population of Na1d decreases while those ofNa1_new1 and Na1c increase. The same trend is observed for AMOEBA_mod.2, although with smaller magnitude.

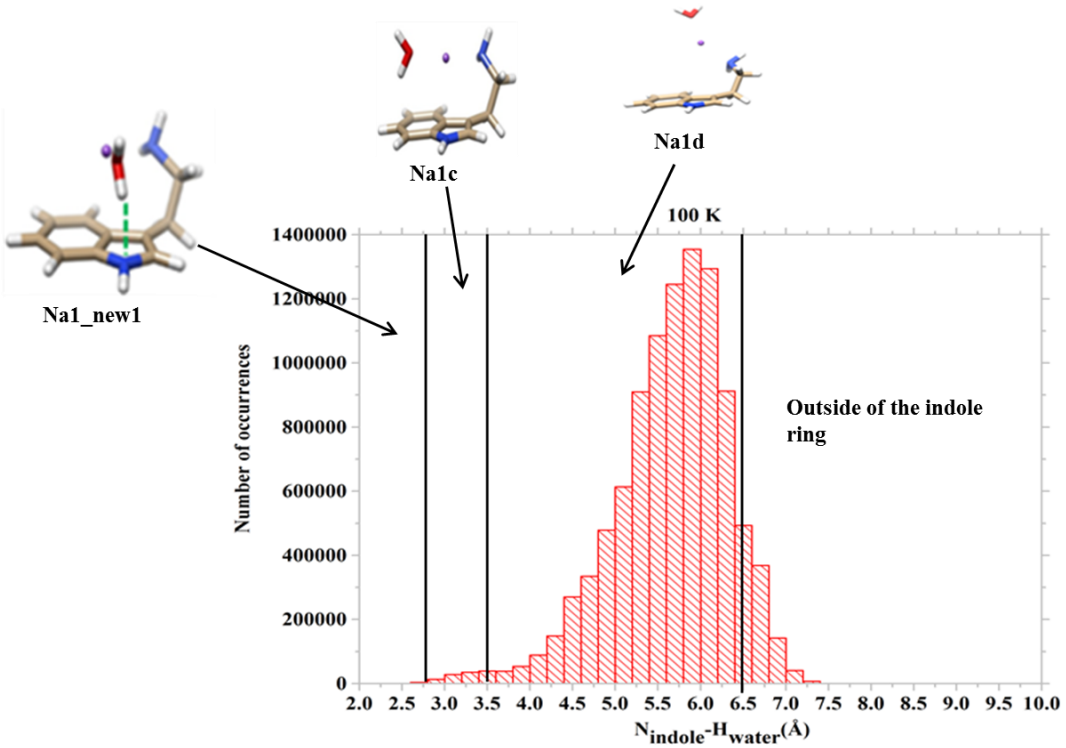


Figure 4 Histogram for N_{indole}-H_{water} distance over 5 trajectories, corresponding to Figure 3a (AMOEBA_initial).

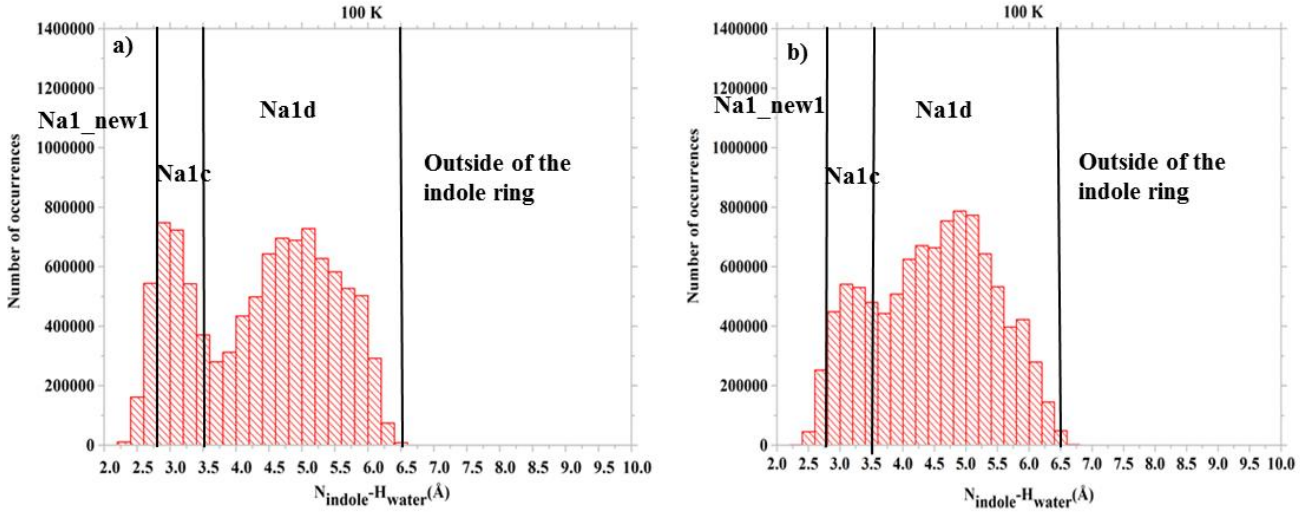


Figure 5 Histograms of occurrences of N_{indole}-H_{water} distance over 5 trajectories. a) corresponds to Figure 3b (AMOEBA_mod.1) while b**)** corresponds to Figure 3c (AMOEBA_mod.2).

Table 5 Relative MP2 and AMOEBA energies of $Na^+(Tryp)(H_2O)$ conformers ($kJ.mol^{-1}$).

Structures	MP2/cc-pVTZ//MP2/6-31G (d,p)	AMOEBA_ initial	AMOEBA_mod.1 Global charge = 0.00	AMOEBA_mod.2 Global charge = 1.00
Na1a	0.0	0.0	0.0	0.0
Na1b	3.8	7.3	7.6	2.1
Na1c	-5.4	-7.5	-8.6	-9.8
Na1_new	-6.8	-10.5	-14.5	-16.5
Na1_new1	1.6	-7.5	-8.6	-9.8
Na1e	32.1	32.0	34.6	-16.5

In order to assign the variation of water band intensities, contour plots of occurrence as a function of C-C-C-O_{water} dihedral angle and C_{indole}-H_{water} distance were computed for one representative trajectory of each spectrum. For example, Figure 6 displays a contour plot related to the spectrum in Figure 3a (AMOEBA_initial), with the two structural parameters shown on the right. The C-C-C-O_{water} dihedral angle (in orange) represents the position of water molecule relative to the indole ring and the C_{indole}-H_{water} distance (in green) discriminates between conformers with similar CCCO dihedrals. Thus, a contour plot describes the water position above or away from the indole, and the populations of the different structures as well. A dihedral angle between 110 and 55° corresponds to the water molecule above the ring as in e.g. Na1c or Na1d while other values are associated to water being near the indole plane as in e.g. Na1e. Distances in the 2.0-2.8 and 2.8-3.5 Å ranges correspond to Na1_new1 andNa1c, respectively while 3.5-7.5 and 7.5-9.5 Å ranges correspond to Na1d and the structures with the water molecule far away from the indole, respectively. All these distance and dihedral angle ranges were taken from structures found during MD trajectories. Figure 6 clearly indicates that the population of Na1d is very high while that of Na1c is much smaller. This figure is correlated with the population histogram of Figure 4. The contour plots related to the spectra

(AMOEBA_mod.1 and AMOEBA_mod.2**)** in Figures 3b and 3c are reported in Figure 7. The latter indicates that in both cases there is significant population of both Na1c and Na1d with a higher population for Na1d (see also the histograms of Figure 5). Band assignment is found to be more straightforward from such contour plots with two structural parameters than from histograms. This representation is therefore used in the following discussion.

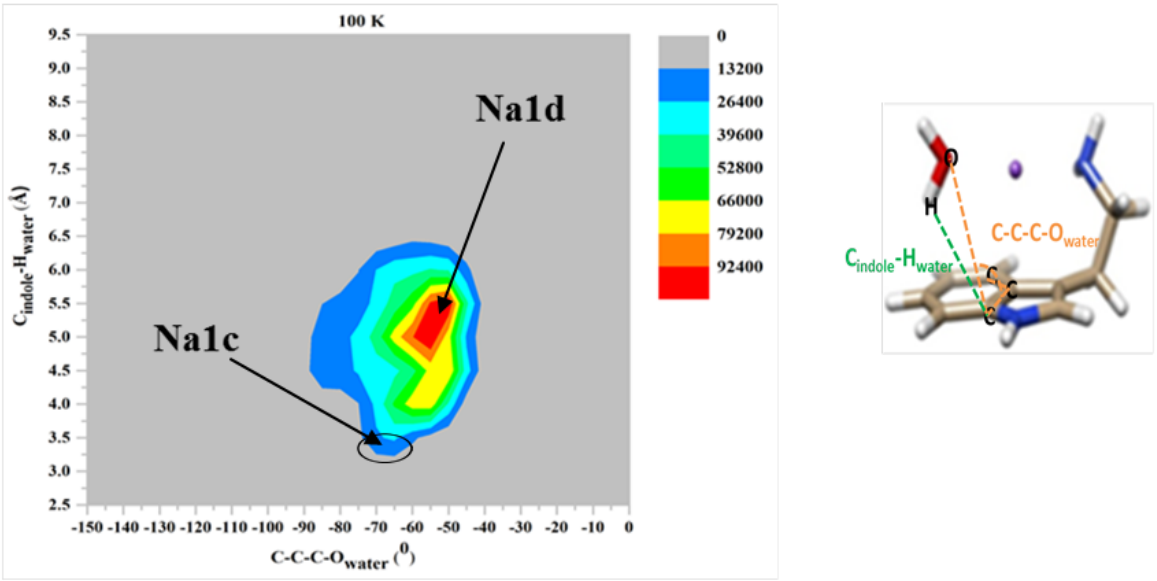


Figure 6 Contour plot of occurrence of C-C-C-O_{water} dihedral angle and C_{indole}-H_{water} distance, corresponding to Figure 3a (AMOEBA_initial). The color code and occurrence scale are given on the right side of the figure (i.e. red color represents high population).

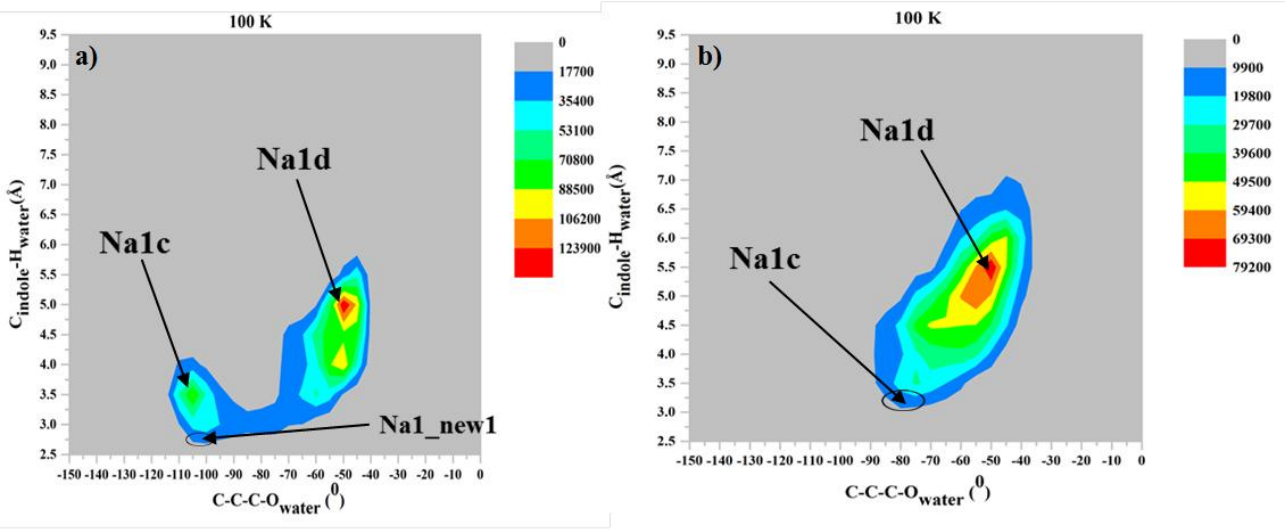


Figure 7 Contour plots of occurrence of C-C-C-O_{water} dihedral angle and C_{indole}-H_{water} distance. a) corresponds to Figure 3b (AMOEBA_mod.1) while b) corresponds to Figure 3c (AMOEBA_mod.2).

4.1.3. Relationship between structures and spectra

The relative abundances of the 3 conformers Na1c, Na1d and Na1_new1 during MD simulations and the analysis of the time evolution of structural parameters can be correlated with the differences obtained in IR band frequencies and intensities. The intense water bands at 3713 and 3612 cm^{-1}in Figure 3a (AMOEBA_initial) are mainly due to the asymmetric and symmetric O-H stretches from Na1d and from a small contribution of Na1c (see Figure 6). The very little indole N-H stretching band near 3500 cm^{-1} in Figure 3a is slightly red shifted (~6 cm^{-1}) and much more intense in Figure 3b (AMOEBA_mod.1). This is due to the population of Na1_new1 (see Figures 5a and 7a) which has a strong hydrogen bond between water and the indole N. In addition in Figure 3b, the indole N-H band (near 3500 cm^{-1}) is also a doublet however not well separated, due to a mixing of Na1c and Na1_new1.

The water band intensities change from Figures 3a to 3b due to a change in the populations of different structures. The contour plot in Figure 7a (spectrum in Figure 3b) shows

that water bands must arise mainly from Na1c and Na1d while the band near 3733 cm^{-1} is due to asymmetric stretch in Na1d and free single O-H stretch in Na1c. The band near 3615 cm^{-1} corresponds to a combination of π…H-O hydrogen bond from Na1c and symmetric stretches from Na1c and Na1d.The low intensity of water bands in Figure 3b is due to the very high population of the Na1c structure while the high intensity in Figures 3a and 3c is due to the high population of Na1d conformer. The populations of the 3 conformers tend to be more even for AMOEBA_mod.2 (Figure 5b) but the trends remain unchanged, making spectral assignments straightforward (Figure 3c). When several structures are explored during the simulations, the O-H bands are broadened and thus less intense (Figure 3b and 3c). The change in band intensities of the NH_2 stretches near 3250-3400 cm^{-1} and the O-H in the 3600-3700 cm^{-1} range from Figures 3b to 3c are again due the change in the populations of Na1c, Na1d and Na1_new (see Figure 7). Both spectra obtained with modified multipoles AMOEBA_mod.1 and AMOEBA_mod.2 (Figures 3b and 3c**)** are in reasonably good agreement with the experimental spectrum and these parameter sets also provide good energetics especially for the low energy conformers (see Table 5).

For AMOEBA_mod.2 (global charge = 1.0), the highest energy conformer Na1e switches to another structural minimum named "Na1_new". The latter, not identified before, was found to be the most stable at both AMOEBA and MP2 levels (Table 1). The Na1c and Na1_new conformers differ only by the orientation of the amino group in the ethylamine side chain of tryptamine. The IR spectrum was nevertheless computed from MD simulations starting from Na1_new, using the AMOEBA_mod.2 multipoles; it is displayed in Figure 8a. It is similar to that in Figure 3c (starting from Na1c) except for the symmetric NH_2 stretching band at 3280-3350 cm^{-1}. The contour plot in Figure 9 shows that the populations of Na1d-likestructures and

the structures with the water molecule away from the indole are both very high. The symmetric stretch of NH_2 band is a doublet in Figure 8 due to the mixing of Na1_new and the analogous of Na1d.

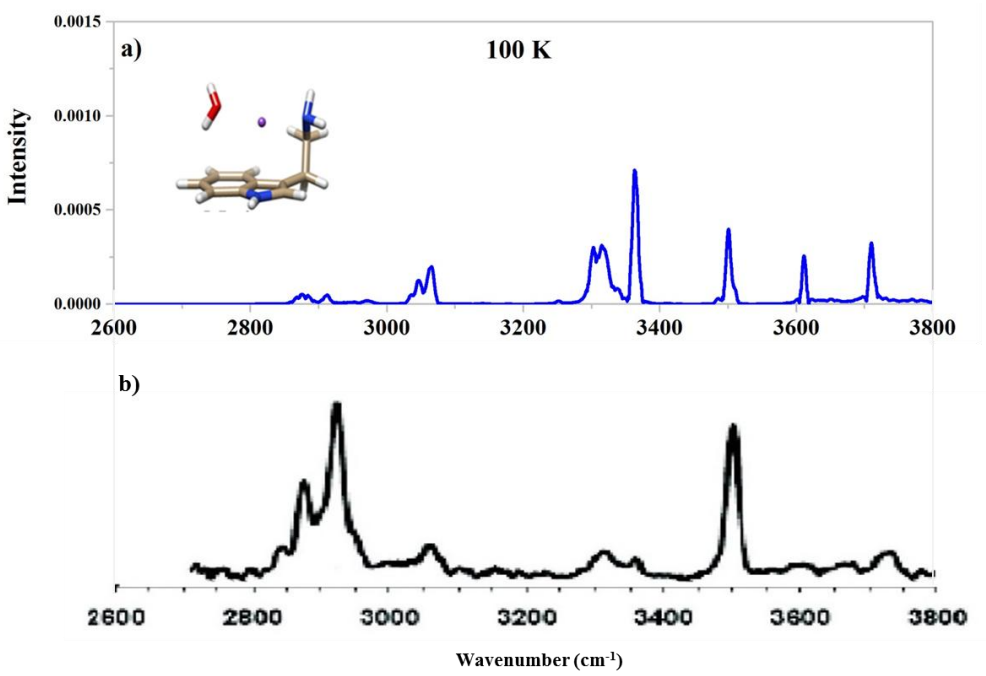


Figure 8 (a) Calculated IR spectrum of of Na^+(Tryp)(H_2O) from MD simulation at 100 K, starting from Na1c_new as shown, using the AMOEBA_mod.2 multipole set(global charge =1.00). (b) Experimental IRPD spectrum from reference 15.

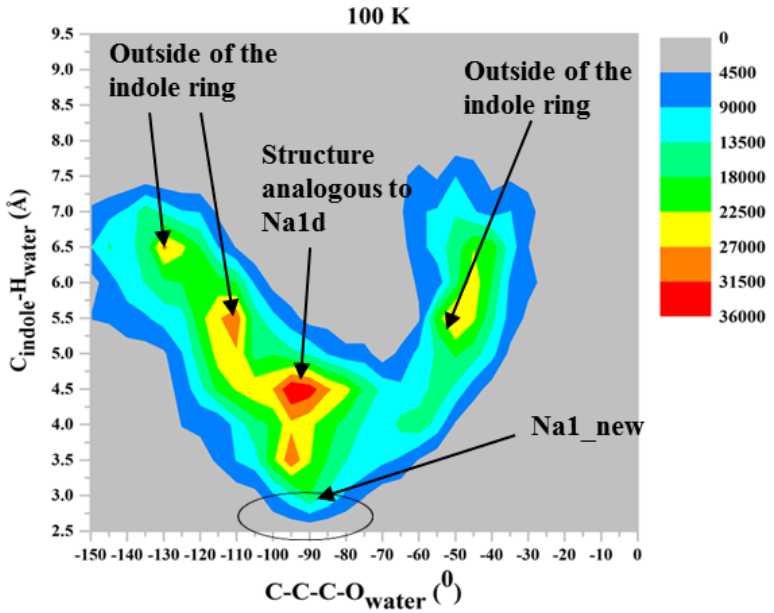


Figure 9 Contour plot of occurrence of C-C-C-O_{water} dihedral angle and C_{indole}-H_{water} distance from one trajectory using the AMOEBA_mod.2 multipole set, corresponding to Figure 8.

4.2. Analysis of IR spectra for 5 independent trajectories

The results described above indicate that water band intensities may be significantly influenced by the multipole set and by the starting structure as well. In order to get a clear picture of these variations, the IR spectra obtained from 5 different trajectories at 100 K from the AMOEBA_mod.2set of multipoles are shown in Figure 10 along with the average of the 5 spectra (Figure 10, bottom). The starting structure is Na1a, however a quick structural change occurs in all cases, to Na1c and Na1d types of conformation. Consequently, the average spectrum is similar to that obtained starting from Na1c described previously, except for water bands which are more intense. This is again due to different populations of structures, as described in the Appendix (Figure A1). This bias of different populations can be avoided by averaging several spectra; it was checked that averaging 5 or 10 spectra did not lead to any significant difference for both frequencies and intensities.

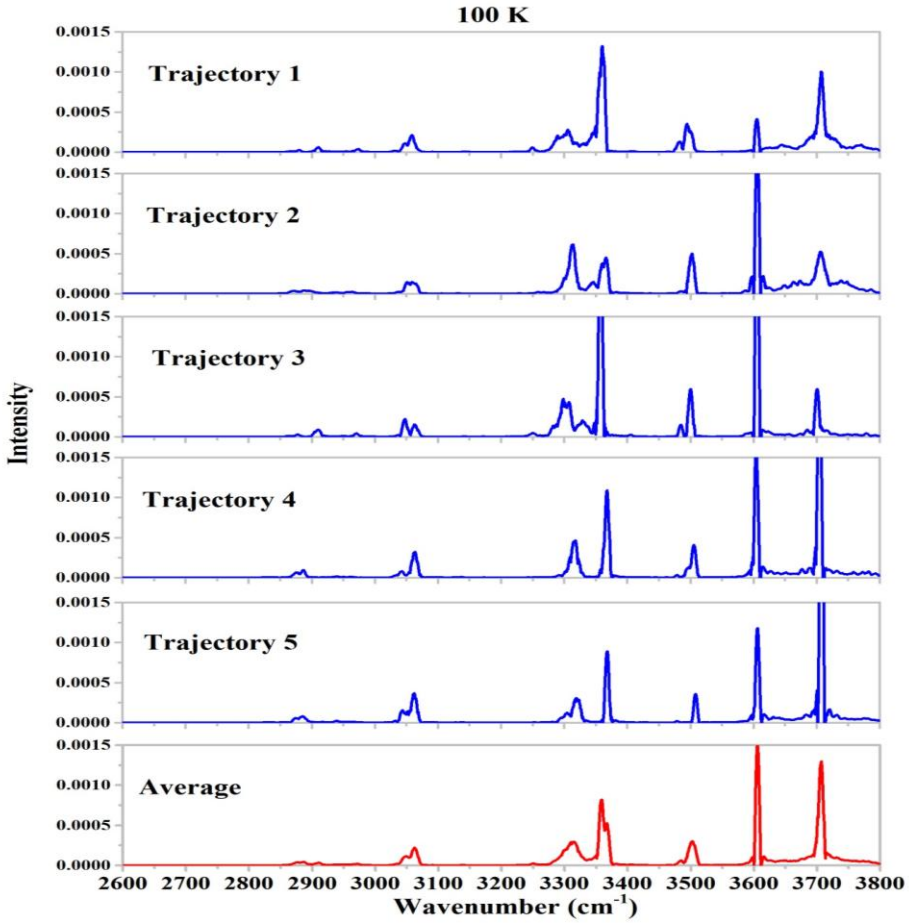


Figure 10 Calculated IR spectra from 5 trajectories plus average spectrum (bottom, red) from MD simulation at 100 K, starting from Na1a with the AMOEBA_mod.2 multipole set.

4.3. Temperature effects

With the aim to apprehend if there are frequency shifts and band broadenings with increasing temperature, IR spectra have been computed at 100, 200 and 300 K starting from Na1a; they are shown in Figure 11. As mentioned above, the starting structure changes quickly to Na1c and Na1d-type conformations as shown by an analysis of the trajectories at various temperatures (see Figures A2-A4, appendix). The mean distance between the sodium ion and amino nitrogen is ca. 2.5 Å and the magnitude of the fluctuation is small (ca. 0.15 Å) for the three temperatures which indicates that a strong electrostatic interaction is maintained. The Na-O_{water} distance also shows small fluctuations (~0.15 Å), which indicates that during the simulation, the water molecule

always follows the sodium ion whatever the temperature. The variation of the N_{amino}-C-C-C dihedral angle of the ethylamine side chain attests that there is no side chain conformation exchange at 100 K while few exchanges appear at 200 K. At 300 K, several exchanges occur for all trajectories.

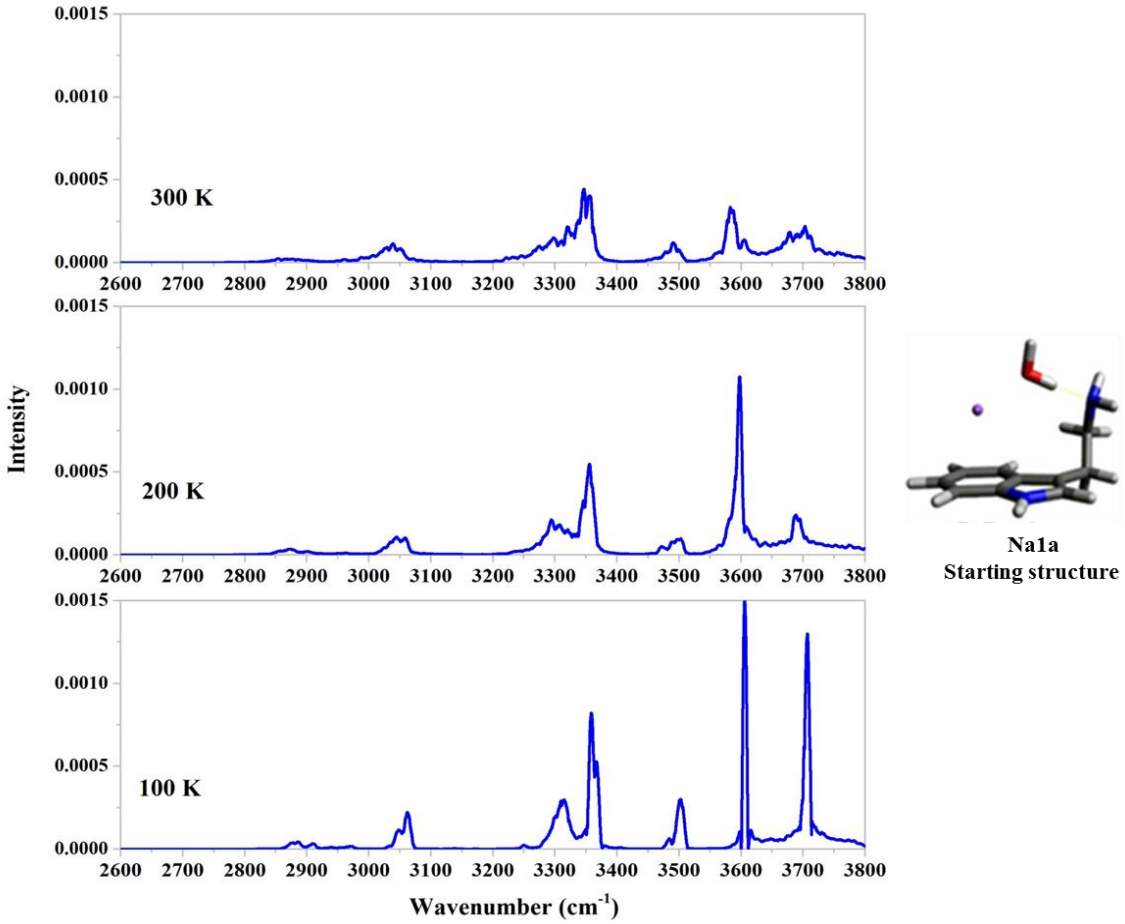


Figure 11 Calculated IR spectra from MD simulation at 100, 200 and 300 K, starting from Na1a, using the AMOEBA_mod.2 multipole set.

IR spectra starting from Na1a at 100, 200 and 300 K are presented in Figure 11. The spectra are obtained at 200 and 300 K indicate that frequency changes with temperature are very small. As described above the exchange of side chains occurs at 200 and 300 K but not at 100 K. Therefore except for water bands, the spectra at 200 and 300 K are similar to that at 100 K with slightly higher intensity for NH_2 symmetric stretch and indole NH stretch bands at 100 K. However significant peak broadening occurs with increasing temperature. In order to describe this broadening contour plots, analogous to those discussed above for 100 K trajectories, are shown in Figure 12 at 200 and 300 K. It can be seen that at these temperatures, the water molecule remains for long times away from the indole π system, thus with both O-H bonds free. In addition at 200 K, the population of Na1d-type structures is much higher than that of Na1c. Thus again different populations of structures lead to changes in water band intensities.

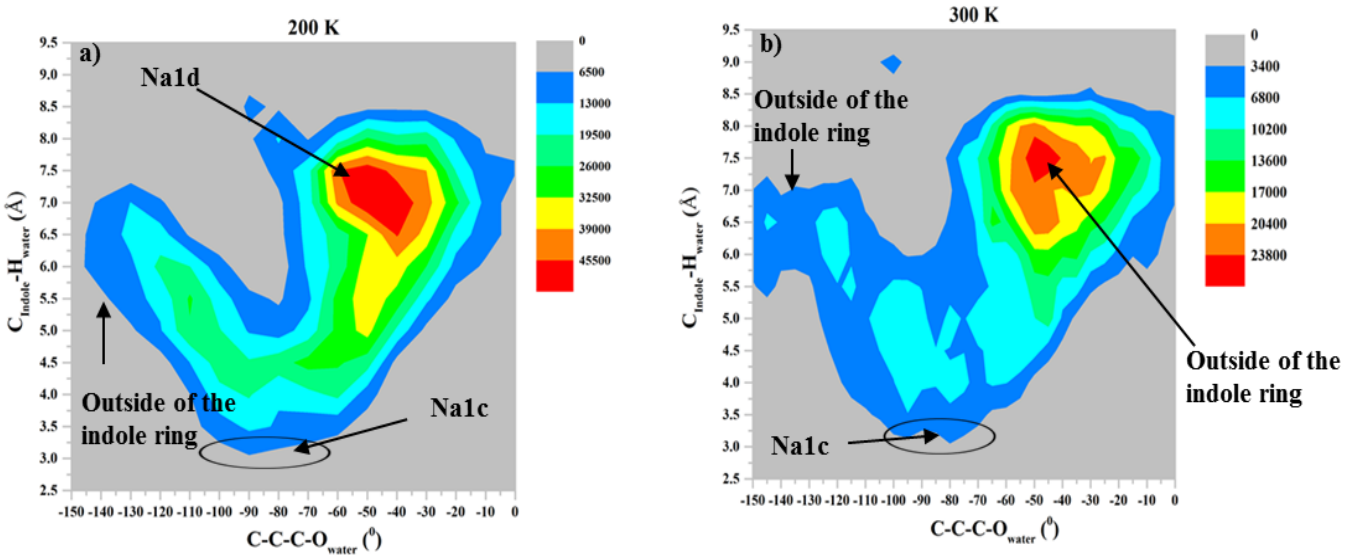


Figure 12 Contour plot occurrence of C-C-C-O_{water} dihedral angle and C_{indole}-H_{water} distance from one trajectory at (a) 200 K (b) 300 K, corresponding to spectra in Figure 11.

4.4. IR spectra for $Na^+(Tryp)(H_2O)_2$

IR spectra starting from the Na2c conformer of $Na^+(Tryp)(H_2O)_2$ were calculated at 100, 200 and 300 K; they are displayed in Figure 13 for comparison with the experimental IRPD spectrum[15]. Lisy et al. described Na2c as a combination of Na1c and Na1d ; combining the spectra for the two structures of $Na^+(Tryp)(H_2O)$ allowed them to account for all experimental features of $Na^+(Tryp)(H_2O)_2$. Indeed, Na2c contains a water molecule which is involved in a π-type hydrogen bond (as in Na1c) and the second water molecule interacts only with the sodium ion (as in Na1d). The experimental spectrum of $Na^+(Tryp)(H_2O)_2$ is similar to that of $Na^+(Tryp)(H_2O)$ with the same features in the N-H (near 3500 cm^{-1}), O-H (OH…π hydrogen bonding, symmetric, and free/asymmetric stretches near 3580, 3640, and 3710 cm^{-1}, respectively), NH_2 (near 3350 cm^{-1}) stretching regions as well as the same C-H (2800-3300 cm^{-1}) signatures as in the $Na^+(Tryp)(H_2O)$ spectrum. The calculated spectra of $Na^+(Tryp)(H_2O)_2$ starting from Na2c (Figure 13, top) and of $Na^+(Tryp)(H_2O)$ are also similar, except for water bands. Contour plots using the same two structural parameters as above (C-C-C-O_{water} dihedral angle and C_{indole}-H_{water} distance) are used to assign the variation of water band intensities (Figure 13) at 100 K. Figure 14a indicates a high population of Na1c and a low population of Na1d for the first water molecule. However, the second water molecule is mainly outside of the indole ring (Figure 14b). This can explain contributions of both Na1c and Na1d conformers in the spectrum. In addition, the asymmetric O-H stretch near 3710 cm^{-1} is an incompletely resolved doublet (~ 4 cm^{-1}), due to the combination of asymmetric stretches from Na1c (~ 3710 cm^{-1}) and Na1d (~ 3714 cm^{-1}). Furthermore, in the calculated spectrum, the O-H stretch near 3600 cm^{-1} is a well separated doublet (~ 12 cm^{-1}) due to a combination of symmetric O-H stretches from Na1c, Na1d (~ 3607 cm^{-1}) and π…H-O stretch from Na1c(~ 3619 cm^{-1}) while in the experimental

spectrum these two bands are present at 3640 and 3580 cm^{-1}, respectively. In the calculated spectrum at 100 K, the water band positions are in slightly less good agreement with experiments. This may be due to the absence of charge transfer term in the AMOEBA force field. At 200 and 300 K (Figure 13, top) all bands become broad but frequency changes are very small compared to 100 K.

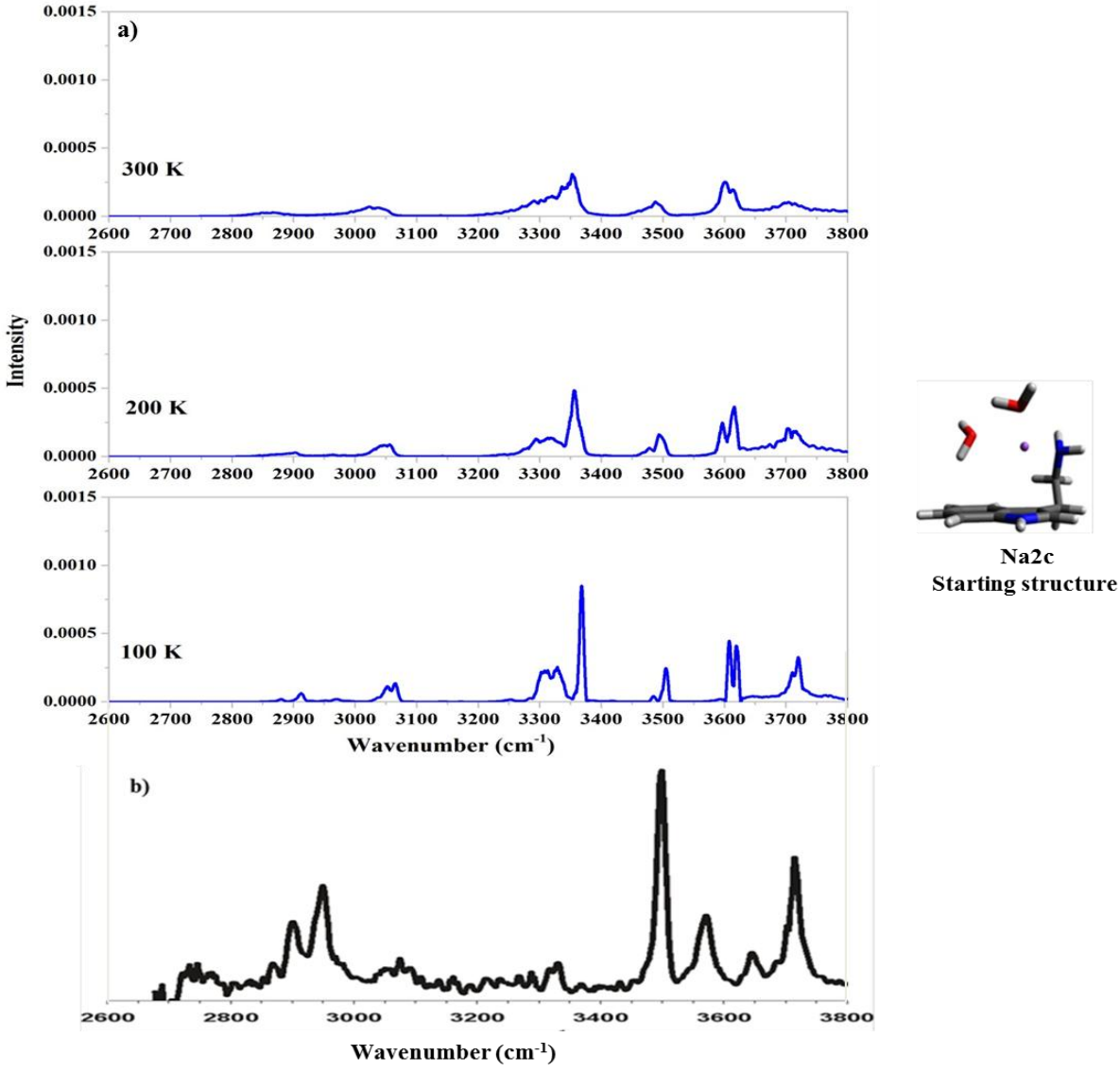


Figure 13 (a) Calculated IR spectra of $Na^{+}(Tryp)(H_2O)_2$ from MD simulation at 100, 200 and 300 K, starting from Na2c, using the AMOEBA_mod.2 multipole set. (b) Experimental IRPD spectrum taken from ref.15.

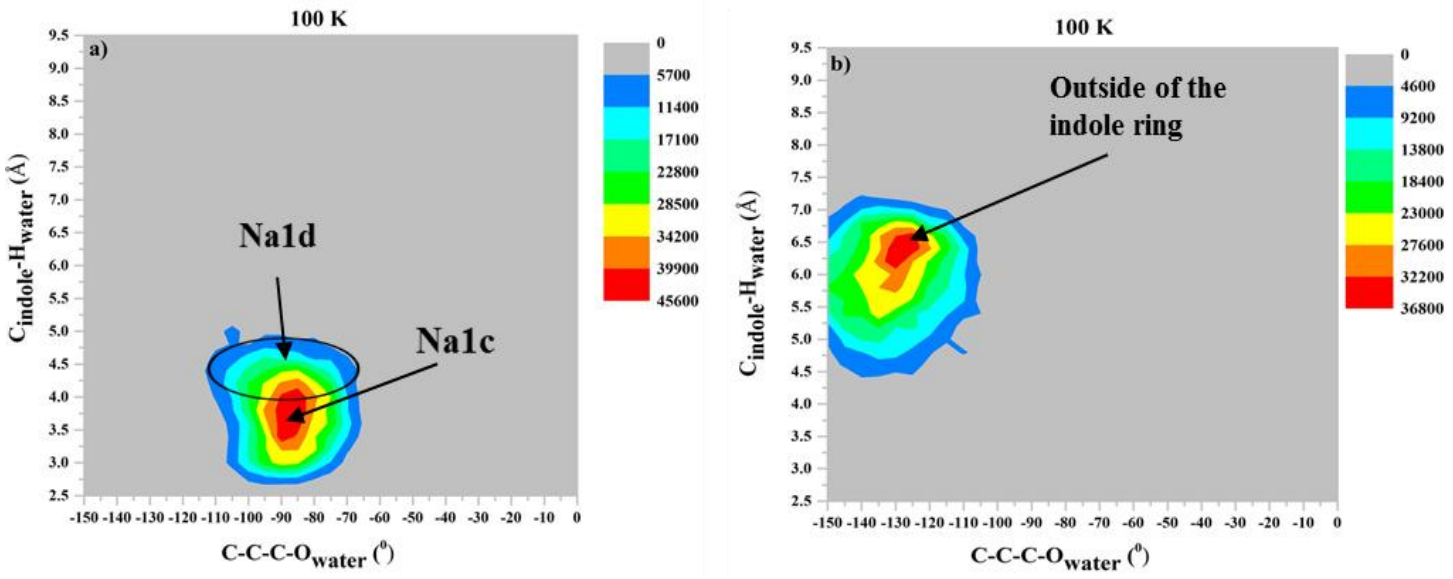


Figure 14 Contour plot of C-C-C-O_{water} dihedral angle vs. C_{indole}-H_{water} distance from one trajectory for (a) first and (b) second water molecules, for spectrum at 100 K in Figure 13 (AMOEBA_mod.2).

4.5. Structures and IR spectra using GEM multipoles

The Gaussian Electrostatic Model (GEM) is an alternative method to calculate the distributed multipoles to be used in the electrostatic part of the potential energy[18-20]. GEM was developed based on the density fitting formalism (Coulomb fitting)[21-22] where Gaussian auxiliary basis sets centered on specific sites of the molecule are used to expand the electron density. The electrostatic contribution can be determined from any level of QM theory that produces a relaxed one-electron density matrix by employing density fitting. A lambda parameter in the "Tikhonov regularization"[22] is used to control noise in the fit. For our purpose, GEM distributed multipoles (GEM-DM) use the same local frame definition than for AMOEBA. Cisneros showed a successful application of GEM-DM in AMOEBA in which multipoles were calculated for various systems such as water and the alanine dipeptide[24]. A comparison with GDMA multipoles

emphasized good agreement for interaction energies, electrostatic potentials and radial distribution functions.

Here, we present the extraction of new GEM-DM multipoles for neutral tryptamine. Densities were obtained at the MP2(full)/aug-cc-pVTZ, MP2(full)/cc-pVTZ and MP2(full)/aug-ccVTZ(-f) levels, the latter being obtained by deleting f-orbitals functions. Then density fitting was done with A2 and A4 auxiliary basis sets[25-26] for the various levels. The local frame definition for tryptamine was taken from the available local frame definition of 3-ethylindole in AMOEBA, analogous to the tryptophan one. The GEM-DM multipoles extracted using both aug-cc-pVTZ and aug-cc-pVTZ(-f) basis sets show that the atomic charges, particularly charges on carbon atoms, are very large in comparison with GDMA atomic charges for neutral tryptamine (Columns "a" and "b" in Table 6). Consequently, the multipoles are not able to reproduce the geometries by complexation with the sodium ion and the water molecule (Figure A5, Appendix).

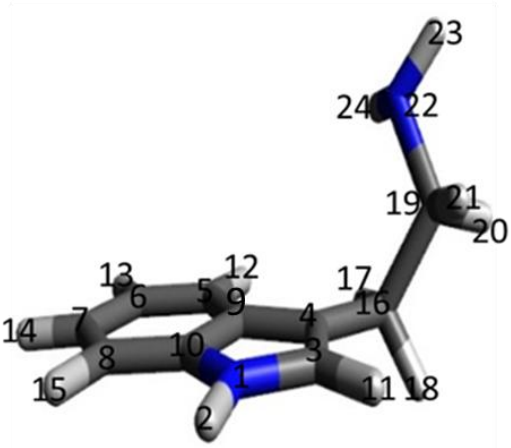
Furthermore, the use of the MP2(full)/cc-pVTZ density provided a clear improvement of carbon charges (Column "c"), but not enough to obtain a good geometry of the Na^+(Tryp)(H_2O) complex. Several tests have been performed to improve these results. The first attempt was to change the lambda parameter in Tikhonov regularization used to converge the multipoles. However, the problems of large carbon charges remained leading to a wrong geometry. Second, these multipoles were fitted to the electrostatic potential using "Potential" routine in Tinker, but without a clear improvement on charges and optimized structures. Few changes in the local frame definition led to improvements of the multipoles. Finally, manual adjustment of some hydrogen charges were done to decrease their too negative values, and this was balanced by a change of the N atom charge of the NH_2 group. These "final" multipoles (Table 6, column d)

produced better geometries for the system. However, the structure Na1a switches to Na1c during geometry optimization (Figure A5, Appendix).

From these multipoles, one AMOEBA/DACF spectrum starting from Na1c was computed at 100 K and cross-compared with the spectrum obtained using GDMA multipoles (AMOEBA_mod.2) and the experimental spectrum (see Figure 15). The GEM-DM spectrum (Figure 15, red) is similar to the GDMA one (Figure 15, blue), except for the position of the NH_2 bands which are blue shifted by ca. 50 cm^{-1}. Furthermore, the band intensities of the C-H stretch frequencies (2800 – 3100 cm^{-1}) are in better agreement with experiments as compared to the GDMA spectrum. However, the indole N-H band (ca. 3500 cm^{-1}) is very low in intensity in disagreement with the experimental spectrum. In overall, the spectrum obtained using GEM mulipoles is in less good agreement with experiment than the one obtained from a GDMA multipole set even with several charge adjustments. Consequently, these results suggest that the GEM-DM method needs further developments to be efficient for the computation of IR spectrum by molecular dynamics.

Table 6 Atomic charges obtained for tryptamine from GEM-DM and GDMA methods.

Atoms		GEM charges				GDMA charges
		a	b	C	D	
1	N	-0.20695	-0.40223	-0.25807	-0.25807	-0.19721
2	H	0.05002	0.05003	0.03778	0.03778	0.26049
3	C	0.72818	0.75682	0.36111	0.36111	0.14476
4	C	0.10567	0.05019	-0.36260	-0.36260	-0.31677
5	C	0.57166	0.60203	-0.02961	-0.02961	-0.14329
6	C	0.82676	0.83208	0.43295	0.43295	-0.02711
7	C	0.42210	0.42564	0.47186	0.47186	-0.10209
8	C	0.28737	0.08980	-0.03078	-0.03078	-0.14871
9	C	-2.43542	-2.42772	-0.43238	-0.43238	-0.28076
10	C	0.23202	0.60266	0.40821	0.40821	0.16941
11	H	0.02727	0.03472	0.00797	0.00797	0.15243
12	H	-0.05177	-0.04352	-0.08111	-0.04361	0.11876
13	H	-0.09828	-0.09970	-0.12030	-0.08280	0.10335
14	H	-0.08516	-0.08842	-0.11776	-0.08026	0.11038
15	H	-0.03594	-0.03097	-0.06151	-0.02401	0.11946
16	C	0.06432	0.04004	0.07053	0.07053	-0.15554
17	H	-0.09574	-0.09774	-0.09886	-0.06136	0.08574
18	H	-0.09574	-0.09774	-0.09886	-0.06136	0.09114
19	C	0.16598	0.17112	0.31265	0.31265	0.06543
20	H	-0.11657	-0.11986	-0.12354	-0.08604	0.09802
21	H	-0.11657	-0.11986	-0.12354	-0.08604	0.01662
22	N	-0.29745	-0.26233	-0.25224	-0.55224	-0.56848
23	H	0.07712	0.06748	0.04405	0.04405	0.20326
24	H	0.07712	0.06748	0.04405	0.04405	0.20326



a = MP2(full)/aug-cc-pVTZ
b = MP2(full)/aug-cc-pVTZ(-f)
c = MP2(full)/cc-pVTZ
d = Final multipoles used for spectrum

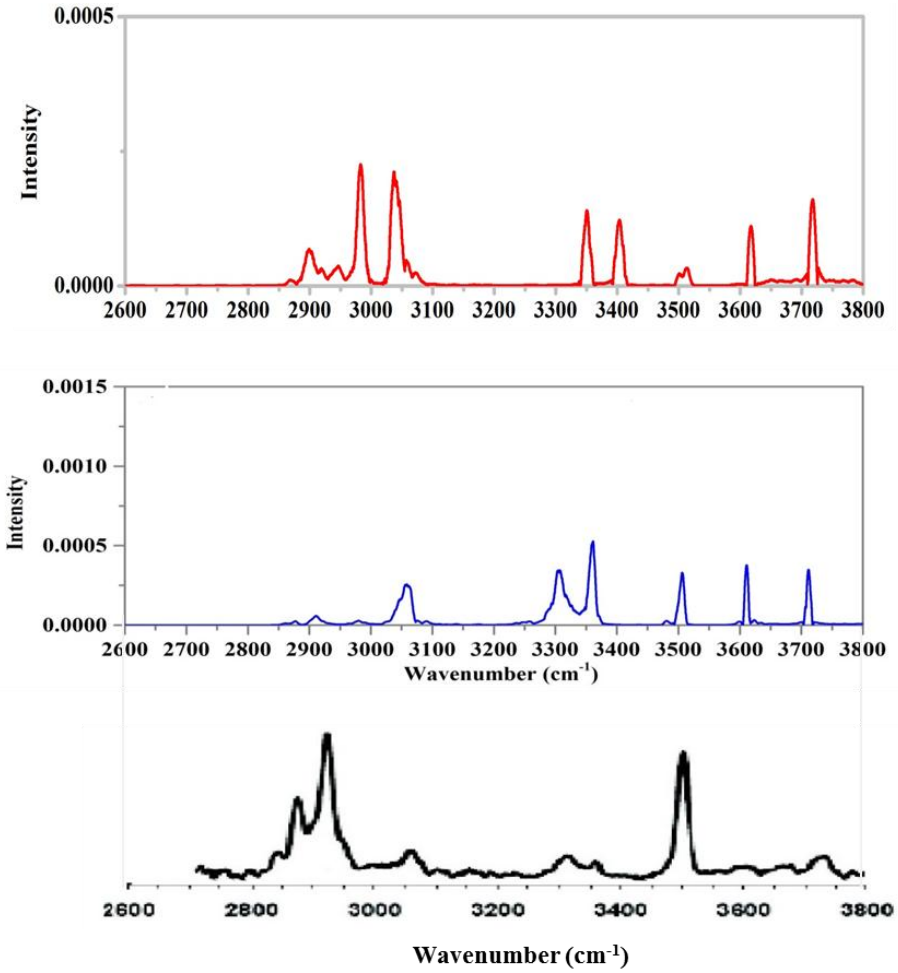


Figure 15 Calculated IR spectra starting from Na1c conformer at 100 K using GDMA (AMOEBA_mod.2) (blue) and GEM-DM (red) multipoles (Table 6, column "d") in comparison with experimental spectrum (black) taken from ref.15.

5. Conclusions

We provide structural and dynamical properties as well as IR spectra of $Na^+(Tryp)(H_2O)$ and $Na^+(Tryp)(H_2O)_2$, by quantum chemistry calculations and molecular dynamics simulations. We have extended the AMOEBA polarizable force field to the tryptamine molecule to provide potential energy surfaces in close agreement with *ab initio* reference data. Atomic multipoles were extracted with the distributed multipole analysis using the GDMA package to refine

electrostatics, and validated against both geometric and energetic *ab initio* data in the gas phase. IR spectra were obtained for various low energy structures and compared with the experimental spectra obtained in Lisy et al. Our results show that the AMOEBA force field is able to reproduce various tryptamine conformers including most low energy ones. The spectrum obtained initially from GDMA multipoles not being in good agreement with experimental intensities, we have found that modifying atomic charges enables much improved agreement in terms of both band positions and intensities. Therefore, our studies clearly indicate that atomic multipoles, especially atomic charges, have very high impact on computed IR spectra. We have also been able to characterize temperature effects on IR spectra. During molecular dynamics simulations we found that various structural changes occur depending on temperature, leading to spectra which differ sometimes from those calculated with different starting structures. This underlines the importance of time frame for computing IR spectra for flexible molecules from MD simulations, an issue which sets a clear distinction between classical and quantum energy surfaces in MD approaches, as QM methods are currently limited to ca. 50 ps. Finally we have also obtained IR spectra using GEM multipoles, however we found that GDMA multipoles provide the best results so far for IR spectrum as well as for geometries. The accuracy achieved with AMOEBA demonstrates the potential of this force field for the realistic modeling of hydrated tryptamine and other flexible cluster ions.

6. References

(1) von Helden, G.; Compagnon, I.; Blom, M. N.; Frankowski, M.; Erlekam, U.; Oomens, J.; Brauer, B.; Gerber, R. B.; Meijer, G. *Phys. Chem. Chem. Phys.* **2008**, *10*, 1248.

(2) Ebata, T.; Hashimoto, T.; Ito, T.; Inokuchi, Y.; Altunsu, F.; Brutschy, B.; Tarakeshwar, P. *Phys. Chem. Chem. Phys.* **2006**, *8*, 4783.

(3) Yongjun, H.; Da, X. *Prog. Chem.* **2009**, *21*, 1592.

(4) Armentrout, P. B.; Rodgers, M. T.; Oomens, J.; Steill, J. D. *J. Phys. Chem. A* **2008**, *112*, 2248.

(5) Rodgers, M. T.; Armentrout, P. B.; Oomens, J.; Steill, J. D. *J. Phys. Chem. A* **2008**, *112*, 2258.

(6) Bush, M. F.; Oomens, J.; Saykally, R. J.; Williams, E. R. *J. Phys. Chem. A* **2008**, *112*, 8578.

(7) Bush, M. F.; Forbes, M. W.; Jockusch, R. A.; Oomens, J.; Polfer, N. C.; Saykally, R. J.; Williams, E. R. *J. Phys. Chem. A* **2007**, *111*, 7753.

(8) Kapota, C.; Lemaire, J.; Maitre, P.; Ohanessian, G. *J. Am. Chem. Soc.* **2004**, *126*, 1836.

(9) Ma, J. C.; Dougherty, D. A. *Chem. Rev.***1997**, *97*, 1303.

(10) Hoyau, S.; Norrman, K.; McMahon, T. B.; Ohanessian, G. *J. Am. Chem. Soc.* **1999**, *121*, 8864.

(11) Mousseau, D. D.; Butterworth, R. F. *J. Neurochem.* **1994**, *63*, 1052.

(12) Carney, J. R.; Zwier, T. S. *J. Phys. Chem. A* **2000**, *104*, 8677.

(13) Zwier, T. S. *J. Phys. Chem. A* **2001**, *105*, 8827.

(14) Nicely, A. L.; Miller, D. J.; Lisy, J. M. *J. Am. Chem. Soc.* **2009**, *131*, 6314.

(15) Nicely, A. L.; Lisy, J. M. *J. Phys. Chem. A* **2011**, *115*, 2669.

(16) Vaden, T. D.; Lisy, J. M.; Carneqie, P. D.; Pillai, E. D.; Duncan, M. A. *Phys. Chem. Chem. Phys.* **2006**, *8*, 3078.

(17) Shi, Y.; Xia, Z.; Zhang, J.; Best, R.; Wu, C.; Ponder, J. W.; Ren, P. *J. Chem. Theory Comput.* **2013**, *9*, 4046.

(18) Piquemal, J-P.; Cisneros, G. A.; Reinhardt, P.; Gresh, N.; Darden, T. A. *J. Chem. Phys.* **2006**, *124*, 104101.

(19) Cisneros, G. A.; Piquemal, J-P.; Darden, T. A. *J. Chem. Phys.* **2005**, *123*, 044109.

(20) Cisneros, G. A.; Piquemal, J-P.; Darden, T. A. *J. Chem. Phys.* **2006**, *125*, 184101.

(21) Boys, S. F.; Shavit, I. *A Fundamental Calculation of the Energy Surface for the System of Three Hydrogen Atoms*; AD212985, NTIS: Springfield, 1959.

(22) Dunlap, B. I.; Connolly, J. W. D.; Sabin, J. R. *J. Chem. Phys.* **1979**, *71*, 4993.

(23) Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. *Numerical Recipes in Fortran77; the Art of Scientific Computing, 2nd ed.; Cambridge University Press: New York, 1992; pp 799−803.*

(24) Cisneros, G. A. *J. Chem. Theory Comput.* **2012**, *8*, 5072.

(25) Andzelm, J.; Wimmer, E. *J. Chem. Phys.* **1992**, *96*, 1280.

(26) Godbout, N.; Andzelm, J. *DGauss, Version 2.0, 2.1, 2.3, 4.0; Computational Chemistry List, Ltd.: Columbus, OH, 1999. The file that contains the A1, A2, and P1 auxiliary basis sets can be obtained from the CCL WWW site at http://www.ccl.net/cca/data/basis-sets/ DGauss/basis.v3.html.*

Appendix

Discussion about the analysis of IR spectra for 5 independent trajectories

The asymmetric O-H stretch near 3600 cm^{-1} has lower intensity than the symmetric stretch near 3700 cm^{-1} while the reverse holds for trajectories 2 and 3; both intensities are similar in trajectories 4 and 5. A detailed analysis of the trajectories shows that during the simulations, water and sodium arrangements are exactly like in either Na1c or Na1d, with changes only in side chain orientation. When the water molecule is oriented as in Na1c conformer, one O-H bond is involved in a π….H-O interaction and the other O-H bond is free. In the structures with the water molecule oriented outside of the indole ring and in the Na1d conformer, the two O-H bonds are free. The contour plots in Figure A1 (in Appendix) clearly indicate that the population of Na1c-type structure with hydrogen bonded O-H is higher than the Na1d-type one for trajectory 3. The populations are reverse for trajectory 1 that leads to change of water bands intensities. For trajectory 4, there is a high population of structures with the water molecule outside of the indole ring, then, both water bands are highly intense. The conclusion is that, starting from the same structure, spectra calculated from different starting velocities may have different intensities due to change of populations.

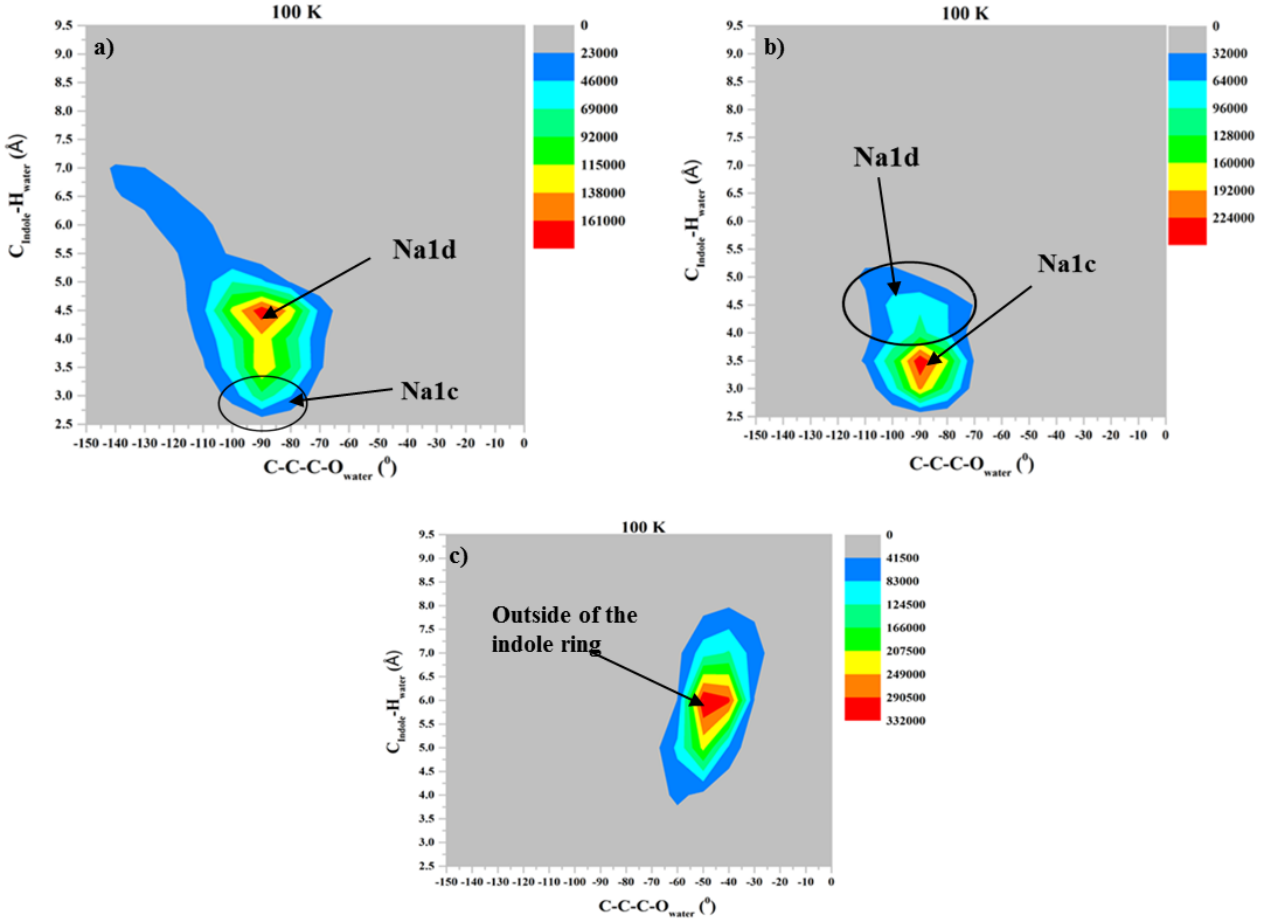


Figure A1. Contour plot of **C-C-C-O$_{water}$** dihedral angle vs. **C$_{indole}$-H$_{water}$** distance from **(a)** trajectory 1, **(b)** trajectory 3 and **(c)** trajectory 4 in relation with Figure 10 (AMOEBA_mod.2).

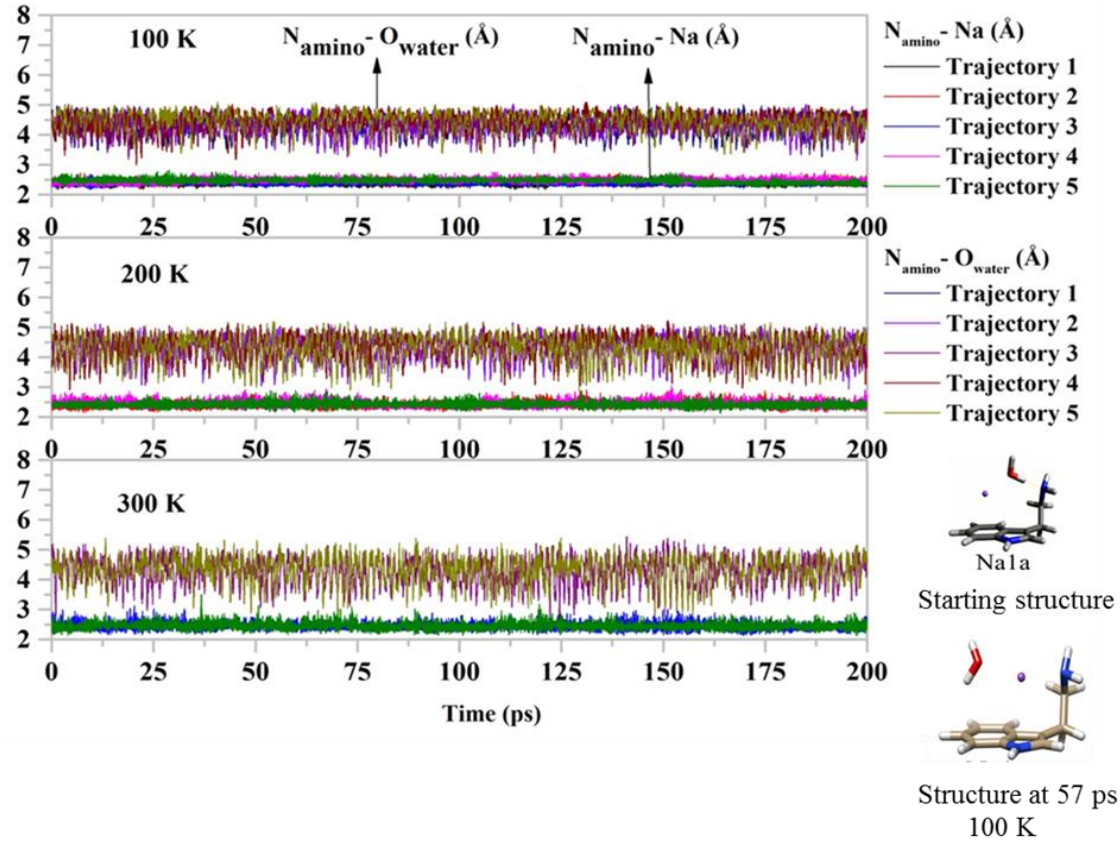


Figure A2. Time evolution of N_{amino}-O_{water} and N_{amino}-Na distances at each temperature (100, 200 and 300 K) from 5 trajectories.

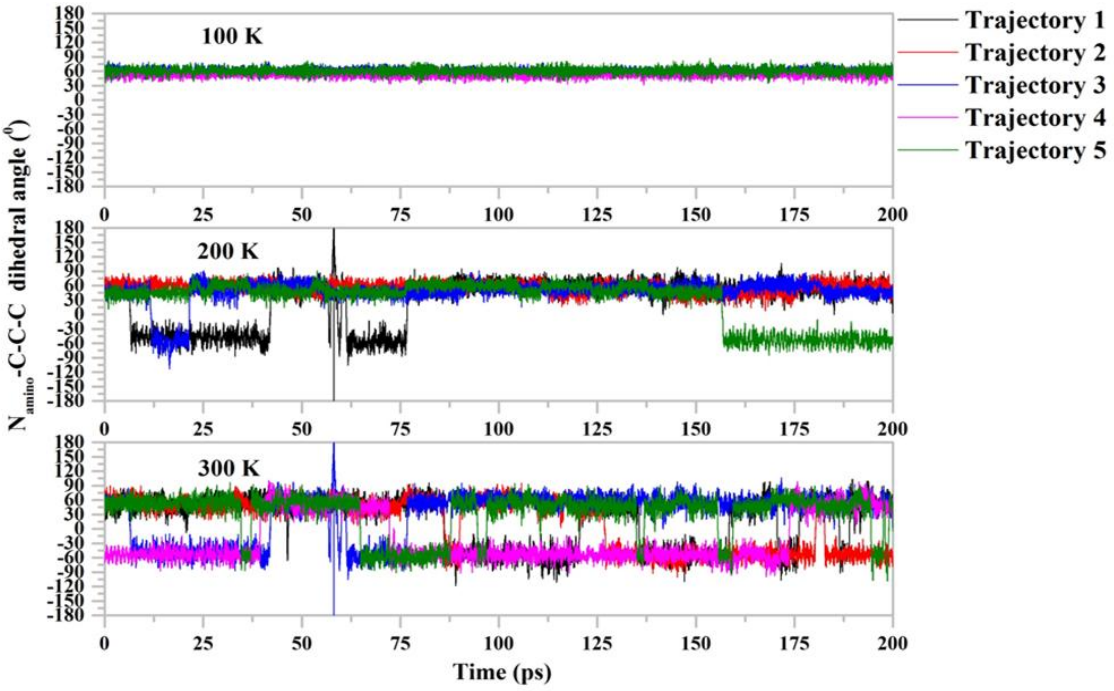


Figure A3. Time evolution of N_{amino}-C-C-C dihedral angle of ethylamine side chain at each temperature (100, 200 and 300 K) from 5 trajectories.

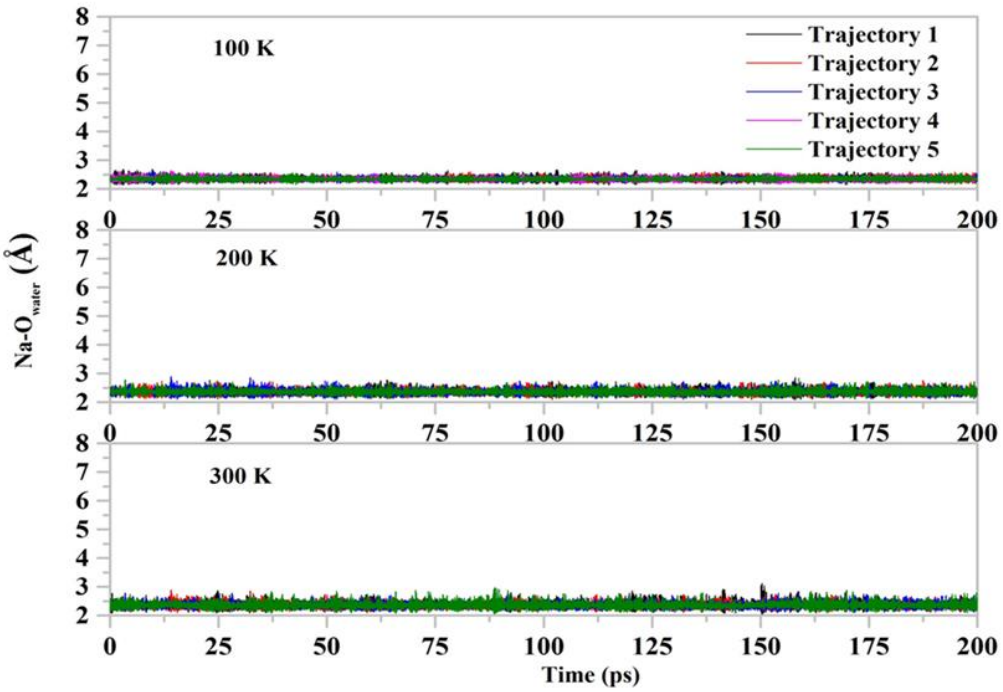


Figure A4. Time evolution of Na-O_{water}distance at each temperature (100, 200 and 300 K) from 5 trajectories.

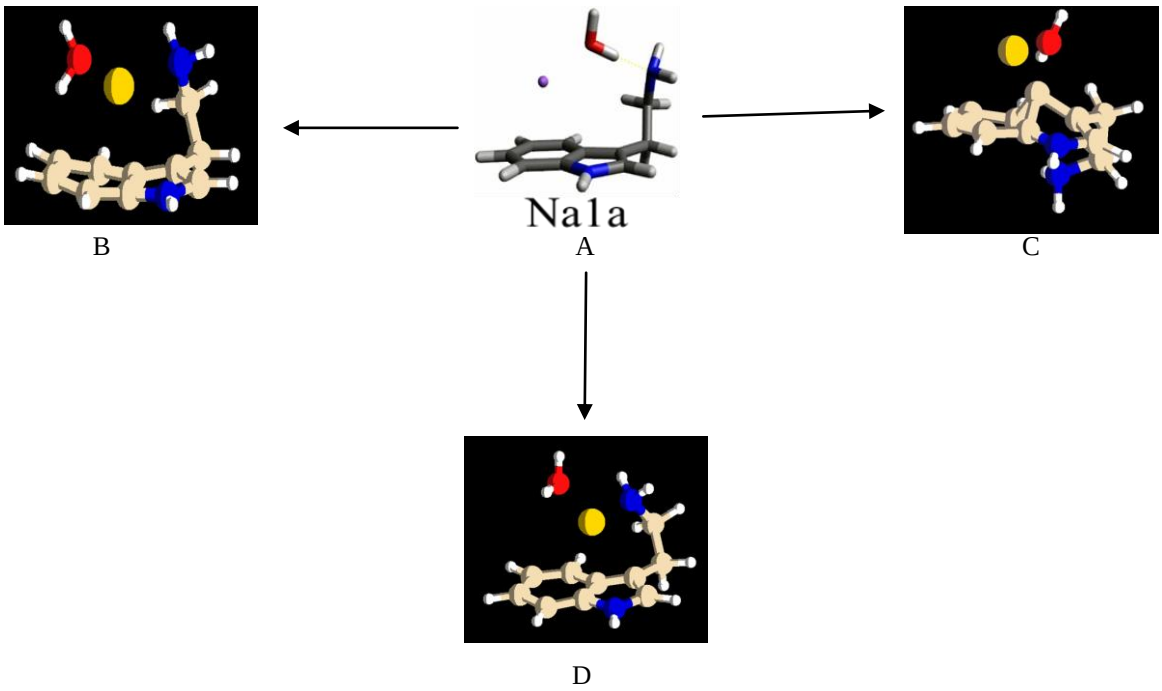


Figure A5. "A" is the MP2 optimized structure. **"B"** is obtained from the charge sets "a" and "c" in Table 6. **"C"** is obtained from the charge set "b" in Table 6. **"D"** is obtained from the charge set "d" in Table 6.

Chapter 5

IR spectra of a hydrated divalent sulfate ion:

$SO_4^{2-}(H_2O)_n$

1. Introduction

Besides its wide importance in inorganic chemistry of molecules and materials, especially in the form of metal salts, the sulfate dianion SO_4^{2-} also has great importance through its hydrates. The latter are found to be a key component in the environment as well as in several biological contexts. They are present in the atmosphere as in aerosol particles[1,2], resulting from the fossil fuel and biomass combustion. This leads to increase the acidity of the atmosphere and produce the acid rain byproduct. These aerosols also have both direct and indirect effects on the global climate such as atmospheric cooling, atmospheric warming, and cloud condensation nuclei[3,4]. For example, the hydrated sulfate dianions play a key role in the homogeneous nucleation of ice particles by sulfate aerosols in the upper troposphere[5]. In biological context, SO_4^{2-} is an important cell nutriment and the fourth most abundant anion in mammalian plasma. It is essential for many metabolic and cellular processes in vivo[6]. Ammonium sulfate is commonly used to purify proteins through selective precipitation. This technique uses this and other salts based on the work by F. Hofmeister more than a century ago. The ranked ability of salts to precipitate proteins, known as the "Hofmeister series", depends upon both ion hydration and upon the way ions interact with proteins. It is still a matter of active research and hot debate[7]. A first step towards understanding the chemistry of sulfate in solution is the knowledge of the structure and dynamics of bare sulfate in water. To this end, sulfate−water clusters $SO_4^{2-}(H_2O)_n$ in the gas phase are appropriate prototype systems. In the recent past, their properties have been characterized using both mass spectrometric and computational tools.

Electrospray ionization (ESI) provides easy access to gaseous sulfate hydrates with varying numbers of water molecules. Cluster size can be controlled by water evaporation, generated by collision induced dissociation (CID) or IR photo-dissociation (IRPD, IRMPD when multiple

photons are involved, or BIRD when photons are emitted by a closed ion chamber or "blackbody"). Numerous experimental studies have been carried out to characterize the energetic of water binding in sulfate hydrates with up to ca. 12 water molecules[8-10]. Numerous computational studies have been performed as well. The most comprehensive and accurate are those of Head-Gordon et al.[11-13] in which very extensive explorations of the potential energy surfaces, using the AMOEBA force field, were carried out for clusters with 3-7 water molecules[11]. Then, extensive assessment of performances of DFT and WFT methods was done[13] with estimates of CCSD(T)/CBS as the reference level[12]. Structural features were studied experimentally by IR spectra and computationally. The first spectra were recorded by Asmis et al. at 17 K for $SO_4^{2-}(H_2O)_{3-24}$ in the 500-1800 cm^{-1} region[14,15] and interpreted in terms of mostly symmetrical structures, such as 6C (see Figure 1 below). Subsequent spectra recorded at 130 K in the 2700-3800 cm^{-1} range[16] led to disagreement with the previous assignment for $SO_4^{2-}(H_2O)_6$, suggesting involvement of water-water hydrogen bonds as in 6E. This new assignment is in agreement with subsequent computational results[11-13]. Competition between sulfate-water and water-water hydrogen bonds is probably a leading factor in shaping the structures of clusters of growing size; structures with the sulfate near the surface have been invoked[17]. Larger clusters, containing 36-80 and then up to ca. 250 water molecules, were then investigated by IR spectroscopy. They unveiled new features attributed to the existence of free, dangling O-H bonds on the cluster surface. These new results form the core motivation for the present work. O'Brien et al.[18] reported spectra at 130 K for selected sizes in the n=36-80 range, between 3000 and 3800 cm^{-1} (see Figure 10 below). The spectra mainly consist in a very broad feature extending from ca. 3200-3700 cm^{-1}, with a maximum gradually red-shifted with increasing size, from ~3480 cm^{-1} for n=36 to ~3400 cm^{-1} for n=80, the latter value nearly matching that for bulk water. This can

then be safely attributed to the stretching modes of water molecules in a bulk-like environment, i.e. surrounded by other water molecules with which they form two hydrogen bonds as an acceptor and two as a donor. Less predictably, a second, much weaker and much narrower band appears at 3710 cm^{-1} around n=45, which can be attributed to free O-H bonds pointing outwards on the cluster surface. This was subsequently found to exist for clusters of a number of other ions with charges ranging from -1 to +3, at sizes 35-37[19] and 245-255[20]. Simple electrostatic models were found to account reasonably well for the observed spectra. The number of free O-H bonds appears to be larger for cations than for anions of a given size, and they appear at smaller sizes for cations. However the sulfate results were published first[18] and this is the reason why we started working on large hydrated sulfate clusters.

It is clear that quantum chemical approaches are inappropriate for modeling such spectra. First, the molecules are large enough that calculations of accurate IR spectra with conventional DFT or WFT methods are currently out of reach. Second, the clusters are highly fluxional even at temperatures as low as 130 K. Thus population of multiple structures is expected, requiring an extensive exploration of the potential energy surface and weighted averaging. Third, and maybe most importantly regarding IR spectra, fluxionality should be included in the vibrational calculation itself, not only in the structure exploration. Finally, as experiments carried out at several temperatures are beginning to appear (see chapters 3 and 4), molecular dynamics appears to be the method of choice. For these reasons, relatively long simulations are required for both structure exploration and computation of IR spectra. This cannot be currently carried out at the quantum mechanical level where simulations are limited to ca. 50 ps.[17] Vibrational SCF (VSCF) calculations are another alternative to the computation of accurate IR spectra, in particular to include anharmonicities. Such calculations were carried out by Gerber et al, with and without

anharmonic couplings between modes, for small clusters with n=1-3[17]. In the spectral range available at the time (540-1850 cm^{-1}), it was found that anharmonic corrections were only a small fraction of the error due to the potential function. Prediction of much larger corrections in the O-H stretching region was very interesting, however the computations are heavy enough that the largest case investigated with VSCF was n=3, while experiments in the 3000-3800 cm^{-1} region start with n=6.

Computing spectra using DACF and a classical force field thus appears to be the most appropriate procedure for large clusters. Yet most classical force fields fail to describe the electrostatics and polarization terms correctly (especially for charged, gaseous species), leading to useless structures and energetics. In such an instance, using the DACF/AMOEBA methodology described in previous chapters was very appealing. In the following, we first calibrate the performance of AMOEBA against *ab initio* reference values. We then describe the tools which were used to explore the potential energy surfaces for several cluster sizes of $SO_4^{2-}(H_2O)_n$: n = 25, 55, 80 and 100. The results of these explorations are described in the next section, to generate appropriate conditions to compute IR spectra with the DACF method, as described in the last section.

2. Quantum chemical calibration of AMOEBA for hydrated sulfate ions

2.1. Reference quantum chemistry calculations

In a recent paper Head-Gordon et al.[11] reported various low energy structures for $SO_4^{2-}(H_2O)_{n=3-7}$. We selected a few low energy structures for n = 3, 5 and 6 from their results. The structures are characterized as (n,s,w) by the total number of water molecules (n), the number of sulfate-water hydrogen bonds in the first solvation shell (s) and the number of water-water

hydrogen bonds (w). As an example, the nomenclature of structure 3C in Figure 1 is (3.4.2) which implies that this cluster contains 3 water molecules, 4 sulfate-water and 2 water-water hydrogen bonds.

Geometry optimizations and frequency calculations were performed for this selected set at the MP2/6-311++G(d,p) level using the Gaussian09 software package. Final energy calculations were performed at the MP2/aug-cc-pVTZ//MP2/6-311++G(d,p) level. Similar calculations were performed at B3LYP and M06 levels and we found MP2 to be the best level in agreement with the previous calibration study by Head-Gordon et al.[12,13]. All of these MP2 calculations were carried out to validate the force field parameters.

A selection of structures and energetics are shown in Figure 1 and Table 1. The energetics in Table 1 show, in agreement with Head-Gordon et al.[11], that the electronic energy differences between the lowest energy conformers are as low as a few $kJ.mol^{-1}$. MP2 and DFT results (not shown) are generally in good agreement, however slightly larger differences are obtained for some high energy conformers. Basis set extension from 6-311G++(d,p) to aug-cc-pVTZ changes energy differences only slightly, by less than 2 $kJ.mol^{-1}$ in most cases. Very similar behavior was obtained for relative enthalpies but relative free energies are sometimes rather different from relative electronic energies and enthalpies, due to entropy effects. This effect is analogous although slightly less strong at the B3LYP and M06 levels; there was no report of free energies in the work by Head-Gordon and coworkers[11] to which we could compare. With the small energy differences, these entropy effects change the isomer ordering for all three sizes considered. In any case, it is the relative electronic energies which will be used to validate the force field parameters.

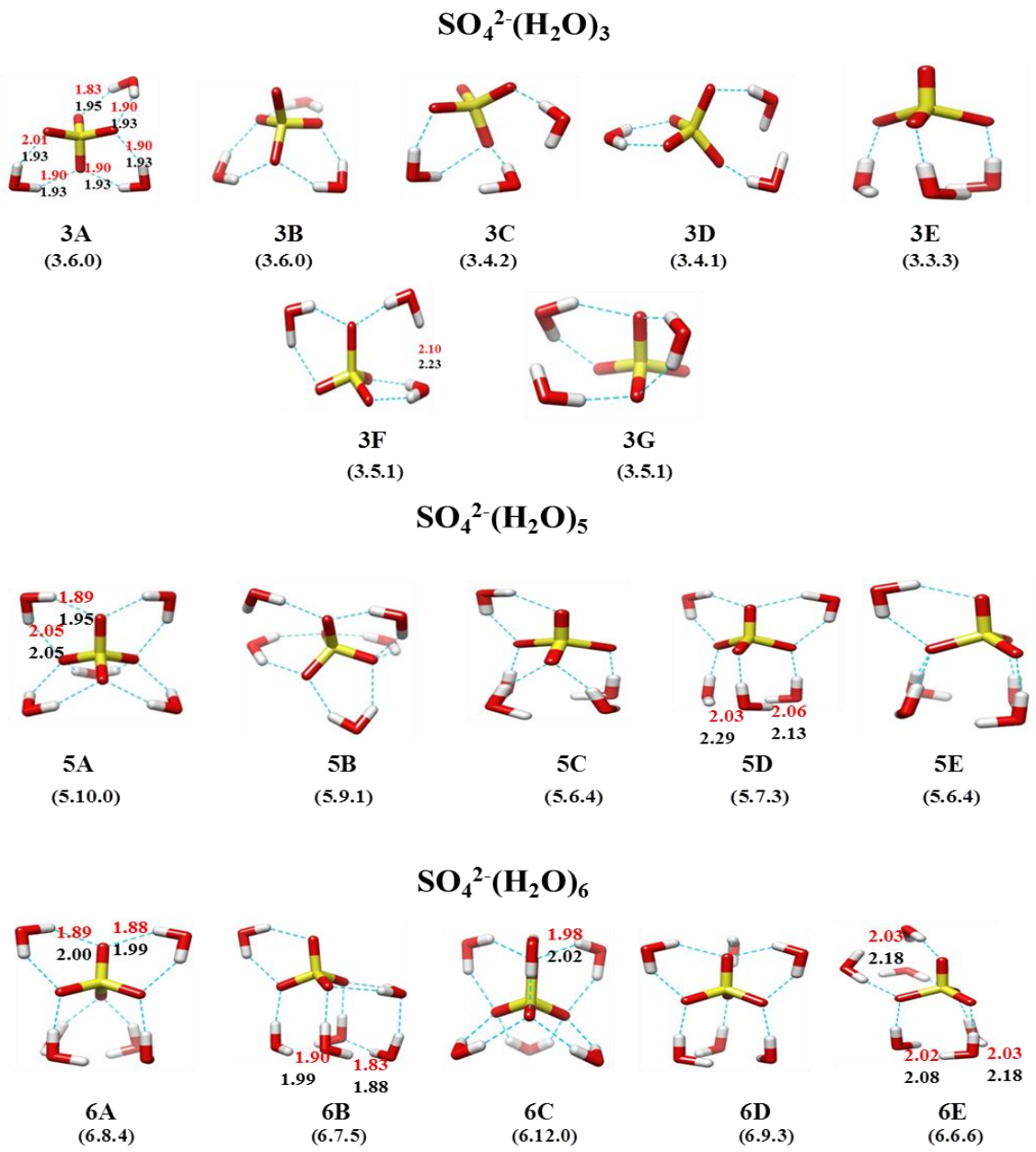


Figure 1 Optimized distances (Å) at the AMOEBA (in red) and MP2 (in black) levels, for selected $SO_4^{2-}(H_2O)_n$ clusters, n=3,5,6.

Complex (n, s, w)[a]	MP2/b1[b] (ΔE)	MP2/b1 (ΔH_{298K})	MP2/b1 (ΔG_{298K})	MP2/b2[c]//MP2/b1 (ΔE)	MP2/b2//MP2/b1 (ΔH_{298K})	MP2/b2//MP2/b1 (ΔG_{298K})
$SO_4^{2-}(H_2O)_3$						
3A (3.6.0)	0.00	0.00	0.00	0.00	0.00	0.00
3B (3.6.0)	-0.65	-0.22	0.22	1.18	1.61	2.05
3C (3.4.2)	11.93	12.51	17.62	13.68	14.26	19.38
3D (3.4.1)	11.19	10.99	11.4	10.52	10.33	10.74
3E (3.3.3)	4.93	6.81	16.86	4.02	5.89	15.95
3F(3.5.1)	2.12	2.94	7.33	2.59	3.42	7.81
3G(3.5.1)	8.42	9.11	13.96	9.33	10.02	14.87
$SO_4^{2-}(H_2O)_5$						
5A (5.10.0)	0.00	0.00	0.00	0.00	0.00	0.00
5B (5.9.1)	-3.57	-1.66	5.13	-3.26	-1.35	5.45
5C (5.6.4)	-3.74	2.23	25.18	-5.16	0.81	23.76
5D (5.7.3)	-9.63	-4.40	12.89	-8.85	-3.63	13.67
5E (5.6.4)	-1.76	3.38	25.64	-3.67	1.47	23.74
$SO_4^{2-}(H_2O)_6$						
6A (6.8.4)	0.00	0.00	0.00	0.00	0.00	0.00
6B (6.7.5)	2.25	2.54	6.13	-2.35	-2.06	1.52
6C (6.12.0)	13.84	5.58	-23.96	14.13	5.87	-23.67
6D (6.9.3)	-1.56	-3.53	-11.08	0.53	-1.44	-8.99
6E (6.6.6)	1.41	1.15	2.19	-3.02	-3.28	-2.24

Table 1 Relative energies (ΔE), thermal enthalpies (ΔH) and thermal free energies (ΔG) of $SO_4^{2-}(H_2O)_n$ (n = 3, 5, 6). All energies in $kJ.mol^{-1}$. The lowest energy structure in each box is indicated in orange.

a: nomenclature of structures as described in the text
b: b1 stands for 6-311++G(d,p)
c: b2 stands for aug-cc-pVTZ
d: ΔH_{298K} stands for the relative thermal enthalpies with 298 K thermal corrections at the MP2/b1 level
e: ΔG_{298K} stands for the relative thermal free energies with 298 K thermal corrections at the MP2/b1 level

2.2. AMOEBA parameters and performance

Head-Gordon et al.[11] have published AMOEBA parameters for the sulfate ion, extracted from MP2 calculations. We have used these parameters in all of our calculations. For the water molecule, we have used alternative AMOEBA parameters as described in the previous chapters.

To validate the AMOEBA sulfate-water parameters we have obtained geometries and energies using AMOEBA, to be compared to MP2 energies as gathered in Table 2. The AMOEBA energies correspond to optimized geometries with the AMOEBA force field starting from MP2 structures. These results suggest that AMOEBA is able to reproduce almost all low energy structures: only 3G could not be found as it switches to 3E. For almost all low energy structures, the AMOEBA and MP2 energy differences are found to be within 5 $kJ.mol^{-1}$; significant differences, up to 7 $kJ.mol^{-1}$, were obtained only in very few cases. Larger differences have been signaled by Head-Gordon et al.[11] in some cases. In agreement with them, we find that structures with more sulfate-water and fewer water-water hydrogen bonds are disfavored at the AMOEBA level. Some water-water and sulfate-water distances at the AMOEBA and MP2 levels are given in Figure 1. With differences smaller than 0.1 Å for distances of ca. 2 Å, they suggest that AMOEBA is able to reproduce MP2 geometries satisfactorily. Thus we have used these sulfate parameters unchanged for our calculations on larger hydrated sulfate clusters.

Complex (n, s, w)[a]	MP2/b2[b]//MP2/b1[c] (ΔE)	AMOEBA (ΔE)
$SO_4^{2-}(H_2O)_3$		
3A (3.6.0)	0.00	0.0
3B (3.6.0)	1.2	3.7
3C (3.4.2)	13.7	9.0
3D (3.4.1)	10.5	7.7
3E (3.3.3)	4.0	-0.1
3F(3.5.1)	2.6	6.9
3G(3.5.1)	9.3	-0.1
RMSD	4.78	
$SO_4^{2-}(H_2O)_5$		
5A (5.10.0)	0.0	0.0
5B (5.9.1)	-3.3	2.9
5C (5.6.4)	-5.2	-12.3
5D (5.7.3)	-8.9	-10.8
5E (5.6.4)	-3.7	-11.2
RMSD	5.45	
$SO_4^{2-}(H_2O)_6$		
6A (6.8.4)	0.0	0.0
6B (6.7.5)	-2.4	-2.2
6C (6.12.0)	14.1	19.6
6D (6.9.3)	0.5	5.5
6E (6.6.6)	-3.0	-7.4
RMSD	3.86	

Table 2 Relative energies (ΔE) at MP2 and AMOEBA levels, of $SO_4^{2-}(H_2O)_n$ (n = 3, 5, 6). All energies in kJ.mol^{-1}.The lowest energy structure in each box is indicated in orange

a: nomenclature of structures as described in the text

b: b2 stands for aug-cc-pVTZ

c: b1 stands for 6-311++G(d,p)

Global RMSD error (kJ.mol^{-1})

3. Computational tools for MD explorations

3.1. Starting points for MD explorations

We chose various starting points of sulfate ion hydrated with 25, 55, 80 and 100 water molecules for MD exploration, as a preliminary to IR spectra simulations. For n = 25, we have constructed two starting points, 25A and 25B where structure 25A contains a free O-H bond while structure 25B has none. They were constructed by adding manually 25 water molecules around a sulfate ion, using the molecule builder and visualizer Avogadro[21,22]. For n = 55, a starting structure was built starting from 25A to which 30 more water molecules were added using Avogadro. For n = 80, two starting points 80C and 80D were constructed. Structure 80C was built from a spherical water cluster of 365 water molecules which is available in Tinker. We first removed 280 water molecules from the surface, resulting in a 85 water cluster. SO_4^{2-} ion was then inserted at the center of mass of this cluster by further removing 5 water molecules using the "xyz-edit" program in Tinker. On the other hand, to construct the starting structure 80D, we have manually added 80 water molecules around the sulfate ion using Avogadro. After optimization with AMOEBA, structure 80C is close to being spherical while 80D is strongly non spherical, providing some initial diversity. For n = 100, the starting structure is built from a spherical water cluster of 100 molecules which is available online at http://www1.lsbu.ac.uk/water/strand.html#str. This structure is that of a water nanodroplet trapped in a polyoxomolybdate[23]. It is nearly spherical and it contains sufficient space at the center of mass to accommodate a sulfate ion, as it forms three concentric shells and the smallest is a dodecahedron. All structures were then optimized with AMOEBA. At this stage, all resulting structures contain free O-H bonds except structure 25B.

For large clusters the various water-water hydrogen bonding patterns are nicely described by Williams and coworkers[24,25]. As an example, different hydrogen bonding patterns are shown in Figure 2. The term acceptor (A) labels a water molecule that accepts a hydrogen bond from another water molecule and the term donor (D) refers to a water molecule that donates a hydrogen bond to another water molecule. The "AAD" and "AD" hydrogen bonding patterns shown in Figure 2 lead to free OH bonds, also called dangling OH bonds. Most internal water molecules are "AADD", double hydrogen bond donors and acceptors as in bulk water.

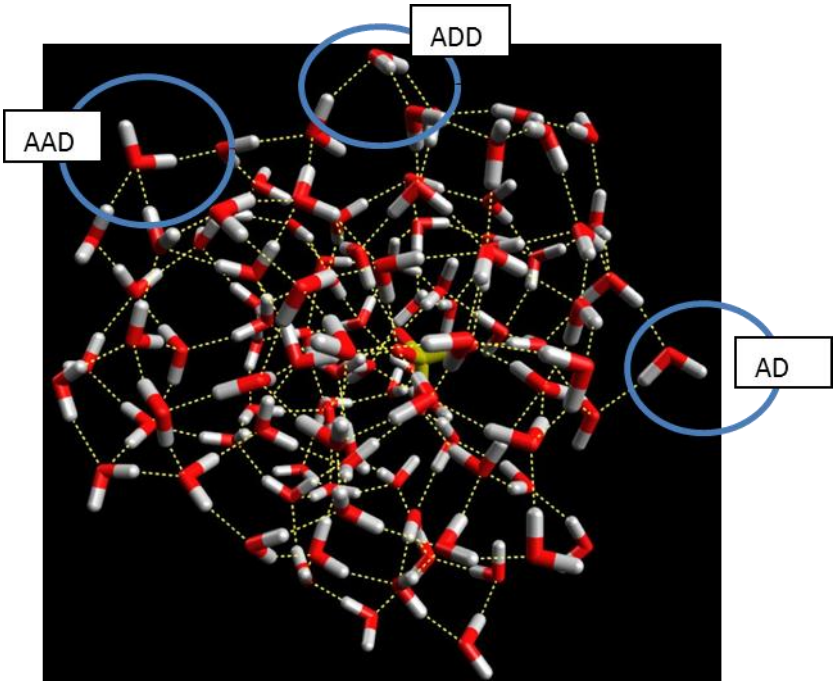


Figure 2 Various water-water hydrogen bonding patterns at the surface of a hydrated sulfate cluster for 80 water molecules.

3.2. Simulation conditions

The starting structures for n = 25, 55, 80 and 100 described above are not equilibrated structures. Therefore, MD exploration is necessary to obtain equilibrated structures and more importantly to generate extensive conformational sampling, as necessary for such highly flexible clusters. To this purpose, from each of these starting points we run up to 3 ns of trajectories at 200 K with a time step of 1 fs; structures were saved every 0.1 ps. Extension to 5 ns was also done in a few cases to check convergence. All simulations except for n=25 were carried out under spherical boundary conditions, as described below.

After MD exploration, the IR spectra were obtained from MD simulations using the AMOEBA force field. In this approach, IR spectra were computed by taking the Fourier transform of the autocorrelation function over time of the total dipole moment as described in chapter 2. The main difference with exploratory trajectories is the reduced time step of 0.1 fs to capture the high frequency motions reliably.

3.3. Spherical boundary conditions

In order to carry out simulations on large clusters at finite temperatures, evaporation must be avoided. For the sizes mentioned above, only n=25 was found to retain all water molecules for long durations at 200 K. For larger sizes, evaporation occurred. This can be corrected by using a spherical van der Waals wall around the cluster, added to the force field energy expression, as implemented in the Tinker suite of programs. The wall radius must be chosen carefully, as an unduly small value may introduce significant bias especially at the cluster surface, which is our main focus in this work. Conversely, too large a radius value would be inefficient against evaporation. All clusters except n = 25 were restrained by a van der Waals wall with a fixed

buffer distance of 2.5 Å outside the specified wall radius. The potential has a strong repulsive part of r^{-12} form and a weak attractive part of r^{-6} form, as in a typical Lennard-Jones potential, however with very different coefficients:

$$e = \frac{a}{r^{12}} - \frac{b}{r^6} \qquad a = 2048.0 \quad b = 64.0 \quad (a, b\ in\ kcal.mol^{-1})$$

$$r = r_{wall} + r_{buffer} - r_i r_{buffer} = 2.5$$

To determine an appropriate wall radius r_{wall} for each cluster, we first ran 1 ns simulations at 100 K for various wall radii with time step 1 fs and structure saving every 0.1 ps. For example, for cluster 80D, simulations were carried out with several wall radii ranging from 9 to 20 Å. A detailed trajectory analysis (using scripts written with the help of Florian Thaunay at LCM) was then carried out and the distributions of water molecules were obtained as a function of average distances between the center of mass of the cluster and O atoms. In Figure 3, histograms are presented for radii ranging from 9 to15 Å. Note that the distributions in Figure 3 are obtained as a function of average distances of O atoms from the center of mass rather than for individual snapshots. With 9 Å, several water molecules are found to be outside of the wall radius, showing this value to be too small. This is corrected with a radius of 10 Å, however several molecules are close to the radius value. Since an H atom pointing outwards lies ca. 1 Å further away than the O atom from the center of mass, avoiding significant population in the buffer region led us to choose a secure wall radius of 12 Å for this cluster size. Using larger values brings no further advantage, as shown in Figure 3 for a radius of 15 Å.

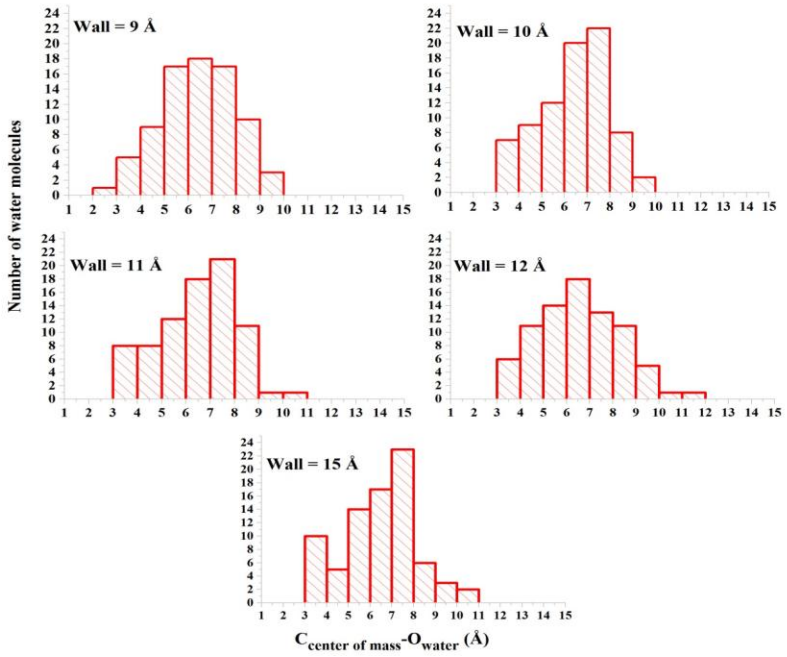


Figure 3 Distribution of the numbers of water molecules in simulation at 100 K starting from structure 80D of $SO_4^{2-}(H_2O)_{80}$. These numbers are drawn as a function of average distance between the center of mass of the cluster and the O atoms.

With a wall radius of 12 Å, further simulations were carried out for 5 ns at both 100 and 200 K. The distribution of water molecules during these simulations are displayed in Figure 4, showing that in neither case there are water molecules reaching the wall radius limit of 12 Å. A more "individual" look is available with the fluctuations of each water molecule during these simulations, as shown in Figures 5 and 6at 100 and 200 K, respectively. The green lines correspond to the average distances which were used to obtain the histograms in Figure 4 and the standard deviations are shown in black. As expected, fluctuations are larger at 200 than at 100 K. Individual movements for the water molecules with high fluctuations especially at 200 K were checked; they always remain in the cluster. Therefore, finally we choose 12 Å as wall radius for this cluster. By using the same sets of calculations and analysis we choose the same wall radius of 12 Å for cluster 80C and for n =100. For n = 55 we chose a smaller wall radius of 10 Å.

Finally, we have kept the attractive coefficient of 64 kcal.mol^{-1} in the wall form as implemented in Tinker, as there was no significant difference when 32 or 16 kcal.mol^{-1} was used.

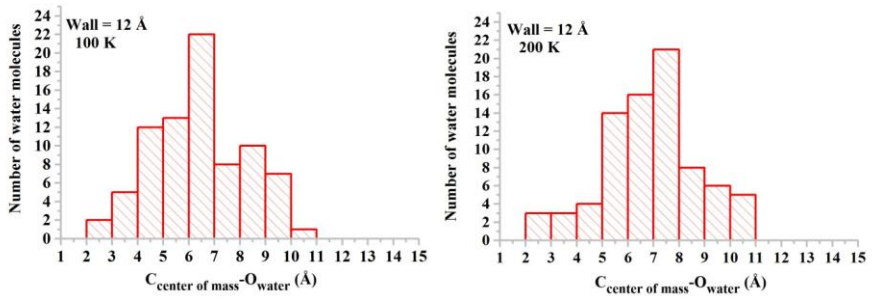


Figure 4 Distribution of the numbers of water molecules in 5 ns simulations starting from structure 80D of $SO_4^{2-}(H_2O)_{80}$. These numbers are drawn as a function of average distance between the center of mass and the O atoms.

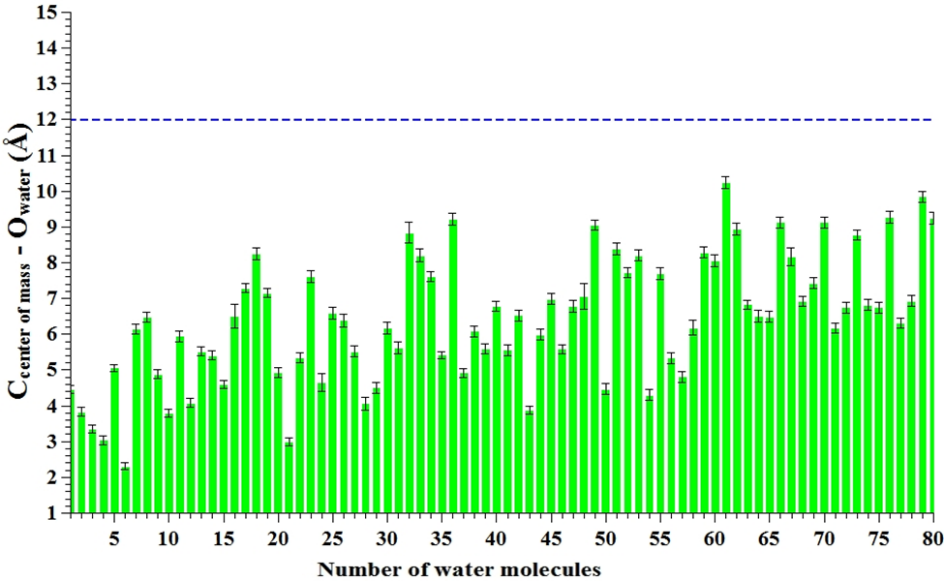


Figure 5 Average distances between the center of mass and the O atom of each water molecule (green line) with fluctuations (black, standard deviation) for 5 ns simulation at 100 K and wall radius of 12 Å.

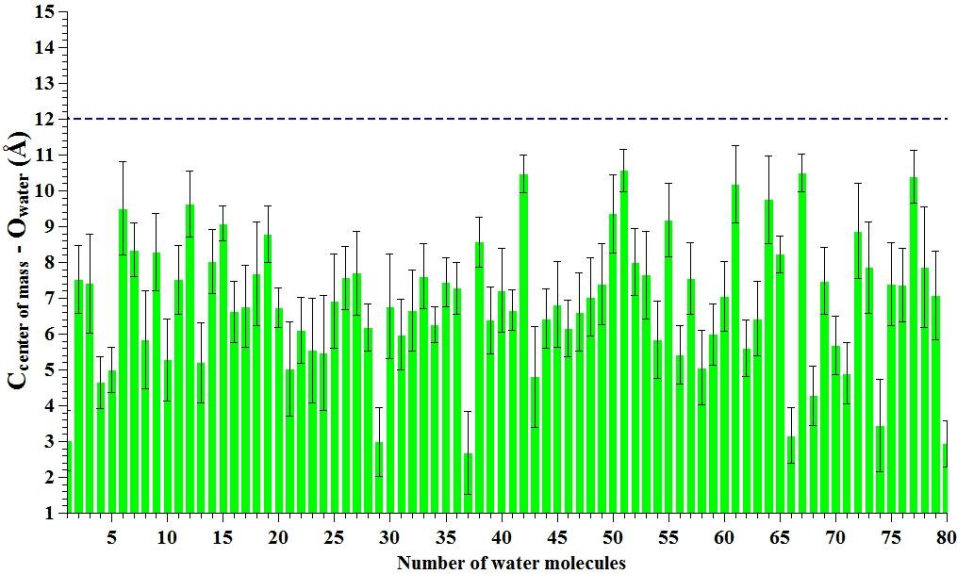


Figure 6 Average distances between the center of mass and the O atom of each water molecule (green line) with fluctuations (black, standard deviation) for 5 ns simulation at 200 K and wall radius of 12 Å.

3.4. Calculation of the number of free O-H bonds

Since the main goal of this work is to characterize cluster surfaces in terms of free O-H bonds, the number of such O-H bonds has been calculated during simulations. Hydrogen bonding interactions were considered absent if the non-bonded H---O distances were all larger than a threshold. These numbers were calculated for all 10^4 saved snapshots during 1 ns simulations at 200 K, starting from structure 25A of $SO_4^{2-}(H_2O)_{25}$. The histograms in Figure 7 display the numbers of occurrences as a function of the numbers of free O-H bonds, for various threshold values from 2.5 to 3.5 Å. The 3.0 Å value appears to be appropriate, as it is consistent with the non-bonded H---O distances manually collected for the outer part of the structure of $SO_4^{2-}(H_2O)_{100}$ after 1 ns MD, all lying between 3.1 and 3.5 Å. The same conclusion was obtained for structure 80C (n=80). Therefore, we choose the 3.0 Å threshold to compute the numbers of free O-H bonds. It is clear that this value is much larger than that for a typical hydrogen bond.

This is because the O-H---O angles were not taken into account to characterize hydrogen bonding, which would require intractable data processing.

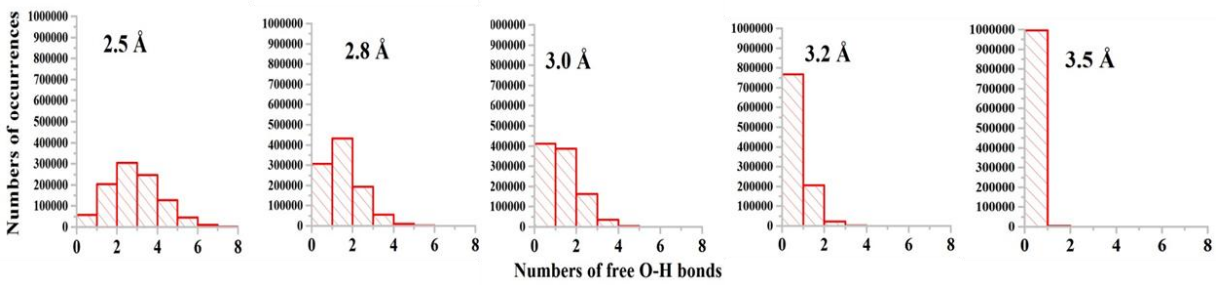


Figure 7 Numbers of occurrences as a function of numbers of free O-H bonds with different hydrogen bonding thresholds for structure A of $SO_4^{2-}(H_2O)_{25}$. MD simulations of 1 ns at 200 K.

4. MD exploration of the potential energy surfaces of hydrated sulfate clusters

With the previous tools at hand, it is now possible to run simulations with conditions appropriate to hydrated sulfate clusters. Prior to being able to calculate IR spectra, it is necessary to generate equilibrated structures, with the number of free O-H bonds in addition to the usual energy criterion. First, MD explorations of 1 ns were carried out for each of the starting structures described above. The magnitudes of potential energy fluctuation of 20-40 kcal.mol^{-1} in the last 0.1 ns, depending upon cluster size and starting structure, indicate incomplete equilibration. Therefore, another set of 1 ns simulations were carried out, starting from the final point of the first ns runs. For these subsequent simulations, the variations of the numbers of free O-H bonds are shown as a function of simulation time (e.g. Figure 8, right). We have also drawn histograms displaying the numbers of occurrences as a function of the number of free O-Hs during simulations (e. g. Figure 8, left).

Figure 8 displays the trajectories starting from structures 25A and 25B. The histograms of numbers of occurrences and time evolution graphs for the numbers of free O-H bonds are very similar for both structures, with 1 and 2 free O-H bonds as the most likely numbers, and little difference between 25A and 25B. For 55 water molecules, Figure A1 in appendix (right) suggests that the structure is almost equilibrated with some fluctuation at ca. 0.4 ns. In addition, the histogram in Figure A1 in appendix reflects that the populations of 4 and 5 free O-H bonds are large. For n = 80, the two structures 80C and 80D are also found to yield similar results, with populations of 7 and 8 free O-H bonds found to be the most likely (see Figure A2 in appendix).

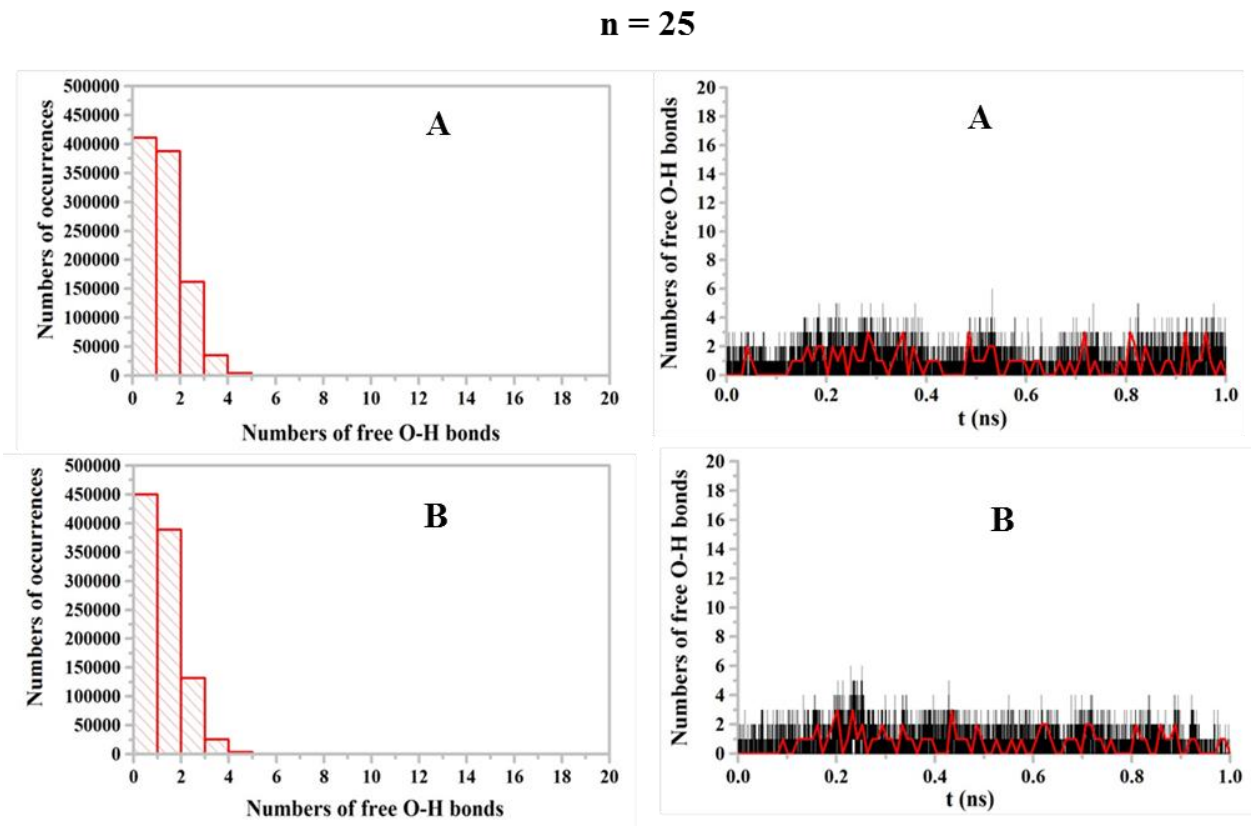


Figure 8 Free O-H bonds in structures 25A (top) and 25B (bottom) of $SO_4^{2-}(H_2O)_{25}$. In each case the left part displays the numbers of occurrences as a function of number of the free O-H bonds while the right part shows the variations of the numbers of free O-H bonds as a function of simulation time (1 ns) at 200 K. The red line corresponds to short-time averaged values.

For n = 100 (Figure 9), the time evolution graph of free O-H bonds suggests that the starting structure is equilibrated and the histogram clearly displays that populations of 9 and 10 free O-H bonds are the most probable. These results suggest that the present equilibration conditions are efficient enough to avoid memory effects of the starting structures. Thus the final structures from these MD simulations may be used as starting points for IR spectra simulations. We have also run yet another 1 ns of simulation, this time using both final positions and velocities of the previous runs. Trajectory analysis for these simulations is shown in Figure A3 in appendix. The histograms clearly reflect that the numbers of free O-H bonds remain almost the same with little differences between structures 25A and 25B in comparison with previous MD exploration.

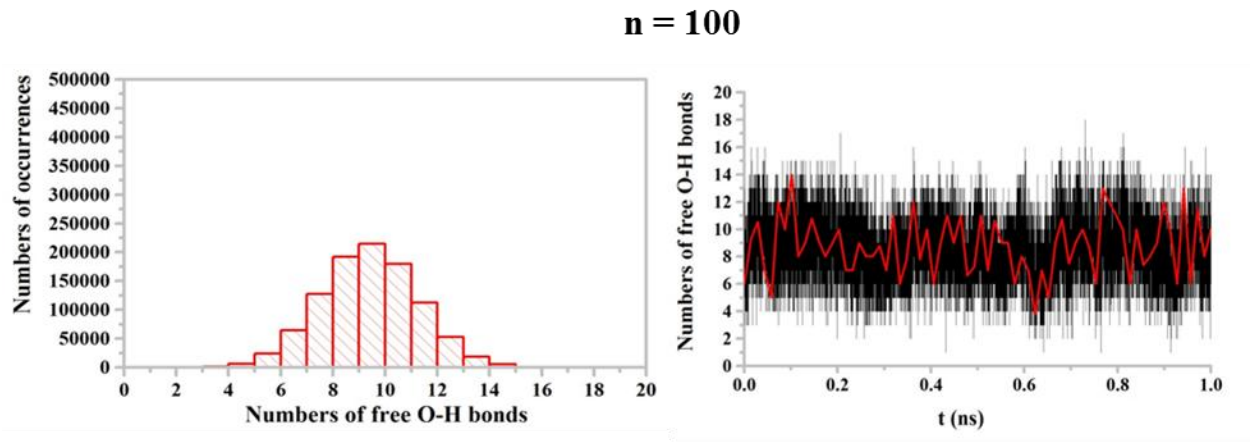


Figure 9 Free O-H bonds in $SO_4^{2-}(H_2O)_{100}$. The left part displays the numbers of occurrences as a function of number of the free O-H bonds while the right part shows the variations of the numbers of free O-H bonds as a function of simulation time (1 ns) at 200 K. The red line corresponds to short-time averaged values.

5. IR spectra of hydrated sulfate ions

5.1. Experimental results

As mentioned above, experimental IR spectra of hydrated sulfate cluster ions were published in recent years. These spectra were recorded at low temperature of 130 K, in the 2900-3800 cm^{-1} range, which includes all O-H stretching frequencies. Since these frequencies depend very strongly upon the environment (mostly hydrogen bonds and charged sites), it is expected that this wavenumber range can provide clues about cluster structures. While small sizes allow to study the competition between sulfate-water and water-water bonds[16], spectra for clusters including 36-80 water molecules revealed a new and relatively unexpected feature: one band near 3710 cm^{-1}, which can only be assigned to O-H bonds that are free of hydrogen bonding[18], and whose relative intensity augments with cluster size. These spectra are shown in Figure 10.

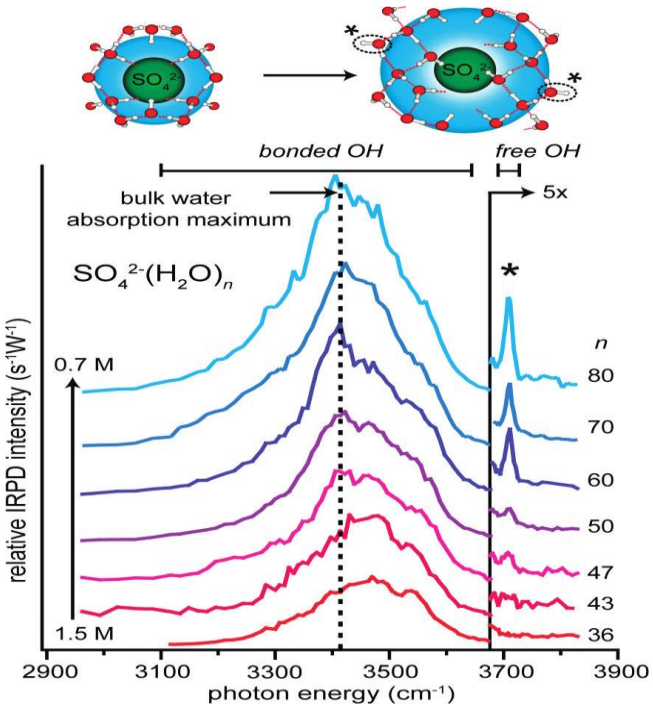


Figure 10 IRPD spectra of $SO_4^{2-}(H_2O)_n$, n=36-80,recorded at 130 K. Taken from ref.18.

It can be seen that the band near 3710 cm^{-1} is only observed in the spectra for n> 43, suggesting that for n=36 and 43, there is no free O-H bond. The appearance of this band indicates that intrinsic water-water interactions begin to dominate over the ion-induced structure for cluster sizes greater than 43. Indeed, a similar band exists in SFG spectra of pure water[26], indicating that some molecules at the air-water interface have free O-H bonds, pointing away from the liquid.

The main feature of these spectra is a broad and intense band observed within 3100-3700 cm^{-1} for all cluster sizes, which corresponds to hydrogen bonded O-H groups. The red end and the maximum of this band are both red shifted when the cluster size increases. Modeling the structure of such clusters is rather difficult, due to their large size and high fluxionality. With the computational tools previously described in this chapter, this is the challenge we aimed to meet in this work.

5.2. Computation of IR spectra

We have computed MD spectra of sulfate ion with 25, 55, 80 and 100 water molecules at several temperatures ranging from 100 to 300 K. All starting structures described above were used (i.e. two structures, 25A and 25B for n=25, and two structures, 80C and 80D, for n=80; one structure for each of n=55 and n=100). The simulations started with a non productive 50 ps period to reach the desired temperature, followed by 200 ps of productive propagation with a time step of 0.1 fs to sample the fastest O-H stretching vibration accurately, as discussed in chapter 2. The resulting MD spectra are provided in Figures 11-15in the 3300-3800 cm^{-1} range. The frequencies at band maxima are provided in Table 3 together with experimental frequencies.

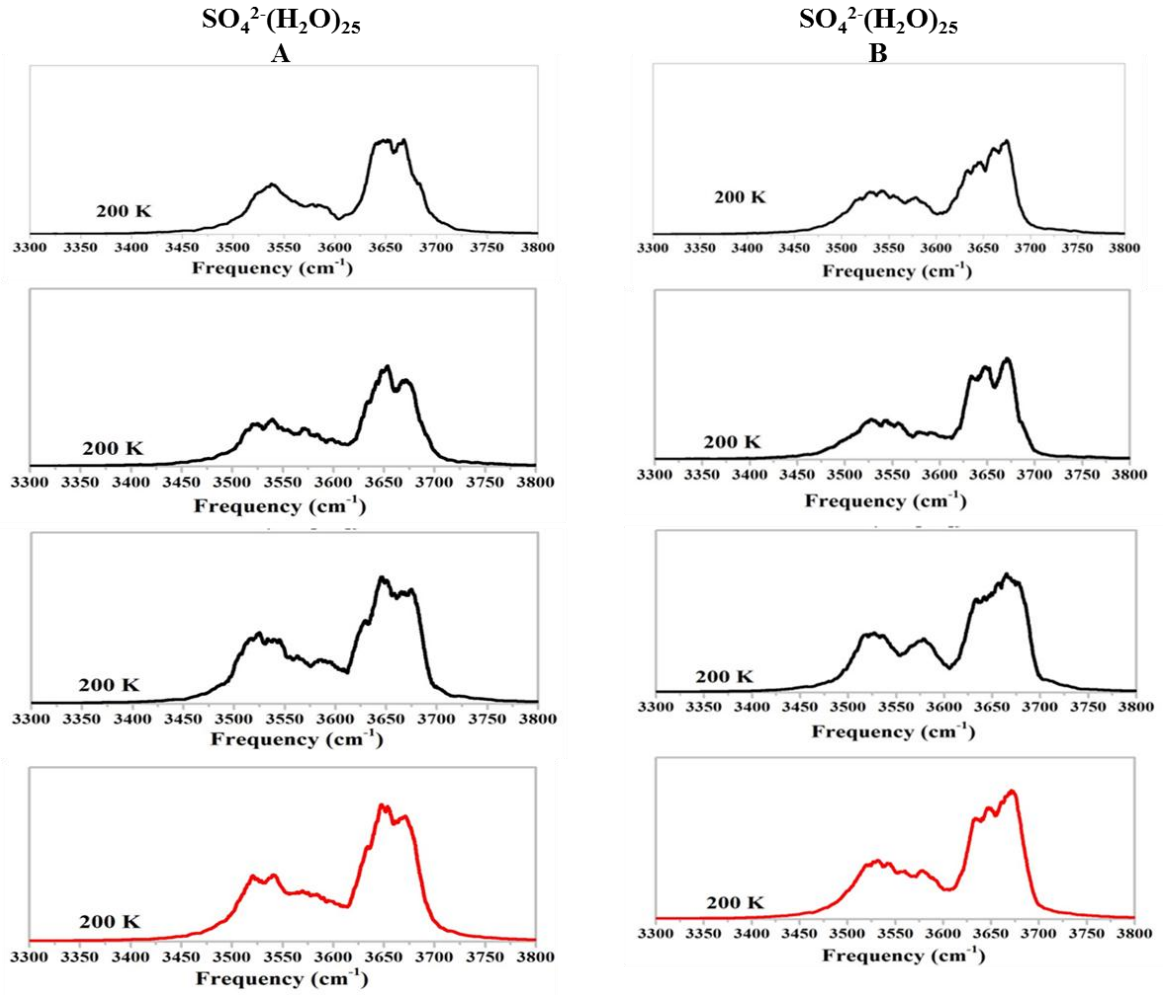


Figure 11 Computed MD spectra for structures 25A (left) and 25B (right) of $SO_4^{2-}(H_2O)_{25}$ at 200 K. In each column the first three spectra (black) are calculated following 1, 2 and 3 ns of equilibration. The bottom spectrum (red) is their average.

In Figure 11 for $SO_4^{2-}(H_2O)_{25}$, spectra computed with varying degrees of pre-equilibration are shown for both starting structures. It is clear that equilibration is a significant issue and that averaging may be a useful tool. Similar conclusions can be drawn from analogous calculations for other cluster sizes (Figures A4 and A5in the appendix). Thus all spectra shown below were computed after 2 ns of equilibration and as averages of three runs. The spectra shown in Figure 12, again for both starting structures of $SO_4^{2-}(H_2O)_{25}$, were computed at 100, 150 and 200 K. It appears that both series are very similar, indicating that equilibration has been efficient enough

to avoid memory effects. At low temperatures there is a very low intensity band at 3730 cm^{-1}, corresponding to free O-H bonds at 100 and 150 K for both starting structures 25A and 25B.This very low intensity comes along with the very low populations of free O-H bonds, as described above. In addition, broad and intense bands appear in the 3400-3700 cm^{-1} region, corresponding to hydrogen bonded O-H stretching which may arise from both the water-water and sulfate-water hydrogen bonds.

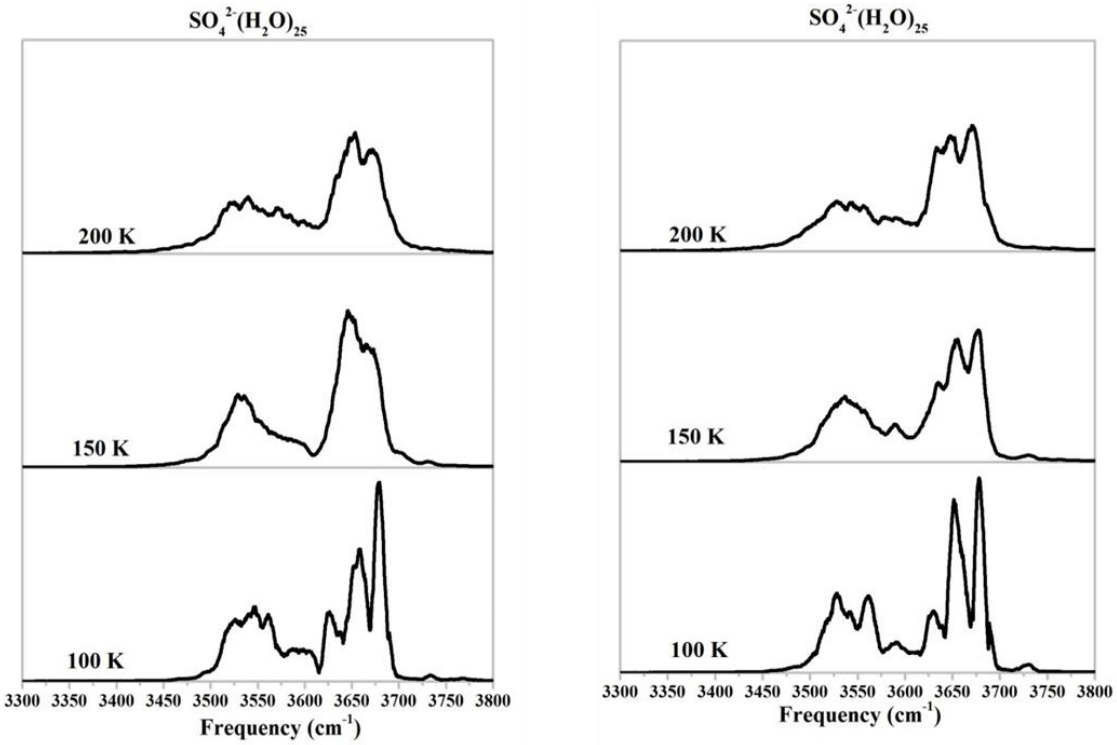


Figure 12 Computed MD spectra for structures 25A (left) and 25B (right) of the $SO_4^{2-}(H_2O)_{25}$ at 100, 150 and 200 K after 2 ns of MD exploration.

Low temperature spectra are far from smooth, indicative of hampered dynamics with little internal energy. It is interesting to note that smoother spectra are generally obtained at 200 K. This appears to be general for all cluster sizes, see Figures 13-15. Another general feature is that in the hydrogen bonded O-H stretching region (3400-3730 cm^{-1}), the major effect of temperature

is band broadening with small shift of band maxima. However computed frequencies are strongly blue shifted compared with experimental frequencies in this region (Table 3).

The computed MD spectra obtained for sulfate ion with 55 water molecules are provided in Figure 13 at four temperatures. Spectral features at 100, 150 and 200 K are similar to those of $SO_4^{2-}(H_2O)_{25}$ spectra, with some new features.

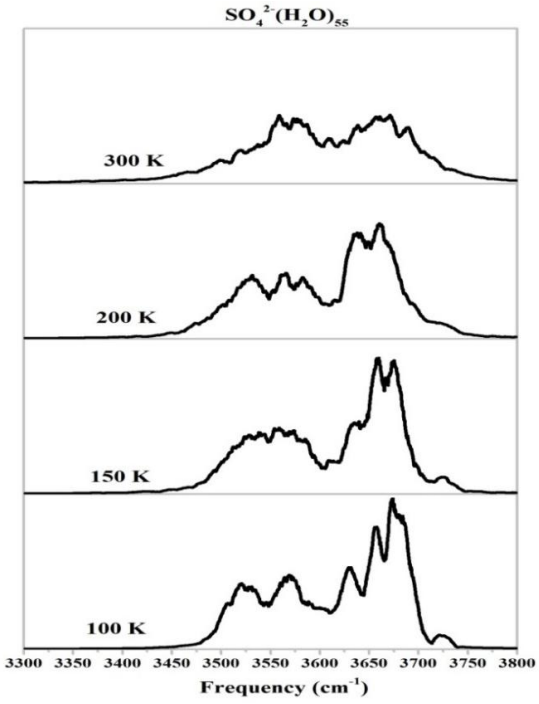


Figure 13 Computed MD spectra for $SO_4^{2-}(H_2O)_{55}$at 100, 150, 200 and 300 K after 2 ns of MD exploration.

There was no free O-H band visible at 200 K for $SO_4^{2-}(H_2O)_{25}$ while this band is obtained for $SO_4^{2-}(H_2O)_{55}$ however it is not well separated. The free O-H band intensities for n=55 at 100 and 150 K are higher than those for n=25 (see Figures 12 and 13). These higher intensities are due to the higher numbers of free O-H bonds for n=55 than for n=25 as described above. Furthermore,

at 300 K, the bands are broader compared to 200 K. Intensities and frequencies for n=55 are in slightly better agreement with experiments than those for n=25.

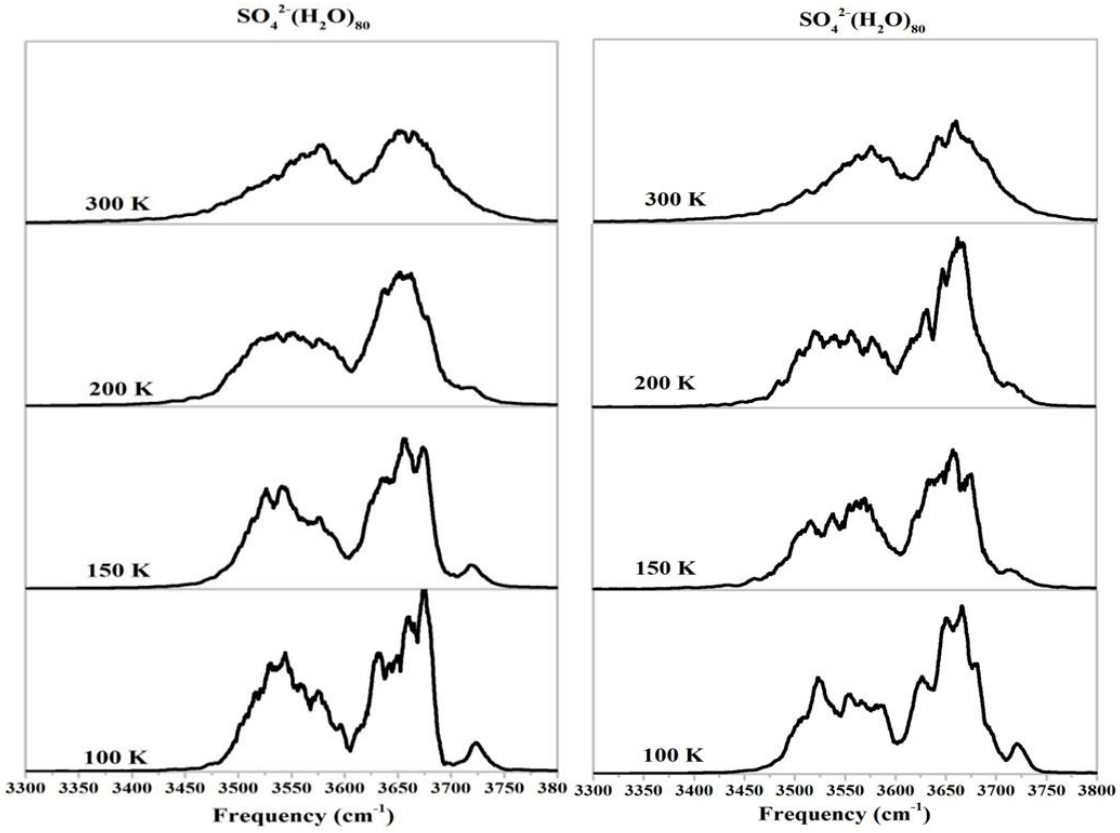


Figure 14 Computed MD spectra for the structures 80C (left) and 80D (right) of $SO_4^{2-}(H_2O)_{80}$ at 100, 150, 200, 300 K after 2 ns of MD exploration.

For n = 80, the spectra obtained for starting structures 80C and 80D are displayed in Figure 14. As for n=25, comparison of spectra arising from 80C and 80D reveals no significant differences. The free O-H band is slightly more intense than for smaller clusters, in agreement with the larger number of free O-H bonds. Its frequency at 100 K (3723 cm^{-1}) shifts to 3713 cm^{-1} at 150 and 200 K. The hydrogen bonded O-H stretching bands are in the 3400-3755 cm^{-1} range, again strongly blue shifted relative to experiment.

The computed MD spectra for 100 water molecules are shown in Figure 15. The same trends are observed as for the smaller clusters. At this size, a significant number of free O-H bonds generate a well identifiable peak at low temperatures, however absorption appears to become continuous in this spectral range above 200 K.

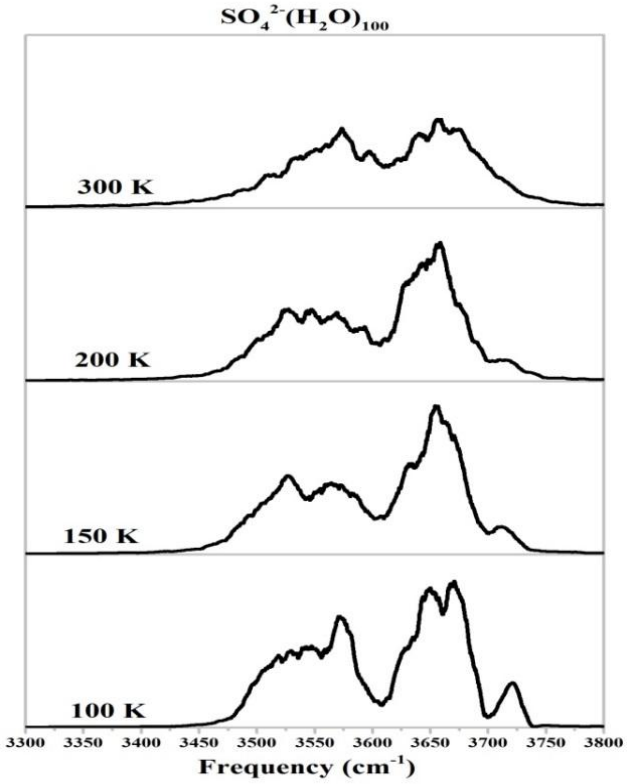


Figure 15 Computed MD spectra for $SO_4^{2-}(H_2O)_{100}$ at 100, 150, 200 and 300 K after 2 ns of MD exploration.

5.3. Comparison to experiments

Computed frequencies for the various sizes at several temperatures are gathered in Table 3. The results can be summarized as follows: at 100 K, free O-H stretches are computed to appear at 3730 cm^{-1} for the smaller clusters, with gradual red shifts down to 3723 cm^{-1} for the larger ones. At 150 and 200 K, changes are again very limited, with red shifts down to 3713 cm^{-1} for the larger sizes. These values compare favorably with the experimental frequency of 3710 cm^{-1} observed for sizes from n=47 to 80 at 130 K. The bonded O-H massif is computed to extend from ca. 3450 to 3715 cm^{-1}, with a slight red shift with increasing size, to ca. 3400-3700 cm^{-1} for n=100. This shift is consistent with the experimental spectra, which show red-end shifting from ~3200 cm^{-1} for n=36 to ~3100 cm^{-1} for n=80 and red shifting of the maximum from 3480 to 3420 cm^{-1}, a value which matches that of the maximum for bulk water. Clearly these red shifts are underestimated in the present computations; this point will be addressed below.

In addition, we observe a systematic two-component shape for this massif, with a minimum lying near 3600 cm^{-1}, while experimental spectra show a limited intensity decrease near 3510 cm^{-1} for n=36, which gradually disappears beyond n=60. However the general peak shape remains non symmetrical (see Figure 10). Spectral fitting of bands recorded for still larger clusters (n=245-255) indicates that this massif is much better fitted by two Gaussians than it is by only one[20]. The 3500-3600 region is attributed to "water molecules that are less strongly H-bonded than those at lower frequencies"[20]. It is clear that if one component was arising from O-H stretches related to water-water hydrogen bonds and the other to the water-sulfate, their relative intensities would change dramatically with cluster size since the sulfate first hydration shell remains nearly unchanged while water-water interactions become more numerous for larger sizes. This is clearly not the case, thus a more detailed analysis is required. However band

assignment of DACF spectra is far from straightforward, especially for modes which are strongly delocalized over a number of water molecules. Analysis of static spectra for some selected structures indicates that for n=25 and 55, the lower frequency portion is made of symmetric O-H stretches, comprising water molecules engaged in either water-water or water-sulfate types of hydrogen bonds. The higher frequency part is made of the corresponding antisymmetric modes. For larger clusters, both components spread over wider frequency ranges, leading to a significant overlap region for n=80 and extended mixing for n=100, so that assignment to symmetric or antisymmetric modes is no longer possible. This explanation is consistent with the gradual disappearance of the deep in the experimental spectrum.

$SO_4^{2-}(H_2O)_n$	Temperature	AMOEBA	
		Free O-H	Bonded O-H
n = 25	100 K	3730	3475-3700
A	150 K	3730	3450-3715
	200 K		3450-3730
n =25	100 K	3730	3475-3700
B	150 K	3730	3450-3715
	200 K		3450-3715
	100 K	3725	3450-3715
n = 55	150 K	3725	3450-3715
	200 K	3725	3450-3715
	300 K		3400-3755
	100 K	3723	3450-3700
n = 80	150 K	3715	3450-3700
C	200 K	3715	3450-3700
	300 K		3400-3755
	100 K	3723	3450-3705
n = 80	150 K	3713	3450-3705
D	200 K	3713	3450-3705
	300 K		3400-3755
	100 K	3723	3450-3700
n = 100	150 K	3713	3400-3700
	200 K	3713	3400-3700
	300 K		3400-3755
Experimental	130 K	3710	3100-3700

Table 3 Computed free O-H stretching frequencies and hydrogen bonded O-H stretching massif maximum frequencies. Experimental frequencies are taken from ref. 18.

5.4. Improvement of hydrogen bonding

It is clear that the frequency range of the bonded O-H stretching massif does not extend enough to the red, compared to experiment. This is likely to be due to insufficient treatment of hydrogen bonds in the AMOEBA force field. The most red-shifted bands can be tentatively attributed to those O-H bonds which are directly interacting with the sulfate, as ionic hydrogen bonds are expected to be very strong. However this cannot explain the shift of the maximum of

the bonded OH massif. The latter must arise from the slow convergence of the water-water interaction network towards bulk characteristics, since it tends to be less influenced by the sulfate with growing cluster size. Thus the AADD type molecules, typical of bulk water, are relatively few in the smaller cluster studied herein, but become the majority in the larger ones. It is the description of these, plus that of the sulfate first shell, which must be improved. One clear weakness of AMOEBA is the absence of charge transfer, an interaction which is known to participate to hydrogen bonding[27]. An alternative could be the SIBFA force field[28] in which charge transfer is explicitly accounted for, however it lacks intramolecular flexibility, which is required to compute IR spectra.

Thus we have resorted to a simple modification of AMOEBA, in the anharmonic part of the stretching potential energy of the O-H bond. This term is expected to be accurate for isolated water since its experimental IR spectrum is very well matched, provided that angular corrections to the standard AMOEBA parameters are included. However the energy variation with bond distance is too stiff to account for environment effects. In its standard MM3 version, the bond stretching term reads[29]:

$$\boldsymbol{U_{bond}} = 71.94 * K_b\,(b - b_0)^2[1 - 2.55(b - b_0) + \frac{7}{12} * 2.55^2(b - b_0)^2]$$

where a single constant of standard value -2.55 is used to define both the cubic and quartic corrections. In order to retain the MM3 formalism for intramolecular interactions in AMOEBA, we have only modified the value of this anharmonic constant. To provide a direct test of this simple idea, we have recomputed IR spectra of $SO_4^{2-}(H_2O)_6$ and $SO_4^{2-}(H_2O)_{80}$with a value of either -3.55 or -4.55 for this constant. The results are shown in Figure 15, where structure 6B (6.7.5) was the starting point for n=6, and the non spherical 80D structure was used for n=80.

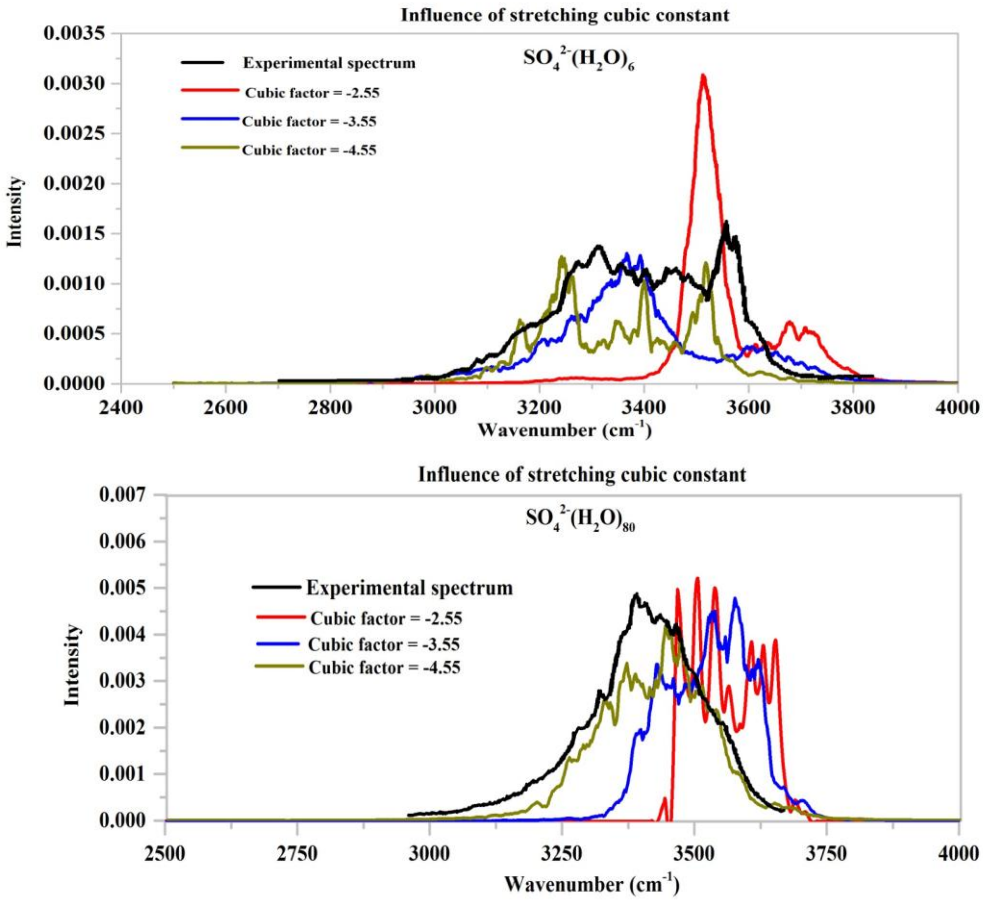


Figure 16 Computed IR spectra of $SO_4^{2-}(H_2O)_6$ and $SO_4^{2-}(H_2O)_{80}$ with several O-H bond stretching anharmonic constants.

It is immediately apparent that this correction has a profound effect on the bonded O-H frequencies. For n=6, the -4.5 value may overshoot somewhat, however the peculiar band structure is probably due to overlimited structure sampling. Since this cluster is mostly made of sulfate-water interactions, this cluster size is probably not the best for tuning the stretching potential. On the other hand the computed spectrum for n=80, in which water-water interactions dominate, leads to better agreement with experiment. Thus, although this simple correction may not capture all physical components of the hydrogen bond interaction, it provides satisfactory results at least in the present case. Inspection of bonded O-H distances did not reveal significant

change with the modified stretching potential, indicating that this change may be used for other cases. It is expected that analogous quality would be obtained for clusters of other hydrated ions in this size range.

6. Conclusions

In this chapter we have provided the first modeling of IR spectra of large hydrated sulfate clusters, with up to 100 water molecules, at finite temperature. This required calibration of force field performance, construction of starting structures, determining a proper confining wall, general simulation conditions and conformational sampling. Then using the DACF/AMOEBA procedure, we were able to simulate IR spectra of $SO_4^{2-}(H_2O)_n$ ions with 25, 55, 80 and 100 water molecules. In agreement with experiments, we find that the number of free O-H bonds, pointing towards the exterior of the cluster, grows with cluster size from negligible to significantly strong. This is clearly reflected in the computed IR spectra with a narrow band at 3730-3710 cm^{-1} depending upon cluster size and temperature, in good agreement with the experimental frequency of ~3710 cm^{-1}. There is also qualitative agreement with a broad band associated with hydrogen bonded O-H bonds, although the position and structure of this band require further improvement. The former has been taken care of by modifying the anharmonicity constant of the O-H bond stretching term in the force field.

Altogether, these results allow, for the first time, the calculation of IR spectra of large and highly fluxional clusters. They provide a new avenue to study the structure and dynamics of ion hydration much beyond the first two solvation shells. This is the main perspective of this work, since analogous experimental spectra have been obtained for ions with charges ranging from -2 to +3, showing significant differences[20]. Extension to other solvents can also be envisaged.

7. References

(1) Finlayson-Pitts, B. J.; Pitts Jr., J. N. *Chemistry of the upper and lower atmosphere – theory, experiments and applications*; Academic Press, San Diego, **2000**.

(2) Twomey, S. *Atmospheric aerosols*; Elsevier, New York, **1977**.

(3) Pincus, R.; Baker, M. B. *Nature* **1994**, *372*, 250.

(4) Abbatt, J. P. D.; Benz, S.; Cziczo, D. J.; Kanji, Z.; Lohmann, U.; Mohler, O. *Science* **2006**, *313*, 1770.

(5) Ramanathan, V.; Crutzen, P. J.; Kiehl, J. T.; Rosenfeld, D. *Science* **2001**, *294*, 2119.

(6) Tallgren, L. *Acta Med. Scand. Suppl.* **1980**, *640*, 1.

(7) Cremer, P. S.; Jungwirth, P. *Nature Chem.* **2014**, *6*, 261.

(8) Blades, A. T.; Kebarle, P. *J. Phys. Chem. A* **2005**, *109*, 8293.

(9) Wong, R. L.; Williams, E. R. *J. Phys. Chem. A* **2003**, *107*, 10976.

(10) Wang, X-B.; Wang, L-S. *Annu. Rev. Phys. Chem.* **2009**, *60*, 105.

(11) Lambrecht, D. S.; Clark, G. N. I.; Head-Gordon, T.; Head-Gordon, M. *J. Phys. Chem. A* **2011**, *115*, 11438.

(12) Lambrecht, D. S.; McCaslin, L.; Xantheas, S. S.; Epifanovsky, E.; Head-Gordon, M. *Mol. Phys.* **2012**, *110*, 2513.

(13) Mardirossian, N.; Lambrecht, D. S.; McCaslin, L.; Xantheas, S. S.; Head-Gordon, M. *J. Chem. Theory Comput.* **2013**, *9*, 1368.

(14) Zhou, J.; Santambrogio, G.; Brümmer, M.; Moore, D. T.; Wöste, L.; Meijer, G.; Neumark, D. M.; Asmis, K. R. *J. Chem. Phys.* **2006**, *125*, 111102.

(15) Miller, Y.; Chaban, G. M.; Zhou, J.; Asmis, K. R.; Neumark, D. M.; Gerber, R. B. *J. Chem. Phys.* **2007**, *127*, 094305.

(16) Bush, M. F.; Saykally, R. J.; Williams, E. R. *J. Am. Chem. Soc.* **2007**, *129*, 2220.

(17) Wan, Q.; Spanu, L.; Galli, G. *J. Phys. Chem. B* **2012**, *116*, 9460.

(18) O'Brien, J. T.; Prell, J. S.; Bush, M. F.; Williams, E. R. *J. Am. Chem. Soc.* **2010**, *132*, 8248.

(19) Prell, J. S.; O'Brien, J. T.; Williams, E. R. *J. Am. Chem. Soc.* **2011**, *133*, 4810.

(20) O'Brien, J. T.; Williams, E. R. *J. Am. Chem. Soc.* **2012**, *134*, 10228.

(21) *Avogadro: an open-source molecular builder and visualization tool. Version 1.0.3.* *http://avogadro.openmolecules.net/*.

(22) Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeersch, T.; Zurek, E.; Hutchison, G. R. *J. Cheminform.* **2012**, *4*, 17.

(23) Muller, A. A.; Bogge, H.; Diemann, E. *Inorg. Chem. Common.* **2003**, *6*, 52.

(24) Cooper, R. J.; Chang, T. M.; Williams, E. R. *J. Phys. Chem. A* **2013**, *117*, 6571.

(25) Bush, M. F.; Saykally, R. J.; Williams, E. R. *J. Am. Chem. Soc.* **2008**, *130*, 15482.

(26) Ji, N.; Ostroverkhov, V.; Tian, C. S.; Shen, Y. R. *Phys. Rev. Lett.* **2008**, *100*,096102.

(27) Gresh, N.; Claverie, P.; Pullman, A. *Int. J. Quant. Chem.* **1986**, *29*, 101.

(28) Gresh, N.; Cisneros, G. A.; Darden, T. A.; Piquemal, J-P. *J. Chem. Theory Comput.* **2007**, *3*, 1960.

(29) Allinger, N. L.; Yuh, Y. H.; Lii, J. H. *J. Am. Chem. Soc.* **1989**, *111*, 8553.

Appendix

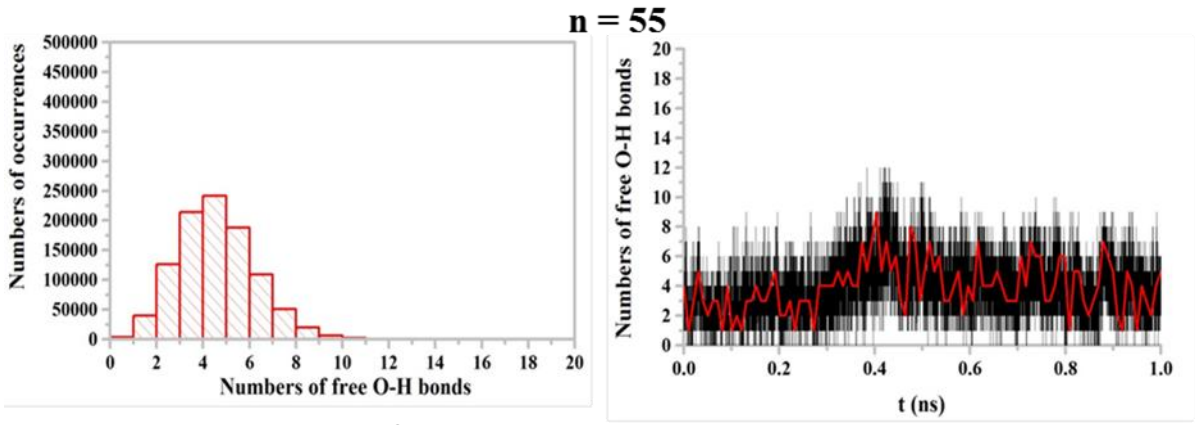


Figure A1 Free O-H bonds in $SO_4^{2-}(H_2O)_{55}$. The left part displays the numbers of occurrences as a function of number of the free O-H bonds while the right part shows the variations of the numbers of free O-H bonds as a function of simulation time (1 ns) at 200 K. The red line corresponds to short-time averaged values.

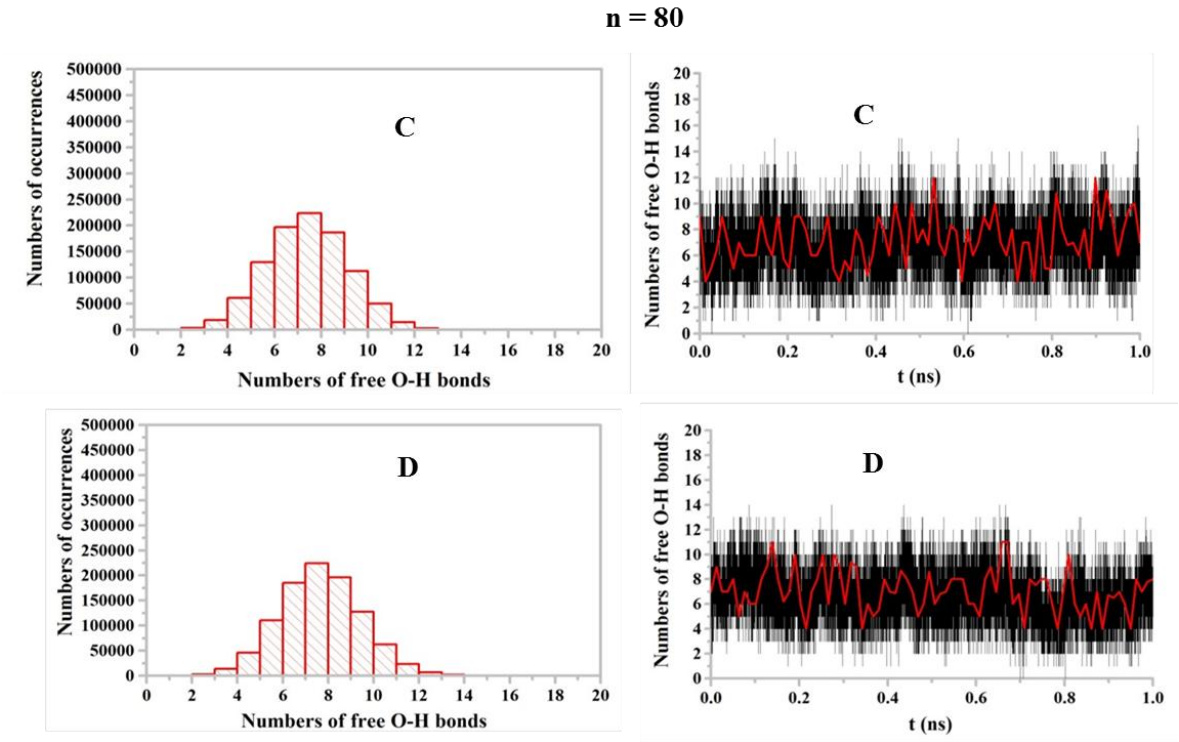


Figure A2 Free O-H bonds in structures 80C (top) and 80D (bottom) of $SO_4^{2-}(H_2O)_{80}$. In each case the left part displays the number of occurrences as a function of number of the free O-H bonds while the right part shows the variations of the numbers of free O-H bonds as a function of simulation time (1 ns) at 200 K. The red line corresponds to short-time averaged values.

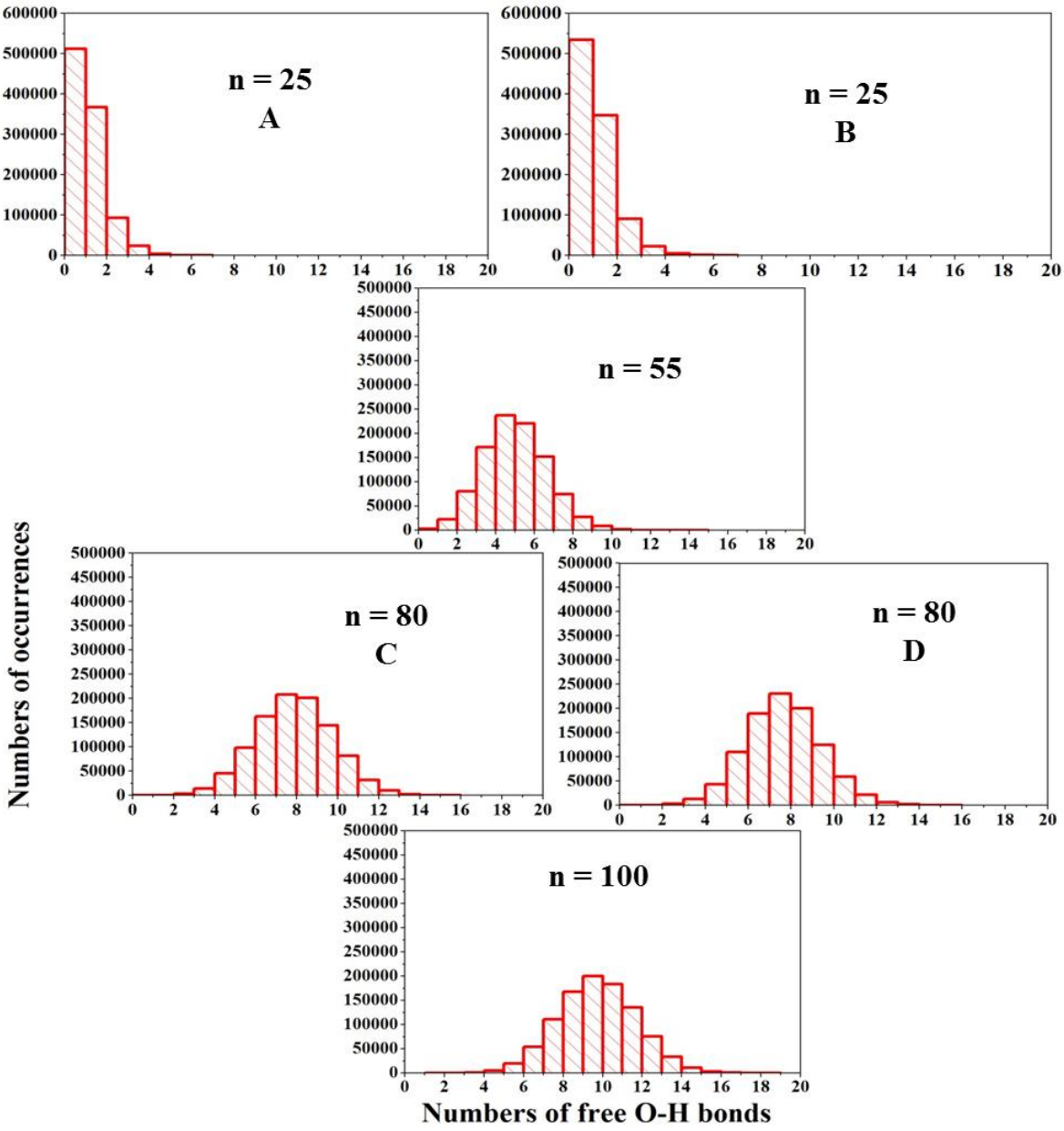


Figure A3 Histograms obtained for 1 ns simulations where the starting points are selected from the final structures of the previous simulations of 1 ns (Figures 8-11) at 200 K. These histograms represent the numbers of occurrences as a function of number of the free O-H bonds.

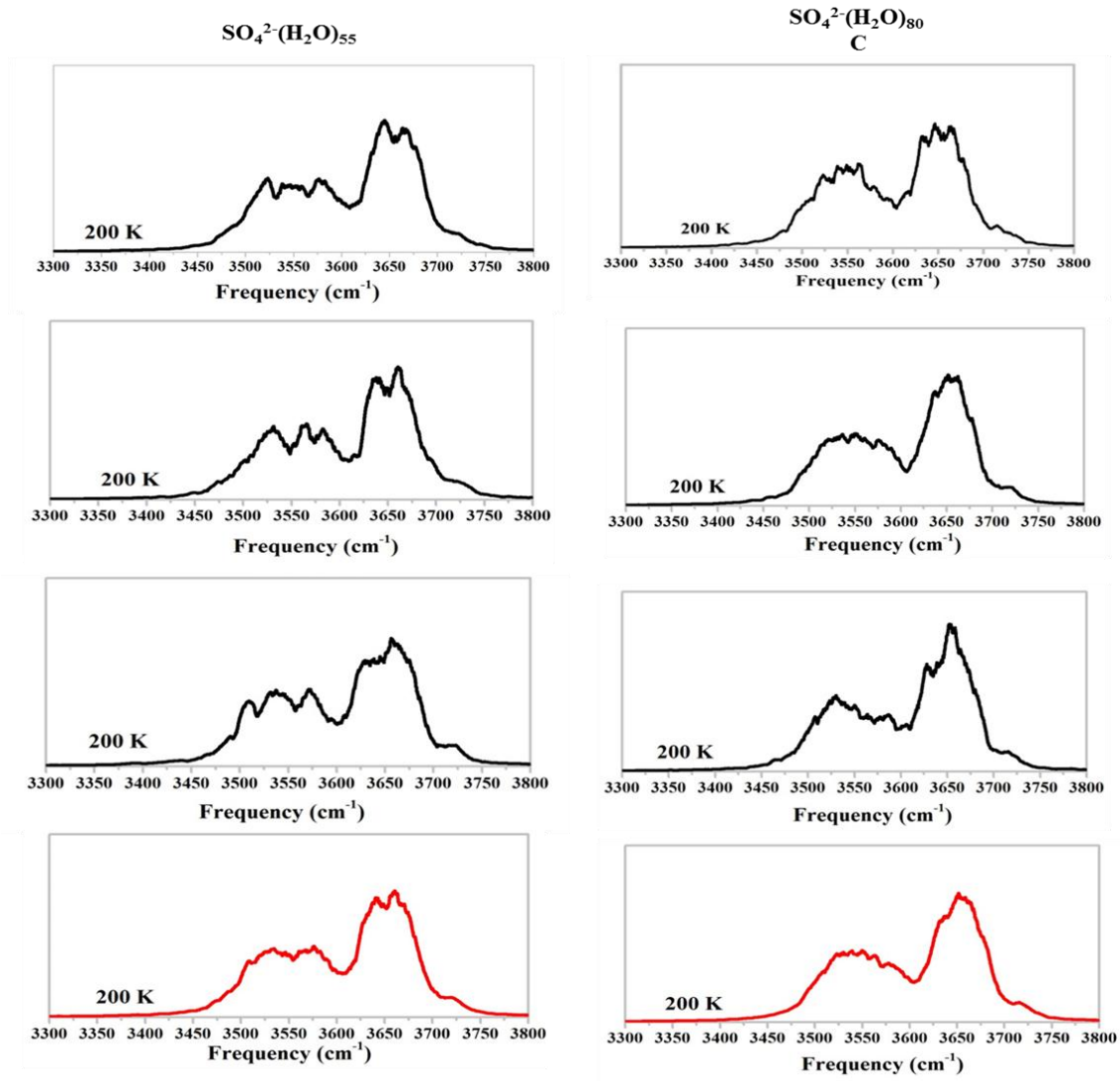


Figure A4 Computed MD spectra for n =55 (left) and for n = 80C (right) at 200 K. In each column the first three spectra (black) are calculated following 1, 2 and 3 ns of equilibration. The bottom spectrum (red) is their average.

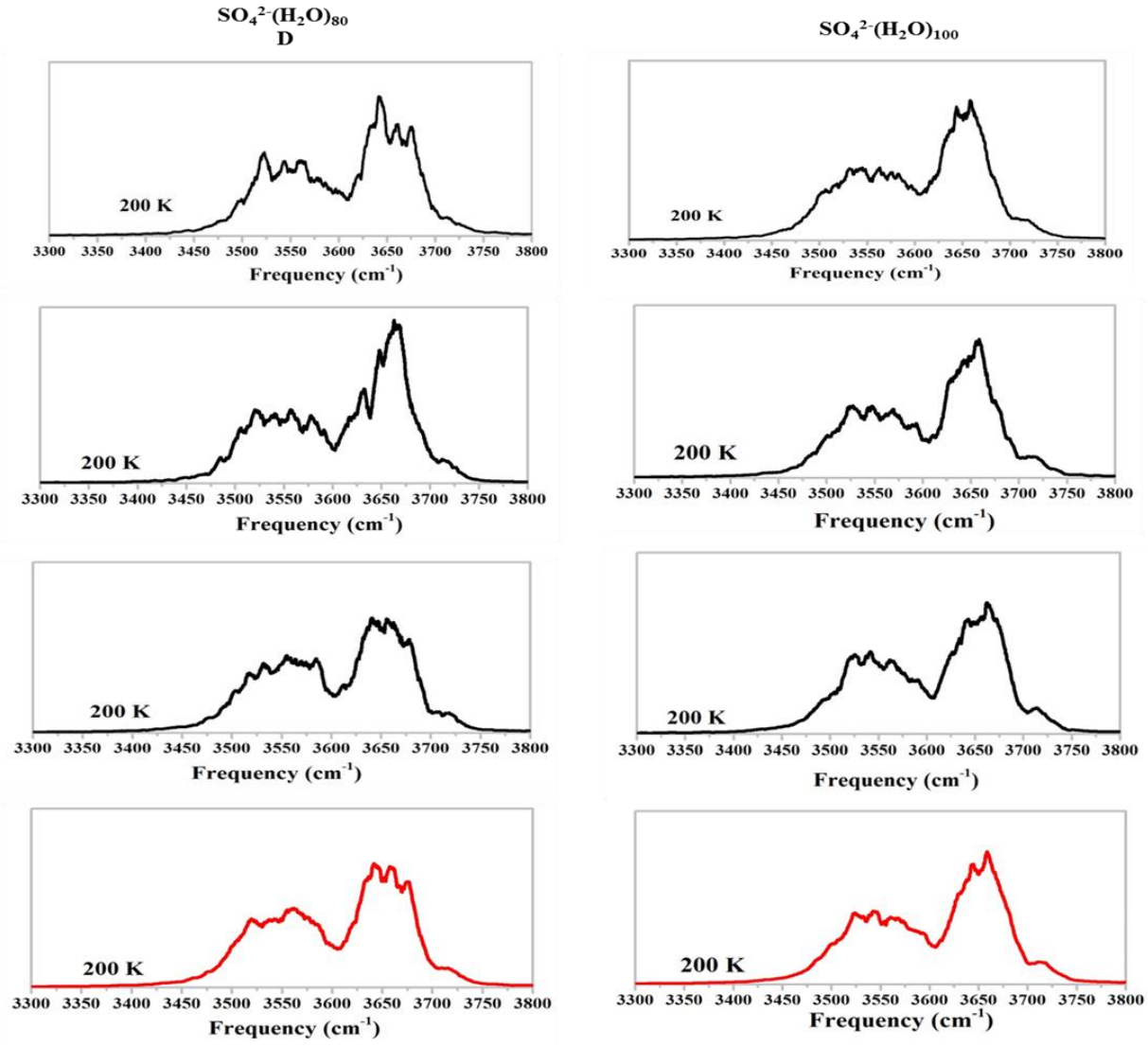


Figure A5 Computed MD spectra for 80D (left) and for n = 100 (right) at 200 K. In each column the first three spectra (black) are calculated following 1, 2 and 3 ns of equilibration. The bottom spectrum (red) is their average.

General conclusions and perspectives

The objective of this work was to gain insight into ion micro-hydration thanks to molecular modeling. Structures and IR spectra of ion-water clusters have been computed in relation with recent gas phase experiments. The micro-hydration of the Zn^{2+}dication, the SO_4^{2-} dianion and the Na^+-tryptamine complex was studied by classical molecular dynamics associated to the AMOEBA polarizable force field. Quantum chemistry calculations at MP2 and DFT levels were also carried out, mostly to provide reference values to calibrate the force field. The AMOEBA force field was chosen for its high-level electrostatics treatment based on a multipolar expansion on each atom, and the explicit inclusion of polarization effects. Both components are essential to treat accurately the hydration of ionic systems. Molecular dynamics was used with the aim to explore extensively the potential energy surfaces and to calculate the IR spectra using the Fourier transform of the Dipole AutoCorrelation Function (DACF).

Special care was taken of intermolecular interactions through the electrostatic parameters of the force field. To this end, the structures and IR spectra of the Na^+-tryptamine complex were studied when attached to one or two water molecules. Atomic multipoles were extracted by using various techniques such as DMA, electrostatic potential fit and GEM-DM. The results show that even if standard procedures are able to reproduce the geometries and relative energies of conformers, obtaining good multipoles for accurate IR spectra is not straightforward. In particular, atomic charges need to be parameterized with caution. Finally, good agreement with experiment could be obtained on band frequencies and intensities. The band attributions have been done by detailed analyses of the MD trajectories used to record the spectra. These investigations have shown that several conformations of $Na^+(Tryp)(H_2O)_{1\text{-}2}$ are explored during the simulations leading to different spectral signatures. Temperature effects on the computed spectra were also pointed out in terms of fluctuations of conformer populations. These results

open the way to the calculation of IR spectra for conformationally flexible systems such as gas phase complexes of amino acids and peptides[1-3].

Micro-hydration of Zn^{2+} and SO_4^{2-} was considered from a different viewpoint, with emphasis on increasing cluster size towards nanodroplets or all the way to bulk hydration. The growing complexity of potential energy surfaces requires extensive exploration, which was tackled by using molecular dynamics with extended time frames. In the case of zinc, reliable coordination numbers (CNs) have been obtained for clusters growing from 6 to 216 water molecules and at temperatures ranging from 200 to 300 K. The results provide a global picture of CN as a function of cluster size, encompassing CNs growing from ca. 5 for small gaseous clusters to 6 in bulk water. This helps bridging the gap between recent experimental values determined in the gas phase[4] and in solution[5,6]. These results are also consistent with a previous AMOEBA study in periodic boundary conditions[7]. Analysis of the simulations highlights possible transitions between several coordination types, depending on the temperature. Finite temperature IR spectra were also computed showing limitations of the force field for the description of bonded O-H modes. Charge transfer was mimicked in a simplistic manner by a decrease of zinc charge. Using this model, IR spectra have been obtained in the 50 – 300 K temperature range, showing a quasi-linear behavior of the bonded O-H frequencies for a small cluster.

Sulfate-water nanodroplets including 25 to 100 water molecules were considered because of intriguing recent IR experiments in the gas phase[8], indicating the appearance of free O-H bonds beyond ca. 45 water molecules. Calculation of IR spectra required a number of preparatory steps: confinement with a spherical repulsive wall to avoid water evaporation, extensive structural sampling to avoid memory effects of the starting points, and calculation of

the number of free O-H bonds at the cluster surface during simulations. This number was found to increase with cluster size, from nearly negligible to ca. 10 in the 25-100 size range. Finite temperature IR spectra were then computed for all sizes in the 100-300 K range. A specific band pertaining to free O-H bonds was found to appear and grow with cluster size at ca. 3720 cm^{-1}, in good agreement with experiments. These are the first reliable computed spectra of such large fluxional systems. They may provide the missing link between spectral feature and the number of free O-H bonds involved. Future developments are required to better describe the vibrations of hydrogen-bonded O-H bonds. The study of water nanodroplets of other ions such as La^{3+}, Ca^{2+} and Cl^- is the obvious next application, since experiments indicate some significant variations with ion charge[9-11].

A realistic picture of the structure, dynamics and spectroscopy of micro-hydrated ions requires a well-grounded model and a large sampling. Molecular dynamics associated to a polarizable force field such as AMOEBA appears to be the method of choice to meet such requirements. However, this work has pointed out some limitations of this model, that we have tried to overcome by extracting new parameters or suggesting modifications to the potential energy expression. In order to improve the computation of IR spectra at finite temperature, developments should focus on the representation of the electrostatic and charge transfer effects. Current implementation of the electron density-based GEM model in AMOEBA for a better treatment of electrostatic, instead of DMA or GEM-DM,[12] appears to be promising in this respect.

References

(1) Durand, S.; Rossa, M.; Hernandez, O.; Paizs, B.; Maitre, P. *J. Phys. Chem. A* **2013**, *117*, 2508.

(2) Kim, J-Y.; Ahn, D-S.; Park, S-W.; Lee, S. *R. Soc. Chem. Adv*. **2014**, *4*, 16352.

(3) Moghaddam, M. B.; Fridgen, T. D. *J. Phys. Chem. B* **2013**, *117*, 6157.

(4) Cooper, T. E.; O'Brien, J. T.; Williams, E. R.; Armentrout, P. B. *J. Phys. Chem. A* **2010**, *114*, 12646.

(5) Chillemi, G.; D'Angelo, P.; Pavel, N. V.; Sanna, N.; Barone, V. *J. Am. Chem. Soc.* **2002**, *124*, 1968.

(6) D'Angelo, P.; Barone, V.; Chillemi, G.; Sanna, N.; Meyer-Klaucke, W.; Pavel, N. V. *J. Am. Chem. Soc.* **2002**, *124*, 1958.

(7) Wu, C.; Piquemal, J-P.; Chaudret, R.; Reinhardt, P.; Ren, P. *J. Chem. Theory Comput.* **2010**, *6*, 2059.

(8) O'Brien, J. T.; Prell, J. S.; Bush, M. F.; Williams, E. R. *J. Am. Chem. Soc.* **2010**, *132*, 8248.

(9) Bush, M. F.; Saykally, R. J.; Williams, E. R. *J. Am. Chem. Soc* .**2008**, *130*, 15482.

(10) O'Brien, J. T.; Williams, E. R. *J. Am. Chem. Soc.* **2012**, *134*, 10228.

(11) Prell, J. S.; O'Brien, J. T.; Williams, E. R. *J. Am. Chem. Soc.* **2011**, *133*, 4810.

(12) Duke, R. E.; Starovoytov, O. N.; Piquemal, J-P.; Cisneros, G. A. *J. Chem. Theory Comput.* **2014**, *10*, 1361.

Abstract

Ion micro-hydration has been modeled by a combination of classical and quantum methods. The structures, dynamics and IR spectra of micro-hydrated cluster ions ranging from 1 to 216 water molecules have been considered with special emphasis on the comparison to recent experimental data. Quantum modeling has been used to provide reference values against which the polarizable AMOEBA force field could be calibrated. Extension of the parameter set was carried out in some cases, in particular when improvement of the electrostatic and polarization energy terms was deemed necessary. Classical molecular dynamics was then used with two different aims: (1) extensive exploration of potential energy surfaces for large cases with numerous low energy minima, for which statistical sampling is required, and (2) computation of IR spectra through Fourier transform of the Dipole moment AutoCorrelation Function (DACF), at various temperatures.

The hydration of Zn^{2+} was studied in detail. While the coordination number (CN) of zinc cation is firmly established to be six in bulk solution, recent experimental and computation studies in the gas phase have unveiled a complex pattern for small clusters with up to 12 water molecules with a CN more near 5 or maybe even 4. Our calculations, performed from 6 to 216 water molecules, allow a precise description of CN evolution with cluster size.

Sodiated tryptamine with one or two water molecules attached was investigated in order to interpret gas phase IR spectra of these species. Particular care was taken of electrostatics - especially atomic charges - and of starting structures in molecular dynamics simulations. It was then possible to investigate temperature-dependent dynamics in detail and to obtain good agreement with IRPD spectra.

Finally, the hydration of the sulfate anion with up to 100 water molecules was analyzed focusing on the existence of dangling O-H bonds on the cluster surface as a function of cluster size. In agreement with IRPD experiments, we find that small clusters are compact with no dangling bond. The latter appear when ca. 25 water molecules are present with a characteristic vibrational band.

Overall, these results underline the requirements for accurate modeling of ion hydration. These involve refined treatment of electrostatic and polarization interactions, extensive sampling of the potential energy surfaces, careful treatment of cluster boundaries including a confining repulsive wall, and long-time simulations to attain converged radial distribution functions. With such developments at hand, it is now possible to obtain a realistic picture of structures and dynamics of large micro-hydrated cluster ions, as can be judged from the comparison to experiment, especially vibrational spectroscopy.

Keywords: molecular dynamics, infrared spectroscopy, polarizable force field, quantum chemistry, ion hydration

Résumé

La micro-hydratation d'ions a été modélisée par une combinaison de méthodes classiques et quantiques. Les structures, la dynamique et les spectres IR d'agrégats ioniques micro-hydratés allant de 1 à 216 molécules d'eau ont été examinés en mettant l'accent sur la comparaison avec des résultats expérimentaux récents. La modélisation quantique a été utilisée pour fournir des valeurs de référence pour la calibration du champ de force polarisable AMOEBA. L'extension du jeu de paramètres a été réalisée dans certains cas, en particulier lorsqu'il a été nécessaire d'améliorer les termes d'interaction électrostatique et de polarisation. La dynamique moléculaire classique a ensuite été utilisée avec deux objectifs distincts: (1) l'exploration extensive des surfaces d'énergie potentielle pour des systèmes de grande taille avec de nombreux minima de basse énergie, pour lesquels un échantillonnage statistique est nécessaire, et (2) le calcul des spectres IR par transformée de Fourier de la fonction d'autocorrélation du moment dipolaire (DACF) à différentes températures.

L'hydratation du cation Zn^{2+} a été étudiée en détail. Alors que le nombre de coordination (CN) de Zn^{2+} est clairement établi à six en solution aqueuse, des études expérimentales et théoriques récentes en phase gazeuse ont mis en évidence un modèle complexe pour les petits agrégats jusqu'à 12 molécules d'eau avec un CN de l'ordre de 5 ou peut-être même 4. Nos calculs allant de 6 à 216 molécules d'eau permettent une description précise de l'évolution du CN en fonction de la taille de l'agrégat.

La tryptamine complexée aux ion et potassium et à une ou deux molécules d'eau a été étudiée afin d'interpréter des spectres IR en phase gazeuse de ces espèces. Une attention particulière a été portée à l'électrostatique, en particulier les charges atomiques, et sur le rôle des structures de départ des simulations de dynamique moléculaire. Il a alors été possible d'étudier en détail les phénomènes dynamiques dépendant de la température et d'obtenir une bonne concordance avec les spectres IRPD.

Enfin, l'hydratation de l'anion sulfate avec un nombre maximum de 100 molécules d'eau a été analysée en mettant l'accent sur l'existence de liaisons O-H libres sur la surface de l'agrégat en fonction de sa taille. En accord avec les expériences IRPD, nous constatons que les petits agrégats sont compacts et sans liaison pendante. Ces dernières apparaissent avec une bande de vibration caractéristique quand environ 25 molécules d'eau sont présentes.

Globalement, ces résultats mettent en évidence les conditions requises pour une modélisation précise de l'hydratation d'ions. Il s'agit notamment de prendre en compte un traitement raffiné des interactions électrostatiques et de polarisation, un échantillonnage étendu des surfaces d'énergie potentielle, le traitement adapté des conditions aux limites des agrégats, y compris d'un mur de confinement répulsif, et des simulations de longue durée pour atteindre la convergence des fonctions de distribution radiale. Avec de tels développements, il est maintenant possible d'obtenir une description réaliste des structures et de la dynamique d'agrégats ioniques micro-hydratés de grande taille, ainsi que leur dépendance en température, comme le montre la comparaison avec l'expérience, en particulier la spectroscopie vibrationnelle.

Mosts-clés: dynamique moléculaire, spectre infrarouge, champ de force polarizable, chimie quantique, hydratation d'ions

Abstract

Ion micro-hydration has been modeled by a combination of classical and quantum methods. The structures, dynamics and IR spectra of micro-hydrated cluster ions ranging from 1 to 216 water molecules have been considered with special emphasis on the comparison to recent experimental data. Quantum modeling has been used to provide reference values against which the polarizable AMOEBA force field could be calibrated. Extension of the parameter set was carried out in some cases, in particular when improvement of the electrostatic and polarization energy terms was deemed necessary. Classical molecular dynamics was then used with two different aims: (1) extensive exploration of potential energy surfaces for large cases with numerous low energy minima, for which statistical sampling is required, and (2) computation of IR spectra through Fourier transform of the Dipole moment AutoCorrelation Function (DACF), at various temperatures.

The hydration of Zn^{2+} was studied in detail. While the coordination number (CN) of zinc cation is firmly established to be six in bulk solution, recent experimental and computation studies in the gas phase have unveiled a complex pattern for small clusters with up to 12 water molecules with a CN more near 5 or maybe even 4. Our calculations, performed from 6 to 216 water molecules, allow a precise description of CN evolution with cluster size.

Sodiated tryptamine with one or two water molecules attached was investigated in order to interpret gas phase IR spectra of these species. Particular care was taken of electrostatics - especially atomic charges - and of starting structures in molecular dynamics simulations. It was then possible to investigate temperature-dependent dynamics in detail and to obtain good agreement with IRPD spectra.

Finally, the hydration of the sulfate anion with up to 100 water molecules was analyzed focusing on the existence of dangling O-H bonds on the cluster surface as a function of cluster size. In agreement with IRPD experiments, we find that small clusters are compact with no dangling bond. The latter appear when ca. 25 water molecules are present with a characteristic vibrational band.

Overall, these results underline the requirements for accurate modeling of ion hydration. These involve refined treatment of electrostatic and polarization interactions, extensive sampling of the potential energy surfaces, careful treatment of cluster boundaries including a confining repulsive wall, and long-time simulations to attain converged radial distribution functions. With such developments at hand, it is now possible to obtain a realistic picture of structures and dynamics of large micro-hydrated cluster ions, as can be judged from the comparison to experiment, especially vibrational spectroscopy.

Keywords: molecular dynamics, infrared spectroscopy, polarizable force field, quantum chemistry, ion hydration