Università degli Studi di Roma
"La Sapienza"

Dipartimento di Chimica

DOTTORATO DI RICERCA IN SCIENZE CHIMICHE
XXII CICLO
ANNO ACCADEMICO: 2008 - 2009

A combined theoretical and experimental investigation of Ion Hydration

Dottoranda:
Valentina Migliorati

Relatore:
Prof.ssa Paola D'Angelo

Correlatore:
Prof. Francesco Gianturco

In memory of my grandparents,
Sabatino and Elvira

Contents

Chapter 1

Introduction

1.1 Ion hydration

Water is the most abundant compound on the surface of earth and, being the principal constituent of all living organism, it is the basis for life on our planet. Consequently, knowledge of the structural and dynamic properties of water is crucial in many problems of physics, chemistry and biology. Even if water has an apparently simple molecular structure, it is a rather complex fluid and shows many distinctive properties which are generally ascribed to the hydrogen bond at molecular level. The structure of water is in fact well described in terms of a dynamical network of hydrogen bonded clusters in which tetrahedral cages play a dominant role. The fundamental dynamical process occurring in water is the formation and breaking of hydrogen bonds which generally take place in the subpicosecond time scale [1].

The structure and dynamics of these hydrogen bonded clusters is modified by changes of temperature and pressure, as well as by the introduction of solutes [2]. In particular, ions in solution strongly distort the structure of surrounding water molecules, as the result of the change in the microscopic balance of intermolecular forces, from that of water-water interactions in the neat solvent to that of ion-water interactions in the resulting solution. Our present understanding of the changes occurring to water in the presence of an ion is based on the scheme introduced by Frank and Wen [3] and Gurney [4] who considered three concentric solvent regions around the ions: the innermost region, the so called first hydration shell, in which the water molecules are strongly oriented by the ion and tend to be carried by the ion as it moves through the solution, the second hydration shell in which the water molecules are only weakly oriented by the ion and finally, in the outermost region far from the ion, the structure of water is generally the same as that of bulk water. Figure 1.1 provides a schematic picture of the hydration spheres of a metal cation having a first solvation shell of six water molecules. This is for example the situation encountered in many 3d metal ions.

Ions have been classified as structure makers and structure breakers according to their ability to induce structuring of water. Small ions with high electric charge are generally structure makers, as the water molecules in the

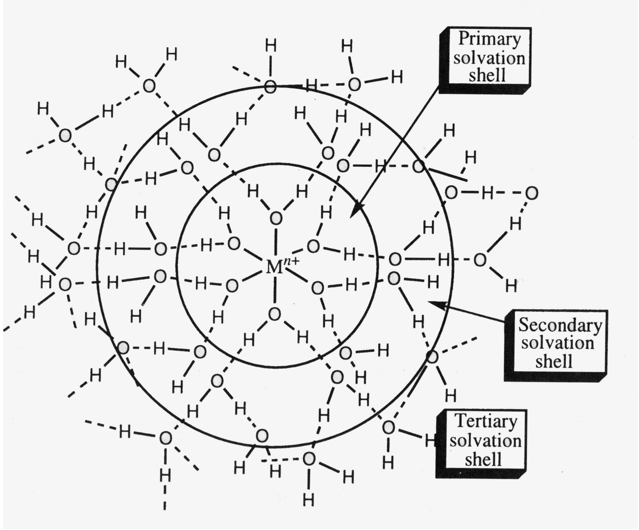

Figure 1.1: Structure of a generic hydrated metal cation in aqueous solution.

first solvation shell are strongly bound to the ion and it is appropriate to think of a well defined ion-water complex. Larger ions instead are often structure breakers as their main influence is the disruption of the hydrogen bond network characteristic of bulk water. All ions are hydrated to varying extents in water but the degree of hydration depends on a number of factor, such as the ionic size and the charge density [1]. Cations are more strongly hydrated in general than anions due to a combination of high positive charge density and a particularly strong interaction with the negatively polarized oxygen atom of water. However, a well defined first hydration shell of water molecules exists also around halide ions, even if the interaction (via the hydrogen atoms) is somewhat weaker [1].

The structural and dynamic properties of the hydration spheres of aqua ions are fundamental to understanding the behaviour of ions in chemical and biological systems and processes. Consequently, a large number of experimental techniques, primarily X-ray and neutron diffraction, have been applied to obtain structural information on the ion-water interaction. For 3d metal transition elements the identification of the primary hydration geometry, usually octahedral six-coordinated, has proved relatively straightforward [2]. For the rest of ions, including anions, alkaline and earth-alkaline cations, information on the hydration structure is not very conclusive, in principle as a result of the higher disordered environments and a general lack of direct information relating to static and dynamic properties of solvent molecules when they coordinate ions [2].

During the last several years it has been shown that X-ray absorption spectroscopy (XAS) is particularly well suited for the investigation of the local solvent structure of ions dissolved in water, due to its atomic selectivity and its sensitivity to dilute solutions. From the analysis of the Extended X-Ray Absorption Fine Structure (EXAFS) it is possible to obtain very accurate ion-water first shell distances. However, in the case of disordered systems, such as ionic solutions, the uncertainty in the coordination numbers determined by the EXAFS analysis is usually too large for a conclusive determination of the geometry of hydration complexes. Conversely, a quantitative analysis of the X-ray Absorption Near Edge Structure (XANES) region, which includes the rising edge and about 200 eV above it, can provide accurate geometrical information on the hydration clusters existing in water. Nevertheless, the characterization of the structural and dynamical properties of ions and water molecules in the hydration spheres is very difficult to be obtained from experimental techniques only, and the combined use of experimental and theoretical methods is essential to obtain reliable information. Among computer simulation techniques, Molecular Dynamics is a powerful tool in the analysis of both static and dynamic properties of solvated ions in solution and has been extensively used in the last decades for the study of aqueous solutions [2].

In this context, the aim of this work is to unveil the detailed structure and dynamics in aqueous solutions of the group 12 divalent cations, i.e. Zn^{2+}, Cd^{2+} and Hg^{2+}, and of one anion (Br^-), using a procedure which combines XAS spectroscopy and Molecular Dynamic simulation techniques.

1.2 Group 12 aqua ions

Zinc is an extremely important biological element and it is the second most abundant element in the human body after iron. The primary role for zinc is to promote hydrolitic processes in the form of a Lewis acid, the most well known examples being the enzymes carboxypeptidase, carbonic anhydrase, and the alcohol dehydrogenase [1]. In aqueous solution the Zn^{2+} ion is well known to be coordinated by six water molecules in an octahedral symmetry common to many 3d transition metals such as Ni^{2+} and Co^{2+}. The existence of a stable octahedral geometry has been observed in many experimental and computational studies [1, 5] and has been recently confirmed using a combined Molecular Dynamics-XAS approach [6, 7].

Even if they belong to the same group of the periodic table, cadmium and mercury have very different biological properties as compared to zinc. In fact both of them are toxic and environmentally hazardous elements and can alter a wide variety of cellular and biochemical processes, as they are able to replace biological Zn^{2+} and Ca^{2+} in enzymes, proteins, and nucleic acids, modifying the normal activity of these species [8]. Structural information on Cd^{2+} and Hg^{2+} hydration complexes is key for understanding their transport in aqueous environments and their interaction with biological molecules and proteins. In particular, it is very important to resolve the ability of Cd^{2+} and Hg^{2+} to mimic ions like Zn^{2+} and Ca^{2+} which show very different coordination properties [9] in biological systems, since Zn^{2+} in protein tends to form well-defined coordination structures generally with four or six ligands, while the Ca^{2+} coordination is characterized by higher coordination numbers and more disordered coordination spheres [1].

The structure of the hydrated Cd^{2+} ion in aqueous solution has been the subject of several X-ray diffraction investigations, where the number of nearest neighbors was assumed "a priori" to be six (as the structures of all isolated hydrated Cd^{2+} ions in the solid state display an octahedral configuration), and which led to consistent values for the Cd-O internuclear distance [2]. Conversely, an investigation by our group combining Molecular Dynamics simulations, EXAFS, and large angle X-ray scattering measurements has suggested a more dynamical picture with a first coordination shell which transits between hexa- and heptahydrated clusters [10, 11]. The existence of such a flexible hydration structure could provide a rationale for the ability of Cd^{2+} to substitute ions having variable coordination geometries in biological systems. Nevertheless, this finding is still controversial since a recent study on the hydration structure of cadmium by ab initio calculations concluded that the hexacoordinated model is more probable [12]. These new results suggest the need of additional experimental and theoretical work to support the claim that cadmium has a variable coordination number in which a heptacoordinated species plays a predominant role.

The Hg^{2+} ion has usually been described as being octahedrally coordinated by water [13], although a degree of uncertainty has been caused by the variety of coordination structures found in solid complexes [1], and by the quite short residence time of water molecules in the first coordination shell [14] (on the order of 1 ns), as compared to divalent first-row transition ions which form octahedral complexes in aqueous solution, indicating a larger flexibility of the Hg^{2+} hydration complex. From a computational point of view, several approaches have been used to investigate the properties of the Hg^{2+} aqua ion. These include quantum mechanical ab initio calculations on $Hg^{2+}(H_2O)_n$ clusters [15], classical Molecular Dynamics simulations (both with simple pair potentials and including three-body corrections) [16], and ab initio quantum mechanics/molecular mechanics (QM/MM) simulations [17], all describing the Hg^{2+} innermost hydration shell as an octahedral complex.

However, the radial distribution function obtained from X-ray diffraction shows an unexpectedly large variation in the Hg-O bond lengths, explained by a pseudo Jahn-Teller effect in the octahedral complex [18]. In a very recent study the Hg^{2+} hydration structure has been investigated by neutron diffraction with isotopic substitution [19]. The conclusion drawn by the authors is that the first solvation shell contains six water molecules, even though the actual experimental determination of the number of deuterium atoms around the Hg^{2+} ion is 13.5 ± 2.1, thus suggesting a higher coordination number. It is clear from these findings that a conclusive description of the structural properties of Hg^{2+} ion in aqueous solution is still lacking.

1.3 Ion hydration in high-density water

The properties of aqueous solutions under pressure are fundamentally important for a wide range of scientific disciplines, such as geoscience, environmental sciences and planetary modeling. At high pressures (and temper-

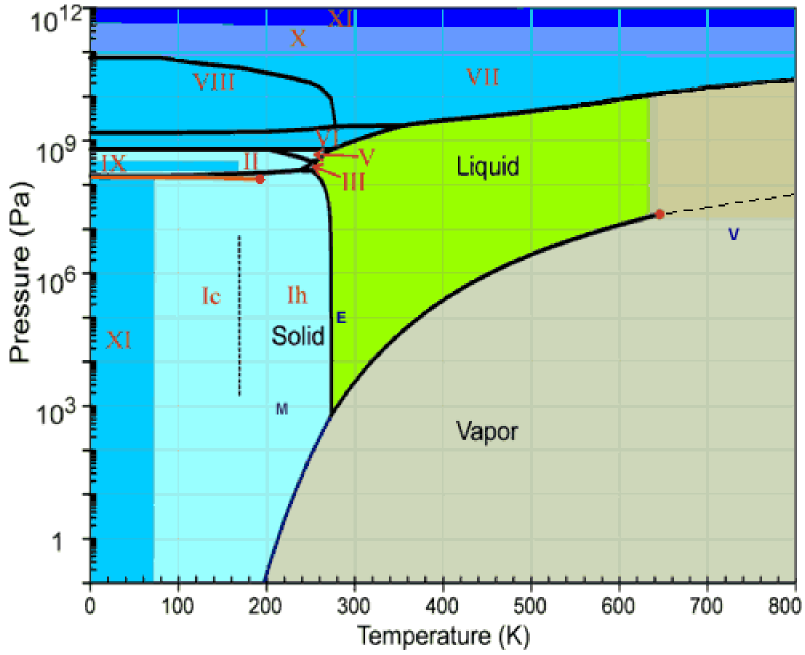

Figure 1.2: The phase diagram of water.

atures) water exists in different liquid or solid phases whose properties are very different from those found at normal conditions. Supercritical water exists at temperatures above 600 K and pressures between 10 MPa and 1 GPa (see Figure 1.2). It is characterized by a low dielectric constant (~ 6 in a typical supercritical regime) and weak hydrogen bonding formation which makes it an excellent solvent for non polar substances. Due to the broad range of applications, such as in the decomposition of organic wastes or in the hydrothermal synthesis, supercritical water has been extensively studied [20, 21, 22]. At pressures between 0.1 and 10 GPa and temperature between 300 and 500 K liquid water is in equilibrium with several forms of ice and its phase in this region is called compressed water (see Figure 1.2). Due to the huge pressure, the phases existing in this zone are very different and much more dense than those found at normal conditions. The trend of the water density as a function of temperature and pressure is shown in Figure 1.3. For a long time, compressed water has not been so extensively studied from an experimental point of view due to the difficulties of reaching this very high pressures and only in these last years it has been possible to make experiments in this pressure range. For this reason compressed water is a system that has recently gained much interest, and the effect of pressure on the structure and hydrogen bond network of compressed water has been addressed by a variety of experimental and theoretical works [23, 24, 25, 26, 27, 28, 29]. Although different conclusions are drawn from these studies, they suggest the possible existence of a high-density state of water (HDW) which at a pressure of few kilobars gradually evolves at the expense of a low-density state of water (LDW) [26]. In HDW the second shell of neighbor molecules is thought to

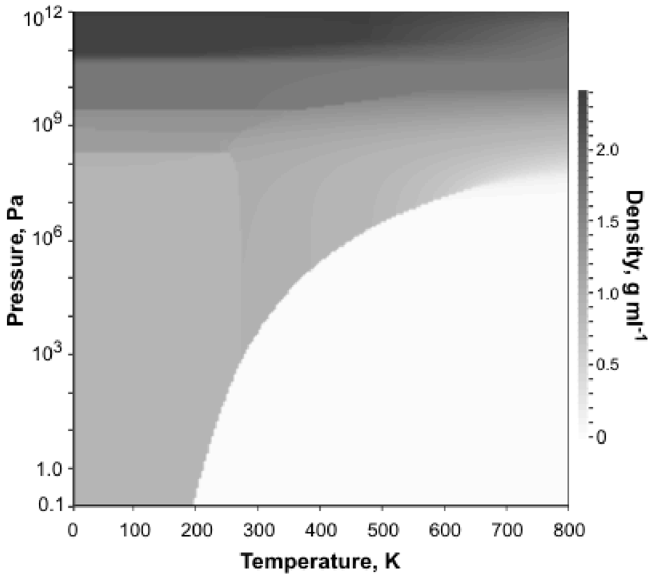

Figure 1.3: Trend of the water density as a function of temperature and pressure.

collapse onto the first one, as a consequence of the rupture of the hydrogen bond network. However, the results of a Molecular Dynamics simulation [28] have shed light on the key role played by interstitial molecules in the second coordination shell, while in a recent neutron diffraction investigation the authors conclude that with increasing density water approaches a local structure common to a simple liquid [29]. Despite the plethora of studies on the properties of pure high-density water, to the best of our knowledge there is only one work in the literature addressing the structural changes of ionic aqueous solutions under high pressure [30]. In this study the transformation of the Rb^+ and Br^- first hydration shells was investigated and dramatic effects in the anion hydration structure have been observed with increasing pressure.

1.4 Halide ions in aqueous solution: the bromide case

While much experimental and theoretical work has been devoted to characterize the structural and dynamical behaviour of cations in water, the hydration properties of halides are still the subject of intense debate [31, 32]. For the bromide ion the hydration number as determined by X-ray and neutron diffraction is 6 in most cases, but as found for other halide ions, it is very difficult to establish the number of water molecules belonging to the first coordination shell due to its diffuse character [2]. The Br-O first shell bond length determined by X-ray diffraction is in the range 3.29-3.40 Å [2]. These scattered results underline the difficulty of defining the halide coordination shells also as a consequence of the fast water exchange between the first and second hydration shells. The experimental residence time is generally estimated to be very short (less than 5 ps) [33]. A K-edge EXAFS study of bromide ions in aqueous solutions has been carried out using classical Molecular Dynamics simulations [34]. The obtained Br-O radial distribution function was used with the integral formulation of the EXAFS equation to simulate the EXAFS theoretical signal which was found to be in very good agreement with the experimental data. While this approach does not allow to get the detailed orientation of water molecules around the bromide ion (also the H atoms were neglected), it did provide an oxygen coordination number (6.9) and first shell distance (3.34 Å), well within the range of previous studies [2]. The same approach was used by Wallen et al. [35] who extended the range of temperatures into the supercritical regime of water while Filipponi et al. [30] investigated the pressure dependence of the Br-O radial distribution function up to 2.8 GPa.

Very recently, time-resolved X-ray absorption spectroscopy has been used to observe the transient species generated by one-photon detachment of an electron from aqueous bromide [36]. A laser pulse of 200 nm has been used to modify the Br^- electronic structure, while X-ray pulses with a duration of 80 ps have been used as probe. The water shell organization has been found to vary upon electronic structure changes but a quantitative analysis of the solvent shell reorganization relies on a clear description of the solvation shell structure around Br^- prior to laser excitation. These results clearly show that the interpretation of experimental data relies heavily on high quality computational simulations [2, 32, 34, 37]. Recently, Raugei and Klein [38] reported Density Functional Theory (DFT)-based Car-Parrinello Molecular Dynamics simulations of a box of 32 water molecules and either one HBr molecule or a Br^- ion. They found a rather structured and asymmetric solvation shell, at variance with the classical Molecular Dynamics results.

Even if many quantum simulations have been carried out on cluster models of water molecules around the halides [37, 39], in general only few simulations (classical, semi-classical or quantum) have been checked against experimental observables. Merkling et al. [37] derived optimized geometries of $[Br(H_2O)_n]^-$ clusters ($1 \leq n \leq 8$) from quantum chemical calculations, and, in order to reproduce the X-ray Absorption Near Edge Structure spectra in a satisfactory way, they had to consider statistical fluctuations, which in their case were obtained from snapshots of Monte Carlo simulations. In limiting their analysis to the first shell, the simulations of Merkling et al. [37] are not sufficient to understand the dynamics of hydration, including exchange with the bulk.

Clearly, more work, both experimental and theoretical, is needed to unveil the structure and dynamics of the bromide ion hydration shells, and the complexity of the problem is underlined by the spread of coordination numbers and distances found in the literature.

1.5 Aim of this work

In this work a detailed investigation of the structure and dynamics of Zn^{2+}, Cd^{2+}, Hg^{2+} and Br^{-} ions in aqueous solution will be carried out combining X-ray absorption spectroscopy and Molecular Dynamics simulations. The solvation properties of Cd^{2+}, Hg^{2+} and Br^{-} ions will be studied at 1 bar and 300 K, since the hydration structure of these ions is still poorly defined even at ambient conditions. Conversely, in the case of the Zn^{2+} ion, which is well known to be octahedral at 1 bar and 300 K [5, 6, 7], the hydration properties will be studied in conditions of very high pressure (up to 2.85 GPa), with a double aim: on the one hand to unveil the structural transformation occurring to water from the low- to the high-density conditions, on the other hand to investigate the pressure effects on the hydration structure of the Zn^{2+} ion.

For the group 12 aqua ions classical Molecular Dynamics will be used as it allows to study dynamic processes occurring in the nanosecond or longer time scale. In fact, the residence time of water molecules in the first hydration shell of Cd^{2+} and Hg^{2+} is of the order of nanoseconds [14] and quantum mechanical Car-Parrinello or QM/MM simulations, which are computationally very expensive and can be used only in short simulations (up to a few hundreds of picoseconds), can be affected by poor sampling if used to investigate transport properties occurring on the nanosecond or longer time scale. On the other hand, classical Molecular Dynamics can be used to obtain simulations several orders of magnitude longer, but the proper choice of the interaction potentials used is always a mandatory prerequisite for a reliable description of the system under investigation. For this reason, in this work it will be further developed and applied a computational procedure for the generation of effective two-body ion-water potentials, by means of ab initio calculations, to be employed in the Molecular Dynamics simulations. The reliability of this methodology has been already assessed by an investigation of the hydration properties of Zn^{2+}, Co^{2+} and Ni^{2+} at ambient conditions [6].

On the contrary, in the case of halide ions classical Molecular Dynamics could not be able to reproduce the hydration structure accurately, since it is difficult in the classical framework to take properly into account polarization effects, which have a strong impact on the halide-water interactions. As a consequence, ab initio Molecular Dynamics will be used in the study of the bromide aqua ion, as it is able to describe the dynamics of the system in a self-consistent fashion, taking into account in a natural way polarization and many-body interactions. Moreover in the case of the Br^{-} ion, the estimated residence time of the first shell water molecules is three orders of magnitude shorter (in the picosecond timescale) as compared to Cd^{2+} and Hg^{2+}, so that sufficiently long ab initio Molecular Dynamics can be carried out to achieve a proper sampling of the phase space of the system.

This thesis is organized as follows. Chapter 2 and 3 describe the theoretical background of electronic structure methods and Molecular Dynamics simulations. In chapter 4 the basic concepts of X-ray absorption spectroscopy are given, and the methods employed in the EXAFS and XANES data analysis are described, with particular emphasis on their application to the study of disordered systems. Chapter 5 addresses the procedure used to generate the ion-water pair potentials, to be used in the classical Molecular Dynamics simulations of group 12 aqua ions. The theoretical and experimental methods employed in the study of group 12 aqua ions are summarized in chapter 6. Chapter 7 and 8 describe all of the results on the hydration properties of Hg^{2+} and Cd^{2+} at ambient conditions, respectively, while the results of our investigation on the Zn^{2+} hydration properties under high pressure will be presented in chapter 9. Chapter 10 describes our combined theoretical and experimental investigation of the bromide ion hydration properties. Finally, chapter 11 summarizes and concludes this thesis.

Chapter 2

Electronic structure methods

In this chapter an introduction to electronic structure methods will be given. Atomic units will be used throughout [40]:

$$e = 1, \; m_e = 1, \; \hbar = 1, \; 4\pi\epsilon_0 = 1$$

where e and m_e are the electron charge and mass, respectively, and ϵ_0 is the dielectric constant of vacuum.

2.1 The Molecular Hamiltonian

According to quantum mechanics the energy and many other properties of a stationary state of a molecule can be obtained by solution of the time-independent Schrödinger equation [41]:

$$\mathbf{H}\Psi = E\Psi \tag{2.1}$$

where $\mathbf{H}$ is the Hamiltonian operator representing the total energy, E is the numerical value of the energy of the state and Ψ is the wavefunction, which depends on the coordinates of all of the particles and contains all the information on the system. The square modulus of the wavefunction, $|\Psi|^2$, is interpreted as a measure of the probability distribution of the particles within the molecule. The Hamiltonian $\mathbf{H}$, like the energy in classical mechanics, is the sum of kinetic $\mathbf{T}$ and potential $\mathbf{V}$ operators:

$$\mathbf{H} = \mathbf{T} + \mathbf{V} \tag{2.2}$$

In the simplest case, i.e. if no external electrostatic or magnetic fields are present and if we restrict the interactions among particles to the Coulomb interactions (neglecting spin-spin and spin-orbit coupling operators), the molecular Hamiltonian for a system composed of M nuclei and N electrons is given by:

$$\mathbf{H} = \mathbf{T}_n + \mathbf{V}_{nn} + \mathbf{V}_{ne} + \mathbf{T}_e + \mathbf{V}_{ee} \tag{2.3}$$

where

$$\mathbf{T}_n = -\sum_{a=1}^{M} \frac{1}{2M_a} \nabla_a^2 \tag{2.4}$$

is the operator of the kinetic energy of the nuclei (M_a is the mass of nucleus a and ∇_a^2 is the Laplace operator acting over the coordinates of nucleus a),

$$\mathbf{V}_{nn} = \sum_{a=1}^{M} \sum_{b>a=1}^{M} \frac{Z_a Z_b}{R_{ab}} \tag{2.5}$$

is the Coulomb repulsion between the nuclei (Z_a is the charge of nucleus a, R_{ab} is the distance between nuclei a and b),

$$\mathbf{V}_{ne} = -\sum_{a=1}^{M} \sum_{i=1}^{N} \frac{Z_a}{r_{ia}} \tag{2.6}$$

is the Coulomb nucleus-electron interactions (r_{ia} is the distance between nucleus a and electron i),

$$\mathbf{T}_e = -\sum_{i=1}^{N} \frac{1}{2} \nabla_i^2 \tag{2.7}$$

is the operator of the kinetic energy of the electrons (∇_i^2 being the Laplace operator acting over the coordinates of electron i) and finally

$$\mathbf{V}_{ee} = \sum_{i=1}^{N} \sum_{j>i=1}^{N} \frac{1}{r_{ij}} \tag{2.8}$$

is the Coulomb repulsion between the electrons (r_{ij} is the distance between electrons i and j).

This Hamiltonian is non relativistic and it is only valid for particles with a velocity much smaller than the velocity of light. This is generally sufficient for light elements (with atomic number Z lower than 36), and also for the valence electrons of heavier elements.

2.2 The Born-Oppenheimer approximation

In order to solve the Schrödinger equation 2.1 we need to use approximate methods. In fact the Schrödinger equation can be solved exactly only for very simple systems, such as the harmonic oscillator or the rigid rotor, and for one-electron systems (hydrogenoid atoms or the H_2^+ molecular ion). The first important approximation used to simplify the solution of the Schrödinger equation is the Born-Oppenheimer approximation, which consists in a separation of the motions of the electrons from that of the nuclei. This approximation is reasonable since the mass of the nucleus is thousands of times greater than that of the electron. The nuclei thus move very slowly with respect to the electrons and the electrons react essentially instantaneously to changes in nuclear positions. Therefore, the electron distribution within a molecular system depends on the positions of the nuclei and not on their velocities or, in other words, the nuclei appear fixed to the electrons and the electronic motion can be described as occurring in a field of fixed nuclei.

In the Born-Oppenheimer approximation the total wavefunction $\Psi(\mathbf{r}, \mathbf{R})$, which depends on the coordinates of all of the nuclei ($\mathbf{R}$) and all of the electrons ($\mathbf{r}$), is written in the form of a product of an electronic wavefunction $\Psi_e(\mathbf{r}; \mathbf{R})$ and a wavefunction $\chi_n(\mathbf{R})$ which describes the motion of the nuclei [41]:

$$\Psi(\mathbf{r}, \mathbf{R}) = \Psi_e(\mathbf{r}; \mathbf{R})\chi_n(\mathbf{R}) \tag{2.9}$$

By means of this ansatz, the total Schrödinger equation 2.1 is decomposed into one equation for the electronic wavefunction:

$$\mathbf{H}_e\Psi_e(\mathbf{r}; \mathbf{R}) = E_e(\mathbf{R})\Psi_e(\mathbf{r}; \mathbf{R}) \tag{2.10}$$

where the electronic Hamiltonian is given by:

$$\mathbf{H}_e = \mathbf{H} - \mathbf{T}_n = \mathbf{T}_e + \mathbf{V}_{ee} + \mathbf{V}_{ne} + \mathbf{V}_{nn} \tag{2.11}$$

and one equation for the wavefunction $\chi_n(\mathbf{R})$ describing the motion of the nuclei:

$$(\mathbf{T}_n + E_e(\mathbf{R}))\chi_n(\mathbf{R}) = E\chi_n(\mathbf{R}) \tag{2.12}$$

The independent variables in the electronic wavefunction $\Psi_e(\mathbf{r}; \mathbf{R})$ are the coordinates $\mathbf{r}$ of the electrons. However, $\Psi_e(\mathbf{r}; \mathbf{R})$ depends also parametrically on the coordinates $\mathbf{R}$ of the nuclei, since the electronic Hamiltonian is a function of the positions of the nuclei. This means that the electronic Schrödinger equation 2.10 has to be solved for a given nuclear geometry. As a consequence, the electronic energy, i.e. the eigenvalue of $E_e(\mathbf{R})$ of the electronic Schrödinger equation, is not a constant, but depends on the nuclear geometry. This geometry dependent electronic energy $E_e(\mathbf{R})$ plays the role of the potential energy in the Schrödinger equation 2.12 for the nuclear motion. It is therefore generally called Potential Energy Surface (PES). Thus in the Born-Oppenheimer picture the nuclei move on a PES which is a solution of the electronic Schrödinger equation.

In most cases the Born-Oppenheimer approximation is a very good approximation. Generally it only breaks down when two (or more) solutions of the electronic Schrödinger equation come close together energetically. In this case one has to go beyond the Born-Oppenheimer approximation, by writing the total wavefunction as a linear combination of products of electronic and nuclear wavefunctions. By inserting this total wavefunction in the Schrödinger equation 2.1, one obtains two terms in the resulting equations which couple, via the ∇_a operators, different electronic states. This terms are the so called *non-adiabatic coupling elements* and, as already pointed out, being important only for systems involving more than one electronic surface with comparable energies, they are completely neglected in the framework of the Born-Oppenheimer approximation.

2.3 Requirements for the Electronic wavefunction

In the following, we will stay within the Born-Oppenheimer approximation and will only be concerned with the electronic Schrödinger equation 2.10. For simplicity we will drop the subscript e and the dependence on the nuclear coordinates from $\Psi_e(\mathbf{r};\mathbf{R})$.

A wavefunction used for describing the electronic structure of a molecular system has to satisfy three fundamental requirements:

- Normalization. As for all quantum mechanical wavefunctions describing stationary states we will assume that $\Psi(\mathbf{r})$ is normalized ($\langle\Psi(\mathbf{r})|\Psi(\mathbf{r})\rangle = 1$), which means that the probability of finding the system somewhere in space is unitary.

- Antisymmetry with respect to the permutation of two electrons. Since electrons are fermions, the solutions of the electronic Schrödinger equation, which can be used to describe electronic systems, have to be antisymmetric with respect to a permutation of any two electrons in the system (Symmetry Postulate). Mathematically speaking, only wavefunctions which belong to the totally antisymmetric representation of the permutation group of electrons are allowed to describe N-electron systems. If we denote a permutation of the electrons i and j by the operator $\mathbf{P}_{ij}$, the antisymmetry requirement can be formulated as:

  $$\mathbf{P}_{ij}\Psi(\cdots,i,j,\cdots) = \Psi(\cdots,j,i,\cdots) = -\Psi(\cdots,i,j,\cdots) \tag{2.13}$$

  The Symmetry Postulate implies the Pauli principle which states that two electrons cannot have the same series of all quantum numbers (if $i = j$ in equation 2.13 then $\Psi \equiv 0$).

- Electronic spin. Since the electronic Hamiltonian does not contain any spin operator, it commutes with the operators of the z-component ($\mathbf{S}_z$) and of the square ($\mathbf{S}^2$) of the total electronic spin:

  $$[\mathbf{H},\mathbf{S}_z] = 0, \quad [\mathbf{H},\mathbf{S}^2] = 0 \tag{2.14}$$

  As a consequence, the electronic wavefunction has to be eigenfunction of $\mathbf{S}_z$ and $\mathbf{S}^2$. A rigorous quantum mechanical treatment of the spin angular momentum is only possible by means of the relativistic Dirac equation, in which four-component electron wavefunctions (spinors) have to be used [42]. However, even in a non-relativistic treatment, one has to introduce the electron spin and the common way to do this is by representing the wavefunction of the single electron ($\phi(\mathbf{r},\mathbf{s})$) as a product of a spatial orbital $\phi(\mathbf{r})$ times a spin function $\eta(\mathbf{s})$:

$$\phi(\mathbf{r},\mathbf{s}) = \phi(\mathbf{r})\eta(\mathbf{s}) \tag{2.15}$$

There are only two possible spin eigenfunctions which are commonly denoted as α and β and are orthonormal. The one-electron wavefunctions defined by equation 2.15 are called spin orbitals or sometimes simply orbitals (even if, in a strict sense, the orbital is only the spatial part $\phi(\mathbf{r})$).

2.4 The Slater Determinant

A closer inspection of the electronic Hamiltonian (equation 2.11) shows that it contains one-electron terms, i.e. terms that depend only on the coordinates of one electron, two-electron terms, and the nuclear repulsion $\mathbf{V}_{nn}$ which is independent of the electron coordinates. It thus can be rewritten as:

$$\mathbf{H} = \sum_{i=1}^{N} \mathbf{h}_i + \sum_{i=1}^{N} \sum_{j>i=1}^{N} \mathbf{g}_{ij} + \mathbf{V}_{nn} \tag{2.16}$$

where the $\mathbf{h}_i$ are one-electron operators that describe the motion of a single electron in the field generated by the nuclei and all of the other electrons:

$$\mathbf{h}_i = -\frac{1}{2}\nabla_i^2 - \sum_{a=1}^{M} \frac{Z_a}{r_{ia}} \tag{2.17}$$

while $\mathbf{g}_{ij}$ are two-electron operators describing the Coulomb repulsion between electrons i and j:

$$\mathbf{g}_{ij} = \frac{1}{r_{ij}} \tag{2.18}$$

If these two-electron terms $\mathbf{g}_{ij}$ were absent, $\mathbf{H}$ would be a sum of terms each depending only on the coordinates of one electron ($\mathbf{V}_{nn}$ is just a constant with respect to the electronic coordinates and it does not cause problems). Then $\mathbf{H}$ would be separable: its eigenfunctions would be just products of eigenfunctions of the one-particle Hamiltonians $\mathbf{h}_i$, and its eigenvalues sums of eigenvalues of $\mathbf{h}_i$. Unfortunately, $\mathbf{H}$ contains the two-electron repulsion terms, and so it is far from being separable.

Nevertheless, we start by constructing an N-electron wavefunction as a product of one-electron wavefunctions:

$$\Pi = \phi_1(1)\phi_2(2)\ldots\phi_N(N) \tag{2.19}$$

Here the serial number 1 stands for the space and spin coordinates of electron number 1, i.e. $\phi_1(1)$ is a spin orbital. Equation 2.19 is interpreted by saying that the first electron “occupies” the spin orbital ϕ_1, the second electron the spin orbital ϕ_2, and so on. However, the product in equation 2.19 does not possess the required antisymmetry property, therefore we have to apply to it an anti-symmetrization operator **A**, which generates a fully antisymmetric wavefunction, defined as [41]:

$$\mathbf{A} = \frac{1}{\sqrt{N!}} \sum_{p} (-1)^p \mathbf{P} \tag{2.20}$$

where **P** is the permutation operator which performs all the permutations of $1, 2, 3 \ldots$ electrons and it is defined as:

$$\sum_{p} (-1)^p \mathbf{P} = \mathbf{I} - \sum_{ij} \mathbf{P}_{ij} + \sum_{ijk} \mathbf{P}_{ijk} + \cdots \tag{2.21}$$

where **I** is the identity operator, $\sum_{ij} \mathbf{P}_{ij}$ performs all binary permutations, $\sum_{ijk} \mathbf{P}_{ijk}$ performs all ternary permutations and so on. Since this anti-symmetrization operator is nothing else than an operator generating a determinant out of a simple product, we can write the N-electron wavefunction in the following form [41]:

$$\Psi = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(1) & \phi_2(1) & \ldots & \phi_N(1) \\ \phi_1(2) & \phi_2(2) & \ldots & \phi_N(2) \\ \vdots & \vdots & \vdots & \vdots \\ \phi_1(N) & \phi_2(N) & \ldots & \phi_N(N) \end{vmatrix}$$

This determinant is called Slater determinant and is built in such a way that over the rows we assign to a single spin orbital i all the N electrons, and on the columns we assign electron j to all the N spin orbitals. Moreover it is generally assumed that the spatial orbitals form an orthonormal set and, due to the orthonormality of the spin functions α and β, also the spin orbitals form an orthonormal set as well.

2.5 The Hartree-Fock method

In the Hartree-Fock method the electronic wavefunction is constructed as a Slater determinant from molecular spin orbitals and the optimum spin orbitals are determined by means of the variational method. The variational principle states that any approximate wavefunction Ψ_t, which satisfies the same set of boundary conditions as imposed upon the exact wave function Ψ, has an energy E_t above or equal to the exact energy of the ground state of a system E_0. The equality holds only if the trial wavefunction is the exact solution.

$$E_t = \langle \Psi_t | \mathbf{H} | \Psi_t \rangle \geq E_0 \tag{2.22}$$

Therefore, in the variational method, by making a trial wavefunction containing a certain number of parameters, it is possible to generate the "best" trial function of the given form by minimizing the energy as a function of these parameters. In the present case, the trial wavefunction is the Slater determinant and the parameters are the molecular spin orbitals.

In order to apply the variational method we have first of all to find an expression for the energy of a Slater determinant (E). By representing the electronic wavefunction with the anti-symmetrization operator, we obtain:

$$\begin{aligned} E &= \langle\Psi|\mathbf{H}|\Psi\rangle \\ &= \langle\Pi|\mathbf{A}^{\dagger}\mathbf{H}\mathbf{A}|\Pi\rangle \\ &= \langle\Pi|\mathbf{H}\mathbf{A}\mathbf{A}|\Pi\rangle \\ &= \sqrt{N!}\langle\Pi|\mathbf{H}\mathbf{A}|\Pi\rangle \\ &= \sum_p (-1)^p\langle\Pi|\mathbf{H}\mathbf{P}|\Pi\rangle \end{aligned}$$

where we have used the properties of the $\mathbf{A}$ operator, which commutes with $\mathbf{H}$ and acting twice gives the same as $\mathbf{A}$ acting once, multiplied by $\sqrt{N!}$. By using equation 2.16 for the electronic Hamiltonian we can write:

$$E = \sum_p (-1)^p \sum_{i=1}^{N}\langle\Pi|\mathbf{h}_i\mathbf{P}|\Pi\rangle + \sum_p (-1)^p \sum_{i=1}^{N}\sum_{j>i=1}^{N}\langle\Pi|\mathbf{g}_{ij}\mathbf{P}|\Pi\rangle + \mathbf{V}_{nn} \tag{2.23}$$

Since $\mathbf{h}_i$ depends only on the coordinates of the electron i, all of the terms in the first summation related to spin orbitals containing the other electrons can be factored out of the integral. Thus, all of the products containing one or more permutations produce integrals between orthogonal functions and go to zero. For instance exchanging the coordinates of electrons 1 and 2 we obtain:

$$\begin{aligned} \langle\phi_1(1)\phi_2(2)\cdots\phi_N(N)|\mathbf{h}_1|\phi_2(1)\phi_1(2)\cdots\phi_N(N)\rangle &= \\ \langle\phi_1(1)|\mathbf{h}_1|\phi_2(1)\rangle\langle\phi_2(2)|\phi_1(2)\rangle\ldots\langle\phi_N(N)|\phi_N(N)\rangle &= 0 \end{aligned} \tag{2.24}$$

As far as the two-electron operator $\mathbf{g}_{ij}$ is concerned, only the identity operator or binary permutations can give a non zero contribution to the energy. From the identity operator we obtain for $\mathbf{g}_{12}$:

$$\begin{aligned} \langle\Pi|\mathbf{g}_{12}|\Pi\rangle = &\ \langle\phi_1(1)\phi_2(2)\cdots\phi_N(N)|\mathbf{g}_{12}|\phi_1(1)\phi_2(2)\cdots\phi_N(N)\rangle = \\ = &\ \langle\phi_1(1)\phi_2(2)|\mathbf{g}_{12}|\phi_1(1)\phi_2(2)\rangle\ldots\langle\phi_N(N)|\phi_N(N)\rangle = \\ &= \langle\phi_1(1)\phi_2(2)|\mathbf{g}_{12}|\phi_1(1)\phi_2(2)\rangle = J_{12} \end{aligned}$$

where J_{12} is the Coulomb integral and represents the classical repulsion between the two $|\phi_1(1)|^2$ e $|\phi_2(2)|^2$ charge distributions. On the other hand, exchanging the coordinates of electrons 1 and 2 by means of the $\mathbf{P}_{12}$ operator we can write:

$$\begin{aligned} \langle\Pi|\mathbf{g}_{12}|\mathbf{P}_{12}\Pi\rangle = &\ \langle\phi_1(1)\phi_2(2)\cdots\phi_N(N)|\mathbf{g}_{12}|\phi_2(1)\phi_1(2)\cdots\phi_N(n)\rangle = \\ = &\ \langle\phi_1(1)\phi_2(2)|\mathbf{g}_{12}|\phi_2(1)\phi_1(2)\rangle\ldots\langle\phi_N(N)|\phi_N(N)\rangle = \\ &= \langle\phi_1(1)\phi_2(2)|\mathbf{g}_{12}|\phi_2(1)\phi_1(2)\rangle = K_{12} \end{aligned}$$

K_{12} is the Exchange integral and does not have a simple classical interpretation as the Coulomb integral. By defining the Coulomb operator $\mathbf{J}_i$ and the Exchange operator $\mathbf{K}_i$ as:

$$\begin{aligned}
\mathbf{J}_i|\phi_j(2)\rangle &= \langle\phi_i(1)|\mathbf{g}_{12}|\phi_i(1)\rangle|\phi_j(2)\rangle &\quad (2.25)\\
\mathbf{K}_i|\phi_j(2)\rangle &= \langle\phi_i(1)|\mathbf{g}_{12}|\phi_j(1)\rangle|\phi_i(2)\rangle &\quad (2.26)
\end{aligned}$$

it is now possible to rewrite the energy as:

$$E = \sum_{i=1}^{N}\langle\phi_i|\mathbf{h}_i|\phi_i\rangle + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}(\langle\phi_j|\mathbf{J}_i|\phi_j\rangle - \langle\phi_j|\mathbf{K}_i|\phi_j\rangle) + V_{nn} \tag{2.27}$$

The aim of the Hartree-Fock method is to determine the set of spin orbitals which makes the energy of equation 2.27 a minimum. The minimization however has to be performed in such a way that the spin orbitals remain orthonormal and this constrained optimization can be carried out by means of the Lagrange multipliers. By defining the Lagrange functional as:

$$L = E - \sum_{i,j=1}^{N}\lambda_{ij}\gamma_{ij} = E - \sum_{ij=1}^{N}\lambda_{ij}(\langle\phi_i|\phi_j\rangle - \delta_{ij}) \tag{2.28}$$

where the constant parameters λ_{ij} are the Lagrange multipliers and γ_{ij} are the constraints, i.e. the orthonormality conditions, we impose that the Lagrange function is stationary with respect to a spin orbital variation (calling for the sake of simplicity the generic spin orbital $|i\rangle$:

$$\begin{aligned}
\delta L &= \delta E - \sum_{i,j=1}^{N}\lambda_{ij}(\langle\delta i|j\rangle + \langle i|\delta j\rangle) = 0 &\quad (2.29)\\
\delta E &= \sum_{i=1}^{N}(\langle\delta i|\mathbf{h}_i|i\rangle + \langle i|\mathbf{h}_i|\delta i\rangle + \\
&+ \frac{1}{2}\sum_{i,j=1}^{N}(\langle j|\mathbf{J_i} - \mathbf{K_i}|\delta j\rangle + \langle\delta j|\mathbf{J_i} - \mathbf{K_i}|j\rangle) &\quad (2.30)\\
&+ \frac{1}{2}\sum_{i,j=1}^{N}(\langle i|\mathbf{J_j} - \mathbf{K_j}|\delta i\rangle + \langle\delta i|\mathbf{J_j} - \mathbf{K_j}|i\rangle)
\end{aligned}$$

By summing up equal terms and defining the Fock operator as $\mathbf{F}_i = \mathbf{h}_i + \sum_j(\mathbf{J}_j - \mathbf{K}_j)$, we obtain:

$$\delta L = \sum_{i=1}^{N}[\langle i|\mathbf{F}_i|\delta i\rangle + \langle\delta i|\mathbf{F}_i|i\rangle] - \sum_{i,j=1}^{N}\lambda_{ij}(\langle\delta i|j\rangle + \langle i|\delta j\rangle) \tag{2.31}$$

and being $\mathbf{F}_i$ an Hermitian operator:

$$\delta L = \sum_{i=1}^{N} \langle \delta i | \mathbf{F}_i | i \rangle - \sum_{ij} \lambda_{ij} \langle \delta i | j \rangle + \sum_{i} \langle \delta i | \mathbf{F}_i | i \rangle^* - \sum_{ij} \lambda_{ij} \langle \delta j | i \rangle^* \tag{2.32}$$

The summations contain all the possible terms $\lambda_{ij} \langle \delta i | j \rangle$, so that we can exchange the index orders i and j in the fourth sum:

$$\delta L = \sum_{i=1}^{N} \langle \delta i | \mathbf{F}_i | i \rangle - \sum_{i,j=1}^{N} \lambda_{ij} \langle \delta i | j \rangle + \sum_{i} \langle \delta i | \mathbf{F}_i | i \rangle^* - \sum_{i,j=1}^{N} \lambda_{ji} \langle \delta i | j \rangle^* \tag{2.33}$$

The third and fourth terms in the equation 2.33 are the conjugate complex of the first and second term, respectively. Since the variations of $|i\rangle$ and $|i\rangle^*$ are independent, the terms in the equation 2.33 concerning $|\delta i\rangle$ and $|\delta i\rangle^*$ have to be zero separately:

$$\begin{aligned} \mathbf{F}_i |i\rangle &= \sum_{j=1}^{N} \lambda_{ij} |j\rangle \\ \mathbf{F}_i |i\rangle^* &= \sum_{j=1}^{N} \lambda_{ji} |j\rangle^* \end{aligned} \tag{2.34}$$

The Λ matrix with elements λ_{ij} is Hermitian, and if we diagonalize it by means of a unitary transformation (so that $\lambda_{ij} = \epsilon_i \delta_{ij}$), we obtain the Hartree-Fock equations [41]:

$$\mathbf{F}_i |i\rangle' = \epsilon_i |i\rangle' \tag{2.35}$$

where $|i\rangle'$ are the transformed spin orbitals and they are called the canonical orbitals. The Hartree-Fock equations are a set of pseudo eigenvalue equations: the Fock operator $\mathbf{F}_i$ in fact depends, via the Coulomb and Exchange operators, on the spin orbitals. Therefore, to solve the Hartree-Fock equations an iterative procedure has to be used. The Fock operator $\mathbf{F}_i$ is a monoelectronic operator which incorporates the kinetic energy of the electron i, the Coulombic interaction with the nuclei and the average repulsion with the other electrons. As a consequence, the Hartree-Fock method is a mean field approximation based on an independent particles model in which each electron moves in an effective potential due to the nuclear attraction and to the average repulsion generated by the other electrons [42]. The motion of electrons is thus completely uncorrelated. It is noteworthy to observe that the exact solution of the Schrödinger equation for an interacting particle systems (the electrons in our case) can by no means be a product, even if suitably symmetrized, of one-particle wavefunctions. However, the representation of the electronic wavefunction as a Slater determinant offer an easy interpretation, since nothing seems more natural than the idea that a certain electron is attributed to a certain one-electron function or, stated differently, that an electron is occupying a certain spin orbital.

Slater determinants are eigenfunctions of the N-electron spin operator $\mathbf{S}_z$ with an eigenvalue M_S equal to one-half of the difference between the number of α and β electrons:

$$\mathbf{S}_z\Psi = M_S\Psi \tag{2.36}$$

$$M_S = \frac{1}{2}(N_\alpha - N_\beta) \tag{2.37}$$

However, except for certain cases, Slater determinants are not eigenfunctions of the N-electron spin operator $\mathbf{S}^2$. One exception is the situation of a closed shell system, in which there is an even number of electrons $N = 2n$ and exactly n spatial orbitals are doubly occupied. Fortunately, the stable ground states of most molecules belong to this case. Whenever a single Slater determinant is not an eigenfunction of $\mathbf{S}^2$, one can construct eigenfunctions of this operator as linear combinations of Slater determinants, in which the same spatial orbitals are occupied but with different spin orientations. Such linear combinations are called configuration state functions (CSF).

Although the Hartree-Fock theory has been formulated in the early 1930ths, for a long time numerical calculations could only be performed for atoms. The reason was that numerical solutions of equation 2.35 were only possible for small highly symmetric systems, like atoms and diatomic molecules. The possibility of using the Hartree-Fock method for molecular calculations came in 1951 when Hall and Roothaan independently proposed to expand the Hartree-Fock orbitals into a set of atom centered basis functions [41]. We will return to a detailed description of basis sets at the end of this chapter, but for now we simply assume that the unknown Hartree-Fock spin orbital ϕ_i is expanded into a finite set of m known functions χ_ν:

$$\phi_i = \sum_{\nu=1}^{m} c_{\nu i}\chi_\nu \tag{2.38}$$

Since the basis functions are known, the spin orbital ϕ_i is completely determined if the expansion coefficients $c_{\nu i}$ are calculated. Thus the Hartree-Fock equations 2.35 may be written as:

$$\mathbf{F}\sum_{\nu=1}^{m} c_{\nu i}\chi_\nu = \epsilon_i \sum_{\nu=1}^{m} c_{\nu i}\chi_\nu \tag{2.39}$$

and by premultiplying for the generic basis function $\langle\chi_\mu|$ and integrating we obtain the Roothaan-Hall equations [40]:

$$\sum_{\nu=1}^{m}(F_{\mu\nu} - \epsilon_i S_{\mu\nu})c_{\nu i} = 0 \qquad \mu = 1, \ldots m \tag{2.40}$$

where:

$$F_{\mu\nu} = \langle \chi_\mu \left| \mathbf{F} \right| \chi_\nu \rangle \tag{2.41}$$

$$S_{\mu\nu} = \langle \chi_\mu \left| \chi_\nu \right. \rangle \tag{2.42}$$

All the m equations may be collected in a matrix notation:

$$\mathbf{FC} = \mathbf{SC}\epsilon \tag{2.43}$$

The **S** matrix contains the overlap elements between basis functions, while the **F** matrix contains the Fock matrix elements which consist of two parts, i.e. integrals involving the one-electron operators, and a sum over the occupied spin orbitals of coefficients multiplied with two-electron integrals involving the electron-electron repulsion operator. Obviously, also in this case the problem requires an iterative procedure to be used, since the **F** matrix depend on the solutions (the expansion coefficients). This procedure starts from an initial guess of the coefficients, then the **F** matrix is formed and diagonalized. The new set of coefficients is used for calculating a new **F** matrix and so on until the new set of coefficients used for constructing the **F** matrix is equal, within a certain threshold, to those resulting from the diagonalization and self-consistency is reached. Due to the iterative procedure, this method is called Self Consistent Field (SCF) method.

The Roothaan Hall equations were derived for closed shell systems, i.e. for systems in a singlet state with a total spin of zero. In such systems each spatial orbital is doubly occupied by two electrons with different spin functions and the resulting wavefunction is called a Restricted Hartree-Fock (RHF) wavefunction. For open-shell systems, i.e. systems in a higher spin state with unpaired electrons, the RHF description is unsuitable and other treatments are needed. If there are no restrictions on the form of the spatial orbitals, and different spatial orbitals are assigned to α and β electrons, the wavefunction is called Unrestricted Hartree-Fock (UHF) wavefunction. On the other hand, open shell systems can also be described by imposing that the maximum number of spatial orbitals suitable for the spin requirements is doubly occupied (Restricted Open-shell Hartree-Fock ROHF).

2.6 Post-SCF Methods

The Hartree-Fock method provides only an approximation to the exact solution of the electronic Schrödinger equation. The difference between the energy obtained by the Hartree-Fock method with an infinite basis set (Hartree-Fock limit) E_{HF} and the exact energy E_0 (always in a non-relativistic treatment) is called the electronic correlation energy E_{corr} [41]:

$$E_{corr} = E_0 - E_{HF} \tag{2.44}$$

Physically the correlation energy is due to the fact that the motion of electrons is correlated, while in the Hartree-Fock method the instantaneous electron-electron interaction is replaced by an average interaction. On average the electron are further apart than how the Hartree-Fock theory describes. The correlation between electrons with the same spin is called the Fermi correlation and it is accounted for to some extent by using an antisymmetric wavefunction which satisfies the Pauli principle. On the other hand, the correlation between electrons with opposite spin (Coulomb correlation) is completely neglected in the Hartree-Fock method. The methods that partially calculate the correlation energy and thus go beyond the Hartree-Fock theory are called post-SCF methods. Since the Hartree-Fock theory usually gives 99% of the correct energy, post-SCF methods generally use the Hartree-Fock wave function as a starting point for subsequent improvements.

2.6.1 Configuration Interaction

The conceptually simplest method for accounting for correlation effects is the method of Configuration Interaction (CI). Due to the electron interactions, the full solution of the Schrödinger equation cannot be expressed in terms of a single electron configuration, i.e. a unique assignment of electrons to orbitals, and it is necessary to use wavefunctions that represent more than a single electron configuration. As a consequence, in the CI methods, starting from the Slater determinant of the Hartree-Fock theory, a series of additional excited Slater determinants are generated by replacing spin orbitals that are occupied in the Hartree-Fock determinant by spin orbitals that are unoccupied. Consider a system comprising N electrons described at the Hartree-Fock level with a basis set of m ($m \geq N$) basis functions. The solution of the Hartree-Fock equations will provide N occupied spin orbitals $(\phi_i, \phi_j, \phi_k, \ldots)$ which contribute to the Hartree-Fock wavefunction, and $m - N$ unoccupied (or virtual) orbitals $(\phi_a, \phi_b, \phi_c, \ldots)$. Excited Slater determinants are built by replacing one, two, … N occupied orbitals in the Hartree-Fock wavefunction by one, two, … N virtual orbitals. This corresponds to excite electrons from orbitals at lower energies to orbitals at higher energies. In the Full-CI method all the possible excitations are included and the wavefunction is written as a linear combination of the Hartree-Fock determinant (we call it now Ψ_0) and all of the possible excited determinants $\{\Psi_s\}$ [41]:

$$\Psi = a_0\Psi_0 + \sum_{s>0} a_s\Psi_s \tag{2.45}$$

where the summation runs over all the possible excitations and the $\{\Psi_s\}$ are orthonormal. The expansion coefficients a_s can be determined, in an analogous way to the HF case, with the Lagrange method, obtaining a system of secular equations:

$$\sum_s (H_{ts} - E_i\delta_{ts})a_{si} = 0 \qquad t = 1, 2 \ldots \tag{2.46}$$

where $H_{ts} = \langle\Psi_t| H |\Psi_s\rangle$ are the elements of the CI matrix and E_s is the energy of the electronic state s. The variational problem is thus transformed into solving a set of CI secular equations that can be written also in a matrix notation as:

$$(\mathbf{H} - E\mathbf{I})\mathbf{a} = 0 \tag{2.47}$$

The Full-CI method is, in principle, an exact method. This means that if an infinite basis set was used, Full-CI would generate the exact solution of the electronic Schrödinger equation (always within the Born-Oppenheimer approximation and in a non-relativistic treatment). It is possible to demonstrate that the CI matrix is a highly sparse matrix, in which non-zero elements are found only between determinants which differ by zero, one or two spin orbitals. Nevertheless, the number of excited determinants grows factorially with the size of the basis set, and this makes the Full-CI method unfeasible for all but the very smallest systems. As a consequence, in practical applications truncated CI expansions are used which include only single excitations (CI with Singles or CIS), only double excitations (CI with Doubles or CID), single and double excitations (CI with Singles and Doubles or CISD), single, double and triple excitations (CI with Singles, Doubles and Triples or CISDT), and so on.

2.6.2 Møller Plesset perturbation theory

A different approach to the CI methods to deal with the correlation energy is provided by the perturbation theory. Suppose that we want to solve the electronic Schrödinger equation:

$$\mathbf{H}\Psi_k = E_k\Psi_k \tag{2.48}$$

The perturbation theory is based upon dividing the Hamiltonian of the system under consideration ($\mathbf{H}$) in two parts:

$$\mathbf{H} = \mathbf{H}_0 + \lambda\mathbf{V} \tag{2.49}$$

where λ is a parameter which measures the entity of the perturbation, $\mathbf{V}$ is a perturbation operator and $\mathbf{H}_0$ is the Hamiltonian of a problem whose solutions are known (we call them $\Psi_k^{(0)}$ and suppose that they form a complete and orthonormal set and that their corresponding energies $E_k^{(0)}$ are non-degenerate). If $\mathbf{V}$ is a small perturbation as compared to the unperturbed Hamiltonian $\mathbf{H}_0$, we can expand the energy and wavefunction of the perturbed system in powers of λ:

$$\Psi_k = \Psi_k^{(0)} + \lambda\Psi_k^{(1)} + \lambda^2\Psi_k^{(2)} + \lambda^3\Psi_k^{(3)} + \ldots \tag{2.50}$$

$$E_k = E_k^{(0)} + \lambda E_k^{(1)} + \lambda^2 E_k^{(2)} + \lambda^3 E_k^{(3)} + \ldots \tag{2.51}$$

By substituting these expansions in the Schrödinger equation 2.48, it is possible to obtain corrections to the unperturbed energy and wavefunction at any order of approximation. In particular it is possible to demonstrate that the first ($E_k^{(1)}$) and second order ($E_k^{(2)}$) corrections to the energy are:

$$
\begin{aligned}
E_k^{(1)} &= \langle \Psi_k^{(0)} | \mathbf{V} | \Psi_k^{(0)} \rangle && (2.52) \\
E_k^{(2)} &= \sum_t \frac{\langle \Psi_k^{(0)} | \mathbf{V} | \Psi_t^{(0)} \rangle \langle \Psi_t^{(0)} | \mathbf{V} | \Psi_k^{(0)} \rangle}{E_k^{(0)} - E_t^{(0)}} \\
&= -\sum_t \frac{|\langle \Psi_k^{(0)} | \mathbf{V} | \Psi_t^{(0)} \rangle|^2}{E_t^{(0)} - E_k^{(0)}} && (2.53)
\end{aligned}
$$

The first order correction is just the matrix element of the perturbation $\mathbf{V}$ on the unperturbed wavefunction $\Psi_k^{(0)}$, while the second order correction contains matrix elements of $\mathbf{V}$ among $\Psi_k^{(0)}$ and all of the other unperturbed states $\{\Psi_{t \neq k}^{(0)}\}$. The weight of these contributions to the second order correction is inversely proportional to the energy difference between the two states ($E_t^{(0)} - E_k^{(0)}$).

The perturbation theory is absolutely general and can be applied to any kind of system. In order to apply it to molecular systems for calculating to some extent the correlation energy, the unperturbed Hamiltonian operator $\mathbf{H}_0$ is taken as a sum over the Fock operators:

$$\mathbf{H}_0 = \sum_{j=1}^{N} \mathbf{F}_j \tag{2.54}$$

where N is, as usual, the number of electrons. In this particular case the perturbation theory is called Møller-Plesset perturbation theory [41]. The eigenfunctions of $\mathbf{H}_0$ are the Hartree-Fock determinant, which is the ground state wavefunction, and all of the substituted determinants that represent the excited states. Remembering that the electronic Hamiltonian for a molecular system, by neglecting the term $\mathbf{V}_{nn}$ which is just a constant, is given by:

$$\mathbf{H} = \sum_{i=1}^{N} \mathbf{h}_i + \sum_{i=1}^{N} \sum_{j>i=1}^{N} \mathbf{g}_{ij} \tag{2.55}$$

we obtain that the perturbation operator $\mathbf{V}$ in this specific case is given as:

$$\mathbf{V} = \sum_{i=1}^{N} \sum_{j>i=1}^{N} \mathbf{g}_{ij} - \sum_{i=1}^{N} \sum_{j=1}^{N} (\mathbf{J}_j(i) - \mathbf{K}_j(i)) \tag{2.56}$$

By inserting this form of $\mathbf{V}$ in the equation 2.52 we can calculate the first order correction $E_0^{(1)}$, and if we sum up $E_0^{(0)}$ and $E_0^{(1)}$ we reobtain the Hartree-Fock energy E^{HF}:

$$E_0^{(0)} + E_0^{(1)} = \langle \Psi_0^{(0)} | \mathbf{H}_0 | \Psi_0^{(0)} \rangle + \langle \Psi_0^{(0)} | \mathbf{V} | \Psi_0^{(0)} \rangle = \langle \Psi_0^{(0)} | \mathbf{H} | \Psi_0^{(0)} \rangle = E^{HF} \tag{2.57}$$

Therefore, the first correction to the energy due to the correlation is the second order correction, which can be obtained from equation 2.53 by rewriting it for the ground state as:

$$E_0^{(2)} = \sum_t \frac{\langle \Psi_0^{(0)} | \mathbf{V} | \Psi_t^{(0)} \rangle \langle \Psi_t^{(0)} | \mathbf{V} | \Psi_0^{(0)} \rangle}{E_0^{(0)} - E_t^{(0)}} = - \sum_t \frac{|\langle \Psi_0^{(0)} | \mathbf{V} | \Psi_t^{(0)} \rangle|^2}{E_t^{(0)} - E_0^{(0)}} \tag{2.58}$$

The terms in the summation contain matrix elements of the perturbation between the ground state wavefunction and excited wavefunctions. Since $\mathbf{V}$ is a two electron operator, all the integrals between the Hartree-Fock wavefunction and wavefunctions with triple and higher excitations go to zero; it can be demonstrated that also matrix elements between $\Psi_0^{(0)}$ and wavefunction with single excitations are zero. Thus the only non-zero elements contain double excited wavefunctions denoted with Ψ_{ij}^{ab} (the notation indicates the promotion of electrons i and j to the virtual orbitals a and b), and the second order correction to the energy may be written as [41]:

$$\begin{aligned} E_0^{(2)} &= \sum_{i,j<i}^{occ} \sum_{a,b<a}^{vir} \frac{\langle \Psi_0^{(0)} | \mathbf{V} | \Psi_{ij}^{ab} \rangle \langle \Psi_{ij}^{ab} | \mathbf{V} | \Psi_0^{(0)} \rangle}{E_0^{(0)} - E_{ij}^{ab}} \\ &= \sum_{i,j<i}^{occ} \sum_{a,b<a}^{vir} \frac{|\langle \phi_i \phi_j | \mathbf{g}_{12} | \phi_a \phi_b \rangle - \langle \phi_i \phi_j | \mathbf{g}_{12} | \phi_b \phi_a \rangle|^2}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b} \end{aligned} \tag{2.59}$$

where the first and second summations run over all occupied and virtual orbitals, respectively; E_{ij}^{ab} is the energy of the Ψ_{ij}^{ab} determinant, and the i and j subscripts denote occupied orbitals, while a and b refer to virtual orbitals. The $\mathbf{g}_{12}$ notation indicates integrations over all coordinates for electron 1 and 2 in the occupied and virtual orbitals. Since the numerator is always positive and the denominator is always negative, the second order correction is always negative. The Møller-Plesset perturbation theory method which uses the second order correction to the energy is called MP2 and it is usually a good method to include electron correlation, as it can account for 80% - 90% of it with a computational cost slightly higher than a Hartree-Fock calculation. Higher order corrections to the energy can be calculated with a similar procedure. It is noteworthy to observe than Møller-Plesset method may underestimate the energy since it is not a variational method.

2.7 Density Functional Theory

2.7.1 The Hohenberg-Kohn Theorems

An alternative approach to the Hartree-Fock methods (or post-SCF techniques) to solve approximately the Schrödinger equation is provided by Density Functional Theory (DFT). The basis for DFT is the proof by Hohenberg and Kohn that the ground state electronic energy of a system is completely determined by its electron density $\rho(\mathbf{r})$. In fact, the first Hohenberg-Kohn theorem [43] states that: “For any system of interacting particles in an external potential $V(\mathbf{r})$, the external potential is determined uniquely, except for a constant, by the ground state particle density.” For a system of M nuclei and N electrons the electronic Hamiltonian operator is:

$$\mathbf{H} = \mathbf{T}_e + \mathbf{V}_{ne} + \mathbf{V}_{ee} + \mathbf{V}_{nn} \tag{2.60}$$

(the last term in the Born-Oppenheimer approximation is a constant and can be neglected) and the external potential is the potential generated by the nuclei, $\mathbf{V}_{ne}$. It is seen that the Hamiltonian operator is uniquely determined by the number of electrons and by $\mathbf{V}_{ne}$, and so, according to the first Hohenberg-Kohn theorem, by the electron density. This means that there exists a one-to-one correspondence between the electron density of a system and its energy or, in other words, that the energy is a unique functional of the electron density, $E[\rho]$, and can be written as:

$$E[\rho] = T_e[\rho] + E_{ne}[\rho] + E_{ee}[\rho] = F_{HK}[\rho] + E_{ne}[\rho] \tag{2.61}$$

where the Hohenberg-Kohn functional has been defined as:

$$F_{HK}[\rho] = T_e[\rho] + E_{ee}[\rho] \tag{2.62}$$

In the DFT the complicated N-electron wavefunction used in wavefunction based methods (Hartree-Fock or post-SCF) is thus replaced by the much simpler electron density $\rho(\mathbf{r})$.

Using the electron density as a parameter, the second Hohenberg-Kohn theorem [43] provides a variational principle analogous to that in wave mechanics. Given an approximate electron density ρ' (assumed to be positive definite everywhere) which integrates to the number of electrons N, the energy given by this density is an upper bound to the exact ground state energy (provided that the exact functional is used):

$$E_0[\rho'] \geq E_0[\rho] \tag{2.63}$$

2.7.2 The Kohn-Sham approach

The challenge of DFT is to give an explicit expression to the Hohenberg-Kohn functional $F_{HK}[\rho]$. With reference to the Hartree-Fock theory, the $E_{ee}[\rho]$ term in the $F_{HK}[\rho]$ functional may be divided into a Coulomb and Exchange part, $J[\rho]$ and $K[\rho]$, so that the energy functional becomes:

$$E[\rho] = T_e[\rho] + E_{ne}[\rho] + J[\rho] + K[\rho] \tag{2.64}$$

The $E_{ne}[\rho]$ and $J[\rho]$ functionals are given by their classical expressions:

$$E_{ne}[\rho] = \sum_{a=1}^{M} \int \frac{Z_a \rho(\mathbf{r})}{|\mathbf{R}_a - \mathbf{r}|} d\mathbf{r} \tag{2.65}$$

$$J[\rho] = \frac{1}{2}\int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}' \tag{2.66}$$

$T_e[\rho]$ and $K[\rho]$ are unknown. Early attempts at deducing functionals for the kinetic and exchange energies considered a non-interacting uniform gas (Thomas-Fermi theory [44, 45]) but resulted in a very poor representation for the kinetic energy.

The foundation for the use of DFT methods in computational chemistry has been the introduction of orbitals by Kohn and Sham. The basic idea in the Kohn and Sham theory is splitting the kinetic energy functional into two parts, one that can be calculated exactly, plus a small correction term. To this end, Kohn and Sham introduce in their formalism an hypothetical reference system of non-interacting electrons having the same electron density of the real system. For this non-interacting system the Hamiltonian is separable:

$$\mathbf{H} = \sum_{i=1}^{N} \mathbf{h}_i = \sum_{i=1}^{N}(-\frac{1}{2}\nabla_i^2) + \sum_{i=1}^{N} \mathbf{v_s}(\mathbf{r}_i) \tag{2.67}$$

where $\mathbf{h}_i$ are single particle hamiltonians and $\mathbf{v_s}(\mathbf{r}_i)$ are external potentials adjusted so that the same density is obtained for the reference and the real systems. The exact solution of the Schrödinger equation (for the non-interacting system) is given by a Slater determinant composed of molecular orbitals $\{\phi_i\}$ which solve the N one-particle eigenvalue equations:

$$\mathbf{h}_i\phi_i = \epsilon_i\phi_i \tag{2.68}$$

The exact kinetic energy functional is given by:

$$T_s = \sum_{i=1}^{N} \langle\phi_i|(-\frac{1}{2}\nabla_i^2)|\phi_i\rangle \tag{2.69}$$

The key to the Kohn-Sham theory is thus the calculation of the kinetic energy under the assumption of non-interacting electrons. In reality the electrons are interacting and equation 2.69 does not provide the total kinetic energy. The remaining small correction term to the kinetic energy is absorbed into an exchange-correlation term and the total energy of the real interacting system is written as [46]:

$$E_{DFT}[\rho] = T_s[\rho] + E_{ne}[\rho] + J[\rho] + E_{xc}[\rho] \tag{2.70}$$

where $E_{xc}[\rho]$ is given by:

$$E_{xc}[\rho] = (T_e[\rho] - T_s[\rho]) + (E_{ee}[\rho] - J[\rho]) \tag{2.71}$$

$E_{xc}[\rho]$ contains the difference between $T_e[\rho]$ and $T_s[\rho]$, which may be considered as the kinetic correlation energy, and the non-classical part of $E_{ee}[\rho]$, that contains both exchange and potential correlation energy.

The major problem in DFT is deriving suitable formulas for the exchange-correlation term. If this functional was available, the problem would be, as already encountered in the Hartree-Fock theory, to determine a set of orthogonal orbitals which minimize the energy. The orbital orthogonality constraint may be enforced by the Lagrange method (in analogy with the Hartree-Fock):

$$L[\rho] = E_{DFT}[\rho] - \sum_{i,j=1}^{N} \lambda_{ij}(\langle\phi_i|\phi_j\rangle - \delta_{ij}) \tag{2.72}$$

Requiring the variation of $L[\rho]$ to vanish provides a set of equations involving an effective one-electron operator ($\mathbf{h}_{KS}$), similar to the Fock operator in wave mechanics:

$$\mathbf{h}_{KS}\phi_i = \sum_{j=1}^{N} \lambda_{ij}\phi_j \tag{2.73}$$

where:

$$\mathbf{h}_{KS} = -\frac{1}{2}\nabla^2 + \mathbf{V}_{eff} \tag{2.74}$$

$\mathbf{V}_{eff}$ is given by:

$$\mathbf{V}_{eff}(\mathbf{r}) = \mathbf{V}_{ne}(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + \mathbf{V}_{xc}(\mathbf{r}) \tag{2.75}$$

We can choose a unitary transformation which makes the matrix of Lagrange multipliers diagonal, obtaining a set of canonical Kohn-Sham orbitals. The resulting pseudo-eigenvalue equations are the Kohn-Sham equations [46]:

$$\mathbf{h}_{KS}\phi_i = \epsilon_i\phi_i \tag{2.76}$$

Since the $\mathbf{h}_{KS}$ operator depends on the total density, the determination of the orbitals involves an iterative procedure. Moreover, in practical applications, the unknown Kohn-Sham orbitals are expanded in a set of basis functions, analogously to the Hartree-Fock method.

Although it is clear that there are many similarities between the wave mechanics Hartree-Fock theory and DFT, there is an important difference. If the exact $E_{xc}[\rho]$ was known, DFT would provide the exact total energy including electron correlation. Since this is not the case, it is crucial to any application of DFT the approximation of the unknown exchange and correlation potential. Many different functionals have been proposed in the literature, using for instance Local Density Approximation (LDA), where the functional depends only on the density at the coordinate where the functional is evaluated, Generalized Gradient Approximation (GGA), which includes the electron density gradient, as well as hybrid functionals, such as B3LYP, which include the exact exchange energy calculated by Hartree-Fock theory [41].

2.8 Basis Sets

One of the approximations inherent in essentially all electronic structure methods is the introduction of a basis set in which the molecular orbitals are expanded:

$$\phi_i = \sum_{\nu=1}^{m} c_{\nu i} \chi_\nu \tag{2.77}$$

where the χ_ν are the m basis functions and the constants $c_{\nu i}$ are the expansion coefficients. Expanding an unknown function, such as a molecular orbital, in a set of known functions is not an approximation if the basis set is complete. However, in this case a complete basis set means that an infinite number of functions must be used, which is impossible in actual calculations. An unknown molecular orbital can be thought of as a function in the infinite coordinate system spanned by the complete basis set. When a finite basis set is used, only the components of the molecular orbital along the coordinate axes corresponding to the selected basis can be represented. Obviously, the smaller is the basis, the poorer will be the representation. The two types of basis functions commonly used in electronic structure methods are Gaussian functions and plane waves.

2.8.1 Gaussian basis sets

One of the very few examples of an exactly solvable quantum-mechanical problem is the hydrogenoid atom, i.e. the system formed by a nucleus of charge Z and only one electron (H, He^+, Li^{2+}, . . .), and all the atomic-centered basis sets in quantum chemistry try to resemble the particular solution of this problem. The electronic eigenfunctions of the Hamiltonian operator for an hydrogenoid atom are given by [42]:

$$\Psi_{n,l,m}(r,\theta,\varphi) = C_{n,l} e^{\frac{-Zr}{n}} \Big(\frac{Zr}{n}\Big)^l L^{2l+1}_{n-l-1} Y_{l,m}(\theta,\varphi) \tag{2.78}$$

where $C_{n,l}$ is a constant, n, l and m are quantum numbers, L^{2l+1}_{n-l-1} is a generalized Laguerre polynomial which is of $n - l - 1$ order and the function $Y_{l,m}(\theta,\varphi)$ is a spherical harmonic. The idea that the one-electron solutions of the molecular problem in many-electron atoms are not so different from the hydrogenoid atom solutions, together with the chemical intuition that atoms-in-molecules are not very different from isolated atoms, is what mainly drives the choice of functions similar to this $\Psi_{n,l,m}(r,\theta,\varphi)$ in the basis sets. All the differences between the "real" spin orbitals and the hydrogenoid orbitals will be accounted for by the variational procedure aim at determining the expansion coefficients in equation 2.77.

In this context, the first type of functions centered on the atomic nuclei that has been used is the Slater-type orbital (STO) [40]:

$$\chi_{\zeta_a,n_a,l_a,m_a,}(r,\theta,\varphi) = N^{STO} r^{n-1} e^{-\zeta_a r} Y_{l_a,m_a}(\theta,\phi) \tag{2.79}$$

where N^{STO} is a normalization constant, ζ_a is an adjustable parameter on which the orbital spatial extent depends, $Y_{l_a,m_a}(\theta,\phi)$ is a spherical harmonic. The integers l_a and m_a can be considered quantum numbers, since, due to the fact that the only angular dependence is in $Y_{l_a,m_a}(\theta,\phi)$, the STO defined above is still a simultaneous eigenstate of the one-electron angular momentum operators $\mathbf{l}^2$ and $\mathbf{l}_z$ (with the origin placed at nucleus a). Conversely, the parameter n_a should be regarded as a principal quantum number only by analogy, since, on the one hand, it does not exist a monoatomic Hamiltonian whose exact eigenfunctions it could label and, on the other hand, only the leading term of the polynomial in equation 2.78 has been kept in the STO. The STO basis sets have good physical properties, such as that they have a cusp for $r = 0$, and they decay at an exponential rate when $r \to \infty$, as the hydrogenoid atom solutions. Moreover, the fact that they do not present radial nodes (due to the absence of the non-leading terms of the Laguerre polynomial in equation 2.78) can be solved by making linear combinations of functions with different values of ζ_a. Despite their being good theoretical candidates as basis sets, STO have a serious computational drawback since four-center integrals that have to be computed during the variational procedure cannot be calculated analytically with STO, and this precludes the use of STO in practical ab initio calculations of large molecules.

A major step to overcome these difficulties that has revolutionized the whole field of quantum chemistry was the introduction of Cartesian Gaussian-type orbitals (cGTO) [40]:

$$g^{cGTO}_{\zeta,l_x,l_y,l_z}(x,y,z) = N^{GTO} x^{l_x} y^{l_y} z^{l_z} e^{(-\zeta r^2)} \tag{2.80}$$

where N^{GTO} is a normalization constant, x, y, z are the cartesian coordinates (centered on the nucleus) and the integers l_x, l_y and l_z are generally called orbital quantum numbers. By using GTO there is the enormous computational advantage since all the integrals appearing in SCF theory can be calculated analytically. This makes possible to use a much larger number of functions to expand the one-electron spin orbitals if GTO are used, partially overcoming their bad short- and long-range behaviour (they do not have a cusp for $r = 0$ and fall off too rapidly for $r \to \infty$). To remedy the fact that the angular behaviour of the Cartesian GTOs in equation 2.80 is somewhat hidden, they may be linearly combined to form Spherical Gaussian-type orbitals (sGTO):

$$g^{sGTO}_{\zeta,n,l,m}(x,y,z) = N^{GTO} Y_{l,m}(\theta,\phi) r^{(2n-2-l)} e^{(-\zeta r^2)} \tag{2.81}$$

First, the cGTOs that are combined to make up a sGTO must have all the same value of $l = l_x + l_y + l_z$ and this sum of the three orbital quantum numbers in a particular Cartesian GTO is typically referred to as the angular momentum of the function. Moreover, like atomic orbitals, cGTOs with $l =$ 0, 1, 2, 3, 4,... are called s, p, d, f, g, ..., respectively. For a given $l > 1$, there are more Cartesian GTOs ($(l+1)(l+2)/2$) than spherical ones ($2l + 1$), and from the $(l+1)(l+2)/2$ linear combinations that can be formed using the cGTOs of angular momentum l, only the angular part of $2l + 1$ of them turns out to be proportional to the spherical harmonic $Y_{l,m}(\theta, \phi)$. The rest of them are proportional to spherical harmonic functions with a different value of the angular momentum and this "spurious components" can be removed from the calculation. GTO basis functions are usually used as linear combinations (contractions) of Gaussian functions (primitive Gaussians) instead of single function, and when they have to describe core electrons often contain many primitives in order to better approximate the cusp behavior of hydrogenoid functions near the nucleus [41].

An important concept in quantum chemistry is that of the atomic shells that are defined analogously to those of the hydrogen atom, so that each electron is regarded as filling the multi-electron atom orbitals according to Hund's rules. For instance, the occupied shells of carbon are defined to be 1s, 2s and 2p. Each shell may contain $2(2l + 1)$ electrons if complete, where $2l + 1$ accounts for the orbital angular momentum multiplicity and the factor 2 for that of electron spin. According to the number of basis functions used to describe each atomic shell, basis sets can be classified as minimal if only $2l+1$ functions are employed for each (completely or partially) occupied shell, double zeta (DZ) if $2 \times (2l + 1)$ functions are used for each occupied shell, triple zeta (TZ) if $3 \times (2l + 1)$ functions are used, and so on.

Since core electrons are less affected by the molecular environment and the formation of bonds than valence electrons, core and valence shells may differ in their respective "zeta quality", i.e. the basis set may contain a different number of basis functions in each case. Thus, it is very common to use more basis functions for the valence electrons, to give them more flexibility, and the basis set is called in this case split-valence basis set. As the molecular environment is often highly anisotropic, for most practical applications it turns out to be convenient to add polarization functions, which are functions with larger angular momentum as compared to the last occupied shell in the atom. Typically, the polarization shells are single-primitive contractions and they are denoted by adding a P to the end of the previous acronyms (for instance DZP). Moreover, for calculations of charged species (especially anions), where the charge density extends in space and the tails of the distribution are very important to account for the relevant behaviour of the system, it is common to augment the basis sets with diffuse functions, i.e. single-primitive Gaussian shells of the same angular momentum as some preexisting one but with a small exponent. This improvement is usually denoted by adding the prefix aug- to the name of the basis set.

Correlation consistent (cc) basis sets are widely used in the literature for molecular calculations and they are designed in such a way that functions which provide similar amounts of correlation energy are included at the same stage in the basis set, independently of the function type. For example, if a second d function is added, a first f function is added too, since they contribute in a similar way to the correlation energy. Several different sizes of cc basis sets are available and they are known by their acronyms. For instance, cc-pVDZ is the cc polarized valence Double Zeta basis set, while aug-cc-pVTZ, is the augmented cc polarized valence Triple Zeta basis set [41].

2.8.2 Plane waves

A different approach has its root in solid-state theory. Here, the periodicity of the crystalline lattice produces a periodic potential and this imposes the same periodicity on the wavefunction and the electron density of the system. The periodicity of the potential can be expressed as:

$$\mathbf{V}(\mathbf{r}+\mathbf{L}) = \mathbf{V}(\mathbf{r}) \tag{2.82}$$

where $\mathbf{L}$ is a lattice vector of the crystal. According to the Block theorem the electronic wavefunction can be written as a product of a cell-periodic part and a wavelike part [47]:

$$\psi(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}u(\mathbf{r},\mathbf{k}) \tag{2.83}$$

where k is a vector in the first Brillouin zone. The function $u(\mathbf{r},\mathbf{k})$ has the same periodicity of the direct lattice:

$$u(\mathbf{r},\mathbf{k}) = u(\mathbf{r}+\mathbf{L},\mathbf{k}) \tag{2.84}$$

and can be expanded using a basis set of plane waves whose wave vectors, $\mathbf{G}$, are the reciprocal lattice vectors of the crystal:

$$u(\mathbf{r},\mathbf{k}) = \sum_{\mathbf{G}} c(\mathbf{G},\mathbf{k})e^{i\mathbf{G}\cdot\mathbf{r}} \tag{2.85}$$

where $\mathbf{G}$ are defined by $\mathbf{G}\cdot\mathbf{L} = 2\pi m$ for all $\mathbf{L}$, and m is an integer. Therefore, the electronic wavefunction can be written as:

$$\psi(\mathbf{r}) = \sum_{\mathbf{G}} c(\mathbf{G},\mathbf{k})e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} \tag{2.86}$$

As in the case of Gaussian basis sets, in principle, an infinite number of plane waves is required to expand the electronic wavefunction. However, the coefficients $c(\mathbf{G},\mathbf{k})$ for the plane waves with small kinetic energy:

$$T = \frac{1}{2}|\mathbf{k}+\mathbf{G}|^2 \tag{2.87}$$

are typically more important than those with large kinetic energy [47]. For this reason, at each $\mathbf{k}$ point, only $\mathbf{G}$ vectors with a kinetic energy lower than a given maximum cutoff (E_{cut}) are included in the basis. When dealing with many-electron systems, each orbital can be written in the Bloch form (equation 2.83) and thus expanded in plane waves according to equation 2.86. In the framework of DFT, the electron density can also be expanded in plane wave basis sets and the Kohn-Sham equations assume a particularly simple form:

$$\begin{aligned}\sum_{\mathbf{G}'}\Big\{\frac{1}{2}|\mathbf{k}+\mathbf{G}|^2\delta_{\mathbf{G}\mathbf{G}'}+\mathbf{V}_{ne}(\mathbf{G}-\mathbf{G}') &\ + \\ \mathbf{J}(\mathbf{G}-\mathbf{G}')+\mathbf{V}_{xc}(\mathbf{G}-\mathbf{G}')\Big\}c_i(\mathbf{G}',\mathbf{k}) &= \epsilon_i c_i(\mathbf{G},\mathbf{k})\end{aligned} \tag{2.88}$$

where $\mathbf{V}_{ne}(\mathbf{G}-\mathbf{G}')$, $\mathbf{J}(\mathbf{G}-\mathbf{G}')$ and $\mathbf{V}_{xc}(\mathbf{G}-\mathbf{G}')$ are the Fourier transforms of the nucleus-electron, coulomb and exchange correlation potentials. In this form the kinetic energy is diagonal since plane waves are eigenfunctions of the kinetic energy operator [47]:

$$\frac{1}{2}\nabla^2 e^{i\mathbf{G}\cdot\mathbf{r}} = \frac{1}{2}|\mathbf{G}|^2 e^{i\mathbf{G}\cdot\mathbf{r}} \tag{2.89}$$

It is clear from the above discussion that plane wave based calculations imply the existence of a periodic potential. This is appropriate for crystal calculations but very unnatural for isolated molecules. However, plane waves can also be applied to non periodic system using the supercell method. In this method the molecule is placed at a center of a supercell and is periodically repeated. If the supercell is large enough the interactions between molecules in neighbouring cells become negligible. Plane waves have broad application in the field of Molecular Dynamics simulations where, as we shall see in the next chapter, the periodicity is intimately connected to the use of periodic boundary conditions.

It is noteworthy to observe that, at variance with Gaussian basis functions that are centered on the nuclei, plane waves are originless functions, i.e. they do not depend on the positions of the nuclei. This property facilitates force calculations in Molecular Dynamics simulations. This also implies that plane waves are “delocalized” in space and do not favour certain atoms or regions over others, so that they can be considered as an ultimately balanced basis set. Therefore, the only way to improve the quality of the basis set is to increase E_{cut}, i.e. the largest $\mathbf{G}$ vector that is included in the finite expansion (equation 2.86). This approach is completely different and much simpler than the traditional procedures in quantum chemistry that are needed in order to produce reliable Gaussian basis sets [40]. Another appealing feature is that derivatives in real-space are simply multiplications in G-space, and both spaces can be efficiently connected via Fast Fourier Transforms (FFT). On the other hand, a severe shortcoming of plane waves is that there is no way to assign more basis functions into regions of space where they are more needed than in other regions. This is particularly bad for systems with strong inhomogeneities.

It is important to stress that pseudopotential approximation is intimately connected to using plane waves. Pseudopotentials allow for a considerable reduction of the basis set size, taking out of explicit calculations the innermost electrons. Core orbitals feature strong and rapid oscillations due to the Pauli principle, which enforces a nodal structure into the wavefunction by imposing orthogonality of the orbitals. For this reason, we need to use for these electrons a great number of basis functions. However, core electrons are relatively inert, and most of the important chemical properties of atoms and molecules are determined by the interactions of their valence electrons. In practice, in the pseudopotential approximation, only the valence electrons are treated explicitly, while core orbitals are represented by a smooth and nodeless effective potential. Moreover, since most relativistic effects are connected to core electrons, these effects can be incorporated in the pseudopotentials without complicating the calculation of the final system.

2.9 Continuum solvation models

The computational methods described in the previous sections of this chapter deal with the study of molecular properties in vacuum. However, many processes take place in solution and when a molecule interacts with solvent molecules its properties can be very different from those observed in vacuum. Therefore, in order to study a solvated system, it is necessary to use a model able to describe the solute-solvent interactions. Solvent models may broadly be divided into two classes: discrete and continuum models [48]. Discrete models include in calculations together with the system of interest (the solute) a certain number of solvent molecules which are treated at a comparable level of theory. On the other hand, continuum models treat the solvent as a continuous medium, and the solute-solvent interactions are described by means of the reaction field of the solvent, i.e. the electric field generated by it. For this reason, and due to the fact that they employ iterative procedures, this methods are generally called Self Consistent Reaction Field (SCRF) methods. Combinations of discrete and continuum models are also possible, for example by considering explicitly the solvent molecules of the first solvation shell, while treating the rest of the solvent by a continuum model.

A brief introduction to the class of continuum models will be given in this section. First of all, it is important to stress that in continuum models it is fundamental the presence of a cavity inside the dielectric medium in which solute molecules are placed. This cavity has to exclude solvent molecules, it has to include most of the solute charge distribution, and it should reproduce as much as possible the shape of the solute molecule. The cavity is usually constructed by interlocking spheres centered on each solute atom, with radius equal or proportional to the van der Waals radius of the atom. This surface formed by interlocking spheres is called van der Waals surface. Since such a cavity may be quite irregular and may contain regions not accessible to solvent molecules, two additional forms of cavity are also used, the Solvent Excluding Surface (SES) and the Solvent Accessible Surface (SAS) (see Figure 2.1). In both cases the solvent molecule is represented as a sphere with a volume equal to the van der Waals volume. The positions spanned by the center of a solvent sphere rolling on the van der Waals solute cavity generate the SAS, which is consequently the surface in which the center of the solvent sphere cannot enter. The SES represents instead the whole surface not accessible by the entire solvent molecule, as it is possible to see in Figure 2.1.

In the framework of continuum models, the before mentioned self consis-

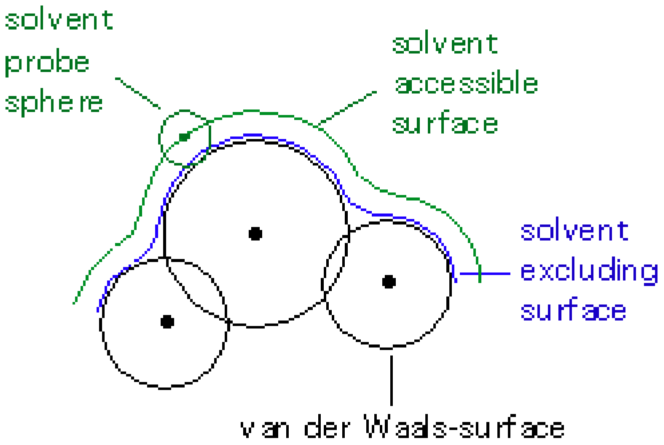

Figure 2.1: Schematic representation of the Solvent Accessible Surface (SAS) and the Solvent Excluding Surface (SES).

tent problem stems from the fact that the solute charge distribution polarizes the solvent continuum medium which, in turn, polarizes the solute charge distribution. Such a problem can be solved only be means of an iterative procedure. The equations that describe the electrostatic properties in continuum models are:

$$\nabla^2 V(\mathbf{r}) \quad = \quad \nabla \cdot \mathbf{E}(\mathbf{r}) \quad = -4\pi\rho_{solu} \quad \textit{inside the cavity} \tag{2.90}$$

$$\nabla^2 V(\mathbf{r}) \quad = \quad \nabla \cdot \mathbf{D}(\vec{r}) \quad = 0 \quad \textit{outside the cavity} \tag{2.91}$$

$$V_{in} \quad = \quad V_{out} \tag{2.92}$$

$$(\mathbf{D} \cdot \mathbf{n})_{in} \quad = \quad \frac{\partial V_{in}}{\partial \mathbf{n}} \quad = \quad \epsilon \frac{\partial V_{out}}{\partial \mathbf{n}} = (\mathbf{D} \cdot \mathbf{n})_{out} \tag{2.93}$$

where ρ_{solu} is the solute charge distribution, $\mathbf{n}$ is the normal to the cavity surface, ϵ is the dielectric constant of the solvent and $\mathbf{D}(\mathbf{r}) = \epsilon\mathbf{E}(\mathbf{r})$ is the dielectric displacement; *in* and *out* indexes indicate that the corresponding properties are calculated inside or outside the cavity. The electric field $\mathbf{E}(\mathbf{r})$ is associated to the potential $V(\mathbf{r})$ which is the sum of the potential $V_{solu}(\mathbf{r})$ generated by the solute charge distribution ρ_{solu} and the potential generated by the solvent polarization $V_R(\mathbf{r})$:

$$V(\mathbf{r}) = V_{solu}(\mathbf{r}) + V_R(\mathbf{r}) \tag{2.94}$$

There exist several methods which solve these electrostatic equations in different ways. Multipole Expansion methods (MPE), for instance, are based on a multipole expansion of $V_{solu}(\mathbf{r})$ [48]. The simplest version of MPE methods is the Onsager method in which the solute is represented just as a dipole. The main drawback of this method is that it generally uses a spherical cavity which is not suitable to describe complex molecular shapes. Conversely, Apparent Surface Charges (ASC) methods [48] treat the solute-solvent interactions by means of a polarization charge distribution, $\sigma(\mathbf{r})$, displaced on the internal cavity surface. In ASC methods the potential $V_R(\mathbf{r})$ of the previous equation is given by:

$$V_\sigma(\mathbf{r}) = \int_\Gamma \frac{\sigma(\mathbf{s})}{|\mathbf{r}-\mathbf{s}|} d^2 s \tag{2.95}$$

where the subscript σ indicates that the potential is calculated from the charge distribution $\sigma(\mathbf{s})$, the integration is over all the surface area of the Γ cavity and $\mathbf{s}$ is a generic point of it. By defining a charge surface distribution as the source of the electrostatic potential, ACS methods simplify the problem as compared to other formulations in which the source of the potential is all the dielectric medium. Nevertheless, the integration of equation 2.95 is too much demanding and, for this reason, the cavity surface is partitioned in small elements called tesserae (from a few tens to a few hundreds of elements for each atom), which have to be small enough that $\sigma(\mathbf{s})$ can be considered as a constant on each tessera (generally their surface area is comprised between 0.05 and 0.4 $\mathring{\text{A}}^2$). By associating to each tessera k a position s_k and a point charge q_k, it is possible to rewrite equation 2.95 as:

$$V_\sigma(\mathbf{r}) = \sum_k \frac{q_k}{|\mathbf{r}-\mathbf{s_k}|} \tag{2.96}$$

where the summation runs over all the tesserae. The ASC methods mostly used are:

- The Polarizable Continuum Model method (PCM), which is the oldest ASC formulation [49].
- The Conductor like PCM (CPCM), which is very efficient to describe the properties of polar solvents such as water [50].
- The Integral Equation Formalism (IEF)-PCM, which employs Green functions in order to define the potential [51].

The main difference between PCM and CPCM is that while in the first case the solvent is considered as a dielectric medium, in the latter one the solvent is represented as a medium with a dielectric constant going to infinity, i.e. as a conductor. By multiplying the charge distribution obtained with $\varepsilon = \infty$ $(\sigma^*(\vec{s_k}))$ for a suitable function of ε, it is then possible to obtain the charge distribution corresponding to a finite ε value [50]:

$$\sigma(\vec{s_k}) = f(\varepsilon)\sigma^*(\vec{s_k}) \tag{2.97}$$

$$f(\varepsilon) = \frac{\varepsilon - 1}{\varepsilon} \tag{2.98}$$

In ASC methods the surface charge distribution is determined by solving the following matrix equation:

$$\mathbf{AQ} = -\mathbf{V} \tag{2.99}$$

$\mathbf{Q}$ is a column vector that contains the polarization charges, while $\mathbf{V}$ is a column vector containing the solute electrostatic potential and its elements are $V_i = V_{solu}(\mathbf{s}_i)$. $\mathbf{A}$ is a square matrix of TxT dimension (where T is the number of tesserae), which depends on the solvent properties and on the cavity shape. The explicit expression of $\mathbf{A}$ depends on the particular ASC formulation used, and in the case of the CPCM method the matrix elements are given by:

$$A_{ii} = 1.07\sqrt{\frac{4\pi}{a_i}} \tag{2.100}$$

$$A_{ij} = \frac{1}{|\mathbf{s}_i - \mathbf{s}_j|} \tag{2.101}$$

where a_i is the area of the i tessera and the $\mathbf{s_i}$ vector corresponds to the center of the tessera.

As far as the quantum mechanical problem is concerned, the Hamiltonian $\mathbf{H}$ of a molecular system placed in a suitable cavity inside the solvent is the sum of two terms [52]:

$$\mathbf{H} = \mathbf{H^{(0)}} + \mathbf{H^{(R)}} \tag{2.102}$$

where $\mathbf{H}^{(0)}$ is the Hamiltonian of the isolated solute, while $\mathbf{H}^{(R)}$ is the operator associated with the reaction potential $V_R(\mathbf{r})$ plus the interaction between the solute nuclei and polarization charge of the solvent (σ). In order to obtain the $\mathbf{H}$ approximate eigenfunctions Ψ it is possible to use one of the electronic structure methods discussed in this chapter. However, since the solute charge distribution ρ_{solu} depends on the wavefunction Ψ, an iterative procedure is employed which generally starts with a calculation of $\Psi^{(0)}$, the $\mathbf{H}^{(0)}$ eigenfunction; ρ_{solu}, σ and Ψ are then computed and from the new Ψ a new ρ_{solu} is obtained and so on until self consistency is reached.

Chapter 3

Molecular Dynamics Simulations

Molecular Dynamics, strictly speaking, is the simultaneous motion of atomic nuclei and electrons forming molecular entities. A complete description of such a system requires in principle solving the full time-dependent Schrödinger equation including both electronic and nuclear degrees of freedom. This is however a too much expensive computational task which is in practice unfeasible for systems consisting of more than three atoms. In order to study the dynamics of the vast majority of chemical systems several approximations have therefore to be introduced. First of all, it is assumed in Molecular Dynamics with the Born-Oppenheimer approximation (see section 2.2) that the motion of electrons and nuclei is separable, and the electron cloud adjusts instantaneously to changes in the nuclear configuration. As a consequence, nuclear motion evolves on a PES, associated with the electronic quantum state which is obtained by solving the time-independent electronic Schrödinger equation 2.10 for a series of fixed nuclear geometries. In practice, most Molecular Dynamics simulations are performed on the ground state PES.

Moreover, in addition to making the Born-Oppenheimer approximation, Molecular Dynamics treats the atomic nuclei as classical particles whose trajectories are computed using the laws of classical mechanics. This is a very good approximation for molecular systems as long as the properties studied are not related to the motion of light atoms (like the hydrogen atoms, which show quantum mechanical behaviour in certain situations such as tunneling phenomena) or vibrations with frequency ν such that $h\nu > k_B T$.

The potential functions which describe the intermolecular and intramolecular interactions between classical nuclei can be treated at various levels of approximation. In classical Molecular Dynamics the interaction potential is expressed as a simple sum of pair potentials. On the other hand, ab initio Molecular Dynamics computes interactions at a much more fundamental level using electronic structure methods. In the mixed Quantum Mechanical and Molecular Mechanics (QM/MM) methods instead the “important” part of the system, for instance where a chemical reaction is taking place, is treated with electronic structure calculations whereas the rest of the system is described by a classical force field. In this chapter an introduction to the classical Molecular Dynamics is given, followed by a brief description of how classical and quantum mechanics is coupled in ab initio Born-Oppenheimer and Car-Parrinello Molecular Dynamics simulations.

3.1 Classical Molecular Dynamics

In the classical Molecular Dynamics simulations the time evolution of a system composed by M particles (generally atoms) is obtained by solving the Newton's equation of motion step-by-step [53]:

$$M_a \ddot{\mathbf{R}}_a = \mathbf{F}_a \quad a = 1, 2, \ldots, M \tag{3.1}$$

where M_a and $\mathbf{R}_a$ are the mass and the position of particle a and $\mathbf{F}_a$ is the force acting on particle a given by:

$$\mathbf{F_a} = -\frac{\partial V}{\partial \mathbf{R_a}} \tag{3.2}$$

V is a potential energy function which depends on the complete set of $3M$ particle coordinates. In classical Molecular Dynamics the complex potential energy function is represented by a sum of simple functions, called force fields. In these force fields, the interactions are usually divided into bonded and non-bonded [54]:

$$V = V_{bonded} + V_{non-bonded} \tag{3.3}$$

Bonded interactions are written as a sum of various terms:

$$V_{bonded} = V_{bonds} + V_{angles} + V_{dihedrals} + V_{impr-dihedr} \tag{3.4}$$

V_{bonds} describes the stretching between the atoms in the system covalently bonded:

$$V_{bonds} = \sum_{bonds} \frac{1}{2} k_{b_{ij}} (b_{ij} - b_{ij}^0)^2 \tag{3.5}$$

where the covalent bond between atoms i and j is represented by a harmonic potential with force constant k_{bij}, instantaneous distance b_{ij} and equilibrium distance b_{ij}^0. For some systems that require an anharmonic bond stretching potential, other functional forms can be used such as the Morse potential [55]. V_{angles} describes the bond angle vibrations:

$$V_{angles} = \sum_{angles} \frac{1}{2} k_{\theta_{ijk}} (\theta_{ijk} - \theta_{ijk}^0)^2 \tag{3.6}$$

where the bending of the bond angle between a triplets of atoms i, j and k is modeled by a harmonic potential with force constant $k_{\theta_{ijk}}$, instantaneous angle θ_{ijk} and equilibrium angle θ_{ijk}^0. $V_{dihedrals}$ mimics the vibrations of dihedral angle (four-body) interactions and is generally modeled as:

$$V_{dihedrals} = \sum_{dihedrals} \frac{1}{2} k_{\phi_{ijkl}} [1 + \cos(n\phi_{ijkl} - \gamma)] \tag{3.7}$$

where ϕ_{ijkl} is the angle between the ijk and jkl planes. $V_{impr-dihedr}$ represents a special type of dihedral interaction (called improper dihedral) which is used to force atoms to remain in a plane or to prevent transitions to a configuration of opposite chirality (a mirror image) and, in harmonic approximation, is given by:

$$V_{impr-dihedr} = \sum_{impr-dihedr} \frac{1}{2} k_{\xi_{ijkl}} (\xi_{ijkl} - \xi^0_{ijkl})^2 \tag{3.8}$$

where $k_{\xi_{ijkl}}$, ξ_{ijkl} and ξ^0_{ijkl} are the force constant, instantaneous and equilibrium improper dihedral angle, respectively.

As far as the non-bonded interactions are concerned, atoms are represented by charged point particles interacting with each other by parametrized model potentials. One of the most used functional form of $V_{non-bonded}$ is given as:

$$V_{non-bonded} = \sum_{pairs(ij)} (\frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{\epsilon_r R_{ij}}) + \sum_{pairs(ij)} (\frac{C^{(12)}_{ij}}{R^{12}_{ij}} - \frac{C^{(6)}_{ij}}{R^6_{ij}}) \tag{3.9}$$

where the first term describes the Coulomb interactions between all of the atomic partial charges of the system (q_i and q_j are the partial charges of particles i and j placed at a distance R_{ij}), while the second term represents the van der Waals interactions by means of a Lennard-Jones potential which is the most commonly used form.

Aspects and details of the Molecular Dynamics simulation techniques will not be discussed here for the sake of brevity. However, it is noteworthy to summarize just a few points:

- The Newton's equations of motion 3.1 are integrated numerically step-by-step. Many algorithms have been designed to do this which are generally based on a Taylor expansion of the particle positions around the positions at a certain time instant. Among these methods, the most commonly used are the Verlet [56] and the leap-frog [57] algorithms. These two methods have the important property to be time-reversible, like the Newtonian equations of motion [58].

- Periodic boundary conditions [54] are applied in Molecular Dynamics simulations in order to minimize edge effects which may produce artifacts in a finite system. The atoms of the system to be simulated are put into a space-filling box, which is surrounded by translated copies of itself. There are thus no boundaries of the system. However, this imposed artificial periodicity by itself may cause errors, especially when considering properties which are influenced by long-range correlations.

- Long range non-bonded interactions are generally not calculated beyond a certain cutoff distance around each particle, in order to reduce the computational cost of the simulations. However, as far as the Coulomb interactions are concerned, the use of a simple cutoff can introduce serious artifacts and, for this reason, several techniques have been developed for handling long range interactions, the most popular of them being the Ewald summation [59] and the Particle Mesh Ewald methods [60, 61].

- Constraints are often used in Molecular Dynamics simulations, i.e. bonds are treated as being constrained to have fixed length. This is very useful when bonds have very high vibration frequencies and should be treated in a quantum mechanical way rather than in the classical approximation. Moreover, they allow to increase the integration time step and thus to perform longer simulations. The most commonly used constraints methods are the LINCS [62] and SHAKE [63] algorithms.

- When solving the Newton's equations of motion 3.1 the energy is a constant of motion and the simulation is performed in an NVE ensemble. However, it is often more convenient to carry out simulations in other ensembles, such as NVT or NPT. To this end, several approaches have been developed which control the temperature and the pressure of a system. As far as the temperature control is concerned, widely used techniques are the Berendsen [64] and Nosé-Hoover [65, 66] methods, while for the pressure control the Parrinello-Rahman scheme [67, 68] is extensively employed.

3.2 Born-Oppenheimer Molecular Dynamics

We have seen in the previous section that in the classical Molecular Dynamics simulations the interaction potentials are expressed as a sum of pair potentials, represented by suitable functional forms. These potentials are determined in advance, either empirically or based on electronic structure calculations. An alternative approach to treat the interatomic potentials consists in solving the electronic structure problem in each Molecular Dynamics step, given the set of fixed nuclear positions at that instant of time. Thus, the electronic structure part consists in solving the time-independent electronic Schrödinger equation 2.10 for that particular nuclear configuration, while the nuclei are propagated via classical Molecular Dynamics. In such an approach, the time-dependence of the electronic structure is not intrinsic, but a consequence of the nuclear motion.

The equations of motion in the Born-Oppenheimer Molecular Dynamics are given by [47]:

$$M_a \ddot{\mathbf{R}}_a = -\nabla_a min_{\Psi} \langle \Psi | \mathbf{H} | \Psi \rangle \tag{3.10}$$

$$\mathbf{H}\Psi = E\Psi \tag{3.11}$$

where M_a and $\mathbf{R}_a$ are the mass and the position of nucleus a, while $\mathbf{H}$, Ψ and E are the electronic Hamiltonian, wavefunction and energy, respectively. In each Born-Oppenheimer Molecular Dynamics step the minimum of $\langle \mathbf{H} \rangle$ has to be reached, and this means that the electronic wavefunction has to be optimized at each instant of time using a suitable electronic structure method. This may be for instance the Hartree-Fock approximation or the DFT described in the previous chapter.

3.3 Car-Parrinello Molecular Dynamics

The Born-Oppenheimer Molecular Dynamics requires, as already pointed out, that the time-independent electronic Schrödinger equation be solved self-consistently at each Molecular Dynamics step and this task is computationally very expensive. The aim of the Car-Parrinello method is to reduce the computational cost of ab initio Molecular Dynamics simulations. The basic idea of the Car-Parrinello approach is to introduce classical degrees of freedom associated with the electronic wavefunction [69]. These electronic degrees of freedom evolve in time classically along with the nuclei. The "electron dynamics" is the defining feature of the Car-Parrinello approach which exploits the quantum-mechanical adiabatic time-scale separation of the fast electronic and slow nuclear motion (see section 2.2), by transforming it into classical-mechanical adiabatic energy-scale separation in the framework of Molecular Dynamics.

Car and Parrinello postulated a new set of Lagrangians ($\mathcal{L}_{CP}$) like [69]:

$$\mathcal{L}_{CP} = \sum_{a=1}^{M} \frac{1}{2} M_a \dot{\mathbf{R}}_a^2 + \sum_{i=1}^{N} \frac{1}{2} \mu_i \langle \dot{\phi}_i | \dot{\phi}_i \rangle - \langle \Psi | \mathbf{H} | \Psi \rangle + (constraints) \tag{3.12}$$

where:

- $T_n = \sum_{a=1}^{M} \frac{1}{2} M_a \dot{\mathbf{R}}_a^2$ is the kinetic energy of the M nuclei of masses m_a and positions $\mathbf{R}_a$.

- $T_e = \sum_{i=1}^{N} \frac{1}{2} \mu_i \langle \dot{\phi}_i | \dot{\phi}_i \rangle$ is the classical fictitious kinetic energy of the N electronic orbitals ϕ_i. This quantity has no relation with the quantum electron kinetic energy and it does not have any physical meaning. $\mu_i = \mu$ are the fictitious masses or inertia parameters assigned to the orbital degrees of freedom. The units of the mass parameter μ are energy times a square time for reasons of dimensionality.

- $V_e = \langle \Psi | \mathbf{H} | \Psi \rangle$ is the potential energy calculated on the electronic wavefunction Ψ ($\mathbf{H}$ is the electronic Hamiltonian).

- ($constraints$) This term is introduced in order to guarantee the conservation of the orbital orthogonality. The constraints are functions of both the set of orbitals and the nuclear positions.

The corresponding Newtonian equations of motion are:

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{\mathbf{R}}_a} = \frac{\partial \mathcal{L}}{\partial \mathbf{R}_a} \tag{3.13}$$

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{\phi}_i^*} = \frac{\partial \mathcal{L}}{\partial \phi_i^*} \tag{3.14}$$

like in classical mechanics, but here for both the nuclear positions and the electron orbitals. The equations of motions in the Car-Parrinello Molecular Dynamics are thereby given as [47]:

$$M_a \ddot{\mathbf{R}}_a = -\frac{\partial}{\partial \mathbf{R}_a}\langle \Psi|\mathbf{H}|\Psi\rangle + \frac{\partial}{\partial \mathbf{R}_a}(constraints) \tag{3.15}$$

$$\mu \ddot{\phi}_i = -\frac{\partial}{\partial \phi_i^*}\langle \Psi|\mathbf{H}|\Psi\rangle + \frac{\partial}{\partial \phi_i^*}(constraints) \tag{3.16}$$

According to the Car-Parrinello equations of motion the nuclei evolve in time at a certain physical temperature $\propto \sum_{a=1}^{M} M_a \dot{\mathbf{R}}_a^2$, whereas a fictitious temperature $\propto \sum_{i=1}^{N} \mu\langle \dot{\phi}_i|\dot{\phi}_i\rangle$ is associated to the electronic degrees of freedom. If this "electronic temperature" is low, the electronic subsystem is close to its instantaneous minimum energy $min_{\{\phi_i\}}\langle \Psi|\mathbf{H}|\Psi\rangle$, i.e. close to the Born-Oppenheimer surface. Thus, a ground state wavefunction optimized for the initial configuration of the nuclei will stay close to its Born-Oppenheimer surface also during time evolution if it is kept at a sufficiently low temperature. Since the Lagrangian in equation 3.12 is time-independent, there is in the Car-Parrinello approach a conserved energy quantity, E_{cons}, given by:

$$E_{cons} = T_n + T_e + V_e \tag{3.17}$$

However, this constant of motion has no physical meaning. The physical energy of the system (E_{phys}) is instead given by:

$$E_{phys} = T_n + V_e = E_{cons} - T_e \tag{3.18}$$

If $T_e << E_{cons}$, i.e. if the fictitious electron kinetic energy is very small with respect to the total energy, E_{phys} is essentially constant on the relevant energy and time scale. Thus, it behaves approximately like the strictly conserved total energy in classical Molecular Dynamics (with only nuclei as dynamical degrees of freedom), or in the Born-Oppenheimer Molecular Dynamics (with fully optimized electron degrees of freedom) [47]. This implies that the resulting dynamics of the nuclei yields an excellent approximation to microcanonical dynamics. It is clear that the validity of the Car-Parrinello approach depends on the maintenance of adiabaticity between the fictitious electron dynamics and the ionic motion. The adiabatic separation is maintained if T_e, which can be considered as a measure of the deviations from the Born-Oppenheimer surface, is small and performs bound oscillations around a constant, i.e. the electrons do not heat up systematically in the presence of the hot nuclei. This condition in practice is achieved by adjusting the time step of the simulation and the electron fictitious mass [47].

Chapter 4

X-ray absorption spectroscopy

4.1 Introduction

X-ray absorption spectroscopy (XAS) measures the X-ray absorption coefficient, $\mu(E)$, as a function of the X-ray energy of the incident photon, $E = \hbar\omega$ (measured in eV). A XAS spectrum, in which the absorption coefficient is plotted as a function of E, shows an overall decrease of the X-ray absorption with increasing energy, with the exception of very sharp peaks at certain energies (called edges) due to the transitions of core electrons to high energy states (see Figure 4.1). The energies of these peaks correspond to the ionization energies of the core electrons. Each absorption edge is related to a specific atom present in the material and, more specifically, to a quantum-mechanical transition that excites a particular atomic core electron to the free or unoccupied continuum levels. The nomenclature for X-ray absorption reflects the origin of the core electron (see Figure 4.2). K edge refers to the transition that excites the innermost 1s electron, L_1 edge is due to the excitation of the 2s electron, while L_2 and L_3 edges are related to the excitations of the 2p electrons with electronic states $^2P_{1/2}$ and $^2P_{3/2}$, respectively. The transitions are always to unoccupied states with the photoelectron above the Fermi energy, which leaves behind a core hole. The energies of the edges are unique to the type of atom that absorbs in the X-ray and thus themselves are signatures of the atomic species present in a material.

From what has been said, it is clear that XAS, being an atomic probe, places few constraints on the samples that can be studied. All atoms have core level electrons, and XAS spectra can be measured for essentially every element on the periodic table. Moreover, crystallinity is not required for XAS measurements (even if it is also possible to measure XAS spectra of crystalline samples), making it one of the few structural probes available for noncrystalline and highly disordered materials, including solutions, amorphous solids and fluid samples in general. In many cases, XAS measurements can be made on elements of minority and even trace abundance, giving a unique and direct measurement of chemical and physical state of dilute species in a variety of systems. XAS spectra are recorded using the properties of synchrotron radiation, which provides tunable X-ray beams with high brilliance. In this way

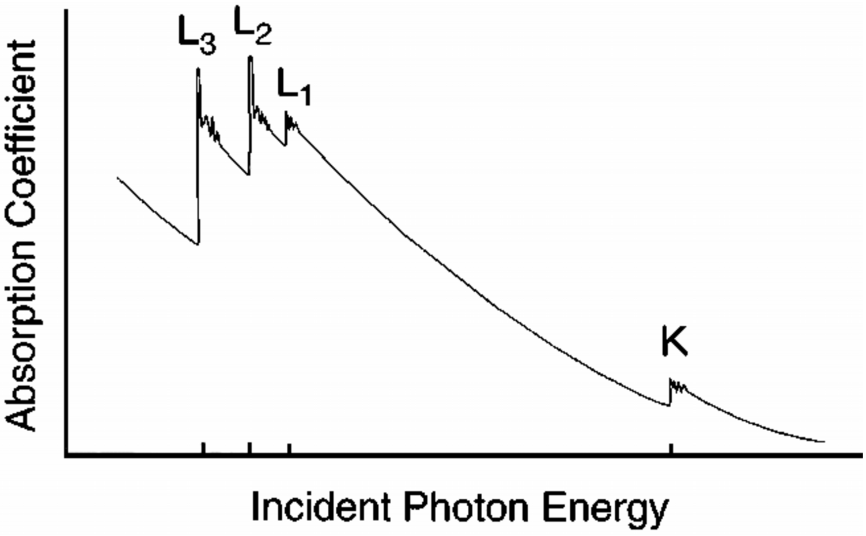

Figure 4.1: Typical XAS spectrum showing the K, L_1, L_2 and L_3 absorption edges.

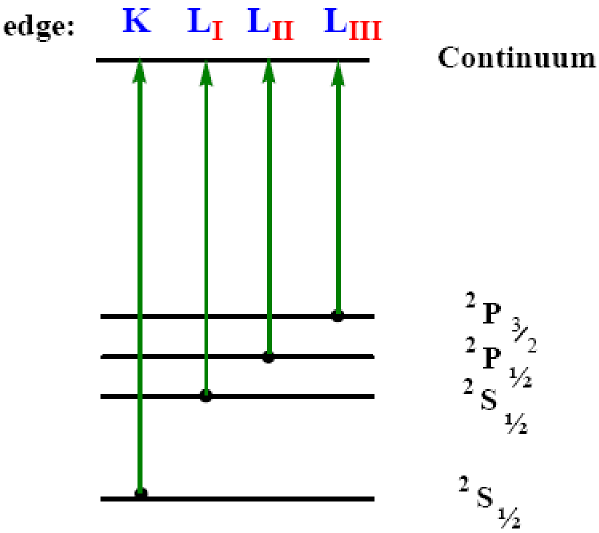

Figure 4.2: Schematic representation of the transitions that take place at the various absorption edges.

it is possible to obtain spectra with a very high signal/noise ratio.

A typical XAS spectrum is shown in Figure 4.1. When the energy of the incident photon, E, is greater than the ionization potential, E_0, of a given electron, this electron is emitted by the photoabsorber atom with a kinetic energy equal to the difference $E - E_0$ and undergoes a scattering process on the nearest atoms. This phenomenon produces a series of wiggles or oscillatory structures above the edge that modulate the absorption, typically by a few percent of the overall absorption cross section. These features contain detailed structural information on the atoms around the photoabsorber, such as interatomic distances and coordination numbers. However, it is important to stress that the information on bond angles and distances that can be obtained from a XAS spectrum is limited in a range of 4-5 Å from the photoabsorber atom [70]. The short-range character of XAS is due to the limited *mean free path* of the photoelectron and to the excited state lifetime (*core hole lifetime*). In fact the high-energy excited photoelectron state is not long lived, but must decay as a function of time and distance and thus cannot probe long-range effects. This decay is due primarily to inelastic losses (i.e. "extrinsic losses") as the photoelectron traverses the sample, either by interacting with and exciting other electrons, or by creating collective excitations (plasmon production). In addition, the intrinsic lifetime of the core-hole state ("intrinsic losses") has to be clearly considered [70]. The net effect is that XAS can only measure the local atomic structure over a range limited by the net lifetime of the excited photoelectron. In this sense, XAS is very different from other techniques such as X-ray diffraction or neutron diffraction, in which also long-range interactions provide a detectable contribution to the experimental spectrum.

A XAS spectrum is conventionally divided into two regions: the X-ray Absorption Near Edge Structure (XANES) up to about 50 eV beyond the absorption edge, and the Extended X-ray Absorption Fine Structure (EXAFS) at higher energies. The border between XANES and EXAFS regions is shown in Figure 4.3. The division of a XAS spectrum is only formal and it is due to the different theoretical treatment and approximations used to calculate the absorption cross section in the two regions. In both cases, a full quantum description of the X-ray absorption phenomenon is not possible, and, as a consequence, approximate models have to be employed. In these models the emitted electron is treated as a quasi-particle (photoelectron) that moves in an effective potential which takes into account both the interaction with the other electrons of the photoabsorber atom and the potential generated by the surrounding atoms. However, as the energy of the photoelectron increases, further approximations can be made to obtain a simpler data analysis protocol. For this reason, historically the first quantitative analyses were made on the high energy part of the absorption spectrum (EXAFS) while the XANES spectra have been analysed for many years only on a qualitative way.

XAS phenomenon was discovered around 1930 but only 40 years later became a structural investigation technique, following the incoming of better X-ray sources in the experiments and the development of theories able to

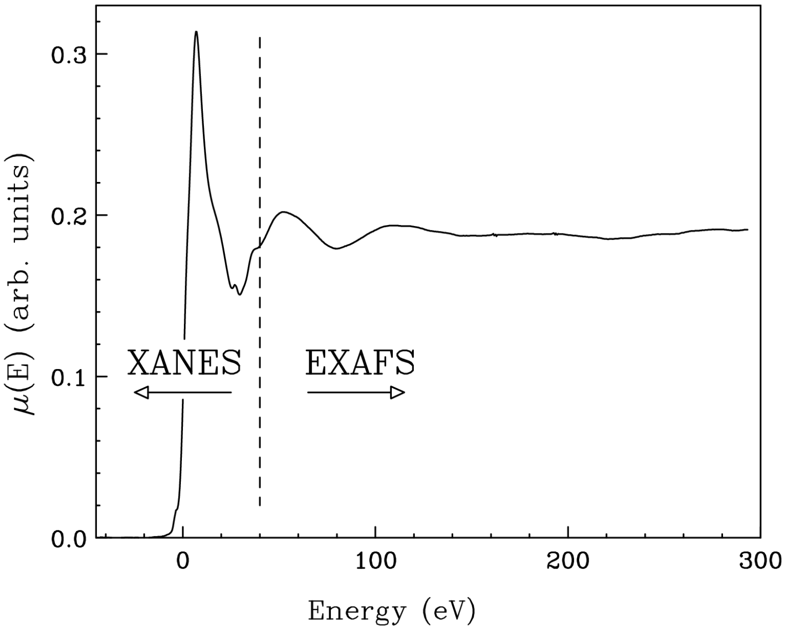

Figure 4.3: Division of the absorption spectrum between XANES and EXAFS regions.

provide a quantitative interpretation of experimental data. The first works were published by Sayers et al. in 1971 and 1975 [71, 72]; in these works a semi-empiric parametrization of the EXAFS signal is proposed for the first time. Later, the work of different groups lead to the development of theories able to explain the many physical phenomena involved in radiation absorption, confirming the validity of Sayers' approach.

Some years ago a unifying scheme of interpretation of the X-ray absorption spectra, based on the Multiple Scattering (MS) theory and valid for the whole energy range, has been developed [73]. An important result of this MS approach, which is based on the Green's function formalism, is that the expression for the absorption cross section ($\sigma(E)$) can be factored in an atomic term ($\sigma_0^l(E)$) depending only on atomic electronic properties, and in a structure factor ($\chi^l(E)$) containing all the structural information on the environment:

$$\sigma(E) \propto \sigma_0^l(E)\chi^l(E) \tag{4.1}$$

The expression for $\chi^l(E)$ obtained within the MS theory is given by [74]:

$$\chi^l(E) = \frac{1}{(2l+1)sin^2\delta_l^0} \sum_m Im[(I + T_aG)^{-1}T_a]^{00}_{lm,lm} \tag{4.2}$$

where I is the unit matrix, G is the matrix describing the spherical wave propagation of the photoelectron from one site to another around the photoabsorber, T is the diagonal matrix describing the scattering process of the photoelectron by the atoms located at the various sites around the photoabsorber and δ_l^0 is the phase shift of the photoabsorbing atom (located at site

0) for angular momentum l. Therefore, the fundamental problem in XAS calculations is the inversion of the matrix reported in equation 4.2. In the high energy region, this problem can be overcome by expanding the matrix inverse in a series in which each term corresponds to the contribution of a scattering path, or in other words, the matrix inverse can be written as a sum over all of the multiple scattering paths [74]. This is possible because in the high energy regime the series is convergent. Conversely, in the XANES region the series does not converge, and the structure factor has to be calculated by the exact matrix inversion. This is one of the most fundamental differences between the theoretical approaches used in the EXAFS and XANES regions of the spectrum. The physical reason of this difference is that in the XANES regime the electron kinetic energy is small and the scattering on the neighbouring atoms tends to be strong, while the effect of the scatterers becomes smaller at higher energies and the photoelectron is only weakly scattered.

In this framework, several data analysis programs have been developed to analyze the experimental data. In this codes two fundamental approximations are generally used. The first one is the so called muffin-tin approximation, in which the potential generated by the atoms surrounding the photoabsorber is spherically averaged inside muffin-tin spheres around each atom, and averaged to a constant in the interstitial region (delimited by a convenient outer sphere enclosing the cluster used in the calculations). The second approximation concerns the choice of the effective optical potential in which the photoelectron moves. The most used approximation is the complex Hedin-Lundqvist energy dependent potential whose imaginary part accounts for extrinsic losses [75]. While this approach is a good approximation in the EXAFS region, in the low-energy regime the complex part of the Hedin-Lundqvist potential introduces an excessive loss in the transition amplitude of the primary channel and thus other approximations are exploited, as we shall see in section 4.3.

In the remainder of this chapter a brief introduction to the techniques used in the EXAFS and XANES data analysis will be given, with particular emphasis on their application to the study of disordered systems.

4.2 EXAFS analysis

The fundamental quantities used in the analysis of EXAFS spectra are defined as follows:

- $\mu(E)$ is the atomic absorption coefficient, defined as the attenuation of the X-ray beam per distance unit, which is proportional to the absorption cross section.
- $\mu_0(E)$ is the absorption coefficient of the isolated atom.
- $k = \frac{p}{\hbar} = \sqrt{\frac{2m_e}{\hbar^2}(E - E_0)}$ is the photoelectron wave number.
- $\chi(k) = \frac{\mu-\mu_0}{\mu_0}$ is the normalized oscillating part of the spectrum, which

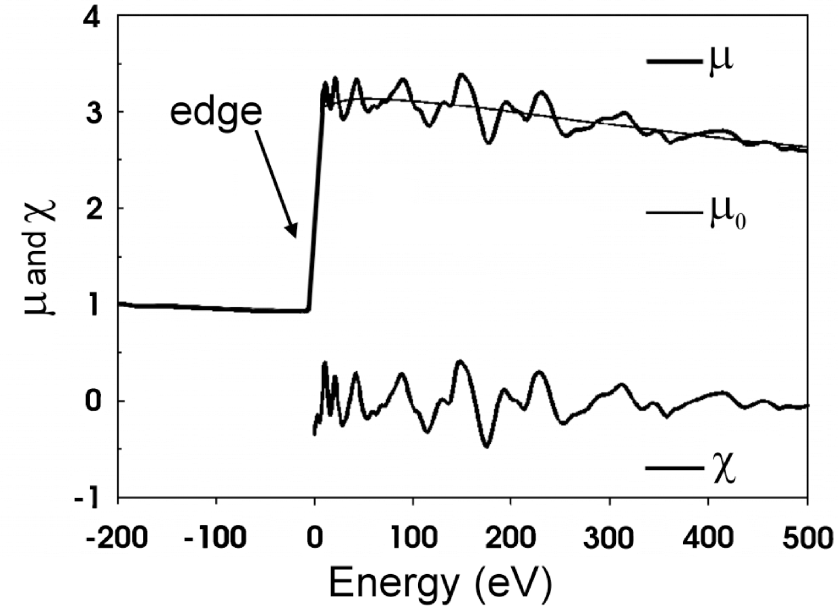

Figure 4.4: Relation between $\mu(E)$, $\mu_0(E)$ and $\chi(E)$.

> is obtained by eliminating the absorption of the isolated atom from the signal and normalizing it to unity.

The relation between these quantities is shown in figure 4.4. $\chi(k)$ contains all the structural information on the system in an analogous way to the structure factor $S(q)$ in diffraction techniques.

Sayers et al. developed a quantitative parametrization of $\chi(k)$ which has become the standard for current EXAFS analysis and it is given by [71]:

$$\chi(k) = S_0^2 \sum_i N_i \frac{f_i(k)}{kR_i^2} \sin(2kR_i + 2\delta_c(k) + \phi_i(k)) e^{-\frac{2R_i}{\lambda(k)}} e^{-2\sigma_i^2 k^2} \tag{4.3}$$

where S_0^2 is a phase reduction factor, the index i is related to the N_i equivalent scattering atoms at distance R_i from the photoabsorber, $\delta_c(k)$ and $\phi_i(k)$ are the phase displacements due to the photoabsorber atom and to the scatterers, respectively. σ_i is the average square fluctuation of the bond distances (or Debye-Waller factor) and contains the structural disorder, and f_i is the diffusion amplitude. The substantial validity of this expression is due to the above mentioned fact that in the high energy range the contributions of the different scattering paths can be factorized and for the special case in which only two-body paths are accounted for, the functional form of equation 4.3 is recovered.

The Debye-Waller factor in equation 4.3 accounts for the fact that, due to the thermal vibrations, the atomic positions oscillate and thus R_i is only the average value of a distance distribution. For low enough temperatures, i.e. in the harmonic approximation limit, this distribution is well approximated by a Gaussian function of width proportional to the Debye-Waller factor (this is the origin of the $e^{-2\sigma_i k^2}$ term in equation 4.3). The analysis of crystalline samples is usually made describing the coordination around the photoabsorber atom using these Gaussian shells. On the other hand, at elevated temperatures or in disordered systems, such as aqueous ionic solutions, the distribution functions become broad and asymmetric towards the large distances, the harmonic approximation is no longer valid and the appropriate description of these systems can be performed in terms of radial distribution functions ($g(r)$). When an asymmetric distribution is present the first peak of the radial distribution functions can be modeled using a set of Gamma functions. These functions are described by an average distance R, a coordination number (N_c), a standard deviation σ, and an asymmetry factor (skewness) $\beta = 2p^{\frac{1}{2}}$. Their general expression is given by:

$$f(r) = N_c \frac{p^{\frac{1}{2}}}{\sigma\Gamma(p)} \left[p + \frac{r-R}{\sigma} p^{\frac{1}{2}}\right]^{(p-1)} e^{-\left[p + \frac{r-R}{\sigma} p^{\frac{1}{2}}\right]} \tag{4.4}$$

where $\Gamma(p)$ is the Euler Gamma function associated to the parameter p.

The EXAFS spectroscopy is particularly suited to the study of the local environment around a photoabsorber atom, such as an ion in aqueous solution, since the EXAFS signal depends only on the distribution functions related to photoabsorber atom; this is one of its main advantages over diffraction techniques where the structure factor $S(q)$ is the superimposition of the $N(N+1)/2$ different distribution functions associated to the N atoms of the systems, and it is very difficult to isolate single contributions. Therefore, in systems like ionic solutions the greater contribution to the structure factor is from bulk water and diffraction techniques can be employed only in rather concentrated solutions (1-2 M), while EXAFS can be used at very low concentrations. Moreover, EXAFS provides values of bond distances with very high accuracy, typically of the order of 0.01 Å, about one order of magnitude greater than the majority of diffraction techniques. On the other hand, EXAFS can only give short range (up to 4-5 Å) information, as it has been already discussed in the previous section, and, in the case of disordered systems, the fitting parameters are often correlated and the EXAFS data analysis can lead to ambiguous result. A strategy to help in the extraction of the structural details contained in the EXAFS spectra of disordered systems is to include independent information derived from computer simulations. In particular, in recent years it has been shown that EXAFS data analysis of ions and molecules in solution can derive strong benefit by using the radial distribution functions calculated from Molecular Dynamics simulations as starting models. In this case the theoretical signal $\chi(k)$, associated for example to the ion-oxygen distribution in an aqueous ionic solution, is expressed as a function of the ion-oxygen $g(r)$ calculated from the Molecular Dynamics trajectories:

$$\chi(k) = \int_0^\infty dr 4\pi r^2 \rho g(r) A(k,r) \sin(2kr + \phi(k,r)) \tag{4.5}$$

where $A(k,r)$ and $\phi(k,r)$ are amplitude and phase functions and ρ is the density of scattering atoms. With such a procedure, it is possible to analyze the EXAFS data using a realistic structural model and including the contribution of the second hydration shell.

The theoretical signal is then compared with the experimental one by minimizing the following function:

$$R_i(\{\lambda\}) = \sum_{i=1}^{N} \frac{\left[\alpha_{exp}(E_i) - \alpha_{mod}(E_i;\lambda_1 \ldots \lambda_p)\right]^2}{\sigma_i^2} \tag{4.6}$$

where N is the number of experimental points E_i, $\{\lambda_i\}$ is the set of p parameters that are optimized and σ_i^2 is the variance associated to each experimental datum $\alpha_{exp}(E_i)$. On the basis of the final value of $R_i(\{\lambda\})$ and of the agreement between the experimental and theoretical spectra, the correctness of the starting $g(r)$ can be evaluated. Therefore, this combined EXAFS-Molecular Dynamics approach on the one hand allows one to verify the reliability of the Molecular Dynamics simulations by comparing the theoretical results with the experimental data, on the other hand provides a useful starting model in the EXAFS analysis of disordered systems. The structural parameters of the starting model can be fitted in order to obtain the better possible agreement with the experimental data, and an accurate description of the first coordination shell can thus be obtained.

It is important to stress that besides the radial (two-body) distribution functions, also information on three-body, four-body ... distribution functions can be obtained from the EXAFS analysis, that are calculated by means of the MS theory as implemented in the GNXAS software package [76]. Obviously, this is possible only when MS processes provide a detectable contribution to the EXAFS experimental spectrum.

4.3 XANES analysis

The XANES region of the spectrum is extremely sensitive to the geometric environment of the absorbing atom and, in principle, an almost complete recovery of the three-dimensional structure can be achieved from it. The possibility to gain structural information from the XANES spectra is extremely important for dilute and biological systems where the low signal-to-noise ratio of the experimental data hampers a reliable analysis of the EXAFS region. Moreover, in the study of disordered systems coordination numbers cannot be accurately determined from the EXAFS data due to their large correlation with the Debye-Waller factors, and for these systems the analysis of the XANES region can thus be essential to address some of the shortcomings of EXAFS. However, the analysis of the low-energy part is much more difficult to be performed and requires the use of heavy time-consuming algorithms to calculate the absorption cross section in the framework of the full MS approach. For this reason, this technique has been for a long time used as a qualitative method and only some years ago a method has been proposed in the literature which performs a quantitative analysis of XANES. In particular a new software procedure, named MXAN, has been developed [77]. This method is based on the comparison between the XANES experimental spectrum and several theoretical calculations performed by varying selected structural parameters associated with a given starting model. Starting from a putative geometrical configuration around the photoabsorber atom, the MXAN package is able to reach the best-fit conditions in a reasonable time, by minimizing a residual function R_{sq} in the space of the structural and non-structural parameters defined as:

$$R_{sq} = n \frac{\sum_{i=1}^{m} w_i \left[(y_i^{\mathrm{th}} - y_i^{\mathrm{exp}}) \varepsilon_i^{-1} \right]^2}{\sum_{i=1}^{m} w_i} \tag{4.7}$$

where n is the number of independent parameters, m is the number of experimental points, y_i^{th} and y_i^{exp} are the theoretical and experimental values of the cross section, ε_i is the error on each experimental point and w_i the statistical weights. The X-ray absorption cross section is calculated using the full MS scheme within the muffin-tin approximation for the shape of the potential. The exchange and correlation part of the potential are determined on the basis of the local density approximation of the self-energy of the photoelectron using an appropriate complex optical potential. The real part of the self-energy is calculated either by the Hedin-Lundqvist energy-dependent potential or by the X_α approximation. However, to avoid over-damping at low energies due to the complex part of the Hedin-Lundqvist potential, the MXAN method can account for all inelastic processes by convolution with a broadening Lorentzian function having an energy-dependent width $\Gamma_{\mathrm{tot}}(E)$ of the form [77]:

$$\Gamma_{\mathrm{tot}}(E) = \Gamma_c + \Gamma_{\mathrm{mfp}}(E)$$

The constant part Γ_c accounts for both the experimental resolution and the core-hole lifetime, while the energy dependent term $\Gamma_{\mathrm{mfp}}(E)$ represents all the intrinsic and extrinsic inelastic processes. The $\Gamma_{\mathrm{mfp}}(E)$ is zero below an energy onset E_s (which, in extended systems, correspond to the plasmon excitation energy), and starts increasing from a given value A, following the universal form of the mean free path in solids [78].

The MXAN procedure has been successfully applied to the study of several systems, both in the solid and liquid state, allowing a quantitative extraction of the relevant geometrical information about the absorbing site [79, 80, 81]. However, in the case of ionic solutions the XANES spectra have been usually computed reducing the system to a single structure since the contribution from molecules and arrangements instantaneously distorted cannot be calculated using the analysis standard methods. A promising strategy to overcome this problem is to analyze the XANES spectra using the microscopic dynamical description of the system derived from Molecular Dynamics simulations. In this framework, we have developed a computational procedure which uses MXAN and Molecular Dynamics simulations to generate a configurational averaged XANES spectrum and we have applied it to the study of ionic aqueous solutions.

Chapter 5

Ab initio generation of cation-water effective potentials

5.1 Introduction

As already mentioned in the Introduction chapter, in this work the structural and dynamic properties of the Zn^{2+}, Cd^{2+}, and Hg^{2+} aqua ions have been studied by means of classical Molecular Dynamics simulations. In particular, Zn^{2+} aqueous solutions were investigated at normal pressures and in high-density water, while Cd^{2+} and Hg^{2+} aqueous solutions have been simulated at ambient conditions, since the hydration structure of these ions are still poorly defined even at 1 bar and 300 K. This study has been carried out further developing and applying a general procedure which combines ab-initio calculations, classical Molecular Dynamics simulations and X-ray absorption spectroscopy. The validity of this methodology has been already assessed by an extensive study on aqueous solutions of Zn^{2+}, Co^{2+} and Ni^{2+} at ambient conditions [6, 7].

The general scheme of the procedure used in this work is shown in Figure 5.1. The first step of the method consists in the generation of a Potential Energy Surface (PES) of the M^{2+}-H_2O system by means of ab initio calculations. Then a two-body ion-water potential is calculated by fitting the ab initio energies with an analytical function and this potential is subsequently used to carry out the Molecular Dynamics simulations. Finally, the theoretical results obtained by the analysis of Molecular Dynamics simulations are compared with the XAS experimental data.

Molecular Dynamics simulations are a powerful tool for the investigation of structural and dynamic properties of ionic aqueous solutions, but a proper choice of the interaction potentials among all of the atoms of the system is always a mandatory prerequisite for obtaining reliable results. Much effort has been made to generate water-water pair potentials which have been determined either empirically or have been fitted on the results of electronic structure calculations. As a result of this effort many water models are now available in the literature. As far as the ion-water interactions are concerned, the most simple approach to treat them consists in using simple two-body

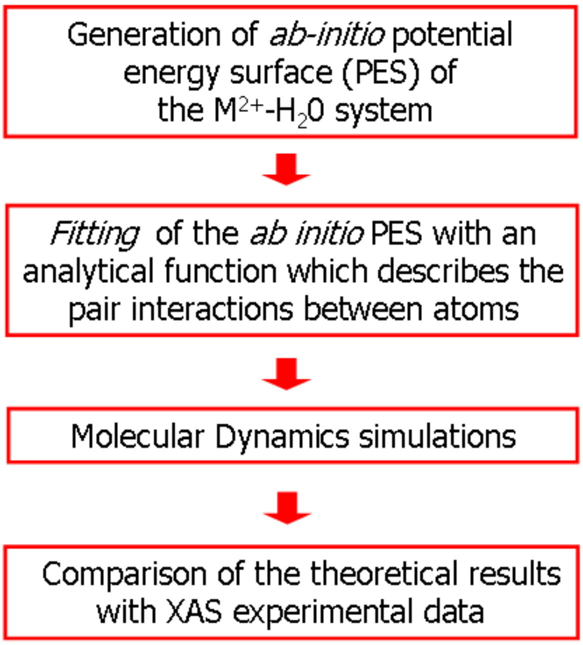

Figure 5.1: General outline of the procedure used in this work.

potentials, implicitly assuming that interactions between the different species are additive. However, these interactions in reality are not additive, and this is particularly true when dealing with doubly or triply charged cations. In this case the non additivity of the pair potentials has two fundamental causes. The first one concerns the long range behaviour of the M^{n+}-H_2O ground state potential: the electronic affinity of M^{n+} (i.e. the n^{th} ionization potential of M) is generally greater than the water ionization potential (which is 12.615 eV) [52]. At great distances the ground state of the system is represented by the $M^{(n-1)+}$-$(H_2O)^+$ charge transfer configuration [82]. Therefore, the long range interaction is repulsive while at short distances this charge transfer is only partial and the ground state of the system is M^{n+}-H_2O. The presence of an avoided crossing between charge transfer and non-charge transfer states can be important in vacuum while it is almost irrelevant in solution, since the $n+$ cation charge is strongly stabilized by the presence of the other water molecules. The many body effects produce thus a qualitative change of the interaction potential form. The second reason of non additivity concerns the strength of the interaction between the ion and the second, third,… nth water molecule. Two fundamental contributions to the ion-water interaction, i.e. the induction and the charge transfer, are in fact non additive. The ion-water interactions tend to be overestimated in simulations based on simple two-body potentials which completely neglect many body contributions, leading to coordination numbers that are too large and ion-water distances that are too short.

The problem of non additivity can be dealt with in several ways. The

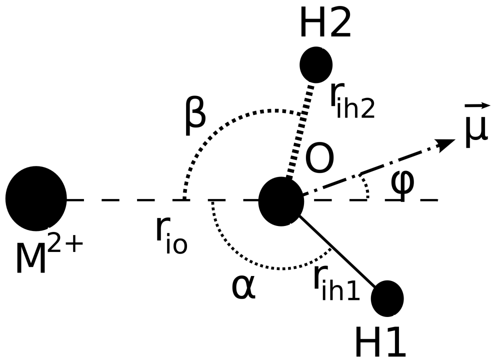

Figure 5.2: Geometric parameters of the Hg^{2+}-H_2O system. r_{io}, r_{ih1}, and r_{ih2} are the ion-oxygen, ion-hydrogen1 and ion-hydrogen2 distances, respectively, while α and β are the M^{2+}-O-H1 and M^{2+}-O-H2 bond angles. The *tilt angle* φ is the angle formed by the M^{2+}-O direction and the water dipole (μ).

more rigorous approach consists in taking into account many body contributions in the simulation, but even including only the three body terms has serious drawback from the point of view of the computational cost, especially if very long simulations are required as in this case. Another strategy is the calculation of the interaction potential between a water molecule and the entire M^{n+}-H_2O complex; this approach is based on the recognition that non additive effects are essentially due to first shell water molecules. However, the limit of this method is that it requires an a priory knowledge of the coordination number (which is generally what one wants to obtain as a result of the investigation). Another possible approach, which has been used in this work, consists in generating an effective two-body potential by including the many-body contributions in an implicit way. This is possible through ab initio calculations which simulate the presence of the solvent by means of the Polarizable Continuum Model (PCM), as it has been described in section 2.9.

5.2 PES generation

5.2.1 Ab initio protocol

For the Zn^{2+} aqua ion the ion-water interaction potential already used to perform the Molecular Dynamics simulation at ambient conditions has been employed, as it has been shown that it is able to provide a good description of Zn^{2+} aqueous solutions [6]. The ab initio calculations were carried out at the RHF level using the LANL2DZ effective potential for the metal ion [83, 84] and the cc-pVTZ basis set for the oxygen and hydrogen atoms of the water molecule [85]. For the Cd^{2+} ion, the ion-water potential previously determined by our group has been used [11]. This potential was developed using the same protocol of that employed in the case of the Zn^{2+} aqua ion.

On the other hand, since Hg^{2+} is far more heavier than Zn^{2+} and Cd^{2+}, the calculations on this cation have been carried out using a post-SCF method

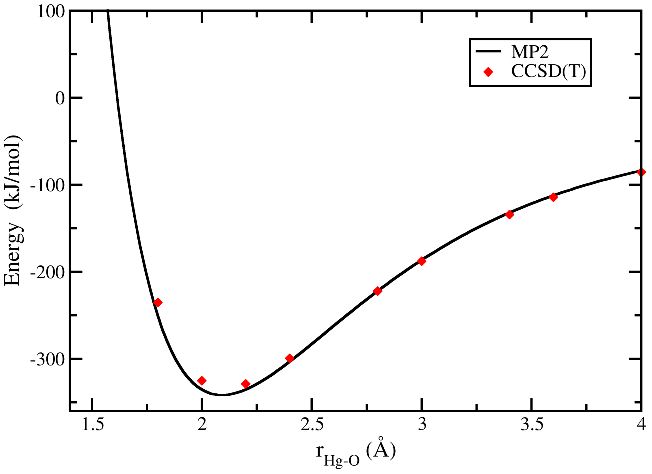

Figure 5.3: Comparison of the MP2 and CCSD(T) binding energies for the Hg^{2+}-H_2O system in its minimum energy configuration.

and more extended basis sets. In particular 2^{nd} order Møller Plesset method (MP2) has been employed (see section 2.6.2), whose validity has been assessed for the description of similar kind of systems [86]. As far as the basis sets are concerned, LANL2 effective core potential (ECP) was retained even in this case to describe the core electrons of the cation, since for Hg^{2+} is very important to account for scalar relativistic effects that are included in this ECP. However, to allow for a greater flexibility, the valence part was replaced with an optimized valence basis set which describes the $5d$, $6s$, $6p$ orbitals and includes multiple d, f and g functions. This basis set has been developed by Prof. Vincenzo Barone of the Normale University of Pisa and it is reported in Table 5.1. The water molecule has been described by the cc-pVTZ basis set, augmented by diffuse s and p functions on the oxygen atom taken from the aug-cc-pVTZ [85]. Test computations revealed that at this level the basis set superposition error (BSSE) is very low (2-3 kJ/mol near the energy minimum), and furthermore have shown that the very reliable CCSD(T) (Coupled Cluster Singles and Doubles with non-iterative account of Triple excitations) and the less CPU demanding MP2 provide essentially equivalent results in the minimum energy configuration of the system, which corresponds to the geometry in which the α angle defined in Figure 5.2 is 127.745° and all of the atoms are coplanar. The results for this analysis are shown in Figure 5.3. Therefore, the computations on the Hg^{2+} aqua ion have been performed by means of the MP2 method. All the ab initio calculations were carried out using the GAUSSIAN03 code [87].

Hg^{2+} Basis set			
Shell type	Number of primitives	Exponents	Coefficients
S	5		
		36.6964560	-0.00010
		5.55024360	-0.00045
		2.96428670	-0.03047
		1.58382950	0.17165
		0.84642317	-0.48393
SP	1		
		0.84642317	1.0 1.0
SP	1		
		0.45238567	1.0 1.0
SP	1		
		0.24179521	1.0 1.0
SP	1		
		0.12923879	1.0 1.0
SP	1		
		0.06907803	1.0 1.0
D	5		
		36.69645600	-0.00097
		5.55024360	-0.00459
		2.96428670	0.00269
		0.24179521	0.09832
		0.12923879	0.02794
D	1		
		1.58382950	1.0
D	1		
		0.84642317	1.0
D	1		
		0.45238567	1.0
F	1		
		1.58382950	1.0
F	1		
		0.45238567	1.0
G	1		
		1.58382950	1.0

Table 5.1: Hg^{2+} valence basis set developed by Prof. Vincenzo Barone of the Normale University of Pisa.

5.2.2 PCM details

The PCM represents the solvent as an isotropic dielectric medium, with a cavity of suitable shape to accommodate the solute (see section 2.9). It is used in this work to include in an implicit way many-body contributions in the ion-water potential. In this context, the non-additivity of intermolecular potential can be seen as a modification of the ion-water pair interaction induced by the environment. According to such a view, an effective pair potential U_{MW} can be determined by performing supermolecule calculations for the ion-water system, in which the influence of the environment, i.e. the surrounding water molecules, is introduced by means of the PCM. The expression of the interaction energy has to fulfil two conditions: it has to contain only terms pertaining to the ion-water pair, and it has to take into account the perturbation due to the environment. Therefore, the Hamiltonian that we have to use is $\mathbf{H}^0$, the Hamiltonian in vacuum, while the wavefunction Ψ should contain the terms originated by the presence of the rest of the solvent, and then it should be an eigenfunction of $\mathbf{H} = \mathbf{H}^0 + \mathbf{H}^\sigma$, where $\mathbf{H}^\sigma$ is the Hamiltonian generated by the polarization charges on the cavity surface:

$$U_{MW} = \langle\Psi|\mathbf{H}^{(0)}|\Psi\rangle_{MW} - \langle\Psi|\mathbf{H}^{(0)}|\Psi\rangle_{W} - \langle\Psi|\mathbf{H}^{(0)}|\Psi\rangle_{M} \tag{5.1}$$

where the M, W and MW subscripts refer to the bare ion, the single water molecule and the mono-hydrated complex. The main sources of non-additivity are the polarization of the water molecule in the electric field of the cation and the electron transfer from water to the cation: both effects give place to binding interactions and both are reduced if the cation is solvated by other water molecules. In the PCM, negative surface charges surround the cation, and their electric field opposes both the polarization of the water molecule and the charge transfer. Thus the PCM gives qualitatively the right behaviour, but the extent of the perturbation, i.e. the difference between the wavefunction calculated in vacuum or with the PCM, depends on some PCM parameters. The most important of these parameters is the shape of the PCM cavity. The cavity for the ion-water system is modeled here as a set of interlocking spheres centered around each atom. For neutral species, the Van der Waals radii are generally employed. For this reason we have used for the oxygen and hydrogen atoms the following Van der Waals radii:

$$\begin{aligned} \rho_O &= 1.68 \text{ Å} \\ \rho_H &= 1.44 \text{ Å} \end{aligned}$$

The choice of the radius for the cation cavity (ρ_m) is somewhat more arbitrary and in order to determine it we have applied a procedure developed by Floris et al. based on an internal consistency criterion [52]. In particular, we choose ρ_m such that the following relationship is verified:

$$U_{WMW} = 2U_{MW} + U_{WW} \tag{5.2}$$

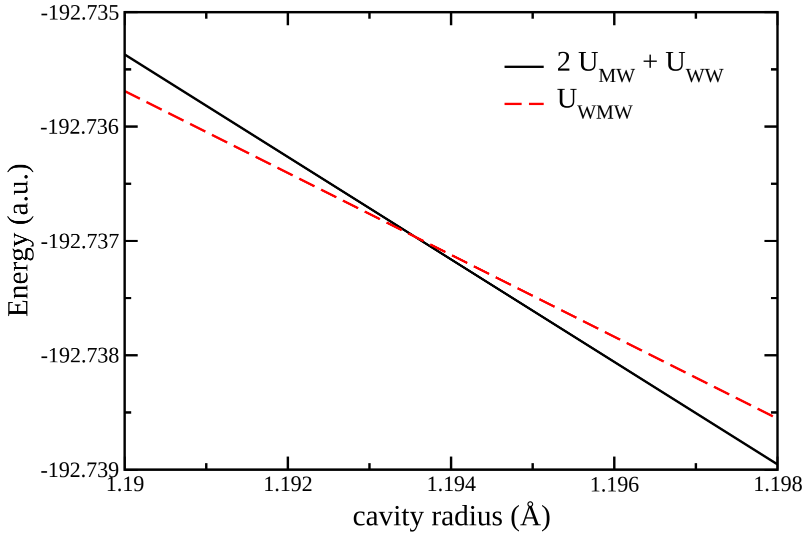

Figure 5.4: Optimization of the cavity radius for the Hg^{2+} ion.

where U_{MW} has been defined in equation 5.1, while U_{WMW} and U_{WW} are given by:

$$U_{WMW} = \langle\Psi|\mathbf{H}^{(0)}|\Psi\rangle_{WMW} - 2\langle\Psi|\mathbf{H}^{(0)}|\Psi\rangle_{W} - \langle\Psi|\mathbf{H}^{(0)}|\Psi\rangle_{M} \quad (5.3)$$

$$U_{WW} = \langle\Psi|\mathbf{H}^{0}|\Psi\rangle_{WW} - 2\langle\Psi|\mathbf{H}^{0}|\Psi\rangle_{W} \quad (5.4)$$

The subscripts refer to the systems on which the calculations are carried out, so WMW is a system composed by two water molecules and one cation, WW is composed by two water molecules and so on. Both sides of equation 5.2 depend on the radius ρ_m (in particular decrease as the radius increases), but U_{MW} is more sensitive than U_{WMW} to this parameter and the equation will be satisfied only for a certain value of ρ_m. In order to determine ρ_m, $\langle\Psi|\mathbf{H}^{(0)}|\Psi\rangle_{WMW}$, $\langle\Psi|\mathbf{H}^{(0)}|\Psi\rangle_{MW}$ and $\langle\Psi|\mathbf{H}^{(0)}|\Psi\rangle_{M}$ have to be computed as a function of ρ_m, by keeping fixed the geometry of the M^{2+}-H_2O and M^{2+}-$(H_2O)_2$ systems. For M^{2+}-H_2O the minimum energy configuration was chosen with an ion-water distance of 2.2 Å. The energy of the M^{2+}-$(H_2O)_2$ system has been instead computed with a face-to face C_{2v} geometry of the two ligands, using always the same value of α and ion-oxygen distance, and with an O-M^{2+}-O angle of 90°, as suggested by Floris et al. in Ref. [52]. Figure 5.4 shows the left- and right-hand sides of equation 5.2 calculated for the Hg^{2+}-water system as functions of the sphere radius on the cation. As it can be seen, the two quantities coincide when ρ_m=1.194 Å. This values of the cavity radius will then be used for the potential energy surface scan. To include solvent effects the Conductor like PCM (CPCM) [50] has been employed (see section 2.9).

5.2.3 Surface scan

A grid of points on the PES of the ion-water system has now to be chosen and the interaction energies related to these configurations have to be calculated. The water geometry was kept fixed to the experimental value during the PES scan (H1-O-H2 angle=104.51°, R_{OH}=0.9575 Å). The grid of points has been obtained by varying the ion-oxygen distance r_{io} from 1.2 Å to 4.0 Å with a step of 0.02 Å, and the α angle in the range of $7.745° \leq \alpha \leq 127.745°$, with a step of 10°, while keeping the ion in the water molecular plane (see Figure 5.2). We have thus obtained 1833 configurations of the system. Then we have performed a series of scans by moving the H2 atom outside from the M-O-H1 plane and towards the cation, i.e. by varying the dihedral angle formed by the M-O-H1 and M-O-H2 planes. Other 330 configurations were obtained for values of the dihedral angle of 150°, 135°, 120°, 90°, 60° and 45°, α angle between 127.745° and 7.745° (step of 30°) and r_{io} distance between 1.7 Å and 2.7 Å (step of 0.1 Å).

A similar grid was used by our group to determine the Cd^{2+}-water effective potential [11], while for the Zn^{2+} ion only configurations whith all the atoms coplanar were chosen since in this case the water molecules are strongly restrained in their movements by the quite strong ion-water interaction [6].

5.3 Pair potential calculation

Once obtained the PES, the two-body ion-water potential is calculated by fitting the ab initio energies with a suitable analytical function. The fitting was carried out using the Newton method as implemented in the SAS statistical package [88]. The Newton iterative method is used to solve systems of non-linear equations and it works by regression of the residuals onto a function of the first and second derivatives of the model with respect to the parameters, until the estimates converge [89]. In contrast to other available non-linear fitting methods (such as the Steepest descent) the Newton one uses the Hessian with respect to the parameters, thus allowing the error minimum to be reached more efficiently and with higher precision.

At this point of the procedure a water model has to be chosen, since the form of the analytical function used for the fitting depends on this particular choice. To perform the Molecular Dynamics simulations of the Cd^{2+} and Hg^{2+} aqua ions we have employed two of the most widely used water models, i.e. the SPC/E [90] and TIP5P [91], while the simulations of Zn^{2+} ion were carried out only with the SPC/E water model. The geometries of the water molecule as described by the SPC/E and TIP5P models are shown in Figure 5.5. In the case of the SPC/E water model, the ab initio single point energies were fitted using the following analytical function for the ion-water interaction:

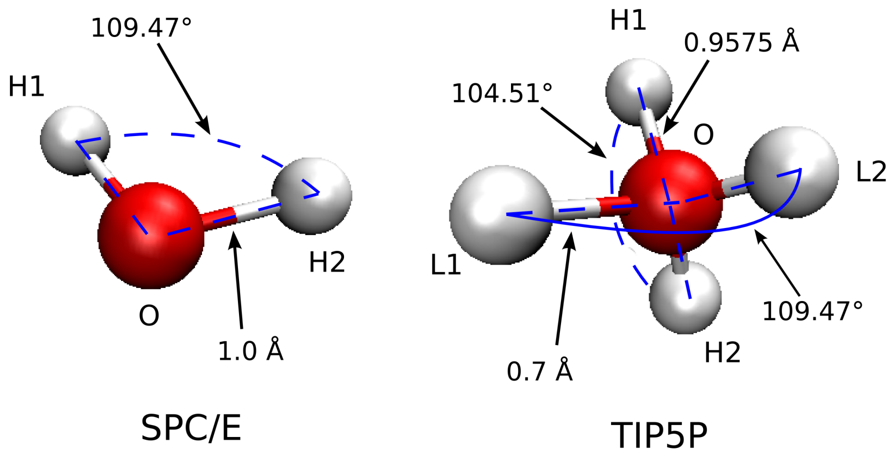

Figure 5.5: Geometry of the water molecule in the SPC/E (left) and TIP5P (right) models.

$$\begin{aligned} V(r) &= \frac{q_i q_o}{r_{io}} + \frac{A_o}{r_{io}^4} + \frac{B_o}{r_{io}^6} + \frac{C_o}{r_{io}^8} + \frac{D_o}{r_{io}^{12}} + E_o e^{-F_o r_{io}} \\ &+ \sum_{ih=ih1,ih2} \frac{q_i q_h}{r_{ih}} + \frac{A_h}{r_{ih}^4} + \frac{B_h}{r_{ih}^6} + \frac{C_h}{r_{ih}^8} + \frac{D_h}{r_{ih}^{12}} \end{aligned} \tag{5.5}$$

where r_{io}, r_{ih1}, and r_{ih2} are the ion-water distances and q_i, q_o, and q_h are the electrostatic charges of the cation (-2 a.u.) and of the oxygen and hydrogen atoms in the SPC/E water model (-0.8476 and 0.4238 a.u., respectively). $A_o, \ldots, F_o$, and $A_h, \ldots, D_h$ are the unknown parameters.

The TIP5P water model has two negative charges centered on two dummy atoms, instead of the oxygen atom, and therefore the two-body potential has been adapted to the following form:

$$\begin{aligned} V(r) &= \frac{A_o}{r_{io}^4} + \frac{B_o}{r_{io}^6} + \frac{C_o}{r_{io}^8} + \frac{D_o}{r_{io}^{12}} + E_o e^{-F_o r_{io}} \\ &+ \sum_{ih=ih1,ih2} \frac{q_i q_h}{r_{ih}} + \frac{A_h}{r_{ih}^4} + \frac{B_h}{r_{ih}^6} + \frac{C_h}{r_{ih}^8} \\ &+ \sum_{iL=iL1,iL2} \frac{q_i q_L}{r_{iL}} + \frac{A_L}{r_{iL}^4} + \frac{B_L}{r_{iL}^6} + \frac{C_L}{r_{iL}^8} \end{aligned} \tag{5.6}$$

where the variables have the same meaning as in equation 5.5; r_{iL1} and r_{iL2} are the ion-dummy atom distances and q_L is the electrostatic charge of the dummy atoms (q_h and q_L are equal to +0.241 and -0.241 a.u., respectively).

The trend of the fitted Hg^{2+} ion-water interaction potential for the SPC/E water model is shown in Figure 5.6. An almost identical result was obtained in the case of the TIP5P model. All the parameters calculated from the fitting procedure are reported in Table 5.2, 5.3 and 5.4 for the Hg^{2+}, Cd^{2+} and Zn^{2+} ions, respectively.

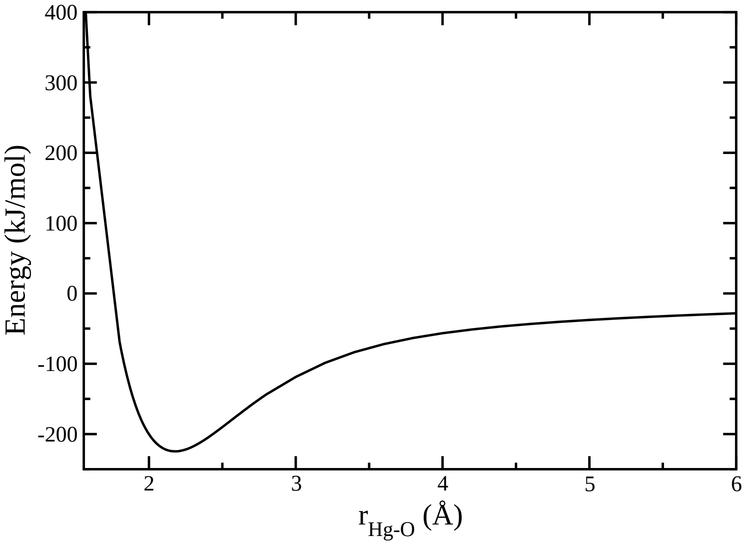

Figure 5.6: Fitted potential energy curve for the Hg^{2+}-SPC/E water interaction as a function of the ion-oxygen distance in the minimum energy configuration of the system.

	SPC/E		TIP5P	
	parameters	std. dev.	parameters	std. dev.
A_o	$1.2712 \cdot 10^{0}$	$4.50 \cdot 10^{-2}$	$0.3718 \cdot 10^{0}$	$6.40 \cdot 10^{-2}$
B_o	$-1.181 \cdot 10^{-1}$	$5.31 \cdot 10^{-3}$	$0.762 \cdot 10^{-1}$	$4.05 \cdot 10^{-3}$
C_o	$8.433 \cdot 10^{-4}$	$7.40 \cdot 10^{-5}$	$-1.021 \cdot 10^{-3}$	$5.52 \cdot 10^{-5}$
D_o	$-1.178 \cdot 10^{-8}$	$3.58 \cdot 10^{-9}$	$4.493 \cdot 10^{-8}$	$2.54 \cdot 10^{-9}$
E_o	$2.130 \cdot 10^{+6}$	$6.86 \cdot 10^{+4}$	$-6.681 \cdot 10^{+5}$	$3.97 \cdot 10^{+3}$
F_o	$3.952 \cdot 10^{+1}$	$3.02 \cdot 10^{-1}$	$2.015 \cdot 10^{+1}$	$2.25 \cdot 10^{-1}$
A_h	$4.408 \cdot 10^{-2}$	$2.07 \cdot 10^{-3}$	$1.169 \cdot 10^{-1}$	$1.85 \cdot 10^{-3}$
B_h	$-5.700 \cdot 10^{-4}$	$4.10 \cdot 10^{-5}$	$-6.900 \cdot 10^{-4}$	$4.10 \cdot 10^{-5}$
C_h	$3.372 \cdot 10^{-6}$	$2.29 \cdot 10^{-7}$	$1.606 \cdot 10^{-6}$	$8.87 \cdot 10^{-8}$
D_h	$-2.348 \cdot 10^{-11}$	$1.88 \cdot 10^{-12}$		
A_L			$1.208 \cdot 10^{-1}$	$5.18 \cdot 10^{-3}$
B_L			$-1.590 \cdot 10^{-3}$	$1.07 \cdot 10^{-4}$
C_L			$6.729 \cdot 10^{-6}$	$6.27 \cdot 10^{-7}$

Table 5.2: Estimated Hg^{2+}-H_2O interaction parameters and relative standard deviations for the SPC/E and TIP5P water models.

	SPC/E		TIP5P	
	parameters	std. dev.	parameters	std. dev.
A_o	$5.577 \cdot 10^{-1}$	$3.89 \cdot 10^{-2}$	$-5.892 \cdot 10^{-1}$	$3.47 \cdot 10^{-2}$
B_o	$-6.130 \cdot 10^{-2}$	$4.54 \cdot 10^{-3}$	$7.895 \cdot 10^{-2}$	$2.86 \cdot 10^{-3}$
C_o	$4.909 \cdot 10^{-4}$	$6.30 \cdot 10^{-5}$	$-8.682 \cdot 10^{-4}$	$3.47 \cdot 10^{-5}$
D_o	$-1.200 \cdot 10^{-8}$	$3.02 \cdot 10^{-9}$	$2.938 \cdot 10^{-8}$	$1.74 \cdot 10^{-9}$
E_o	$1.255 \cdot 10^{+6}$	$5.71 \cdot 10^{+4}$	$6.568 \cdot 10^{+4}$	$3.73 \cdot 10^{+3}$
F_o	$3.951 \cdot 10^{+1}$	$4.29 \cdot 10^{-1}$	$2.224 \cdot 10^{+1}$	$2.17 \cdot 10^{-1}$
A_h	$5.364 \cdot 10^{-2}$	$1.51 \cdot 10^{-3}$	$1.797 \cdot 10^{-1}$	$1.08 \cdot 10^{-4}$
B_h	$2.365 \cdot 10^{-4}$	$2.80 \cdot 10^{-5}$	$-8.816 \cdot 10^{-4}$	$1.40 \cdot 10^{-5}$
C_h	$-1.719 \cdot 10^{-6}$	$1.48 \cdot 10^{-7}$	$1.982 \cdot 10^{-6}$	$4.77 \cdot 10^{-8}$
D_h	$1.097 \cdot 10^{-11}$	$1.09 \cdot 10^{-12}$		
A_L			$1.710 \cdot 10^{-1}$	$3.16 \cdot 10^{-3}$
B_L			$2.340 \cdot 10^{-4}$	$6.50 \cdot 10^{-5}$
C_L			$1.085 \cdot 10^{-5}$	$3.82 \cdot 10^{-7}$

Table 5.3: Estimated Cd^{2+}-H_2O interaction parameters and relative standard deviations for the SPC/E and TIP5P water models.

	SPC/E	
	parameters	std. dev.
A_o	$2.049 \cdot 10^{-2}$	$1.10 \cdot 10^{-3}$
B_o	$2.910 \cdot 10^{-2}$	$1.30 \cdot 10^{-3}$
C_o	$-2.834 \cdot 10^{-4}$	$1.70 \cdot 10^{-5}$
D_o	$8.011 \cdot 10^{-9}$	$8.40 \cdot 10^{-10}$
E_o	$-3.633 \cdot 10^{+4}$	$2.00 \cdot 10^{+3}$
F_o	$2.360 \cdot 10^{+1}$	$1.70 \cdot 10^{-1}$
A_h	$8.377 \cdot 10^{-2}$	$1.80 \cdot 10^{-3}$
B_h	$-1.563 \cdot 10^{-3}$	$6.10 \cdot 10^{-5}$
C_h	$1.244 \cdot 10^{-5}$	$5.30 \cdot 10^{-7}$

Table 5.4: Estimated Zn^{2+}-H_2O interaction parameters and relative standard deviations for the SPC/E water model.

Chapter 6

Methods employed in the study of group 12 aqua ions

6.1 Molecular Dynamics simulations

6.1.1 Molecular Dynamics protocol

The simulations of Zn^{2+}, Cd^{2+} and Hg^{2+} aqua ions were carried out using the GROMACS package version 3.2.1 [92], modified in order to include the ion-water effective pair potentials determined by means of the procedure described in the previous chapter. All the simulations were carried out using the same general protocol. The system was composed by one M^{2+} ion and 819 water molecules (for a total of 2458 atoms) in a cubic box, using periodic boundary conditions. A cutoff of 9 Å was used to deal with non bonded interactions, with the Particle Mesh Ewald (PME) method to treat long range electrostatic effects [60, 61]. A homogeneous background charge has been used to compensate for the presence of the M^{2+} ion [93]. The systems were simulated in a NVT ensemble using the Berendsen method [64] with a coupling constant of 0.1 ps. A time step of 1 fs was employed, saving a configuration every 25 time steps and equilibrating the system for 5 ns before sampling. In the following, the reported simulation time do not include this equilibration phase.

The Cd^{2+} and Hg^{2+} aqua ions were simulated at ambient conditions using both the SPC/E and TIP5P water models in order to evaluate the influence of the water-water interactions on the structural and dynamic properties of the solvation shells of these ions. The estimated residence time of water molecules in the first coordination shell of both Cd^{2+} and Hg^{2+} is in the nanosecond time scale [14]. Therefore, very long simulations are required to properly describe the first hydration shell dynamics of these systems. The Cd^{2+} ion is expected to have a slightly longer residence time than Hg^{2+}, and for this reason simulations of 100 ns were performed for the Cd^{2+} ion, while in the Hg^{2+} case the system was simulated for 60 ns.

Three simulations of Zn^{2+} in aqueous solution were carried out in order to reproduce the properties of the system in three different points of the phase diagram, i.e. 300 K and 0.1 MPa (d = 1006.8 g/l), 300 K and 1.0 GPa (d =1271.2 g/l), 360 K and 2.2 GPa (d = 1345.3 g/l). Different box volumes have been employed in order to reproduce the chosen density, temperature and pressure conditions. In particular, the box sides used are 29.0 Å, 27.1 Å and 26.3 Å, for the simulations at 0.1 MPa, 1 GPa and 2.2 GPa, respectively. Each simulation was carried out with the SPC/E water model for 10 ns. It is noteworthy to observe that the estimated first hydration shell residence time is of the microsecond order of magnitude [14], and in the ambient condition simulation previously published no solvent exchange events were observed [6]. As a consequence, a 10 ns simulation is long enough to describe the structural properties of the Zn^{2+} first hydration shell.

6.1.2 Structural analysis

The structural properties of the M^{2+} aqua ions are described in terms of the metal-oxygen and metal-hydrogen radial distribution functions or g(r)'s [53]:

$$g_{AB}(r) = \frac{\langle \rho_B(r) \rangle}{\langle \rho_B \rangle_{local}} = \frac{1}{N_A \langle \rho_B \rangle_{local}} \sum_{i \in A}^{N_A} \sum_{j \in B}^{N_B} \frac{\delta(r_{ij} - r)}{4\pi r^2} \tag{6.1}$$

where $\langle \rho_B(r) \rangle$ is the particle density of type B at a distance r around particles A, $\langle \rho_B \rangle_{local}$ is the particle density of type B averaged over all the spheres around particles A with radius r_{max} (which is equal to half of the box side), N_A and N_B is the number of A and B particles, and r_{ij} is the distance between i (of type A) and j (of type B) particles. Therefore, an average over time and over the A particles is performed in the $g_{AB}(r)$ calculation. In our case A=M^{2+} and B=O or H atoms, and thus only one A particle is present and only the average over time is performed. The coordination numbers N of the ions were calculated by integration of the radial distribution functions up to the first minimum, using the following relation:

$$N_B = 4\pi\rho_B \int_0^{R_{min}} g_{AB}(r) r^2 dr \tag{6.2}$$

where R_{min} is the first minimum of the $g_{AB}(r)$ and ρ_B is the atomic density of B particles in the system.

Angular distribution functions (a.d.f.) have been calculated to evaluate the geometrical arrangement of water molecules around the M^{2+} ion. In particular, a.d.f. have been computed for three different angles, the angle formed by two different M-O vectors in the first shell (labelled as ψ), the angle formed by the water molecule dipole and the M-O direction (labelled as ϕ), and the angle formed by the normal to the water plane and the M-O direction (labelled as ζ). The definition of these three angles is shown in Figure 6.1 and 6.2.

To describe the distribution of water molecules around the M^{2+} ion, axial-radial 2D density maps (or cylindrical distribution functions) can also

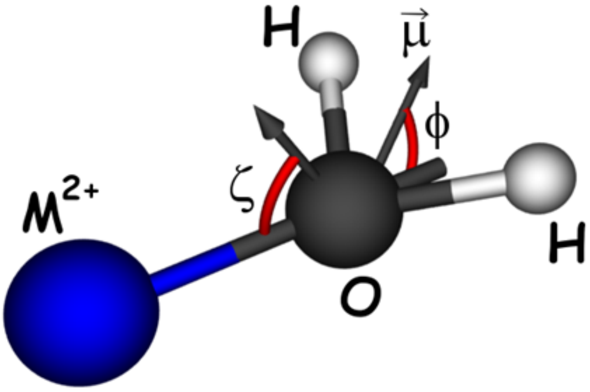

Figure 6.1: Definition of the ϕ and ζ angles used in the structural analysis of the M^{2+} solvation shells.

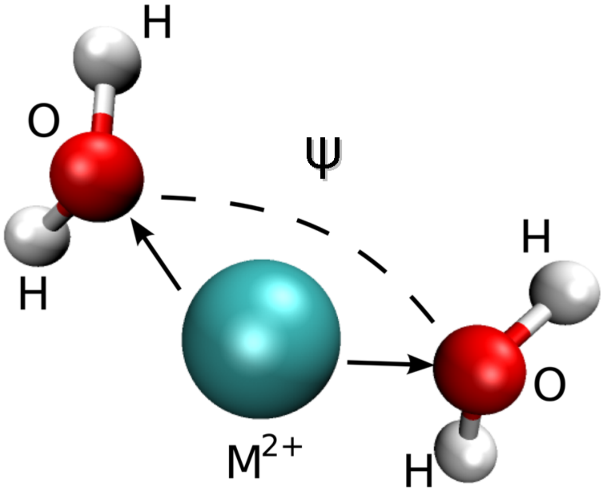

Figure 6.2: Definition of the ψ angle used to evaluate the geometrical arrangement of water molecules in the M^{2+} first hydration shell.

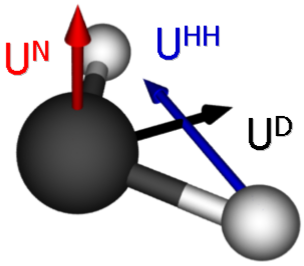

Figure 6.3: Definition of the three vectors used in the study of water reorientational dynamics.

be used. These maps show the distribution of water oxygen atoms with respect to a chosen reference axis containing the ion and the oxygen atom of a selected first shell water molecule. They have been calculated using a radius of 6.0 Å and a grid spacing of 0.1 Å. The structure of the solvation shells can also be visualized using spatial distribution functions (s.d.f.) of the atomic densities. 3D isosurfaces, i.e. surfaces in which the particle (oxygen and hydrogen atom) density is constant, are plotted in a box of 1 nm^3 centered on the ion and using a grid spacing of 0.035 Å [94, 95].

2D density maps and spatial distribution functions have been generated by means of standard GROMACS analysis tools (g_densmap and g_sdf, respectively). Conversely, radial and angular distribution functions have been calculated using in-house written codes.

6.1.3 Dynamic analysis

The mean residence time of water molecules in the ion first hydration shell has been evaluated using the method proposed by Impey et al. in Ref. [96]. This approach is based on the definition of a survival probability function $P_j(t, t_n, t^*)$, which takes the value one if the water molecule j lies within the first hydration shell at both time steps t_n and $t + t_n$ and does not leave the coordination shell for any continuous period longer than t^*, otherwise it is zero. From P_j it is possible to calculate the dynamical hydration number (or survival function) $n_{hyd}(t)$:

$$n_{hyd}(t) = \frac{1}{N_t} \sum_{n=1}^{N_t} \sum_{j} P_j(t, t_n, t^*)$$

where N_t is the total number of steps and the summation goes over all of the water molecules. At long times, $n_{hyd}(t)$ decays in an exponential fashion, with a characteristic correlation time τ which defines the residence time of the water molecule in the shell [96].

The M^{2+} diffusion coefficient, D_M, provides quantitative information on the motion of the ion in the solution and it has been calculated from the mean square displacements using the Einstein relation [58]:

$$D_M = \frac{1}{6} \lim_{t \to \infty} \frac{\left\langle \|\mathbf{r}_M(t) - \mathbf{r}_M(0)\|^2 \right\rangle}{t} \tag{6.3}$$

where $\mathbf{r}_M(t)$ is the position of the ion at time t while $\mathbf{r}_M(0)$ is the ion initial position.

A detailed view of the dynamics of the water molecules surrounding the ion can be obtained using reorientational correlation functions, that are given by [97]:

$$C_l^{\alpha}(t) = \langle P_l(\mathbf{u}^{\alpha}(t) \cdot \mathbf{u}^{\alpha}(0)) \rangle \tag{6.4}$$

where P_l is the l^{th} rank Legendre polynomial, and $\mathbf{u}^{\alpha}(t)$ is an unit vector in a certain direction α at time t. By fitting $C_l(t)$ to an exponential function $C_l(t) = e^{-\frac{t}{\tau_l}}$, it is possible to determine the correlation time τ_l, defined as the rotation time of the $\mathbf{u}^{\alpha}$ vector. For our analyses we used $l = 1$ and three different directions: the normal to the water molecular plane passing through the center of the oxygen atom $\mathbf{u}^N$, a vector along the HH direction $\mathbf{u}^{HH}$, and the water dipole vector $\mathbf{u}^D$ (see Figure 6.3 for the definition of these vectors); $\mathbf{u}^D$ is correlated to the dielectric relaxation rates while $\mathbf{u}^{HH}$ to 1H-1H NMR dipolar relaxation experiments [98]. Using an approach already implemented for pure water, we have employed a mixed integration exponential fit method to evaluate the first rank correlation times, thus minimizing the noise introduced by the slow convergence of the correlation function tail. Direct integration was used in the initial part of the function, while the tail contribution was taken into account by means of an exponential fit. Due to the different behavior of first shell correlation functions caused by the presence of the ion, we have used different time windows for direct integration: a value of 15 ps was employed for $\mathbf{u}^D$, while the $\mathbf{u}^N$ and $\mathbf{u}^{HH}$ correlation functions were explicitly integrated up to 5 ps. As far as the correlation times of bulk water are concerned, the switch value was always 5 ps, the same as used in Ref. [97].

Standard GROMACS tools have been used to calculate the reorientational correlation functions and diffusion coefficients (g_rotacf and g_msd, respectively), while an in-house written code was employed for the determination of residence times of water molecules.

6.2 X-ray absorption spectroscopy

6.2.1 X-ray absorption measurements

A 0.1M Zn^{2+} aqueous solution was obtained by dissolving the appropriate amount of $Zn(NO_3)_2$ in water. X-ray absorption spectra above the Zn K-edge have been measured in transmission mode using the high pressure setup based on the Paris-Edinburgh press developed at the BM29 beam line [99] of the European Synchrotron Radiation Facility (ESRF). Four spectra have been recorded at different pressures and temperatures following the melting curve of the solution phase diagram. In particular spectra were recorded for pressure values of 0.10 MPa, 0.20 GPa, 1.97 GPa and 2.85 GPa.

A 0.1 M Cd^{2+} water solution was prepared by dissolving the appropriate amount of $Cd(NO_3)_2{\cdot}4H_2O$ in freshly distilled water and adding HNO_3 in order to prevent hydrolysis. Cd K-edge X-ray absorption spectra were obtained using the EMBL spectrometer at DESY [100]. Spectra were recorded in transmission mode using a Si(311) double-crystal monochromator detuned to 30% for harmonic rejection [101]. The DORIS III storage ring was running at an energy of 4.4 GeV with positron currents between 70 and 40 mA. The solution was kept in a cell with Kapton film windows and a Teflon spacer of 3 mm.

A 0.1 M Hg^{2+} water solution was obtained by dissolving the appropriate amount of $Hg(ClO_4)_2$ in freshly distilled water that was acified to about pH=1 by adding $HClO_4$ in order to prevent hydrolysis. Hg L_3 XAS spectra were obtained at the X-ray absorption spectrometer BM29 of the ESRF. Spectra were recorded in transmission mode using a Si(311) double-crystal monochromator detuned to 50%.The solution was kept in a cell with Kapton film windows and Teflon spacers of 4 mm. In all the three cases, the counterions ($(NO_3)^-$ for Zn^{2+} and Cd^{2+} and $(ClO_4)^-$ for Hg^{2+}) have been chosen to prevent the formation of ionic pairs in the solutions.

In all the performed measurements, data points were collected for 1 s each, and three spectra were recorded and averaged.

6.2.2 EXAFS data analysis

In the standard EXAFS analysis of disordered systems only two-body distributions are usually included, and the $\chi(k)$ signal is represented by the equation 4.5. As already mentioned in section 4.2, $\chi(k)$ theoretical signals can be calculated by introducing in equation 4.5 the model radial distribution functions obtained from Molecular Dynamics simulations. In all the aqueous solutions studied, both the M-O and M-H g(r)'s obtained from the simulations have been used to calculate the single scattering first shell $\chi(k)$ theoretical signal, as the ion-hydrogen interactions have been found to provide a detectable contribution to the EXAFS spectra of several metal ions in aqueous solutions [7, 34]. Comparison of the total theoretical and experimental $\chi(k)$ signals allows the reliability of the g(r)'s, and consequently of the theoretical scheme used in the simulations, to be checked. In this case, i.e. when a direct comparison between the signal obtained from Molecular Dynamics simulations and the experimental one is performed, the structural parameters are kept fixed during the minimization, while two nonstructural parameters are optimized: S_0^2, which is a many-body amplitude reduction factor due to intrinsic losses, and E_0, which aligns the experimental and theoretical energy scales. On the other hand, the theoretical $\chi(k)$ signal can also be refined against the experimental data in order to obtain the better possible agreement between the two spectra. In this latter case, the fitting is performed by using a least-squares minimization procedure in which structural and non-structural parameters are allowed to float [102]. Since a correct description of the first coordination sphere of hydrated metal complexes has to account for the asymmetry in the distribution of the ion-solvent distances, the M-O and M-H g(r)'s associated with the first coordination shells were modelled with Γ-like distribution functions which depend on four parameters, namely the coordination number N_c, the average distance R, a standard deviation σ, and the skewness (see section 4.2). Moreover, we have calculated the three-body contributions within the first hydration shell from the $g(r_1, r_2, \psi)$ distributions obtained from the Molecular Dynamics simulations. However, only in the case of the Zn^{2+} ion the three-body contributions have been used in the analysis. In fact, in the other studied systems, MS contributions within the first hydration shell have been found to have negligible amplitude, and therefore they have not been considered in the analysis of the EXAFS data.

The EXAFS theoretical signals have been calculated by means of the GNXAS code which uses an advanced theoretical scheme based on the multiple-scattering formalism [76]. Phase shifts, A(k,r) and ϕ(k,r), have been calculated starting from a configuration extracted from the Molecular Dynamics simulation, by using muffin-tin potentials and advanced models for the exchange-correlation self-energy (Hedin-Lundqvist) [75]. The values of the muffin-tin radii used are 0.2 Å and 0.9 Å for hydrogen and oxygen, and 1.2, 1.4 and 1.5 Å for zinc, cadmium and mercury, respectively. Inelastic losses of the photoelectron in the final state have been accounted for intrinsically by complex potential. The imaginary part also includes a constant factor accounting for the core-hole width.

6.2.3 XANES data analysis

The XANES spectra at high energies are strongly broadened by the core-hole width and the structural and electronic details are smeared out from the spectra [103] (the energy of the Cd K-edge is 26711 eV and the core-hole width is 7.28 eV, while the energy of the Hg L_3-edge is 12290 eV with a core-hole width of 5.5 eV). Recently, suggestions to apply core-hole width deconvolution methods to analyze X-ray absorption spectra have appeared in the literature [104] as this treatment largely facilitates the detection of spectral features and the comparison with theoretical calculations. For this reason, the Cd K-edge and Hg L_3-edge raw experimental data have been deconvolved of the whole tabulated core hole width and a Gaussian filter with full width at half maximum of about 3 eV and 1.8 eV for the Cd and Hg edges, respectively, has been applied. The comparisons between the raw and the deconvolved spectra, depicted in Figure 6.4 and 6.5 for Cd and Hg, respectively, show that after deconvolution the threshold region is considerably sharpened with respect to the original spectrum. It is important to underline that even if in principle the deconvolution procedure could introduce small distortions in the experimental data, the advantage of this approach is to avoid the use of the phenomenological broadening function Γ(E) in the calculation of the

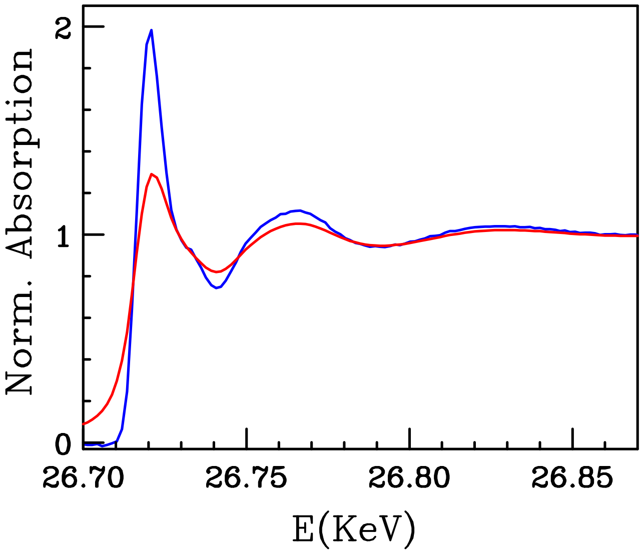

Figure 6.4: Cd K-edge X-ray absorption spectrum of Cd^{2+} aqueous solution before (red line) and after (blue line) deconvolution of a Lorentzian with width of Γ = 7.28 eV.

theoretical spectrum to mimic electronic damping. This method is known to be a shortcoming in the XANES analysis and the deconvolution procedure allows one to overcome, at least partially, this problem.

A standard analysis of the Hg XANES spectra has been performed by carrying out static fits of the data using an average first shell structure and optimizing the Hg-O distances. The XANES data analysis was carried out with the MXAN code [77]. The X-ray absorption cross section was calculated in the framework of the full multiple-scattering scheme within the muffin-tin approximation for the shape of the potential. The real part of the exchange term was calculated using the Hedin-Lundqvist energy-dependent potential, while the inelastic losses are accounted for by convolution with a broadening Lorentzian function (see section 4.3). The constant part Γ_c accounts for the core-hole life time and it has not been included in our calculations as it has been removed from the experimental data, while the energy-dependent term represents all the intrinsic and extrinsic inelastic processes. A least-square fit of the experimental data in the space of the structural and non structural parameters was achieved using the MINUIT routine of the CERN library [105], which minimizes the R_{sq} function defined in equation 4.7. The resolution broadening was taken into account using a Gaussian function.

As already mentioned in section 4.3, we have developed a computational procedure which allows one to analyze the XANES spectra using the microscopic dynamical description of the system derived from Molecular Dynamics simulations. In this work we have applied this methodology to the study of the Cd^{2+} and Hg^{2+} aqua ions. In the first step of our procedure, the XANES spectrum associated with each Molecular Dynamics configuration has been calculated using only the real part of the HL potential, i.e. theoretical spectra do not account for any intrinsic and extrinsic inelastic process, while the damping associated with the experimental resolution is accounted for by con-

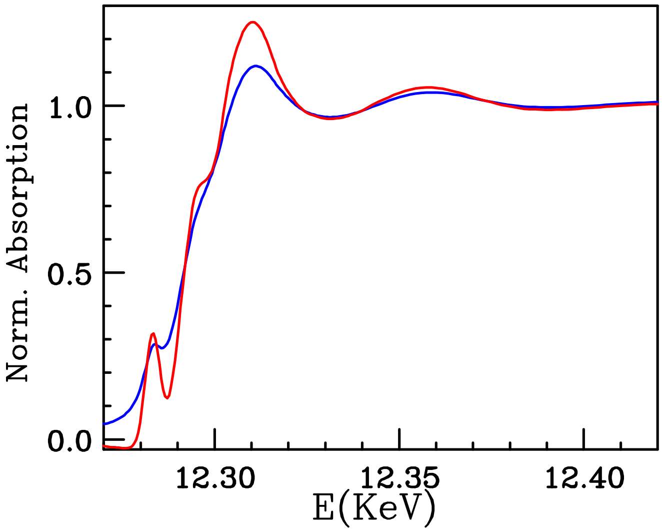

Figure 6.5: Hg L_3-edge X-ray absorption spectrum of Hg^{2+} aqueous solution before (blue line) and after (red line) deconvolution of a Lorentzian with width of $\Gamma = 5.5$ eV.

volution with a Gaussian function with full width at half maximum of 3 eV and 2 eV for Cd^{2+} and Hg^{2+}, respectively. In the second step, to perform a comparison with the experimental data the damping associated with the inelastic processes has to be included in the calculation. To this purpose we have modified the MXAN program in order to read an external theoretical spectrum (the configurational averaged calculated data) and to perform a minimization only in the non-structural parameter space. As in the static fit case, the inelastic losses are accounted for by convolution with a broadening Lorentzian function and does not include the core-hole life-time broadening as it has been previously eliminated from the experimental spectrum. Least-square fits of the XANES experimental data have been performed always by minimizing the R_{sq} function (equation 4.7), but in this case the minimization is carried out only in the non-structural parameter space. Since a slightly different approach has been used for the XANES analysis of the Cd^{2+} and Hg^{2+} aqua ions, all the details on the computational procedure used will be given for each of them separately in section 7.4.2 and 8.3.1.

Chapter 7

Structural and dynamic properties of the Hg^{2+} aqua ion

7.1 Molecular Dynamics results

Two Molecular Dynamics simulations have been carried out for Hg^{2+} using the SPC/E and TIP5P water models. The trajectories started from a six-coordinated first shell configuration, and after a quite long induction time of 630 ps and 755 ps for the SPC/E and TIP5P simulations, respectively, a seventh water molecule entered the first solvation shell of the ion. This configuration with seven water molecules has then remained stable for most of the simulation time. Table 7.1 shows the percentages of the different hydration numbers observed in the SPC/E and TIP5P simulations, together with the longest lifetimes of each coordination complex. The 7-fold coordination is by far the most abundant for both simulations, although a small fraction of frames with $N = 6$ and $N = 8$ is present. This is a quite unexpected result since in the literature the Hg^{2+} aqua ion has been always described as being coordinated by six water molecules.

The Hg-O and Hg-H radial distribution functions obtained using the TIP5P water model are shown in Figure 7.1 (panel A). The two g(r)'s have well-defined first peaks followed by a depletion zone, showing the existence of a quite stable first hydration shell. The first shell maximum of the Hg-O g(r) is at 2.26 Å, and the corresponding hydration number is 6.9 for the TIP5P simulation. The SPC/E first maximum is found at a slightly longer distance

	Total lifetime (%)		Longest lifetime (ns)	
	TIP5P	SPC/E	TIP5P	SPC/E
$N = 6$	4.6	2.6	1.76	0.37
$N = 7$	95.3	96.8	6.05	8.02
$N = 8$	0.1	0.6	0.02	0.14

Table 7.1: Total and longest lifetimes of the Hg^{2+} hexa-, hepta- and octacoordinated complexes for the SPC/E and TIP5P simulations.

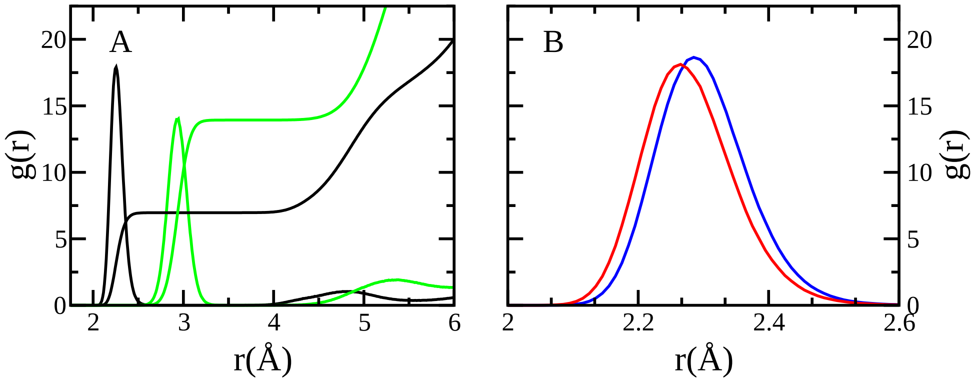

Figure 7.1: Panel A: Hg-O (black line) and Hg-H (green line) g(r)'s and corresponding running integration numbers calculated from the TIP5P simulation. Panel B: Hg-O g(r) first peak for the SPC/E (blue line) and TIP5P water model (red line).

(2.28 Å) with a coordination number of 7. As far as the Hg-H g(r)'s are concerned, the first shell maxima are found at 2.97 and 2.94 Å for the SPC/E and TIP5P models, respectively, with coordination numbers of 13.8 and 14. Figure 7.1 shows the comparison between the Hg-O g(r) first peak obtained from the SPC/E and TIP5P simulations (panel B). To better understand the origin of the small difference, it is useful to analyze the g(r)'s obtained from hexa- and heptacoordinated simulation frames, separately. Inspection of Figure 7.2 shows that the biggest difference is detected for the octahedral clusters (0.03 Å), while the seven-coordinated complexes have similar distance distribution (octahedral clusters' peak positions are at 2.23 and 2.20 Å for the SPC/E and TIP5P simulations, while heptacoordinated ones are at 2.28 and 2.27 Å). Note that the octahedral structure is present only for small percentages of the simulation time in both trajectories, and therefore it has a small effect on the position of the Hg-O g(r) first maxima.

In the following, when only the TIP5P results are shown it means that

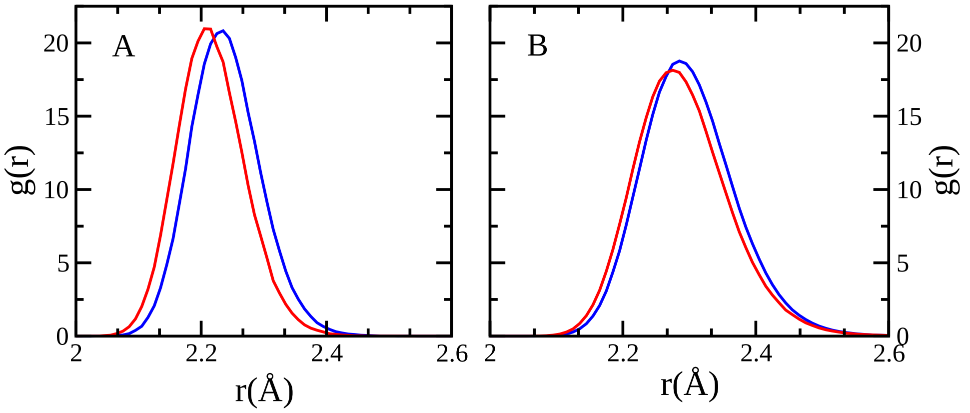

Figure 7.2: Panel A: Hg-O g(r) first peak for the SPC/E (blue line) and TIP5P water model (red line) calculated on six-coordinated frames. Panel B: Hg-O g(r) first peak for the SPC/E (blue line) and TIP5P water model (red line) calculated on seven-coordinated frames.

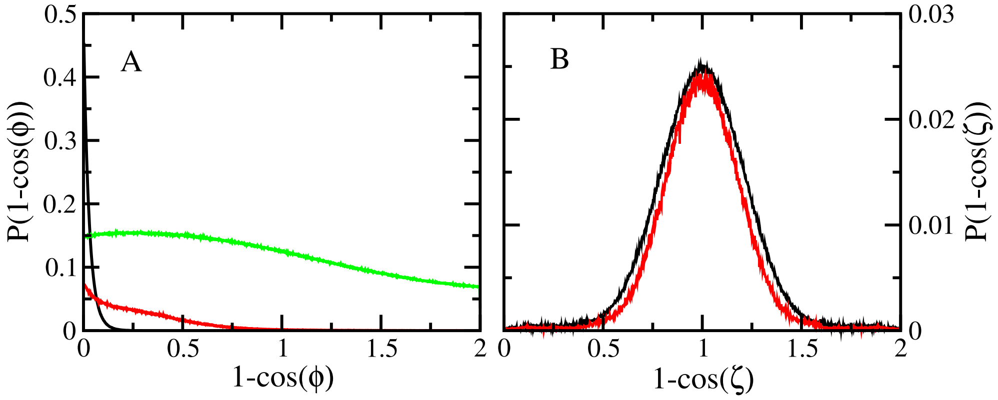

Figure 7.3: Panel A: Angular distribution functions (a.d.f.) of the ϕ angle calculated from the TIP5P trajectory. a.d.f. of first, second hydration shell and bulk water molecules are coloured black, red, and green, respectively. Panel B: a.d.f. of the ζ angle in the first hydration shell calculated from the seven- (black line) and six-coordinated (red line) frames.

identical results were obtained from the SPC/E simulation. Panel A of Figure 7.3 shows the ϕ angular distribution functions calculated for first, second hydration shells and bulk water from the TIP5P simulation. The most probable value of ϕ for the first hydration shell is 0°, corresponding to a configuration in which the water dipole is oriented along the Hg-O direction. In the second hydration shell the peak is always centered at 0° but the distribution is much broader, as expected since the water molecules are less strongly oriented by the ion. At longer distances, up to 10 Å, no preferred orientation can be observed. Since it is not possible to establish whether a water molecule is rotating in its plane or is really tilted using only the ϕ angle distribution, the ζ angular distribution function has been calculated for the first six- and seven-fold coordinated shell (see panel B of Figure 7.3). The distribution peak is found at $\zeta \approx 90°$ for both hexa- and heptacoordinated clusters, indicating that the Hg-O vector is located most of the time in the water molecule plane, with maximum deviations of 30° above and below the plane (the ζ distribution in fact drops to zero at $\zeta \approx 60°$ and $\zeta \approx 120°$). Moreover, the curve associated with the seven-coordinated frames (black line) is slightly broader since in this case the water molecule dipole is more free to rotate. Figure 7.4 shows the ψ angular distribution functions for the hepta- and hexacoordinated clusters obtained from both simulations (even if also in this case the results are almost identical, as it can be seen). For the sixfold coordinated shells the maxima are located at 90° and 180°, as expected for an octahedral hydration complex. The peak maxima for the sevenfold coordinated structure are found at values of about 75° and 145° for both simulations. These values are consistent with a C_2 symmetry of the heptacoordinated complex which corresponds to the minimum energy configuration of the system derived from ab initio optimizations of the heptahydrated cluster, as we shall see in section 7.2, where the O1-Hg-O2 and O1-Hg-O3 angles are about 75° and 145°.

The flexibility of the Hg^{2+} first hydration shell can be highlighted

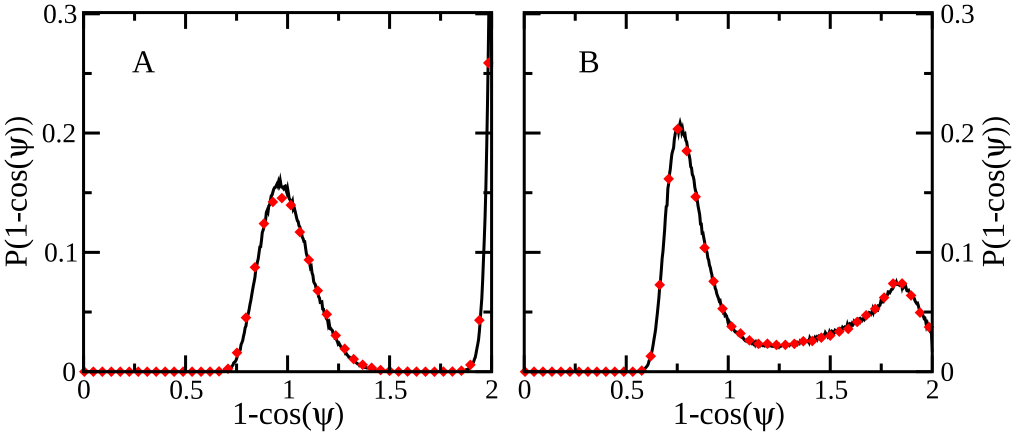

Figure 7.4: O-Hg-O (ψ) angular distribution functions (a.d.f.) calculated from the TIP5P and SPC/E trajectories (black line and red diamonds, respectively). Panel A: a.d.f. calculated from hexacoordinated first shell structures. Panel B: a.d.f. calculated from heptacoordinated first shell structures.

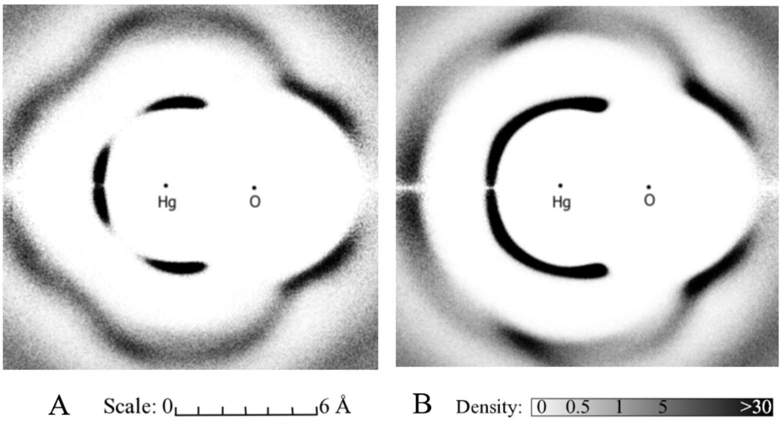

Figure 7.5: Axial-radial 2D density maps of water molecules around a fixed Hg-O axis for the TIP5P simulation. Panel A: Oxygen distribution function calculated from hexacoordinated frames. Panel B: Oxygen distribution function calculated from heptacoordinated frames.

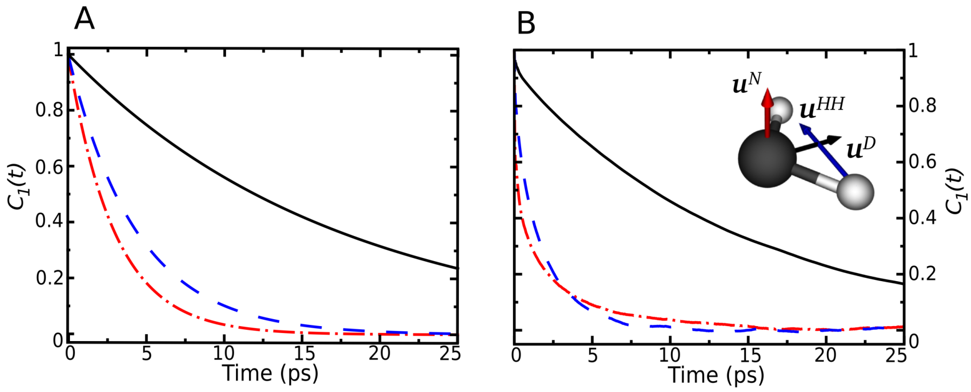

Figure 7.6: First rank reorientational correlation functions $C_1^{\alpha}(t)$ for the water molecules belonging to the Hg^{2+} first hydration shell: $\mathbf{u}^D$ (black solid line), $\mathbf{u}^{HH}$ (blue dashed line), and $\mathbf{u}^N$ (red dot-dashed line). Panel A: Correlation functions obtained using the SPC/E water model. Panel B: Correlation functions obtained using the TIP5P water model.

using axial-radial 2D density maps for the hexa- and heptacoordinated complexes (panels A and B of Figure 7.5, respectively). The map obtained for the six-fold coordinated shell is typical of an octahedral cluster, such as that obtained for the Ni^{2+} hydration shell [106]. On the contrary, in the case of the heptahydrated cluster, the oxygen atoms give rise to a uniform distribution around the chosen axis, showing the higher mobility of the first shell water molecules.

As far as the dynamic properties of the Hg^{2+} aqua ion are concerned, we have calculated the residence time of water molecules in the first solvation shell, the ion diffusion coefficients and the reorentational correlation functions. The computed first hydration shell residence times are 7.4 and 6.8 ns for the SPC/E and TIP5P water model, respectively, of the same order of magnitude of the experimental estimated values (about 1 ns) [14]. The Hg^{2+} ion exhibits a faster translational dynamics in the simulations as compared to the experimental description. In fact, the calculated ion diffusion coefficients are $1.02(0.07)\cdot10^{-5}$ cm^2/s and $1.06(0.08)\cdot10^{-5}$ cm^2/s for the SPC/E and TIP5P, respectively, while the experimental value is of $0.847\cdot10^{-5}$ cm^2/s [107].

Figure 7.6 shows the $C_1^{\alpha}(t)$ functions for SPC/E and TIP5P first hydration shells and Table 7.2 reports the first shell correlation times. The dominant motion is the rotation around the water dipole, as expected. In fact, the reorientation of the water molecular dipole occurs in a time scale almost one order of magnitude bigger, as compared to the rotation of the $\mathbf{u}^N$ and $\mathbf{u}^{HH}$ vectors. This finding is in agreement with the results of our analysis of angular distribution functions and with the results of previous simulations carried out on the Ni^{2+} aqua ion [106]. The two water models show a similar behavior, but the correlation time obtained with the TIP5P water model are lower. It is important to observe that the reorientation of first hydration shell water molecules occurs on the picosecond time scale, while the exchange

	vector	Correlation times
TIP5P	$\mathbf{u}^N$	1.6 (0.2)
	$\mathbf{u}^{HH}$	1.6 (0.2)
	$\mathbf{u}^D$	12.9 (1.4)
SPC/E	$\mathbf{u}^N$	2.6 (0.2)
	$\mathbf{u}^{HH}$	2.8 (0.2)
	$\mathbf{u}^D$	32.7 (1.9)

Table 7.2: Reorientational correlation times (ps) of water molecules belonging to the Hg^{2+} first hydration shell obtained from the TIP5P and SPC/E simulations. Standard deviations are given within parentheses.

of water molecules between six- and seven-coordinated complexes occurs in the nanosecond timescale; thus the breakdown of the solvation shells involves more complicated mechanisms than simple molecular reorientations.

7.2 Ab initio optimizations of the Hg^{2+}-(H$_2$O)$_n$ clusters

As described in the previous section, the results obtained from the Molecular Dynamics simulations has revealed a quite unexpected sevenfold coordination for the Hg^{2+} ion. Since this ion has been always thought to form an octahedral complex in aqueous solution, we have decided to evaluate the relative stabilities of the Hg^{2+}-(H$_2$O)$_6$ and Hg^{2+}-(H$_2$O)$_7$ clusters by means of ab initio optimizations. The calculations have been carried out at the MP2 level of theory both in vacuum and simulating bulk solvent effects by means of the PCM. The optimizations lead to true energy minima (all positive Hessian eigenvalues) for the two clusters, belonging to the T_h and C_2 symmetry group for the six- and seven-coordinated complex, respectively (see Figure 7.7). The absolute energies and main geometrical parameters of these structures are shown in Table 7.3. The stability of the two complexes has been evaluated by comparing the energy of the Hg^{2+}-(H$_2$O)$_7$ cluster, and the energy of a complex obtained with the displacement of one water molecule from the first to the second hydration shell of the Hg^{2+}-(H$_2$O)$_7$ cluster. A structure retaining C_2 symmetry was obtained in the latter case with the seventh water molecule forming two hydrogen bonds with two water molecules of the first coordination sphere, as already found for other metal ions [86]. The structure with one water molecule in the second hydration shell is significantly more stable for the isolated clusters, but the hexa and heptacoordinated species become nearly isoenergetic (ΔE= 0.0003 a.u. = 0.8 kJ/mol) when bulk solvent effects are taken into account by means of the PCM. It is remarkable that

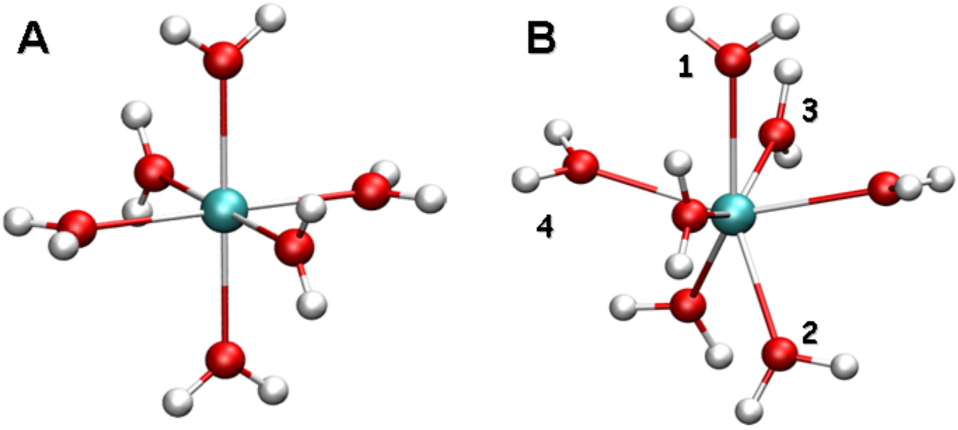

Figure 7.7: Perspective views of the MP2 minimum energy structures for $Hg^{2+}(H_2O)_6$ (panel A) and $Hg^{2+}(H_2O)_7$ (panel B) complexes. The numbers in panel B refer to bond distances and angles reported in Table 7.3.

		VACUUM	PCM
$Hg^{2+}(H_2O)_6$	E(a.u.)	-499.5130	-499.8173
	R_{Hg-O}(Å)	2.34	2.37
$Hg^{2+}(H_2O)_6(H_2O)$	E (a.u.)	-575.8738	-576.1666
$Hg^{2+}(H_2O)_7$	E(a.u.)	-575.8653	-576.1669
	R_{Hg-O1}(Å)	2.43	2.42
	R_{Hg-O1}(Å)	2.41	2.40
	R_{Hg-O3}(Å)	2.34	2.33
	R_{Hg-O4}(Å)	2.44	2.39
	$A_{O1-Hg-O2}$	141.3°	140.4°
	$A_{O1-Hg-O3}$	80.9°	80.3°
	$A_{O1-Hg-O4}$	76.3°	77.6°
H_2O	E(a.u.)	76.3256	76.3382

Table 7.3: Minimum energy and geometrical parameters obtained from the ab-initio optimizations of the Hg^{2+} first hydration shell clusters.

such a strong solvent induced energy variation occurs with only negligible geometry modifications. Inclusion of non potential energy terms (zero point energy, thermal enthalpic and entropic contributions [108]) does not modify in an appreciable way the results. Moreover, in all cases the charge transfer was negligible (maximum charge of 0.02 on a single water molecule).

7.3 EXAFS analysis

$\chi(k)$ theoretical signals have been calculated by means of equation 4.5 starting from the total Hg-O and Hg-H g(r)'s obtained from the SPC/E and TIP5P trajectories, and the structural parameters derived from the Molecular Dynamics simulations were kept fixed during the EXAFS analyses. In this way, the first hydration shell structure obtained from the simulations can be directly compared with experimental data, and the validity of the theoretical framework used in the simulations can be assessed. In the upper panels of Figure 7.8, the comparison between the EXAFS experimental signal and the theoretical curves calculated for the SPC/E and TIP5P simulations are shown. The agreement between the calculated EXAFS spectra and the experimental data is very good in both cases, proving the correctness of the structural results obtained from the two simulations. The S_0^2 value was found to be equal to 0.98, and E_0 was 12285 eV, in both cases. It is important to stress that the theoretical $\chi(k)$ signals calculated from the Hg-O and Hg-H radial distribution functions obtained from the heptacoordinated frames are identical to those calculated from the total g(r)'s, as the EXAFS techniques is not sensitive to the small percentage of hexacoordinated complexes. The Fourier transform (FT) modules of the EXAFS $\chi(k) \times k^2$ theoretical and experimental signals are also shown in Figure 7.8. The FTs have been calculated in the k-range of 2.1-13.5 Å^{-1}, with no phase shift correction applied. The FT spectra show a prominent first shell peak that is mainly due to the Hg-O first shell distance. To gain deeper insight into the structural properties of the Hg^{2+} aqua ion, we have also calculated the EXAFS theoretical signal associated with the Hg-O g(r) calculated from the Molecular Dynamics frames with a first shell octahedral complex. The results of this analysis are shown in the lower panels of Figure 7.8, and, in this case, the agreement between the experimental and theoretical data is not good, for both the SPC/E and TIP5P water models. The same discrepancy is found for the corresponding FTs. In this case, the S_0^2 value was found to be equal to 0.87, and E_0 was 12283 eV, for both minimizations. Therefore, the EXAFS analysis results indicate that the structural and dynamical information derived from the two Molecular Dynamics simulations is basically correct, while the EXAFS signal cannot be reproduced by the first shell octahedral clusters obtained from the two trajectories.

Nevertheless, a definite answer on the Hg^{2+} first shell coordination number cannot be obtained from the EXAFS analysis. In fact, we have carried out three minimizations of the EXAFS signal (in the structural and non-structural parameter space), using a single coordination shell with fixed coor-

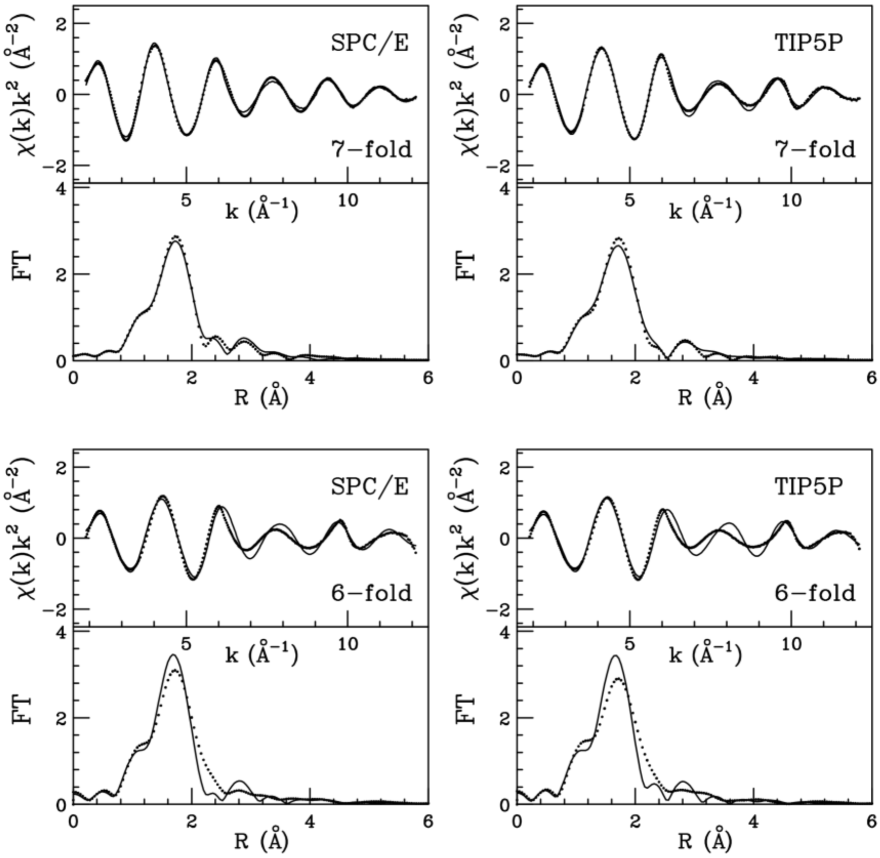

Figure 7.8: EXAFS theoretical signal for the Hg^{2+} ion calculated from the Molecular Dynamics simulations (solid line) and obtained from the experiment (dotted line) for the SPC/E and TIP5P hexa-hydrated and hepta-hydrated first shell clusters. Lower panel: Fourier transforms of the calculated (solid lines) and experimental (dotted lines) signals.

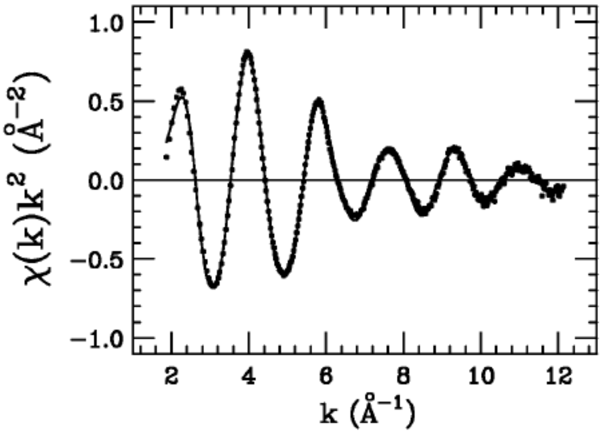

Figure 7.9: Comparison between the experimental k^2 weighted EXAFS data of Hg^{2+} in water (dotted line) and the best fit signal corresponding to an octahedral model (solid line).

dination numbers of 6, 7 and 8 and we have obtained a very similar agreement between the experimental and theoretical signals in the three cases; the Hg-O distance was always 2.32(2) Å with σ^2 values of 0.025, 0.027, and 0.033 $Å^2$, and third cumulant C_3 values of 0.0065, 0.0074 and 0.0096 $Å^3$, for the hexa-, hepta- and octacoordinated models, respectively. Note that the third cumulant C_3 is related to the skewness β through the relation $C_3 = \sigma^3\beta$. The results of the fitting procedure carried out using an octahedral model are shown in Figure 7.9. The coordination number of the Hg^{2+} hydration complex cannot thus be accurately determined from the EXAFS data analysis due to its large correlation with the Debye-Waller factor. Note however that the structural oscillations decrease very rapidly, giving somehow larger Debye-Waller factors than expected for an octahedral coordination complex [7].

7.4 XANES analysis

7.4.1 Static fits

XANES is extremely sensitive to the geometric environment of the absorbing atom as multiple-scattering effects make large contributions to this region of the X-ray absorption spectra. A quantitative analysis of the XANES which includes the rising edge and about 200 eV above it, can address some of the aforementioned shortcomings of EXAFS. To examine the compatibility of the XANES spectrum with the existence of an octahedral complex for the solvated Hg^{2+} ion, we performed a minimization of the experimental data imposing a T_h symmetry. In this fit, only the ion-water ligand distance was allowed to vary, resulting in an Hg-O best-fit distance of 2.29(3) Å and $R_{sq} = 15.7$ (see Table 7.4 and panel A of Figure 7.10). A second minimization has been carried out using an equatorially constrained Jahn-Teller distorted octahedron (see panel B of Figure 7.10). In this case the equatorial Hg-O distances obtained from the fit were equal to 2.24(3) Å, while the two axial distances were found at 2.41(3) Å, with a $R_{sq} = 15.2$ (see Table 7.4). Note that even if the agreement between the experimental and theoretical spectra is not satisfactory, the results of the fitting procedure point towards a structure with a very large difference (0.17 Å) between the equatorial and axial distances. Finally, to assess the compatibility of the XANES with a more disordered hexacoordinated hydration complex, we performed an additional minimization of the experimental data using an unconstrained hexacoordinated cluster, where all of the six ion-oxygen bond distances were refined (panel C of Figure 7.10). In this case Hg-O distances are spread between 2.19 and 2.44 Å, with an average value of 2.31 Å, and R_{sq} is 9.9. Thus all the hexahydrated models lead to a quite poor agreement between the experiment and the calculated spectra, especially in the low-energy region of the spectrum. In particular, the measured XANES spectrum exhibits a shoulder at the main transition edge that could not be reproduced (insets of Figure 7.10), and a poor agreement between the experiment and the calculated model is found also in the high-energy range of

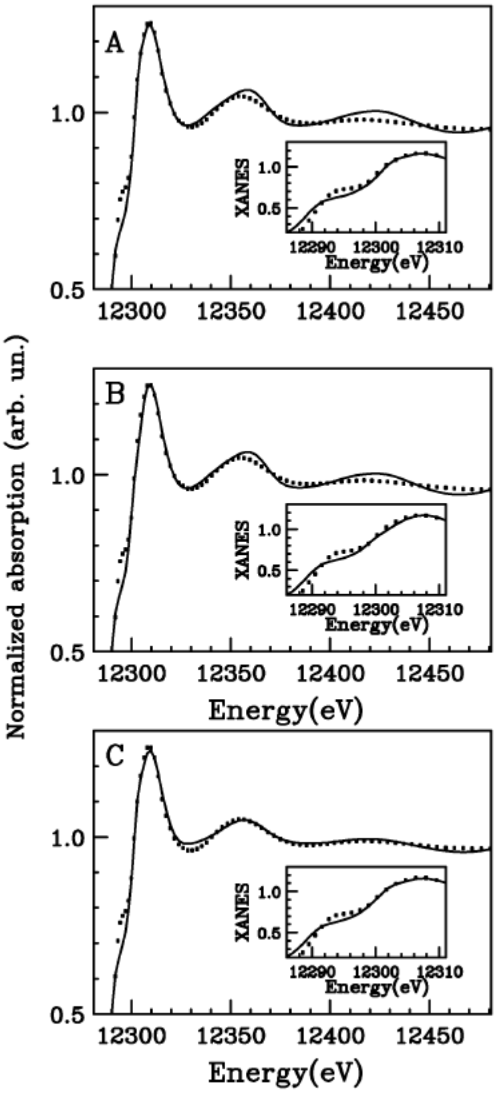

Figure 7.10: Panel A: Comparison of the Hg L_3 deconvolved experimental XANES spectrum of the Hg^{2+} aqueous solution (dotted line), with the best fit theoretical spectrum (solid line) associated with an octahedral cluster. Panel B: Comparison of the experimental XANES spectrum (dotted line), with the best fit theoretical spectrum (solid line) associated with a Jahn-Teller distorted octahedral cluster. Panel C: Comparison of the experimental XANES spectrum (dotted line), with the best fit theoretical spectrum (solid line) obtained from a disordered six-coordinated model.

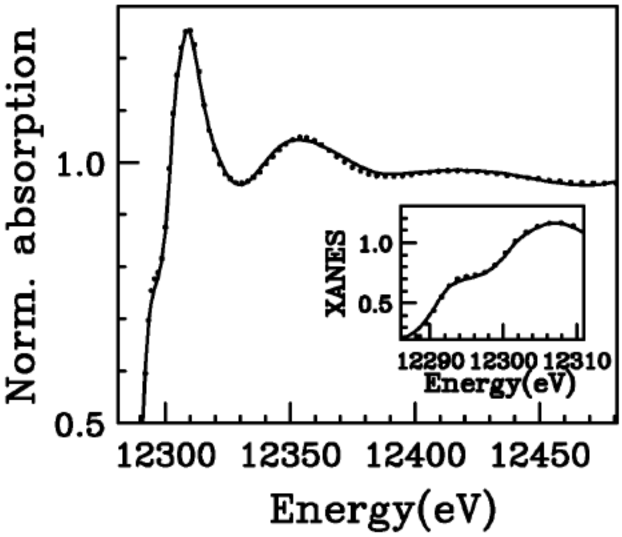

Figure 7.11: Comparison of the Hg L_3 deconvolved experimental XANES spectrum of the Hg^{2+} aqueous solution (dotted line), with the best fit theoretical spectrum (solid line) as obtained from a heptacoordinated complex having a C_2 symmetry.

the spectrum (12350-12600 eV) for the regular and Jahn-Teller distorted octahedra. In the case of the unconstrained hexacoordinated model, the Hg-O distances are quite spread and the structural disorder reduces the amplitude of the signal at high energy, resulting in a better agreement with the experimental data. Nevertheless, the frequency is not completely correct and the shoulder at the main edge is not reproduced (see insets of Figure 7.10). All together these findings suggest that, at variance with all the previously reported results [1, 2, 13, 18], the Hg^{2+} aqua ion does not adopt an octahedral coordination, and it forms a quite flexible hydration complex. Therefore, since our combined QM and Molecular Dynamics results strongly suggest that the hydrated Hg^{2+} aqua ion is characterized by an heptacoordinated first shell structure (whose energy minimum has a C_2 symmetry), we reinterpreted the XANES spectrum in terms of an heptahydrated Hg^{2+} ion. We performed a minimization of the experimental data starting from the ab-initio energy minimum structure of the Hg^{2+}-$(H_2O)_7$ complex, and refining four structural parameters while retaining a C_2 symmetry (see Figure 7.11). The ion-oxygen distances obtained from the minimization procedure are reported in Table 7.4. The Hg-O average distance is 2.31 Å, in agreement with the EXAFS determination, and the R_{sq} value obtained is 1.6. A second fit was performed using an unconstrained heptacoordinated model where all the seven Hg-O distances were refined separately. Also in this case a similar geometry was obtained from the minimization procedure with a Hg-O average distance of 2.32 Å and $R_{sq} = 1.1$. The availability of new degrees of freedom in the minimization procedure does not significantly change the quality of the fit, and it produces a slightly different set of structural parameters, still corresponding to a C_2 symmetry. An additional fit of the experimental data was performed starting from a cluster having seven equal Hg-O distances arranged in a C_2 symmetry and keeping such a symmetry during the minimization by refining a single distance. The optimized Hg-O distance is 2.30 Å and the agreement between the experimental and theoretical data was not good ($R_{sq} = 15.0$) thus showing that the water molecules in the Hg-O heptacoordinated cluster are arranged in a less symmetrical fashion. Attempts to fit the XANES spectrum with a mixture of hexa- and heptahydrated clusters showed that the experimental data can be reproduced only for a strong dominance (higher than 90%) of the heptacoordinated species.

The picture emerging from these tests is both enlightening and intriguing. If a single first shell model cluster is used to fit the XANES data, the best agreement is obtained when a seven-fold coordination within a C_2 symmetry and a quite large distance dispersion (defined as the difference between the maximum and minimum Hg-O bond length) is considered (0.38 Å). However, the dispersion of distances obtained for the heptahydrated cluster from MP2 calculations is much smaller (0.09 Å, see Table 7.3) and this is also the case for the Molecular Dynamics simulations (0.12 and 0.13 Å for the SPC/E and TIP5P trajectories). One might therefore conclude that a proper fit of the XANES data needs a wider distance dispersion than that obtained from the theoretical heptahydrated models. This discrepancy can be due to two main reasons: the need to account for the structural disorder within the first hydration shell in the analysis of the XANES data, and the influence of the second hydration shell on the low energy range of the absorption spectrum. Unfortunately, the effect of the high mobility of the second hydration shell does not allow to build reliable static models including its effect. One must, there-

Parameters	Octahedron	JT distorted Octahedron	Flexible 6coordinated	C_2 7coordinated	Flexible 7coordinated
R_1	2.29	2.24	2.19	2.19	2.17
R_2	2.29	2.24	2.22	2.19	2.17
R_3	2.29	2.24	2.26	2.23	2.30
R_4	2.29	2.24	2.29	2.23	2.30
R_5	2.29	2.41	2.44	2.40	2.39
R_6	2.29	2.41	2.44	2.40	2.45
R_7	-	-	-	2.57	2.45
R_{exp}	0.74	0.71	0.71	0.72	0.72
E_o	12286.5	12286.5	12286.5	12286.5	
R_{sq}	15.7	15.2	9.9	1.6	1.1
n	1	2	6	4	7

Table 7.4: Structural and background parameters obtained from the XANES minimizations. R_i (Å) are the Hg-O distances, R_{exp} (eV) is the resolution broadening, E_0 is the theoretical absorption edge position, R_{sq} is the fit index parameter, and n is the number of independent structural parameters used in the fitting procedure.

fore, resort to analysis of the XANES spectra by means of the microscopic description of the system derived from Molecular Dynamics simulations.

7.4.2 XANES analysis from Molecular Dynamics simulations

As already pointed out in the previous section, in the static fit analysis we found a discrepancy between the theoretical and XANES results for the geometry of the Hg^{2+} heptahydrated cluster. Therefore, we have decided to extend the study of Hg^{2+} aqua ion, using our Molecular Dynamics trajectories to interpret the XANES experimental data. Since the results of the SPC/E and TIP5P simulations are in agreement with each other and essentially both of them describe the Hg^{2+} as being heptacoordinated in aqueous solution, we have used only one simulation in the XANES analysis. In particular, we have computed the XANES spectrum from a set of representative set of geometries extracted from the SPC/E simulation, without carrying out any minimization of the structural parameters.

Computational procedure

In the first step of the analysis, a trajectory containing only the Hg^{2+} ion and its first hydration shell has been extracted from the total Molecular Dynamics trajectory. The first hydration shell has been defined by including all the water molecules separated by the Hg^{2+} ion by a distance shorter than 3.4 Å. From this trajectory we extracted 200 snapshots saved every 12.5 ps. The percentages of hepta- and hexacoordinated first shell clusters of these snapshots were 97% and 3%, respectively, thus resembling the behavior of the total Molecular Dynamics simulation. Seven or six water molecules have been included in the calculation, together with the Hg^{2+} ion. Each snapshot has been used to generate the XANES associated with the corresponding instantaneous geometry, and the averaged theoretical spectrum has been obtained by summing all the spectra and dividing by the total number of Molecular Dynamics snapshots used. At this stage only the real part of the HL potential has been used, i.e. theoretical spectra do not account for any intrinsic and extrinsic inelastic process (see section 6.2.3).

In the second step of the analysis the influence of the second hydration shell on the XANES spectrum of Hg^{2+} has been assessed. In the XANES calculation we included all the water molecules separated from the cation by a distance shorter than 5.2 Å, since water molecules at larger distance have been found to provide a negligible contribution. To this end a trajectory containing the ion and its first two hydration shells has been extracted from the total Molecular Dynamics trajectory and 200 snapshots saved every 12.5 ps have been singled out. Also in this case 200 XANES spectra have been calculated and averaged starting from the instantaneous configurations obtained from the Molecular Dynamics calculations.

An important question when dealing with the computation of spectra from

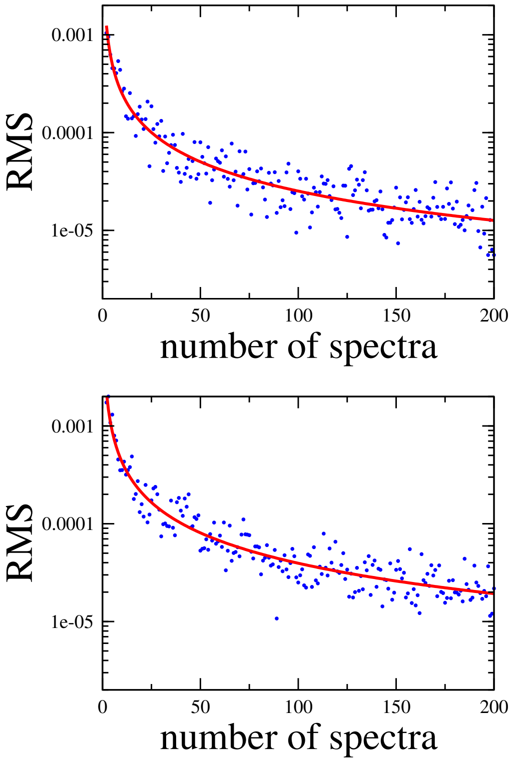

Figure 7.12: Residual function RMS of the XANES averaged spectra as a function of the number of Molecular Dynamics snapshots for the first shell (upper panel) and the first plus second shell (lower panel) analysis.

Molecular Dynamics simulations is to determine the total sampling length that is necessary to have a statistically significant average. To this end we have carried out a statistical treatment of the data. In particular we have calculated a residual function defined as:

$$RMS = \sqrt{\sum_i [\alpha^N(E_i) - \alpha^{N-1}(E_i)]^2} \tag{7.1}$$

where $\alpha^N(E_i)$ is the theoretical spectrum averaged over N snapshots and the sum is extended over all the energy points E_i. A residual value of 10^{-4} was chosen to establish the number of spectra which are necessary to have a statistically significant average. The results of this analysis are shown in Figure 7.12, where the RMS function is plotted against the number of averaged spectra for the first, and first plus second shell theoretical spectra in the upper and lower panels, respectively. As evident from the Figure, 200 configurations are enough to reach convergence in both cases.

Finally, to assess the compatibility of the XANES experimental data with a sixfold coordination around the Hg^{2+} ion we extracted two trajectories from the simulation window around the longest lifetime of the Hg^{2+} hexahydrated clusters. The former trajectory contained only first shell hexahydrated clusters and the latter contained first hexacoordinated clusters plus the second shell (also in this case the second hydration shell includes all the water

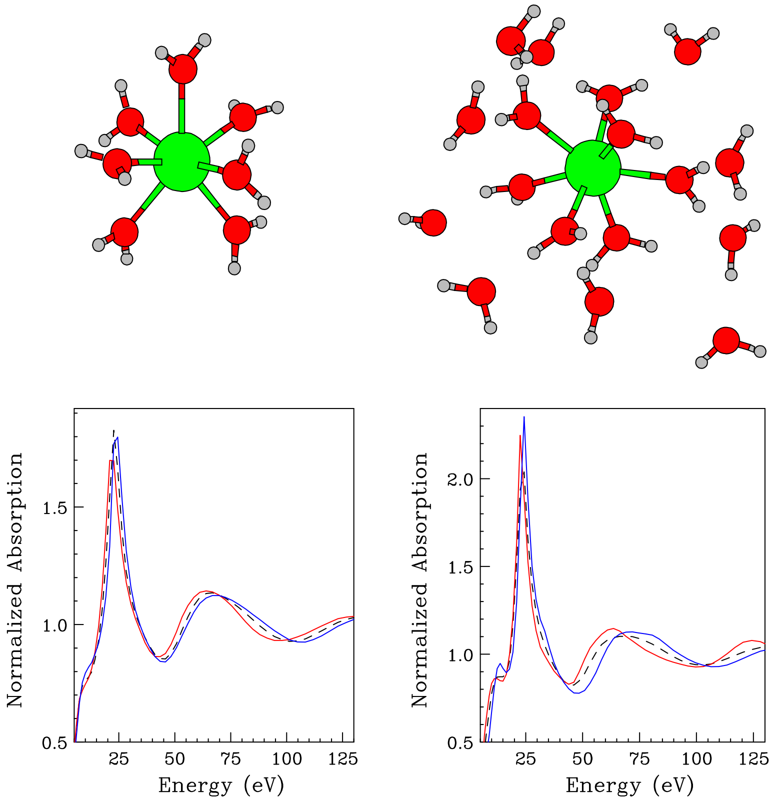

Figure 7.13: Left panel: Comparison of the XANES theoretical spectrum obtained from the Molecular Dynamics average including the first hydration shell (dashed line) and two spectra associated with individual Molecular Dynamics configurations (solid red, and blue line). In the upper panel the Hg^{2+} seven-fold hydration complex is shown. Right panel: Comparison of the XANES theoretical spectrum obtained from the Molecular Dynamics average including the first and second hydration shells (dashed line) and two spectra associated with individual Molecular Dynamics configurations (solid red, and blue line). In the upper panel a typical first and second shell hydration cluster is shown.

molecules at a distance from the Hg^{2+} ion shorter than 5.2 Å). From each trajectory we extracted 200 snapshots saved every 1.85 ps and we calculated two averaged theoretical spectra corresponding to an octahedral first shell coordination with and without the second shell, respectively.

XANES-Molecular Dynamics results

The left panel of Figure 7.13 shows the averaged theoretical spectrum obtained from 200 snapshots (not including intrinsic and extrinsic inelastic process) associated with the Hg^{2+} first shell hydration complex, together with two individual instantaneous structures. The calculated XANES spectra present noticeable differences all along the energy range, showing the sensitivity of XANES to geometrical changes, and the importance of making a proper sampling of the configurational space. As far as the second shell is concerned our Molecular Dynamics simulation suggests the presence of 17 water molecules up to about 5.2 Å from the ion. The XANES total theoretical spectrum obtained from 200 Molecular Dynamics snapshots including both

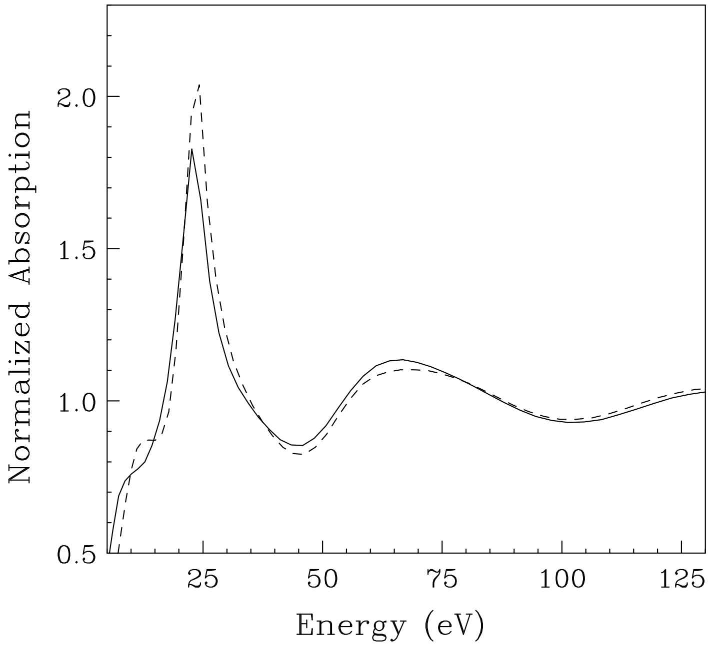

Figure 7.14: Comparison of the theoretical XANES spectrum obtained from the Molecular Dynamics average including the first hydration shell (solid line) and the averaged theoretical XANES spectrum including the first and second shells (dashed line).

the first and second shell is shown in the right panel of Figure 7.13, together with two spectra computed from instantaneous configurations. In this case the individual spectra show more marked differences as compared to the first hydration shell. Considering the large deviation among instantaneous spectra it seems unlikely that a single representative configuration can be used to properly model the experimental data. A deeper insight into the effect of the second hydration shell on the XANES spectrum of Hg^{2+} in aqueous solution can be gained by the direct comparison between the averaged spectrum calculated using first shell only and the first and second shell clusters (see Figure 7.14). Significant differences appear in the low energy region up to about 80 eV from the threshold. In particular the edge intensity is lowered in the spectrum containing only the first shell water molecules which exhibits a different shape in the region around the first minimum. The two spectra become very similar for energy values higher than 80 eV and this finding underlines the insensitivity of the EXAFS technique toward second shell contributions [7]. To assess the reliability of the entire procedure it is necessary to compare the total averaged XANES spectra with the experimental data. To this end all inelastic processes have been accounted for by convoluting the theoretical averaged spectra with a broadening Lorentzian function, and the corresponding E_s and A non-structural parameters have been optimized. The agreement between the experimental and theoretical data has been assessed by the goodness-of-fit parameter (R_{sq}) (equation 4.7). In panel A of Figure 7.15, the experimental XANES data are compared with the averaged theoretical spectrum including the first shell clusters as derived from Molecular Dynamics simulations. The overall agreement of the two spectra is not

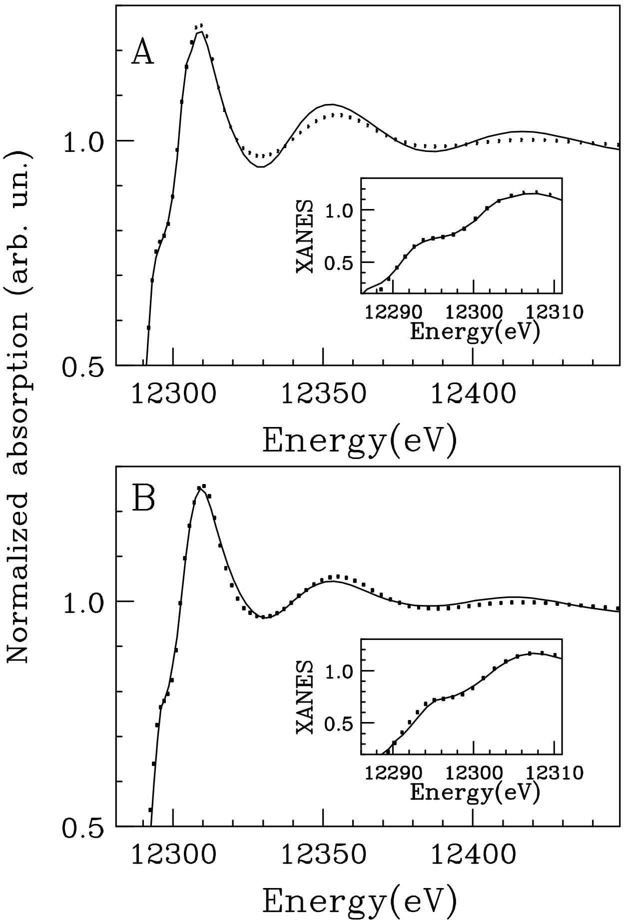

Figure 7.15: Panel A: Comparison between the averaged XANES theoretical spectrum including the first heptacoordinated shell (solid line) and the experimental data (dotted line) of Hg^{2+} in aqueous solution. Panel B: Comparison between the averaged XANES theoretical spectrum including the first heptacoordinated and second shell (solid line) and the experimental data (dotted line) of Hg^{2+} in aqueous solution.

perfect (R_{sq}=6.2) especially for the energy region above 12320 eV. In panel B of Figure 7.15, the averaged theoretical spectrum including both the first and second hydration shells is compared with the experimental spectrum. In this case, the agreement between the experimental and theoretical spectra is very good both in the low- and high-energy regions of the spectrum (R_{sq}=2.1). As evident from the insets of Figure 7.15, the shoulder at the main edge, which is sensitive to the symmetry of the cluster, is perfectly reproduced when a heptacoordinated first shell model as obtained from the Molecular Dynamics simulations is considered. It is important to remark that all the XANES spectra have been calculated using the structural information obtained from the Molecular Dynamics simulations without carrying out any minimization in the structural parameter space. Due to the high sensitivity of the XANES technique towards the structural environment of the photoabsorber this approach is a very strict test on the quality of the potentials used in the Molecular Dynamics simulations, and the almost perfect agreement between the averaged theoretical and experimental XANES spectra proves the reliability of the entire computational procedure.

As mentioned above, in our standard XANES analysis a static heptahydrated first shell cluster has been used to fit the XANES spectrum of Hg^{2+} in aqueous solution. The distance dispersion of the Hg-O first shell distances obtained from the XANES minimization was about 0.26 Å larger than the Molecular Dynamics results and the QM optimized clusters. The results of this dynamical XANES analysis demonstrate that the discrepancy we had found among the Hg-O distance dispersion obtained from static XANES analysis, MP2 calculations and Molecular Dynamics simulations, is mainly due to the lack of the second hydration shell contribution in the XANES calculations and support strongly the need to include statistical structural information on the first and second coordination shell to quantitatively reproduce the experimental data. As previously observed [109], the XANES structures are sensitive to the second shell local geometries and a thorough sampling of the configurational space has to be made to perform a correct analysis.

As a last analysis, we have examined the possibility of reproducing the XANES experimental data using a six-coordinated hydration complex. In panel A of Figure 7.16 we report the comparison between the averaged theoretical spectrum including only the hexacoordinated first shell clusters and the experimental data. In this case the agreement between the two spectra is not satisfactory (R_{sq}=16.3), proving that the Hg^{2+} is not hexacoordinated. Note that also six-coordinated structures with distorted angles, as those obtained from the Molecular Dynamics simulations, are not able to reproduce the experimental data. We analyzed then a trajectory having a hexacoordinated first shell plus the second hydration shell, and the comparison with the XANES experimental spectrum is shown in panel B of Figure 7.16. Also in this case the agreement between the experimental and theoretical spectra is poor (R_{sq}=15.4) and this finding demonstrates that the XANES data cannot be reproduced using a hexacoordinated cluster, even if the second hydration shell is included in the calculation. It is important to stress that the shoul-

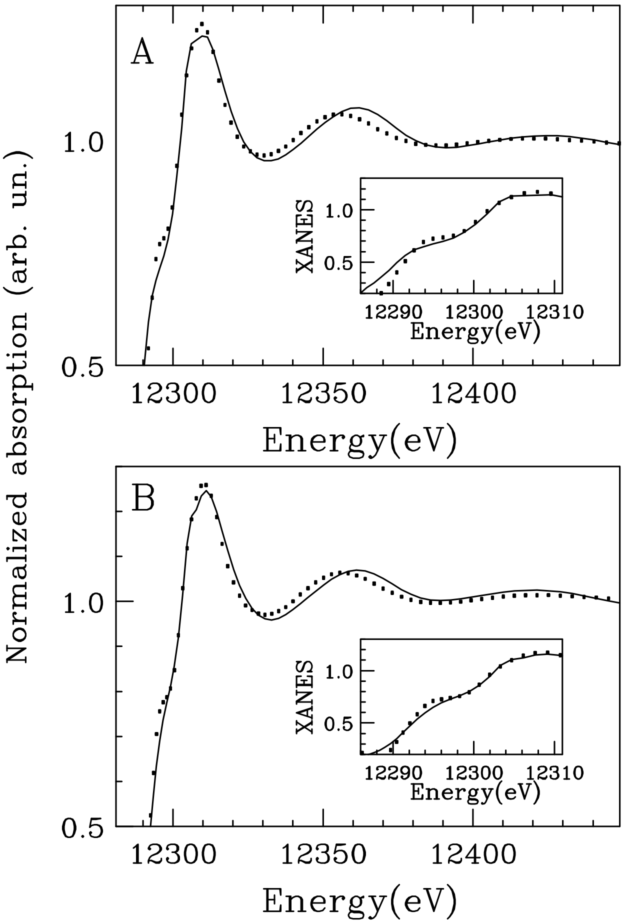

Figure 7.16: Panel A: Comparison between the averaged theoretical XANES spectrum of Hg^{2+} in aqueous solution including the first hexacoordinated shell (solid line) and the experimental data (dotted line). Panel B: Comparison between the averaged theoretical XANES spectrum including the first hexacoordinated and second shell (solid line) and the experimental data (dotted line).

der at the main edge, which is sensitive to the symmetry of the cluster, is never reproduced when a hexacoordinated first shell is considered (see insets of Figure 7.16).

Chapter 8

Structural and dynamic properties of the Cd^{2+} aqua ion

8.1 Molecular Dynamics results

An accurate description of the structural and dynamic properties of the Cd^{2+} aqueous solution has been gained by a thorough analysis of the two Molecular Dynamics simulations carried out with the SPC/E and TIP5P water models (the same models used for the Hg^{2+} aqua ion), previously published by our group [11]. Figure 8.1 shows the Cd-O and Cd-H radial distribution functions and running integration numbers obtained from the SPC/E and TIP5P simulations. In both cases a sharp first peak is observed, followed by a depletion zone which means that a stable first hydration shell is present. The first important result is the non integer first shell coordination number obtained from both trajectories, which means that different hydration structures are present during the simulation. This is quite unexpected, since in the literature the Cd^{2+} ion is always described as six-coordinated in aqueous solution. Conversely, our simulations started from a hexacoordinated first shell configuration around the Cd^{2+} ion, and after 230 and 155 ps for the SPC/E and TIP5P models, respectively, the system went to hepta-coordination. In both cases, a very flexible Cd^{2+} first hydration shell was detected, which transits among coordination numbers of six, seven and, for relatively short times, eight. The percentages of the different hydration numbers observed in the two trajectories and the longest lifetimes of each coordination complex are

	Total lifetime (%)		Longest lifetime (ps)	
	TIP5P	SPC/E	TIP5P	SPC/E
$N = 6$	17.5	65.0	220	2297
$N = 7$	81.7	34.9	815	1048
$N = 8$	0.8	0.1	15	10

Table 8.1: Total and longest lifetimes of the Cd^{2+} hexa-, hepta- and octacoordinated complexes for the SPC/E and TIP5P simulations.

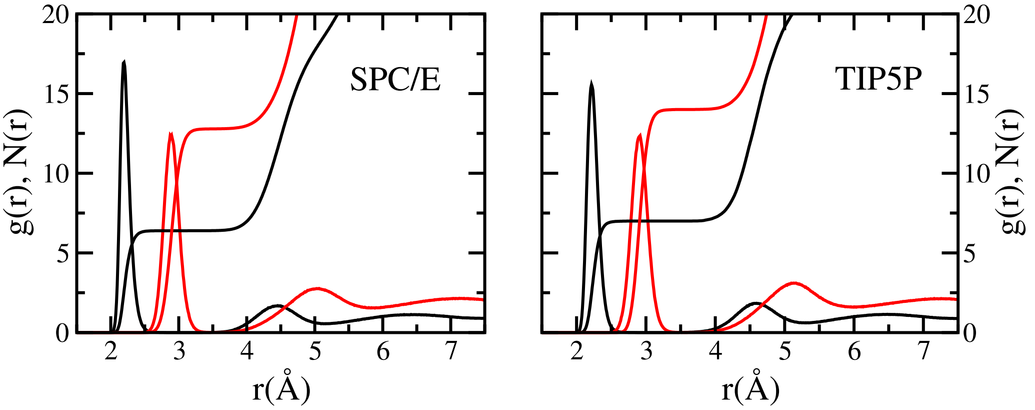

Figure 8.1: Cd-O (black line) and Cd-H (red line) radial distribution functions and corresponding running integration numbers calculated from the SPC/E (left panel) and TIP5P (right panel) simulations.

reported in Table 8.1. The different dynamical behavior between the two simulations is due to the relative higher openness of the TIP5P model which gives rise to a larger number of exchange events between six- and seven-coordinated complexes and increases the stability of the heptahydrated species as compared to the hexahydrated ones. The higher frequency of solvent exchange in the TIP5P simulation is reflected in the residence time of the water molecules in the first solvation shell, calculated with the method of Impey et al. [96] described in section 6.1.3, which is lower in the TIP5P case (the residence times are 3.0 and 1.3 ns for the SPC/E and TIP5P simulations, respectively). However, both results are in agreement with the experimental determination of the water residence time, which is estimated to be in the nanosecond time scale [14]. In the top panel of Figure 8.2 a comparison between the first peak of the Cd-O and Cd-H obtained from the two simulations is shown. As it can be seen, the TIP5P Cd-O first shell maximum is shifted towards larger distances as compared to the SPC/E peak (see also Table 8.2). Moreover, the average coordination numbers of the hydrated Cd^{2+} ion are 6.8 and 6.3 for the TIP5P and SPC/E calculations, respectively, indicating that in the former case the heptacoordinated cluster is the dominant species, while in the latter the octahedral complex is present most of the simulation time (see Table 8.1).

To get a deeper insight into the different behaviour originating from the two water models we have separately analyzed the Molecular Dynamics configurations in which either the hexa- or the heptahydrated species are present. The corresponding Cd-O radial distribution functions are shown in Figure 8.2. In particular, the comparison between the TIP5P and SPC/E g(r)'s associated with the hexacoordinated complexes is shown in the middle panel. The Cd-O g(r) first peak maxima are almost coincident, showing that the geometry of the octahedral complexes is the same in both simulations. A similar result has been obtained for the heptahydrated cluster (bottom panel of Figure 8.2) which has a longer Cd-O first shell distance as compared to the

		N	r_{CdO} (Å)
SPC/E	total	6.3	2.22
	$N = 6$	6	2.19
	$N = 7$	7	2.27
TIP5P	total	6.8	2.25
	$N = 6$	6	2.18
	$N = 7$	7	2.26

Table 8.2: Structural parameters of the Cd-O g(r) first peaks calculated from the two simulations.

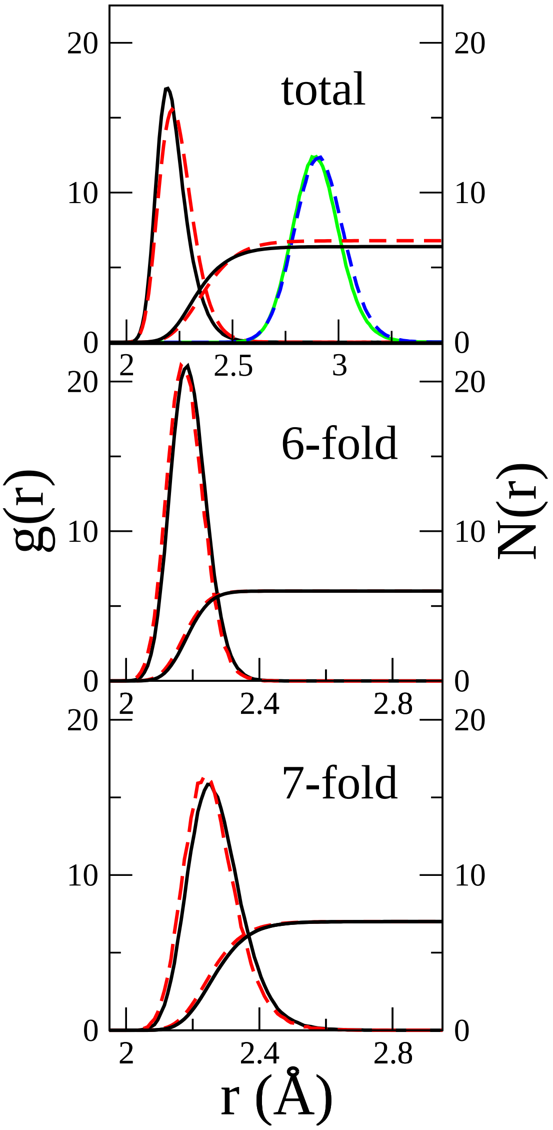

Figure 8.2: Top panel: TIP5P (dashed red line) and SPC/E (black solid line) Cd-O g(r)'s calculated on the total trajectories and corresponding running integration numbers. TIP5P (dashed blue line) and SPC/E (green solid line) Cd-H g(r)'s. Middle panel: TIP5P (dashed red line) and SPC/E (black solid line) Cd-O g(r)'s calculated on hexacoordinated simulation frames and corresponding running integration numbers. Bottom panel: TIP5P (dashed red line) and SPC/E (black solid line) Cd-O g(r)'s calculated on heptacoordinated simulation frames and corresponding running integration numbers.

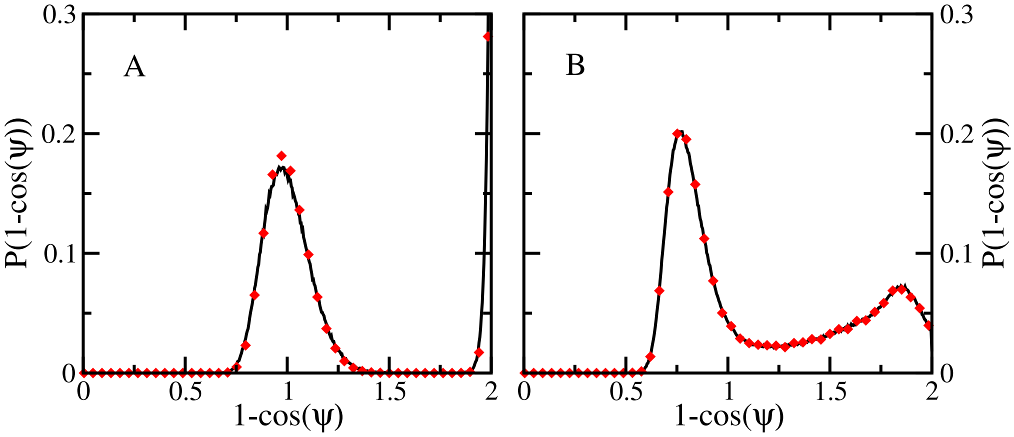

Figure 8.3: O-Cd-O (ψ) angular distribution functions (a.d.f.) calculated from the TIP5P and SPC/E trajectories (black line and red diamonds, respectively). Panel A: a.d.f. calculated from hexacoordinated first shell structures. Panel B: a.d.f. calculated from heptacoordinated first shell structures.

hexacoordinated one. The structural parameters describing all of the g(r)'s are reported in Table 8.2. Therefore, the g(r) results show that the geometry of the Cd^{2+} -(H$_2$O)$_6$ and Cd^{2+}-(H$_2$O)$_7$ clusters is the same regardless the water model used in the simulation, and the differences of the first peak position of the total TIP5P and SPC/E g(r)'s is only due to the different percentage of hexa- and heptacoordinated complexes occurring in the simulations.

A further proof of this finding has been gained looking at the geometrical arrangement of the first shell water molecules around the Cd^{2+} ion for the six and seven-coordinated complexes, obtained from the a.d.f. of the ψ, ϕ and ζ angles plotted as a function of $1 - cos(x)$ ($x = \psi,\ \phi,\ \zeta$). Figure 8.3 shows the ψ angular distribution function for the hexa- and heptacoordinated cluster (panel A and B, respectively), obtained from the TIP5P and SPC/E simulations. The two simulations give identical results and the six-coordinated cluster maxima are located at $1 - cos(x) = 1$ and $1 - cos(x) = 2$,

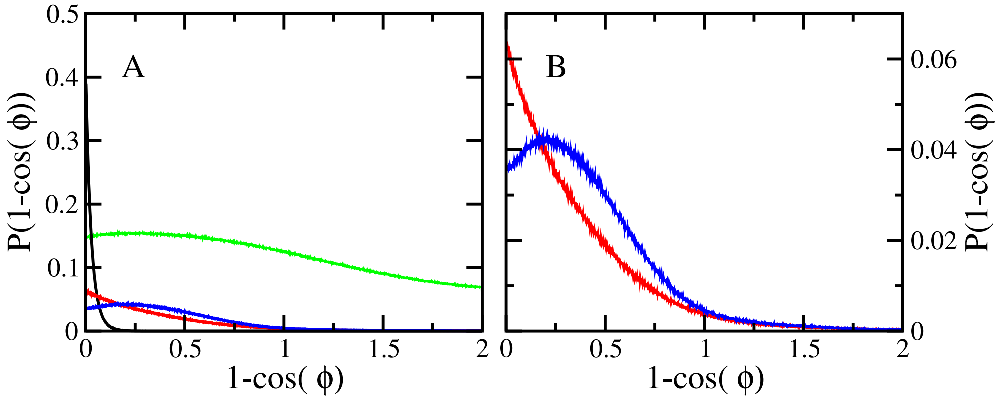

Figure 8.4: Tilt angle (ϕ) angular distribution functions (a.d.f). Panel A: a.d.f calculated for the TIP5P first hydration shell (black line), TIP5P and SPC/E second hydration shell (blue and red line, respectively) and TIP5P bulk water (green line). Panel B: a.d.f. calculated for the TIP5P and SPC/E second hydration shell (blue and red line, respectively).

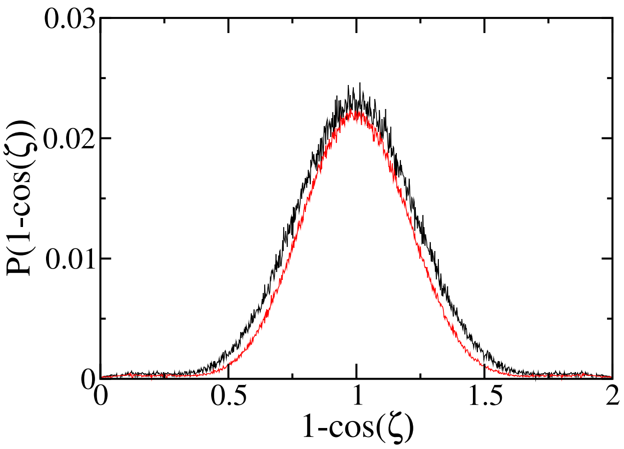

Figure 8.5: ζ angular distribution function calculated from the TIP5P trajectory for hexa- and heptacoordinated first shell structures (red and black line, respectively).

corresponding to ψ values of 90° and 180°, as expected for an octahedral hydration complex. The peak maxima for the seven-coordinated cluster are found at ψ values of about 75° and 145° that are very close to the O-Cd-O angles obtained for the stable minimum at C_2 symmetry from ab initio calculations of the Cd^{2+}-$(H_2O)_7$ cluster in Ref. [11]. Figure 8.4 shows the ϕ a.d.f. calculated for first and second hydration shells, and bulk water. As far as the first hydration shell is concerned, identical distributions are obtained from the two simulations and only the TIP5P one has been reported (panel A). The most probable value of ϕ for the first hydration shell is 0°, corresponding to a water dipole moment vector oriented along the Cd-O direction, and the ϕ distribution for the first shell drops to zero for $\phi \approx 35°$. In the case of the second shell (panel B of Figure 8.4), two different preferred orientations are obtained: the SPC/E distribution function peak is found at $\phi = 0°$, while the TIP5P ϕ distribution function has a well defined maximum at $\phi = 36°$; this effect is caused by the geometry of the TIP5P water model that, having five atoms and being more bulky than SPC/E, does not allow the water molecules to be oriented along the Cd-O direction. Both ϕ distributions drop to zero at $\phi \approx 100°$, a much bigger angle than that of the first shell, indicating less tightly oriented water molecules. At longer distances, all of the angles are allowed and the distribution is very similar to that obtained in the Hg^{2+} case. Since for the distribution of this angle the SPC/E and TIP5P results are almost identical only the TIP5P ϕ distribution function has been reported in Figure 8.4 (panel A, green line). As all of the remaining Molecular Dynamics structural analyses reported in this section produced the same results for the two water models, hereafter we will show only those of the TIP5P model. The ζ angular distribution function was calculated on hexa and heptahydrated clusters and the results are shown in Figure 8.5. Both distributions are centered on $\zeta = 90°$ and the a.d.f. calculated on seven-coordinated frames is slightly broader, as already observed for the Hg^{2+} ion.

Axial-radial 2D density maps obtained from the TIP5P simulation for the six and seven-coordinated complexes are shown in Figure 8.6. The map in

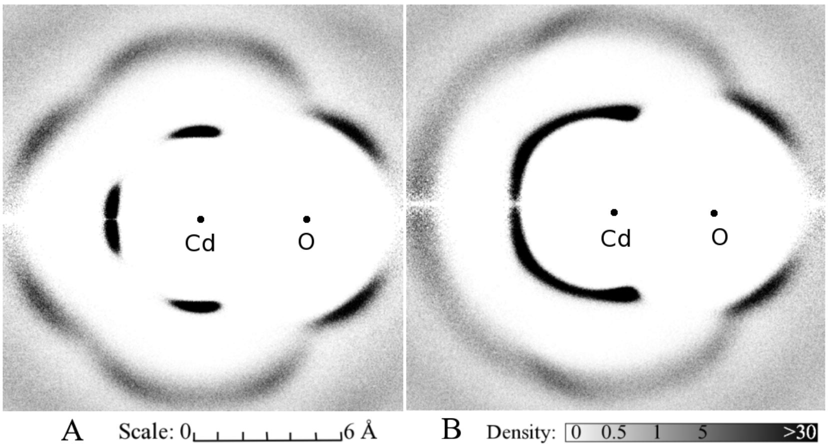

Figure 8.6: Axial-radial 2D density maps of water molecules around a fixed Cd-O axis for the TIP5P simulation. Panel A: Oxygen distribution function calculated from hexacoordinated frames. Panel B: Oxygen distribution function calculated from heptacoordinated frames.

the panel A of Figure 8.6 is typical of an octahedral cluster, while in the case of the seven-coordinated structure the oxygen atoms give rise to a uniform distribution around the selected axis, as already observed for the Hg^{2+} ion. All together these results reinforce the finding that the water model does not affect the first shell hexa and heptahydrated complex structures, while the higher flexibility of the TIP5P model modifies the dynamical behaviour of the system.

To get a deeper insight into the dynamic properties of the Cd^{2+} ion in aqueous solution, we have computed the Cd^{2+} diffusion coefficient. The calculated diffusion coefficients are $0.68(0.08){\cdot}10^{-5}$ cm^2/s and $0.73(0.08){\cdot}10^{-5}$ cm^2/s for the SPC/E and TIP5P simulations, respectively. A slightly better agreement with the experimental value ($0.719{\cdot}10^{-5}$ cm^2/s) [107] is found for the TIP5P water model. The Cd^{2+} ion exhibits a faster translational dynamics in the TIP5P simulation, as compared to the SPC/E trajectory. This result well agrees with the general picture drawn by the TIP5P simulation, characterized by a greater water mobility, resulting in shorter values of the water residence time and a greater number of water exchange events between the first hydration shell and the bulk.

8.2 EXAFS analysis

As shown in section 8.1, the TIP5P and SPC/E simulations of the Cd^{2+} aqua ion present a difference of about 0.03 Å in the position of the Cd-O first shell g(r)'s, while the geometry of the Cd^{2+}-$(H_2O)_6$ and Cd^{2+}-$(H_2O)_7$ clusters is the same. $\chi(k)$ theoretical signals have been calculated using equation 4.5, starting from the total TIP5P and SPC/E Cd-O and Cd-H radial distribution functions, in an analogous way to that used for the Hg^{2+} aqua ion. In the upper panels of Figure 8.7, the comparison between the experimental signal

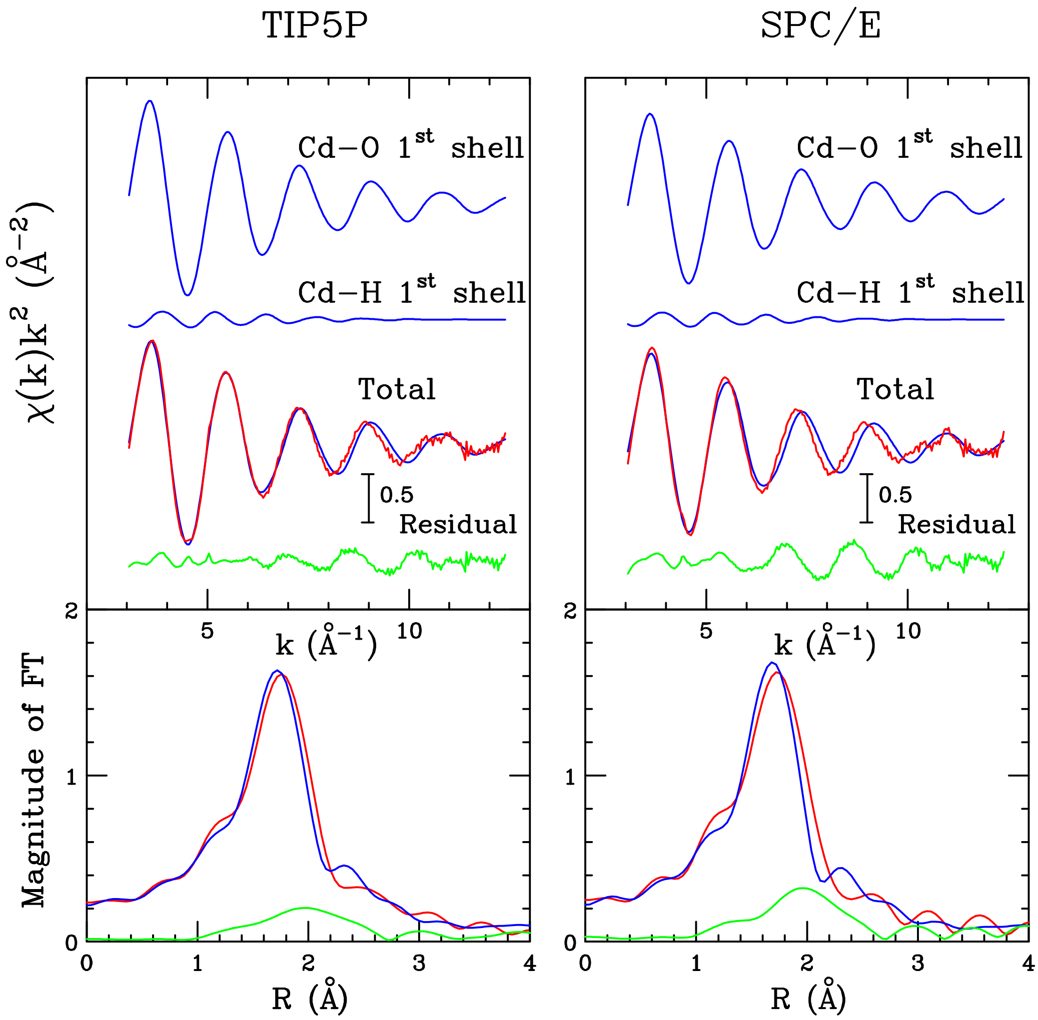

Figure 8.7: Upper panels: Comparison between the TIP5P and SPC/E EXAFS theoretical signals (blue line) calculated from the total Cd-O and Cd-H g(r)'s and experimental data (red line). The residual signals (green line) are also shown. Lower panels: Nonphase-shifted corrected Fourier transforms of the experimental data (red line), of the theoretical signals (blue line), and of the residual curves (green line).

and the theoretical curves is reported for the TIP5P and SPC/E models (left and right panels, respectively). The $\gamma^{(n)}$ signals are shown multiplied by k squared for better visualization. The first two curves from the top of each panel are the Cd-O and Cd-H first shell $\gamma^{(2)}$ contributions. The reminder of the Figures shows the total theoretical contributions compared with the experimental spectra and the resulting residuals. As expected, the dominant contribution to the total EXAFS signal is given by the ion-O first shell signal and, as a consequence, the EXAFS structural information is particularly accurate for the shape of the ion-O $g(r)$'s first peaks, only. Overall, the calculated EXAFS spectra match the experimental data reasonably well in both cases, with R_i values of 0.466 10^{-6}, and 0.147 10^{-5}, for TIP5P and SPC/E, respectively (see section 4.2 for the definition of this index of agreement). Therefore, the structural and dynamical information derived from the two

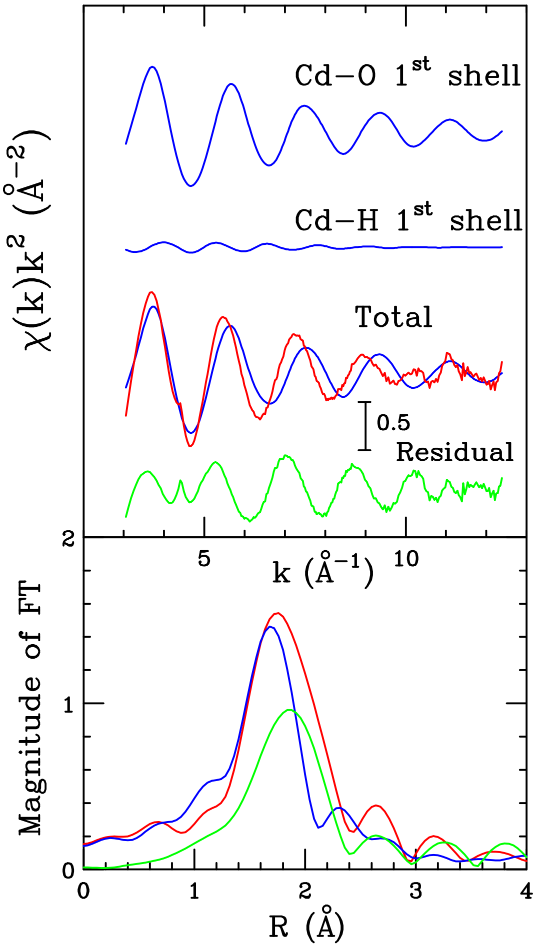

Figure 8.8: Upper panel: Comparison between the TIP5P EXAFS theoretical signal (blue line) calculated from the hexacoordinated Cd-O and Cd-H g(r)'s and experimental data (red line). The residual signal (green line) is also shown. Lower panel: Nonphase-shifted corrected Fourier transforms of the experimental data (red line), of the theoretical signal (blue line), and of the residual curve (green line).

simulations is basically correct. However, the TIP5P $\chi(k)$ theoretical signal is in better agreement with the experimental data, thus showing that this simulation provides a better description of the water structure around the Cd^{2+} ion. In the SPC/E EXAFS analysis the presence of a leading frequency can be clearly identified in the residual curve. This behavior is due to the slightly short value of the average Cd-O first shell distance obtained from the SPC/E as compared to the experimental results. This finding is reinforced by the Fourier transform (FT) moduli of the EXAFS $\chi(k)k^2$ theoretical, experimental and residual signals shown in the lower panels of Figure 8.7. The FT's have been calculated in the k-range 3.1–12.0 $Å^{-1}$ with no phase shift correction applied. A quite good agreement between the FT's of the theoretical and experimental signals has been found for the TIP5P simulation, while the experimental first-neighbour peak is found to be broader and shifted towards longer distances than predicted by the SPC/E simulation. This is a first indication that the EXAFS experimental data are better reproduced by a flexible hydration shell with a high percentage of heptacoordinated species.

In the second step of the analysis we have examined the compatibility of the EXAFS spectrum with the existence of an octahedral geometry of the solvated Cd^{2+} ion. To this end we calculated the $\chi(k)$ theoretical signal associated with the hexacoordinated g(r)'s reported in the middle panel of Figure 8.2. The results are shown in Figure 8.8. Both the frequency and the amplitude of the theoretical curve are in poor agreement with the experimental data. In particular, the Cd-O first shell distance of the hydrated cluster is too short, and the octahedral complex is too stiff as compared to the experimental determination. This is reflected in a different width and peak position of the FT of the $\chi(k)$ signals, as shown in the lower panel of Figure 8.8. However in order to determine in an accurate way the coordination number of the Cd^{2+} ion, one has to go beyond the EXAFS spectroscopy, and resort to the analysis of the low-energy region of the spectrum.

8.3 XANES analysis

As for the Hg^{2+} ion, a quantitative analysis of the XANES data can provide a definitive answer on the hydration properties of the Cd^{2+} ion. In particular, XANES can allow us to determine the percentage of hexa- and heptacoordinated complexes present in aqueous solution. In fact, while for the Hg^{2+} aqua ion a uniform picture emerged from the simulations carried out with different water models, for the Cd^{2+} ion we have obtained from the two trajectories a different behaviour: in the TIP5P case, the dominant species existing in solution is the heptacoordinated one, while the SPC/E simulation increases the stability of the hexacoordinated cluster (see section 8.1). Therefore, we have analysed the XANES spectra starting in this case from both our SPC/E and TIP5P trajectories.

8.3.1 Computational procedure

Since our investigation on the Hg^{2+} aqua ion has shown that the second hydration shell provides a detectable contribution to the XANES spectrum, in the XANES calculations for Cd^{2+} we have directly included both the first and second hydration shells in the analysis. In particular, we have considered all the water molecules separated from the Cd^{2+} ion by a distance shorter than 5.2 Å, since, as already observed for Hg^{2+}, water molecules at larger distance have been found to provide a negligible contribution. Two trajectories containing the Cd^{2+} ion and its first two hydration shells have been extracted from the total simulation in different time ranges when the ion was six- and seven-coordinated. As a result the former trajectory contained first shell hexahydrated clusters only, while the latter contained first shell hepta-coordinated clusters. From each trajectory we extracted 100 snapshots saved every 30 ps. Each snapshot has been used to generate the XANES spectrum associated with the corresponding instantaneous geometry, and the averaged theoretical spectrum has been obtained by summing all the spectra and dividing by the total number of Molecular Dynamics snapshots used.

To evaluate the number of spectra which are necessary to have a statistically significant average, we have calculated the RMS residual function defined in equation 7.1, choosing the same residual value of 10^{-4}, as used

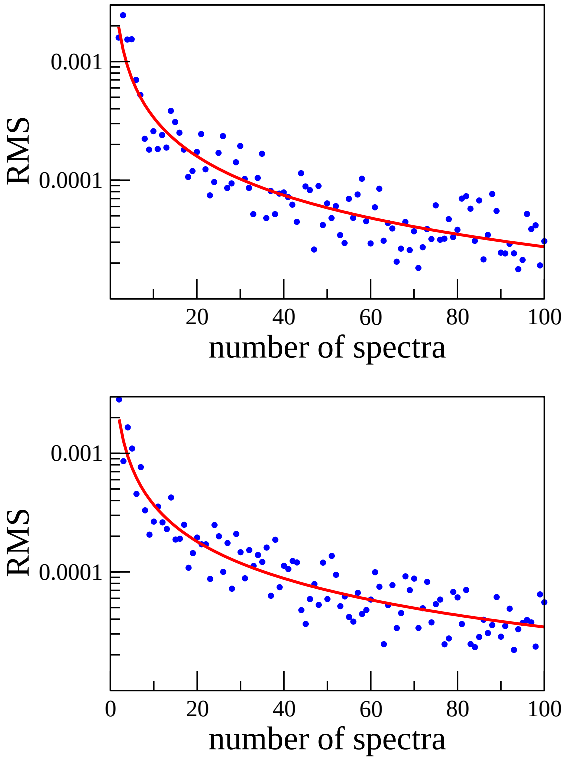

Figure 8.9: Residual function RMS of the XANES averaged spectra as a function of the number of Molecular Dynamics snapshots for hexa- (upper panel) and hepta- (lower panel) coordinated first shell clusters.

in the Hg^{2+} XANES analysis. The RMS function obtained from the TIP5P simulation is simulation is plotted in Figure 8.9 against the number of averaged spectra for the hexa- and heptacoordinated first shell clusters (upper and lower panels, respectively). It is seen that 100 configurations are enough to reach convergence in both cases. The results obtained from the SPC/E simulations are very similar and they are not shown for the sake of brevity.

8.3.2 XANES-Molecular Dynamics results

In the first step of the analysis we have calculated the theoretical XANES spectra associated with the hexa- and heptahydrated complexes. It is useful to repeat that to this end we have extracted from each simulation two trajectories, the former containing first shell hexahydrated clusters only, the latter containing first shell heptacoordinated clusters. In both cases also the second hydration shell was considered including all the water molecules up to 5.2 Å. The corresponding XANES theoretical spectra obtained from 100 Molecular Dynamics snapshots (not including intrinsic and extrinsic inelastic processes) are shown in Figure 8.10. In the upper panel we report the XANES theoretical spectra associated with the octahedral hydration structure for the TIP5P and SPC/E simulations, while the XANES theoretical spectra calculated from the heptacoordinated clusters are shown in the lower panel. As evident from the plots, the spectra obtained from the two simulations are identical. This finding is in line with the results of our Molecular Dynamics simulations which have shown that the structure of hexa- and heptahydrated complexes is the same regardless the water model used in the simulations. Therefore, in the following analysis we can use only one simulation and in particular the TIP5P trajectory has been chosen.

In the second step we exploited the potentiality of the XANES technique to unveil the structural and dynamic properties of the hydration sphere. In particular we were interested in determining the percentage of the heptaaqua and hexaaqua ion existing in solution. To this end we calculated several averaged theoretical spectra with variable percentages of hexa- and heptacoordinated complexes, and the obtained trend is depicted in Figure 8.11. The calculated XANES spectra present noticeable differences all along the energy range, showing the sensitivity of XANES to the geometry of the Cd^{2+} hydration cluster. In particular, the edge intensity is lowered in the spectrum containing only the hexacoordinated complex which exhibits a different shape in the region around 18 eV and a shift of the first maximum around 50 eV. To determine the effective percentages of the two hydration complexes it is necessary to compare the total averaged XANES spectra with the experimental data. To this end, as in the case of Hg^{2+}, all inelastic processes have been accounted for by convoluting the theoretical averaged spectra with a broadening Lorentzian function, and the corresponding E_s and A non-structural parameters have been optimized. The agreement between the experimental and theoretical data has been evaluated also here by the goodness-of-fit parameter (R_{sq}) and in Figure 8.12 the R_{sq} values are plotted against the

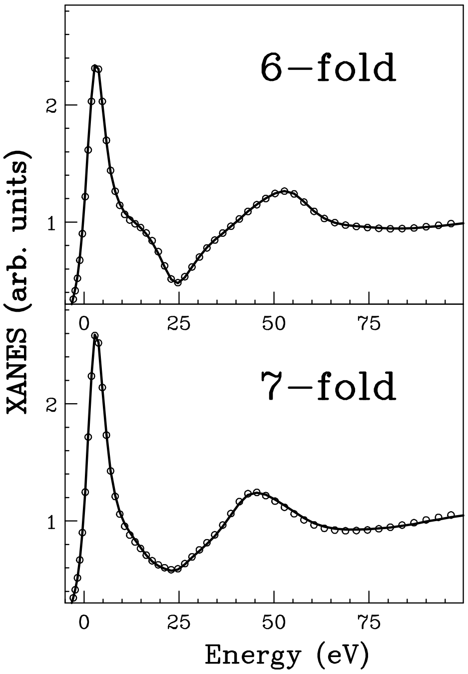

Figure 8.10: Upper panel: Comparison of the averaged XANES theoretical spectra of Cd^{2+} in aqueous solution calculated from the TIP5P (solid line) and SPC/E (dotted line) hexacoordinated simulation frames. Lower panel: Comparison of the averaged XANES theoretical spectra of Cd^{2+} in aqueous solution calculated from the TIP5P (solid line) and SPC/E (dotted line) heptacoordinated simulation frames.

percentage of hexahydrated clusters used in the calculation of the averaged XANES theoretical spectrum. The picture that emerges is quite informative, because the R_{sq} shows a regular trend and the lowest R_{sq} value is obtained for a hexacoordinated cluster percentage of 20%, which corresponds to the result obtained from the TIP5P simulation. Thus, this XANES analysis establishes that the TIP5P simulation provides a very accurate description of the hydration properties of the Cd^{2+} ion. In Figure 8.13, the averaged theoretical spectra corresponding to hexacoordinated cluster percentage of 100% and 20% are compared with the experimental spectrum (panel A and B, respectively). In the former case, as expected, the agreement between the experimental and theoretical spectra is not very good both in the low- and high-energy regions of the spectrum. The intensity of the edge is too small and the position of the first maximun is shifted as compared to the experimental data. Conversely, the averaged theoretical spectrum calculated from the TIP5P trajectory is in excellent agreement with the experimental data in all the energy range (panel B of Figure 8.13). Small discrepancies between the theoretical and experimental spectra are found at about 20 and 70 eV from the edge. A recent investigation on the multielectron photoexcitations affecting the Cd K-edge has been carried out by Kodre et al. [110]. The analysis of the Cd vapour absorption spectrum after natural-width deconvolution has revealed the presence of anomalous features in the atomic background due to multielectron excitation. In particular the onset of the $1s4d$ and $1s4p$

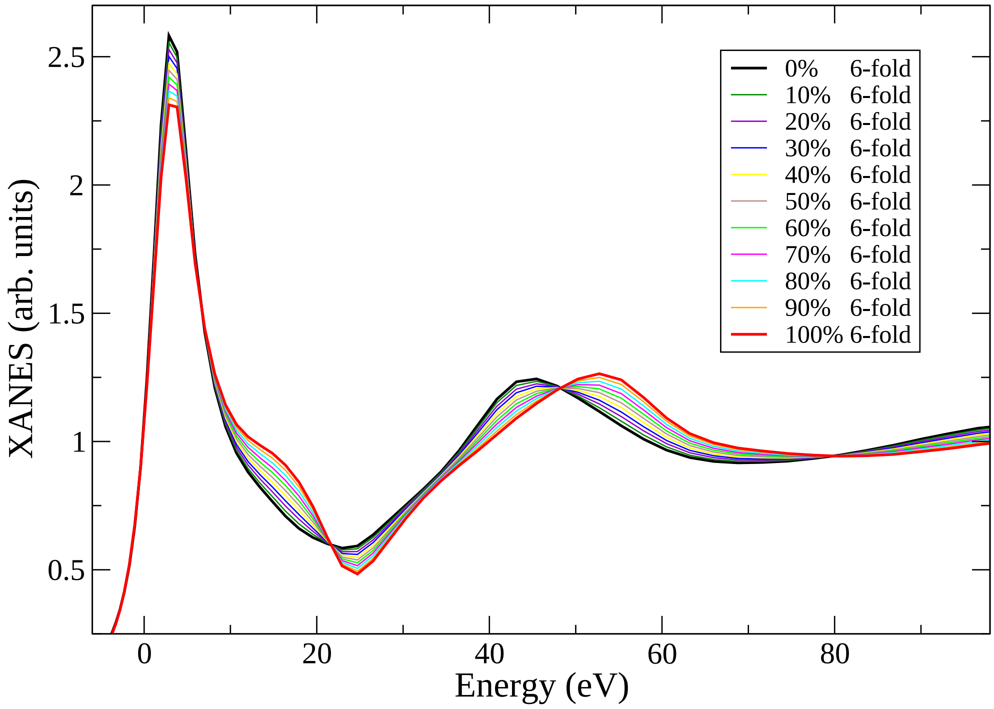

Figure 8.11: Averaged XANES theoretical spectra of the Cd^{2+} aqua ion calculated with different percentages of hexa- and heptacoordinated clusters.

multielectron transitions, was found at about 20 and 70 eV, respectively, and this phenomenon could be responsible for the small discrepancies found in our MXAN analysis, which is carried out in the one-electron approximation. Therefore, according to our XANES analysis, the Cd^{2+} ion in aqueous solution forms a flexible first coordination shell in which an heptacoordinated structure plays a dominant role.

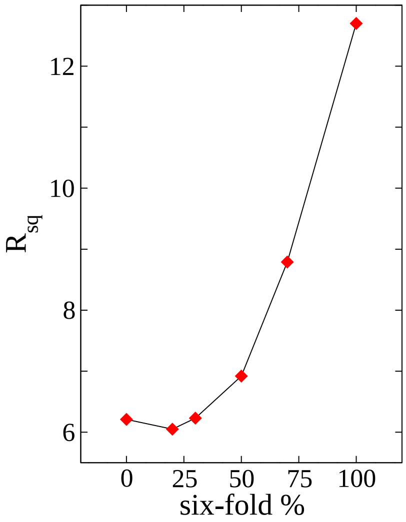

Figure 8.12: Goodness-of-fit (R_{sq}) versus percentage of hexacoordinated complex of the XANES theoretical spectra.

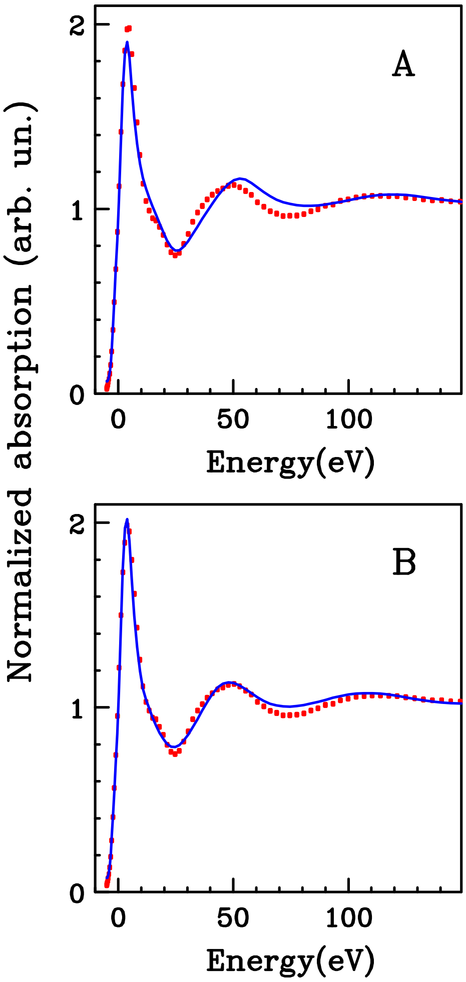

Figure 8.13: Panel A: Comparison between the averaged XANES theoretical spectrum including the hexacoordinated first shell clusters (blue solid line) and the deconvolved experimental data (red dotted line) of Cd^{2+} in aqueous solution. Panel B: Comparison between the averaged XANES theoretical spectrum calculated from the TIP5P trajectory with 80% of hepta- and 20% of hexacoordinated clusters (blue solid line) and the deconvolved experimental data (red dotted line).

Chapter 9

Zn^{2+} aqua ion in high-density water

9.1 Hydration structure from Molecular Dynamics simulations

As already mentioned, the hydration properties of the Zn^{2+} aqua ion have been studied in conditions of very high pressures. The Zn-O and Zn-H radial distribution functions and the corresponding running integration numbers calculated from the three simulations (carried out at pressure values of 0.1 MPa, 1.0 GPa and 2.2 GPa) are shown in Figure 9.1. In all cases a well defined first peak followed by a depletion zone can be observed, indicating the existence of a stable first hydration shell. In all the three simulations, the integration over Zn-O and Zn-H g(r) first peaks gives coordination numbers of six and twelve, respectively. Therefore, the first important result is that the change of thermodynamic conditions does not affect the first coordination shell hydration number. Note that no exchange events between first and second shell water molecules have been observed during the total simulation time in all the three cases, and this is in agreement with the fact that the residence time of water molecules in the first hydration shell of Zn^{2+} is estimated to be in the microsecond time scale (at ambient conditions) [14]. The comparison of the first peak of the Zn-O radial distribution functions is reported in Figure 9.2. As it can be seen, the radial distribution functions show a shortening of the peak position (the first peak distances are 2.06, 2.05 and 2.04 Å for the simulations at 0.1 MPa, 1.0 GPa and 2.2 GPa, respectively), and a broadening of the first peak as pressure increases, meaning that in compressed water the first coordination shell becomes more disordered. The effect of pressure changes is more evident on the structure of the second coordination shell, as shown in Figure 9.3. In this case the shift of the peak positions at shorter distances is about 0.11 Å (the second peak distances are 4.32, 4.26 and 4.21 Å for the simulations at 0.1 MPa, 1.0 GPa and 2.2 GPa, respectively), larger than what observed for the first hydration shell. If we always calculate the hydration number by integration up to the Zn-O g(r) second minimum, we obtain from the three simulations increasing coordination number values

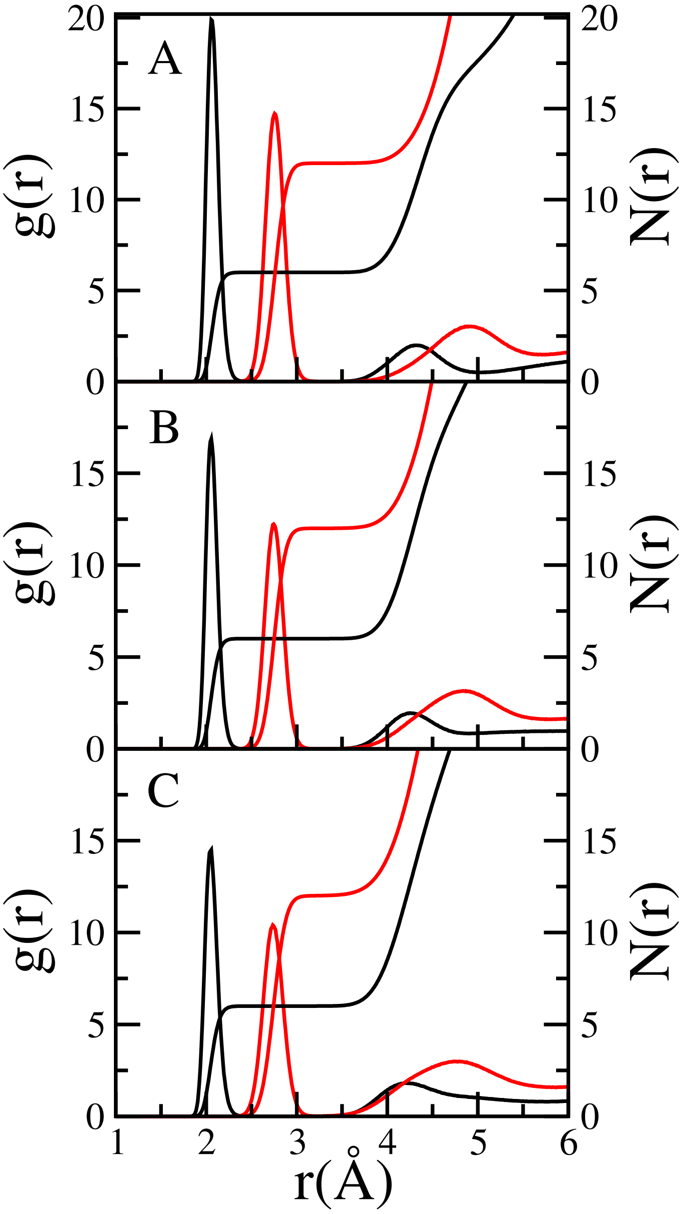

Figure 9.1: Radial distribution functions of the Zn-O (black line) and Zn-H (red line) g(r)'s and corresponding running integration numbers. Panels A, B and C show the distribution functions obtained from the trajectories at 0.1 MPa, 1.1 GPa and 2.2 GPa, respectively.

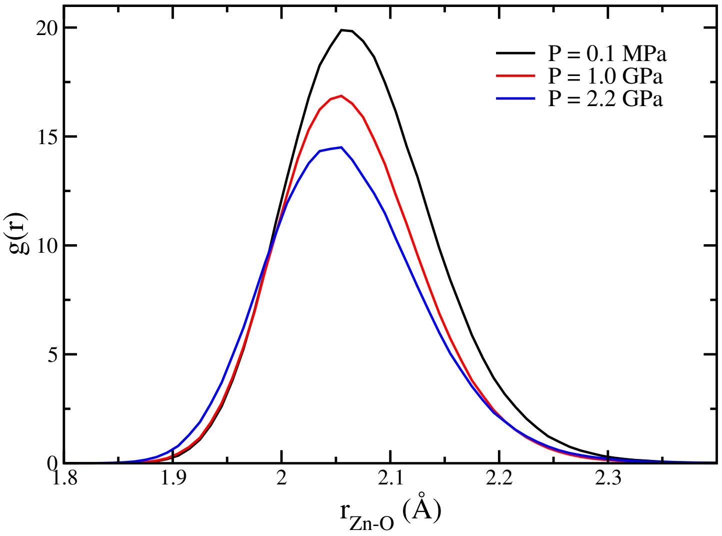

Figure 9.2: Comparison of the Zn-O g(r) first peak obtained from the three simulations.

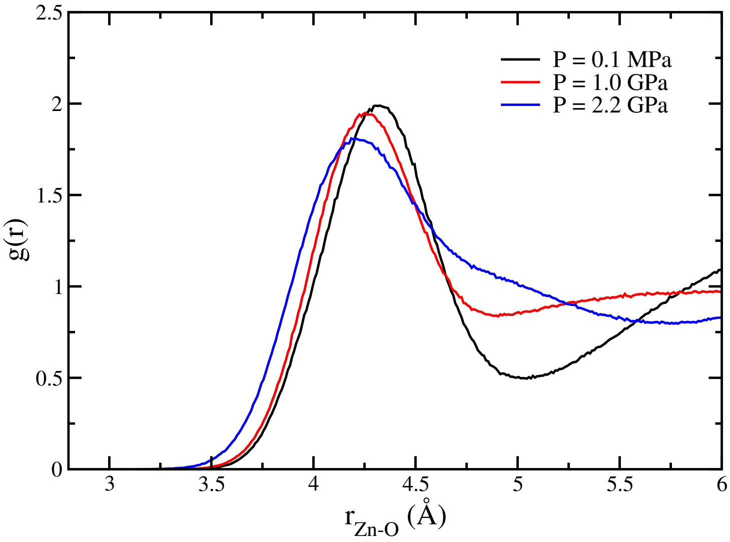

Figure 9.3: Comparison of the Zn-O g(r) second peak obtained from the three simulations.

(they are 10.9, 12.7 and 29.0 for the simulations at 0.1 MPa, 1.0 GPa and 2.2 GPa, respectively). Moreover, the g(r) changes qualitatively increasing the pressure in the zone after 4.5 Å, and at 2.2 GPa the second hydration shell is not only shifted towards the metal ion but expands outwards and is less separated from bulk water.

The a.d.f. obtained from the three simulations of the ϕ, ζ and ψ angles (plotted as functions of $1 - cos(\phi, \zeta, \psi)$) are depicted in Figure 9.4. Panel A shows the a.d.f. of the ϕ angle. The distribution functions calculated from the three trajectories show very sharp peaks at $cos(\phi) = 1$, as the oxygen atoms point towards the ion and wagging movements are very limited. The curves related to the simulation at 0.1 MPa and 1.0 GPa go to zero at $\phi = 30°$ while

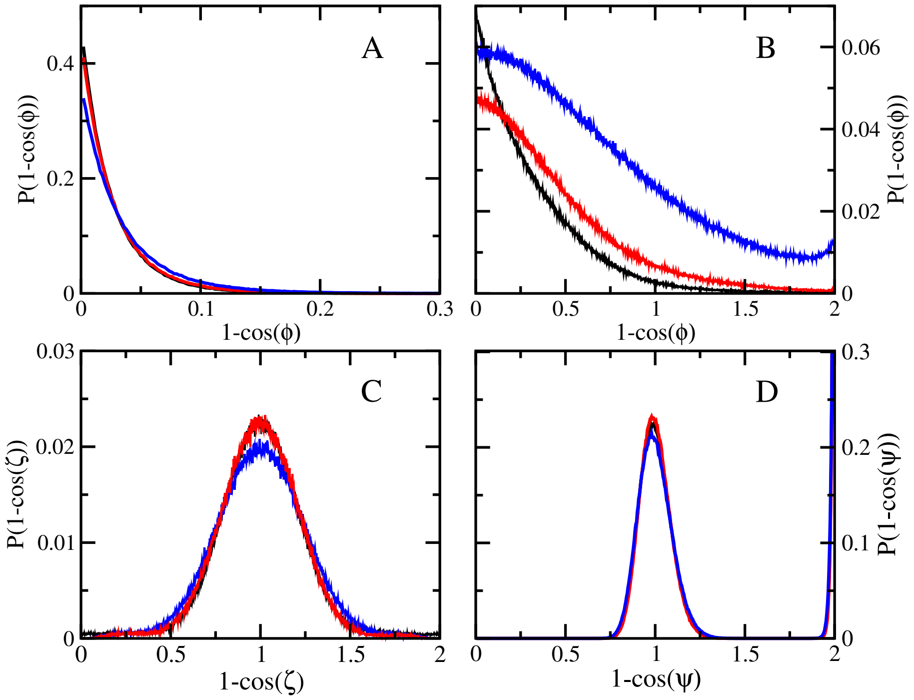

Figure 9.4: Angular distribution functions (a.d.f.) of the ϕ, ζ and ψ angles obtained from the simulations at 0.1 MPa (black lines), 1.0 GPa (red lines) and 2.2 GPa (blue lines). Panel A: First hydration shell a.d.f of ϕ. Panel B: Second hydration shell a.d.f. of ϕ. Panel C: First hydration shell a.d.f. of ζ. Panel D: First hydration shell a.d.f. of ψ.

the a.d.f. obtained at 2.2 GPA drops to zero at $\phi = 38°$ since the increase of pressure allows a small increase of rotational freedom of the first shell water molecules. Analysis of the second coordination shell a.d.f. gives quite different results: the simulations at 1.1 and 2.2 GPa show a change of slope and concavity as compared to the distribution obtained from the trajectory at 0.1 MPa, as the packing of the second hydration sphere forces the water molecules to deviate from the minimum potential energy alignment observed in the first simulation. Moreover, they are much broader, showing that in the compressed solutions wagging movements in the second hydration shell are much more important than at 0.1 MPa. Panel C shows the a.d.f. of ζ angle. In all cases there is a well defined peak centered at $cos(\zeta) = 90°$, showing that the Zn-O vector is located most of the time in the water molecular plane with a maximum deviations at 90° and 120°. The a.d.f.'s for the ψ angle calculated at various pressures on first shell water molecules are shown in Panel D. All of the distribution functions have nearly coincident maxima at $1 - cos(\psi) = 1$ and $1 - cos(\psi) = 2$ (corresponding to ψ values of 90° and 180°), thus showing the existence of a stable octahedral coordination geometry, even at high pressures. Moreover, the a.d.f.'s go to zero at intermediate values, as water molecules are strongly constrained in their positions and large distortions of the octahedral symmetry are not expected.

The three-dimensional structure of the Zn^{2+} first and second solvation shell can be observed looking at the spatial distribution functions calculated from the simulations at 0.1 MPa and 2.2 GPa, that are shown in Figures 9.5 and 9.6. The isodensity surfaces of oxygen and hydrogen atoms are coloured green and red, respectively. It can be clearly seen the octahedral symmetry of the first hydration shell both at 0.1 MPa and at 2.2 GPa, and the compression of the hydration shells going from ambient to the high-density conditions. The torus-shaped distribution of hydrogen atoms in the first hydration shell means that they are free to rotate around the water dipole, in agreement with the a.d.f. analysis. The outer green rings, which correspond to oxygen atoms in the second coordination shell, are much thicker and wider at 2.2 GPa, as more and more water molecules are packed into the second hydration shell and the second coordination shell becomes more disordered and unstructured.

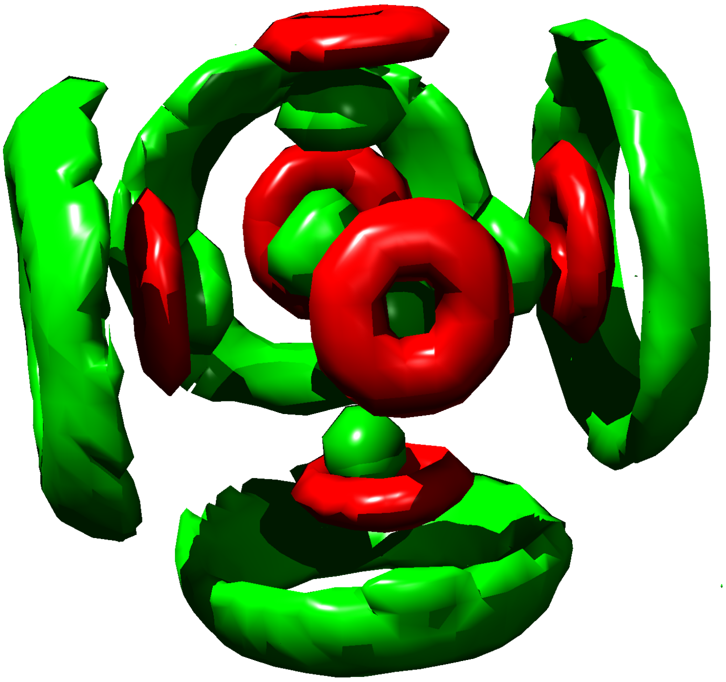

Figure 9.5: Spatial distribution functions of oxygen (green) and hydrogen (red) atoms obtained from the trajectory at 0.1 MPa. The isodensity surfaces at level 2.23 are shown.

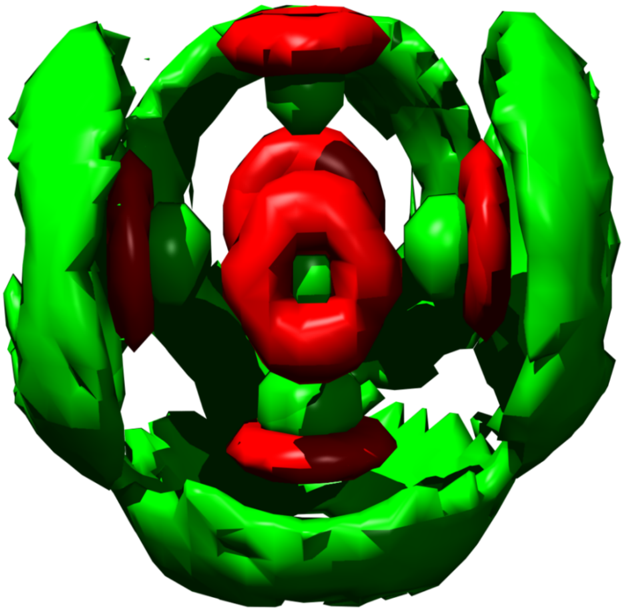

Figure 9.6: Spatial distribution functions of oxygen (green) and hydrogen (red) atoms obtained from the trajectory at 2.2 GPa. The isodensity surfaces at level 2.23 are shown.

9.2 Dynamic properties from Molecular Dynamics simulations

Even if the increase of pressure does not have a huge effect on the structure of the first hydration shell, it has a stronger effect on the dynamics of the water molecules surrounding the ion. This effect can be investigated by calculating the first rank reorientational correlation times of water molecules belonging to the Zn^{2+} first hydration shell and to bulk water. Figure 9.7 shows the trend of $C_1^{\alpha}(t)$ for the trajectories at 0.1 MPa and 2.2 GPa while all of the reorientational correlation times obtained from the three simulations are reported in Table 9.1. First of all, it is useful to compare the behaviour of the different vectors within the same simulation and in all cases the same general trend is observed. As far as the Zn^{2+} first hydration shell is concerned, the dominant motion is the rotation around the water dipole, as expected and as already observed for the Hg^{2+} aqua ion. Moreover, this finding is in agreement with the results of our analysis of angular and spatial distribution functions. The reorientational dynamics of the $\mathbf{u}^N$ vector is the fastest reorientation while the $\mathbf{u}^{HH}$ correlation time is slightly greater than that related to the $\mathbf{u}^N$ vector. In bulk water the $\mathbf{u}^N$ vector rotates faster than $\mathbf{u}^{HH}$ and $\mathbf{u}^D$, showing that water rotation in the bulk is anisotropic. This result has been already observed in Molecular Dynamics simulations of pure water and it is possibly related to the geometry of the water model [97]. By comparing the results of the three simulations, we can clearly see that all of the correlations times, both in the first shell and in bulk water, decrease with increasing pressure. This means that water mobility increases and it is a consequence of the rupture of the hydrogen bond network caused by the pressure increase. Always as a consequence of the disruption of the hydrogen bond network the Zn^{2+} ion diffusion coefficient increases (it is $0.70(0.07){\cdot}10^{-5}$ cm^2/s, $0.91(0.08){\cdot}10^{-5}$ cm^2/s, $1.10(0.08){\cdot}10^{-5}$ cm^2/s for the simulations at 0.1 MPa, 1.0 GPa and 2.2 GPa, respectively). Note that at ambient conditions we have obtained a diffusion coefficient in very good agreement with the experimental determination, which is equal to $0.703{\cdot}10^{-5}$ cm^2/s [107]. We could not find any experimental values of this quantity under high pressure.

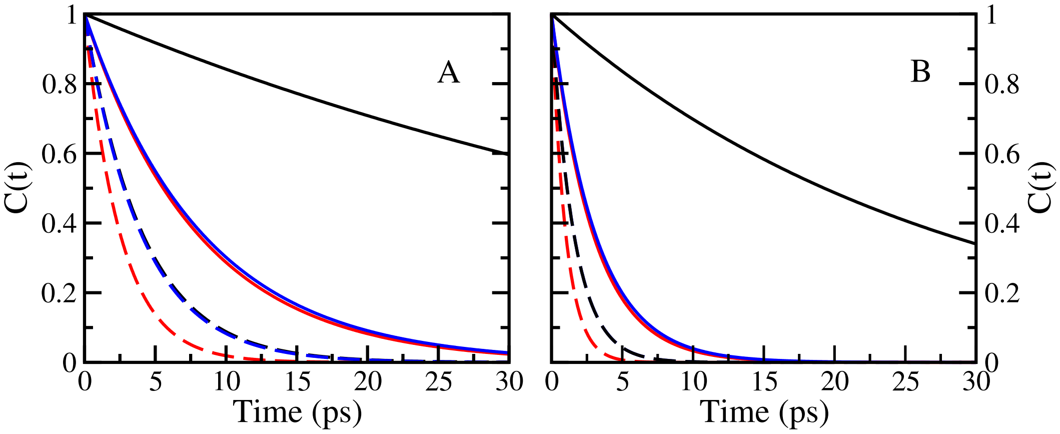

Figure 9.7: First rank reorientational correlation functions $C_1^{\alpha}(t)$ of water molecules belonging to the Zn^{2+} first hydration shell and to bulk water. Continuous lines refer to the ion first coordination shell, while dashed lines to bulk water. Different colors are related to the different vectors: $\vec{u}^D$ (black lines), $\vec{u}^{HH}$ (blue lines), and $\vec{u}^N$ (red lines). Panel A: Reorientational correlation functions obtained at 0.1 MPa. Panel B: Reorientational correlation functions obtained at 2.2 GPa.

	vector	0.1 MPa	1.0 GPa	2.2 GPa
first shell	$\mathbf{u}^N$	7.7 (0.2)	5.0 (0.3)	2.9 (0.3)
	$\mathbf{u}^{HH}$	8.1 (0.3)	5.9 (0.3)	3.0 (0.4)
	$\mathbf{u}^D$	57.5 (5.5)	33.8 (5.0)	27.3 (4.2)
bulk water	$\mathbf{u}^N$	2.5 (0.1)	2.1 (0.1)	1.1 (0.2)
	$\mathbf{u}^{HH}$	3.9 (0.1)	3.3 (0.1)	1.6 (0.2)
	$\mathbf{u}^D$	3.9 (0.1)	3.3 (0.1)	1.6 (0.2)

Table 9.1: Reorientational correlation times (ps) of water molecules belonging to the Zn^{2+} first hydration shell and to bulk water obtained from the three simulations. Standard deviations are given within parentheses.

9.3 EXAFS analysis of Zn^{2+} in high-density water

For the Zn^{2+} ion, the reliability of the Molecular Dynamics simulation carried out at ambient conditions has been already assessed by our group in Ref. [7], by comparing the first shell parameters obtained from the EXAFS data analysis and the structural results of the Molecular Dynamics trajectory. In particular, an excellent agreement between the theoretical and experimental Zn-O first shell distance was obtained.

As far as the trajectories in high-density conditions are concerned, a direct comparison between the structural results of the Molecular Dynamics simulations and EXAFS data analysis is not possible, as the theoretical calculations and the experimental measurements have been carried out at different pressure values. However, the trend with pressure of the structural parameters obtained by fitting the EXAFS experimental signals can be compared with the trend of the theoretical structural results. Figure 9.8 shows the comparison between the experimental EXAFS signals and the fitting results for pressure values of 0.2, 1.97, and 2.85 GPa. Note that both the quality of the experimental data and the agreement with the theoretical curve are excellent in all three cases.

The experimental Zn-O g(r) first peaks obtained from the EXAFS minimizations are shown in Figure 9.9 while the corresponding structural parameters are reported in Table 9.2. The experimental g(r)'s confirm that the Zn^{2+} ion first shell hydration number does not vary with pressure and the shift of the Zn-O g(r) peak position to shorter distances with increasing pressure is

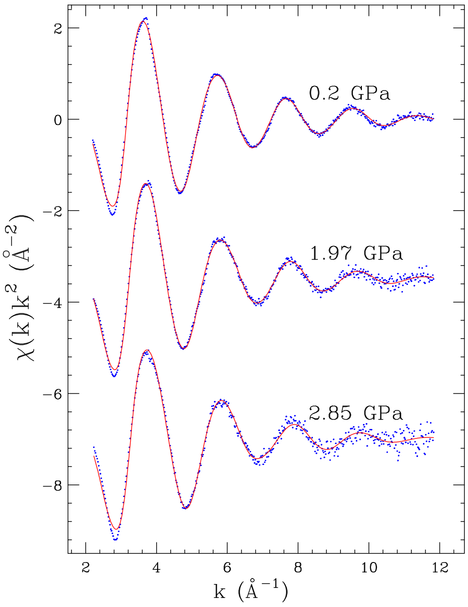

Figure 9.8: Comparison between the experimental (blue dotted line) and theoretical (red solid line) EXAFS signals for the Zn^{2+} ion at different pressure values.

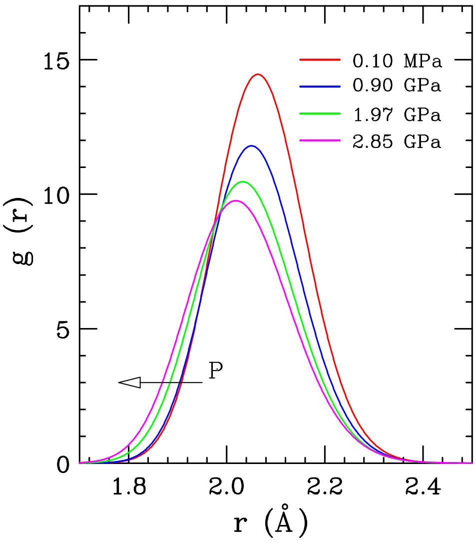

Figure 9.9: Pressure dependence of the Zn-O g(r) first peak obtained from the analysis of EXAFS spectra.

a clear evidence of the first hydration shell compression. The higher mobility of water molecules in the first coordination sphere is highlighted by the trend of the Debye-Waller factors, that increase with increasing pressure (see Table 9.2). Moreover, the shift towards shorter distances of the Zn-O rising edge indicates that the oxygen atoms of the first shell water molecules in compressed water can reach regions of space not normally accessible at ambient conditions. All together, these findings reinforces the theoretical results obtained from our Molecular Dynamics simulations.

Pressure	R(Å)	σ^2(Å^2)	β
0.10 MPa	2.078	0.0087	0.2
0.20 GPa	2.074	0.0088	0.2
1.97 GPa	2.053	0.010	0.2
2.85 GPa	2.042	0.011	0.2

Table 9.2: Structural parameters of Zn-O g(r) first peak obtained from the EXAFS analysis. R is the mean Zn-O distance, σ is the Debye-Waller factor, and β is the asymmetry parameter.

9.4 Water structure under pressure

As already mentioned in the Introduction chapter, the second aim of our study on compressed water is the investigation of the structural transformations occurring to water from the low- to the high-density conditions. To this end, we have carried out three Molecular Dynamics simulations of pure water at the same thermodynamic conditions used in the simulations of the Zn^{2+} ion (and using also the same general Molecular Dynamics protocol described in section 6.1.1).

The first question we address is: *"Does the presence of the ion affect the structure of water?"* To give an answer to this question we have calculated the oxygen-oxygen g(r) in the pure water simulations and in the second hydration shell of the Zn^{2+} ion. More specifically, the O-O g(r)'s in the second hydration shell of the ion means that in equation 6.1 particle A is the oxygen atom of a water molecule in the Zn^{2+} second hydration shell and particles B are the oxygen atoms of all the other water molecules. An in-house written code has been used in which an average over all the oxygen atoms in the ion second coordination shell has been performed. The results of this analysis for the three simulations are shown in Figure 9.10. At ambient conditions the two g(r)'s are very similar. The most important difference is the peak at about 5.5 Å which is found only in the g(r) calculated in the second hydration shell of the ion. However, this peak is only due to the presence of first shell water molecules that are strongly oriented by the ion. This hypothesis is confirmed by calculating the oxygen-oxygen g(r) in the second hydration shell of the Zn^{2+} ion but excluding the first shell water molecules from the computation. The result is also shown in Figure 9.10, and it is clearly seen that the peak in this case disappears. The lower intensity of this g(r) as compared to the previous one is due to the fact that we did not apply any correction for the excluded volume effects caused by the removal of the first hydration shell. The small differences between the g(r)'s calculated in pure water and in the second shell of the ion completely disappear under high pressure and at 2.2 GPa the two g(r)'s become almost identical. All together, these findings show that solvent molecules are not significantly influenced by the presence of the solute, and the impact of the ion on the structure of water does not extend in an appreciable way beyond the first hydration shell. This is an interesting result that contradicts the results of neutron diffraction studies [111, 112] which support the idea that ionic solutes have a strong effect on the microscopic structure of water and, according to the same studies, this effect is similar to the application of high external pressures to pure water. In our opinion the effects of ions on the structure of water surely depends on the particular ion under investigation, but in general cannot be determined by neutron diffraction experiments which uses solute concentrations so high that all the water molecules in practice belong to the first hydration shell of the cation or of the anion.

The second question we address is: *"How does the water structure change with increasing pressure?"* In order to understand this, we compare in Figure

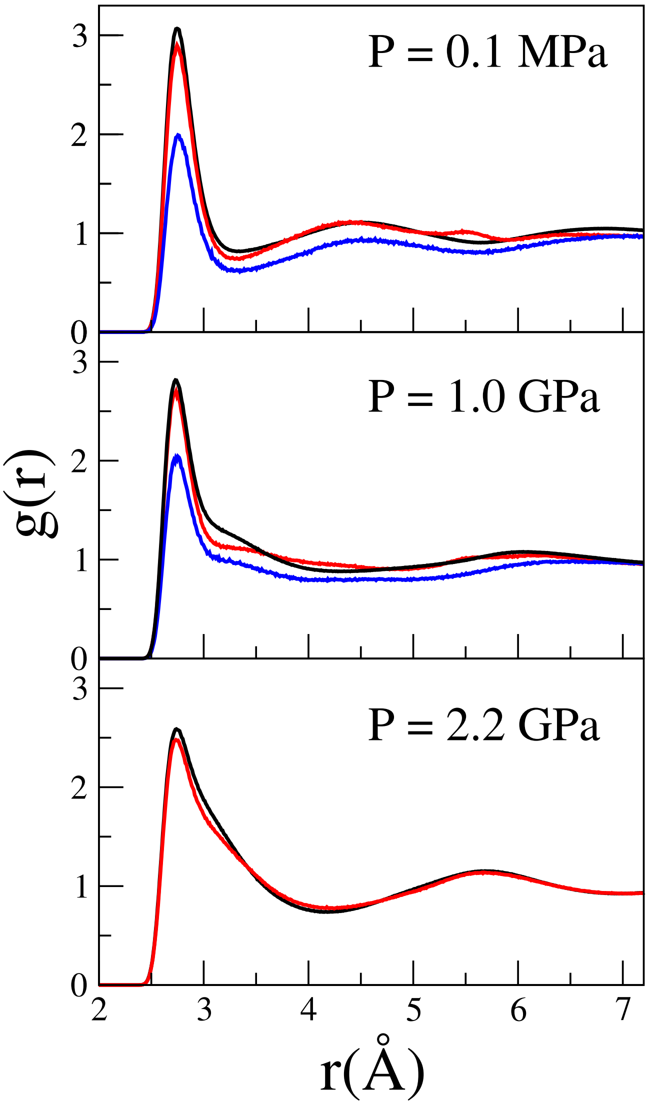

Figure 9.10: Oxygen-oxygen radial distribution functions calculated from the pure water simulations (black lines), and for the second hydration shell of the Zn^{2+} ion including (red lines) or excluding (blue lines) the contribution of the oxygen atoms belonging to the ion first coordination shell.

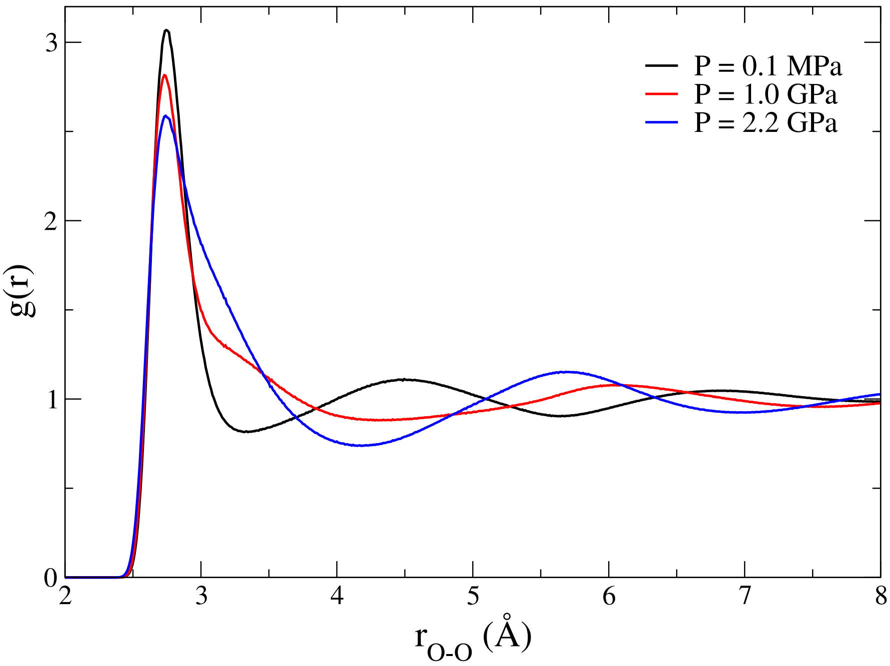

Figure 9.11: Comparison of the O-O g(r) obtained from the Molecular Dynamics simulations of pure water carried out at P=0.1 MPa, 1.0, and 2.2 GPa.

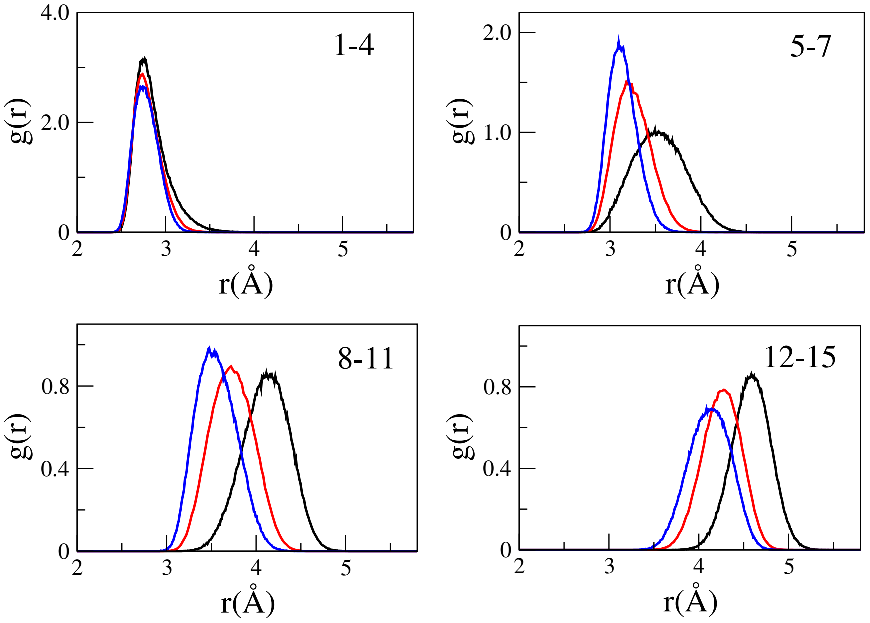

Figure 9.12: Distance distributions of groups of 1^{st} to 4^{th}, 5^{th} to 7^{th}, 8^{th} to 11^{th}, and 12^{th} to 15^{th} oxygen neighbors around a central water molecule for P=0.1 MPa (black), 1.0 (red), and 2.2 GPa (blue).

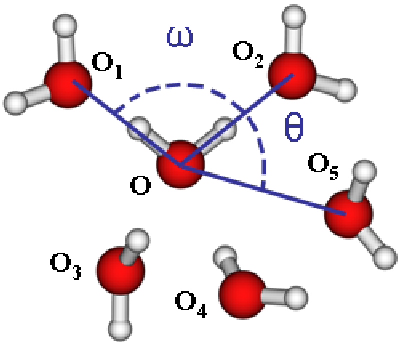

Figure 9.13: Definition of the ω and θ angles used in the study of water structure under high pressure.

9.11 the O-O g(r)'s obtained from the three pure water simulations. Remarkable differences among the oxygen-oxygen g(r)'s are found with increasing pressure. At 2.2 GPa a pronounced broadening of the O-O g(r) first peak, as compared to ambient conditions, is observed. This change is caused by a collapse of the second coordination shell, as a consequence of the rupture of the hydrogen bond network. To get a deeper insight into the structural properties of water under high pressure we have calculated the oxygen-oxygen g(r)'s of distance ranked groups of neighbours with respect to a given central water molecule (see Figure 9.12). The group of first neighbours (1-4) which form the tetrahedral first shell of the water molecule behaves rigidly with an almost density independent distance from the central molecule, in agreement with neutron diffraction results [29]. On the contrary, distances of all the other groups of neighbours decrease with increasing density, as more and more water molecules move towards the central one.

The angular distribution functions of two different angles (whose definition is reported in Figure 9.13) have been also computed using in-house written codes. Figure 9.14 shows the distribution of the oxygen-oxygen-oxygen angle (labelled as ω) calculated on the four closest water molecules around a central oxygen atom. At ambient conditions the distribution is highly peaked at 109°, corresponding to a tethaedral arrangement. The intensity and the position of this peak is only slightly affected by increasing pressure, as the tetrahedral first shell cluster is minimally distorted in the high density conditions. Conversely, the intensity of the peak at 55° significantly increases at higher pressures and the position of the peak moves towards larger angles. This is the angle formed by the central oxygen, a tethraedral neighbouring oxygen and the oxygen of an interstitial water molecule which is not hydrogen bonded to any of the first shell molecules. The trend of ω distributions then suggests that the number of interstitial molecules increases under high pressure and that these molecules collapse into the first coordination shell. This result is reinforced by the analysis of the O_{1-4}-O-O_{5-8} angle (labelled as θ), i.e. the angle about an oxygen atom formed by its first four (1-4) and second

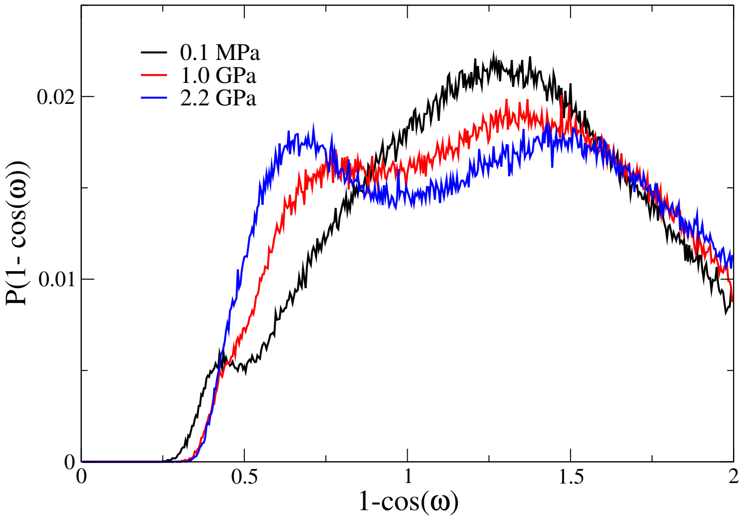

Figure 9.14: Angular distribution functions of the O-O-O angle (ω) calculated on the four closest water molecules around a central oxygen atom obtained from the three Molecular Dynamics simulations of pure water.

four (5-8) neighbours (see Figure 9.15). At ambient conditions two peaks at 48° and 75° are found which merge into a single peak centered at 60° at 2.2 GPa. As already pointed out in Ref. [28], this peak is due to the presence of interstitial molecules. The growth of interstitial molecules with pressure has been also highlighted by computing the average number of hydrogen bonds ($\bar{n}$) formed by the first four (1-4) and second four (5-8) neighbours of a central water molecule. A decrease of $\bar{n}$ from 0.65 at ambient conditions to 0.43 at 2.2 GPa was found, in line with previous results.

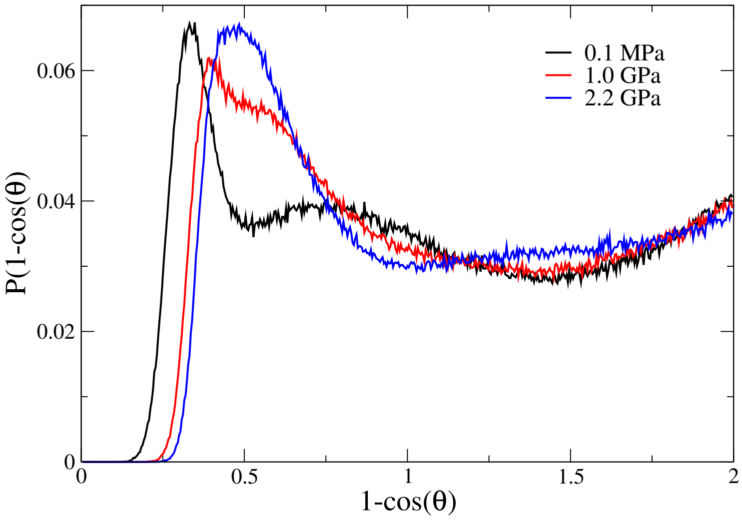

Figure 9.15: Angular distribution functions of the O_{1-4}-O-O_{5-8} angle (θ) obtained from the three Molecular Dynamics simulations of pure water. O_{1-4}-O-O_{5-8} is the angle about an oxygen atom formed by its first four (1-4) and second four (5-8) neighbors.

Chapter 10

Bromide ion hydration properties

In the last part of this work we have studied the hydration properties of one halide, i.e. the bromide ion. Our investigation has been carried out at ambient conditions, since the coordination structure of this anion is still controversial even at 1 bar and 300 K [2], as already pointed out in the Introduction chapter.

In general, halide ion-water interactions, excluding fluoride, are weaker than those of most cations and energetically comparable with the water-water interactions in bulk water. Therefore, the structural properties of the solvent increase significantly their role and as a result a correct description of the halide hydration properties can be obtained only with theoretical methods able to reproduce the delicate compromise among solute-solvent and solvent-solvent interactions.

Moreover, it has been shown that the polarizability of the anion may be one of the most important factor in determining the structure of gas phase clusters [113, 114]. As far as the aqueous solutions are concerned, in most cases polarizable models can provide qualitative predictions in good agreement with available experimental data. However, it is very important to obtain an appropriate value of the halogen ion polarizability since there are no direct measurements of this quantity in aqueous solution, and the available data are usually extrapolations from ionic crystals and salt solutions [115]. As a consequence, the contributions of the polarizability are actually not clearly obtainable.

All together, these findings have prompted us to use in the study of the bromide aqua ion a completely different approach with respect to that employed for the group 12 metal cations. In particular, we have resorted to Car-Parrinello Molecular Dynamics simulations, which are able to describe the dynamics of the system in a self-consistent fashion, taking into account in a natural way polarization and many-body interactions. In the Car-Parrinello approach, any electronic structure method can be used to solve the electronic Schrödinger equation. In our case, we have decided to employ the DFT method, as it is able to include to some extent the correlation energy with a computational cost comparable to that of the Hartree-Fock method, provided that a suitable form of the exchange and correlation potential is chosen. Moreover, plane wave basis sets were used, due to their advantages with respect to Gaussian basis sets when employed in Molecular Dynamics simulations (see section 2.8.2). Besides the quantum mechanical Car-Parrinello Molecular Dynamics simulation, we have performed also a classical Molecular Dynamics simulation of Br^- in aqueous solution, and we have compared our theoretical results with the X-ray absorption spectroscopy experimental data, as in the case of the group 12 aqua ions. With such a procedure it has been possible to identify which method provides the most reliable structural results.

10.1 Computational methods

10.1.1 Classical Molecular Dynamics Simulation

The classical Molecular Dynamics simulation of the bromide ion in aqueous solution has been carried out with the GROMACS package [92] following the general setup employed also in the case of the group 12 metal cations described in section 6.1.1. The OPLS parameters [116] were used for the ion-water interactions, while the TIP3P model [117] was employed to describe the water-water interactions. This water model adopts the experimental molecular geometry and it is widely used in the literature, especially together with the OPLS force fields, as in our case, since it is OPLS-consistent. The system, composed of one bromide ion and 819 water molecules in a cubic box, was simulated for 15 ns, with a time step of 1 fs. The first 5 ns have been used for equilibration and discarded in the following analyses.

10.1.2 Car-Parrinello Molecular Dynamics simulation

The Kohn-Sham formulation of DFT, which is implemented in the CPMD code [118], has been used in the ab initio Molecular Dynamics simulation. The gradient-corrected BLYP functional has been employed [119, 120]. The simulated system consists of one bromide ion and 90 water molecules in a periodic cubic box with 14 Å edge. The Martins-Troullier pseudopotential [121] was used for the bromide ion, while the core electrons of oxygen and hydrogen atoms have been treated using the recently developed DCACP pseudopotentials [122]. Recent calculations using DCACP pseudopotentials in combination with the BLYP functional have shown improved structural and dynamical properties of liquid water [123]. The electronic wavefunctions have been expanded in a plane wave basis set up to an energy cut-off of 70 Ry. The simulation was performed using the Car-Parrinello approach [69], with a fictitious mass associated to the electronic degrees of freedom of 400 a.u.. After an equilibration time of 2 ps, during which thermalization at 300 K has been achieved by a Nosé-Hoover thermostat [65, 66] with a coupling frequency of 1500 cm^{-1}, the equations of motion have been integrated with a time step of 4 a.u., for a total simulation time of 4.4 ps in the NVE ensemble.

It is important to stress that we observed no drift in the electronic kinetic energy during the total simulation time. Note that a homogeneous background charge has been used to compensate for the negative charge of the bromide ion [93].

10.1.3 Structural and dynamic analysis

The structural properties of the bromide ion aqueous solutions were described in terms of radial distribution functions, Br-O and Br-H g(r)'s, as explained in section 6.1.2. Angular distribution functions have been calculated for three different angles: the angle formed by two Br-O vectors in the first shell (labelled as ψ), the angle formed by the water molecule dipole and the Br-O vector (labelled as ϕ), and the angle formed by the Br-O and O-H vectors (labelled as ω).

The mean residence time of water molecules in the first hydration shell has been evaluated, for the classical Molecular Dynamics simulation, using the approach proposed by Impey et al. [96]. This procedure have been described in section 6.1.3. A second approach, called the "direct method" and proposed by Hofer et al. [124], has been also applied. This method scans the whole trajectory for movements of the ligands, either entering or leaving the first coordination shell. Whenever a ligand crosses the boundaries of the shell, its further path is followed, and if its new placement outside or inside the shell lasts for more than a chosen t*, the event is accounted for as a "real" exchange process. Thus t* has the same role as in the Impey procedure. The water residence τ_d is then calculated as:

$$\tau_d = \frac{t_{sim}\overline{n}}{N_{ex}} \tag{10.1}$$

where t_{sim} is the total simulation time, $\overline{n}$ is the average first shell coordination number, and N_{ex} is the number of "real" solvent exchanges between the first hydration shell and the rest of the solvent.

All the structural and dynamic analyses have been carried out using in-house written codes.

10.2 Computational results

Structural arrangements of water molecules around the bromide ion are characterized by the Br-O and Br-H radial distribution functions, and the results obtained from the two Molecular Dynamics simulations are collected in Figure 10.1. In both cases, the presence of a nonzero first minimum in the Br-O g(r)'s and of a nonzero second minimum in the Br-H g(r)'s indicates that the first solvation shell is not well defined, and several exchange events take place between the first and second hydration sphere. Moreover, the Br-H g(r)'s show the presence of two peaks, the former at shorter distances and the latter at longer distances as compared to the Br-O g(r) first maxima, meaning that the first shell water molecules orient only one hydrogen atom towards

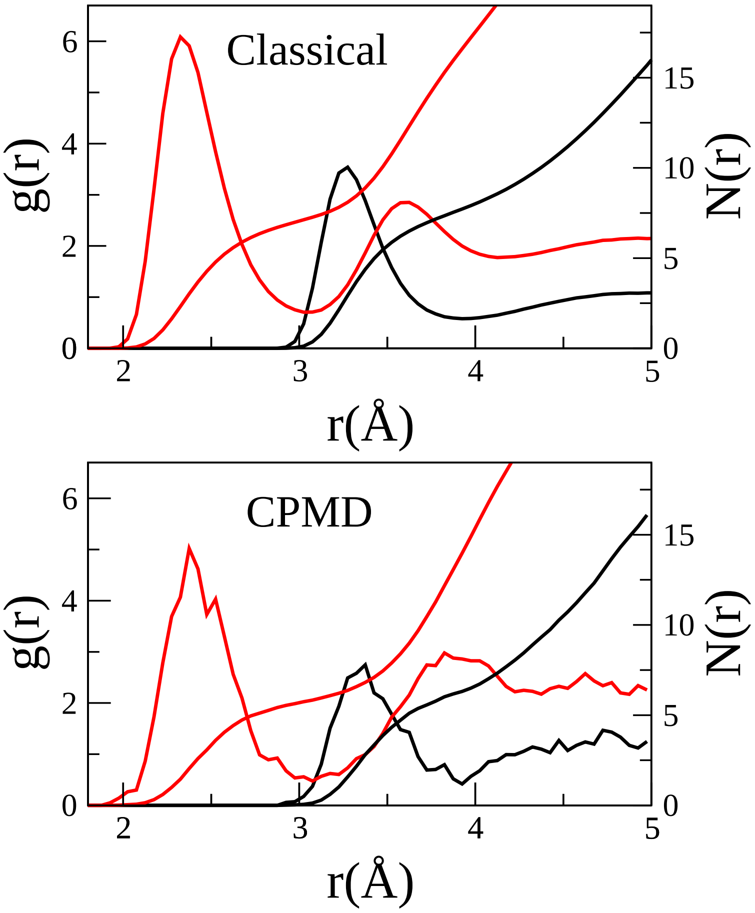

Figure 10.1: Br-O (black line) and Br-H (red line) radial distribution functions obtained from the classical and CPMD simulations. Running integration numbers are also shown.

the bromide ion. This result is not obvious for the classical Molecular Dynamics simulation since in the classical framework the anion is represented by a negative charge having electrostatic and Lennard-Jones interactions with the water molecules. This simple description, which completely neglects the ion polarizability, could have led to an overestimation of the ion-water dipole interactions, giving rise to a symmetric arrangement with the two hydrogen atoms pointing towards the anion almost at the same distances.

The main difference between the classical and the CPMD simulation is the Br-water first shell distances. Inspection of Figure 10.1 reveals that the rising edges and the first peak positions are shifted towards shorter distances in the classical case both in the Br-O and Br-H g(r)'s. The Br-O g(r) first peak distances are indeed 3.27 Å and 3.33 Å for the classical and CPMD simulations, respectively. This result is a first indication that the pair potentials used in the classical approach are more rigid giving rise to a first hydration shell which is more tightly bound to the bromide ion. Note that also the intensity of the classical g(r) peaks is higher as compared to the CPMD one. This overall tendency is reflected in the number of water molecules present in the first hydration shell of the anion as obtained from the running integration numbers of the Br-O g(r)'s that are 7.6 and 6.5 up to the first minima (3.9

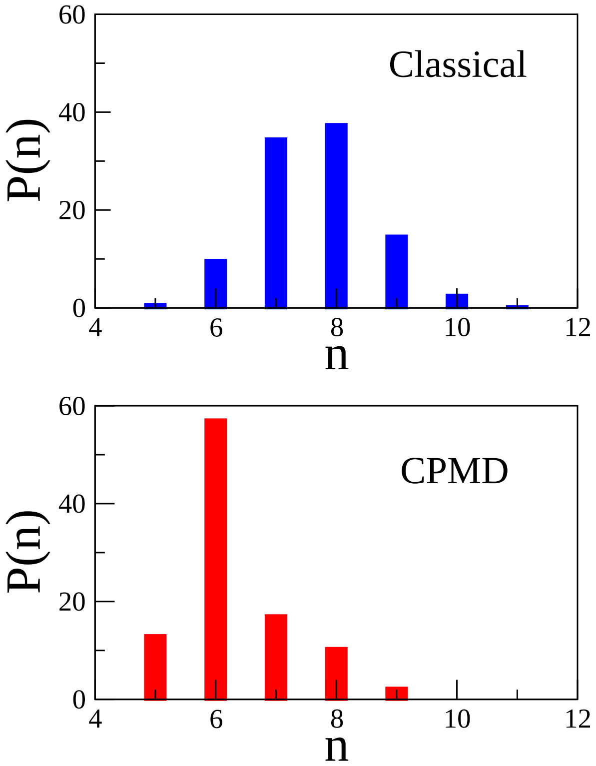

Figure 10.2: Coordination number (n) distribution obtained from the classical and CPMD simulations.

Å in both cases) for the classical and CPMD simulations, respectively (see Figure 10.1). These results are different from those of Raugei and Klein [38] who found a first shell coordination number of 5.1. More accurate information can be obtained by defining an instantaneous coordination number n, as the number of oxygen atoms at a distance shorter than the Br-O g(r) first minimum, and analyzing its variation along the simulations. The coordination number distributions are shown in Figure 10.2. In both trajectories the bromide ion transits among several coordination numbers, but the distributions obtained from the two simulations are very different. The dominant species existing in solution are indeed seven- and eight-fold hydration complexes in the classical Molecular Dynamics simulation, while the six-fold complex in the CPMD simulation plays a dominant role, even if the distribution of coordination numbers fluctuates in the range 5 to 9.

The geometrical arrangement of the first shell water molecules around the bromide ion can be evaluated by looking at the angular distribution functions of the O-Br-O (ψ) angle, plotted in Figure 10.3 as $1 - cos(\psi)$. The ψ distribution obtained from the classical and CPMD trajectories drops to zero for $1 - cos(\psi)$ less than 0.25, showing that ψ angle values smaller than 41° are prohibited in both simulations. As far as the CPMD trajectory is concerned, no clear peaks can be observed in the ψ distribution, showing the absence of a well defined configuration of water molecules around the bromide ion. This finding differs from the results obtained by Raugei and Klein [38] in their CPMD simulation where water molecules occupy well defined positions around the Br$^-$ ion (the most stable first shell configuration has a 4+2 structure where four water molecules coordinate the anion according to a hypothetical square pyramidal geometry and the other two molecules coordinate it on the other side). Conversely, in the classical Molecular Dynamics simulation two broad and low intensity peaks are found at $1 - cos(\psi)$ =0.63 and $1 - cos(\psi)$=1.78 ($\psi = 68.3°$ and $\psi = 141.3°$, respectively), indicating the existence of a more structured coordination sphere. The orientation of a single water molecule in the first hydration shell can be inferred from the distribution function of ϕ angles (Figure 10.4). In both cases the ϕ distributions show a well defined peak and the maxima are located at ϕ=49.5° for the classical simulation and ϕ=55.3° for the CPMD simulation. These distributions are consistent with a nearly linear Br- - -H-O hydrogen bond, in agreement with the results of Raugei and Klein [38] and with previous calculations on anions in water [2]. The broader and lower-intensity ϕ distribution obtained from the CPMD simulation provides a further proof of the higher flexibility of the first hydration shell in agreement with the above mentioned results. The distributions of the ω angle are shown in Figure 10.5. The sharp peak located at 0° indicates that one hydrogen atom of the first shell is strongly bound to the anion in a linear Br- - -H-O configuration. The second hydrogen atom, which is less tightly bound and more free to rotate, is responsible for the second peak which is less intense and broader than the first one. As evident from the inset of Figure 10.5, the maxima of the second peak are found at ω values of 103.2° for the classical and of 109.3° for the CPMD trajectory, coherently with the linear Br- - -H-O configuration. The form of the ω distributions shows once more that the flexibility of the first coordination shell increases going from the classical to the CPMD simulation.

Even if the water molecules do not occupy well defined positions around the ion, as evidenced from the ψ a.d.f. analysis, it is possible to evaluate

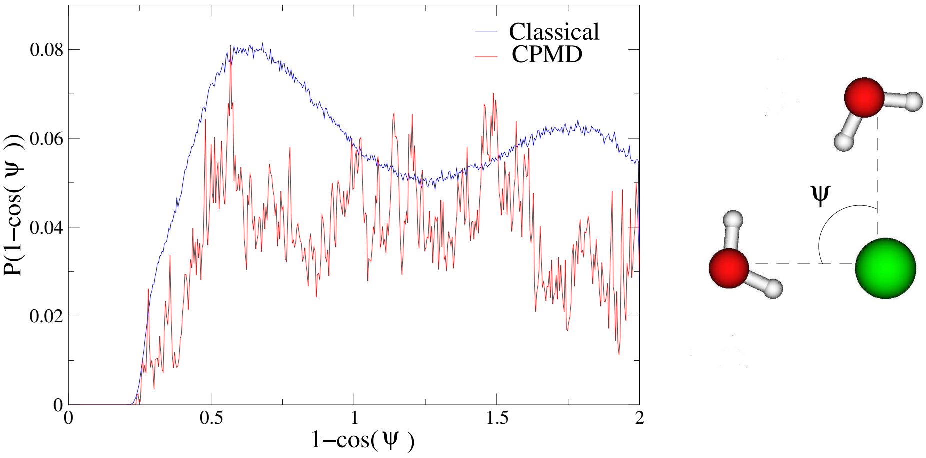

Figure 10.3: O-Br-O (ψ) angular distribution functions obtained from the classical and CPMD simulations.

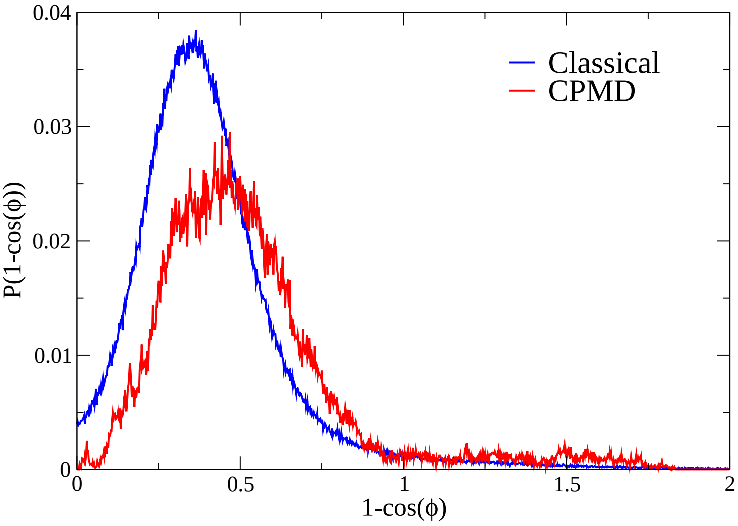

Figure 10.4: ϕ angular distribution functions obtained from the classical and CPMD simulations.

the asymmetry of the first coordination shell by analysing the distance between the center of mass of the first hydration shell cage and the bromide ion R_{cage} (see Figure 10.6). From Figure 10.6 a slight asymmetry of the first hydration shell can be inferred, with the asymmetry being more pronounced in the CPMD trajectory (the R_{cage} distribution peaks are located at 0.40 and 0.58 Å for the classical and CPMD simulations, respectively). The stronger asymmetry found by Raugei and Klein [38] was due to the highly asymmetric 4+2 first shell structure which resulted from their CPMD simulation. The asymmetry of the first coordination is responsible for the presence of an induced net dipole moment on the bromide ion. The anion dipole moment has been calculated from the CPMD trajectory with respect to the anion nucleus position, using the Wannier function centers [125, 126]. The positions of the

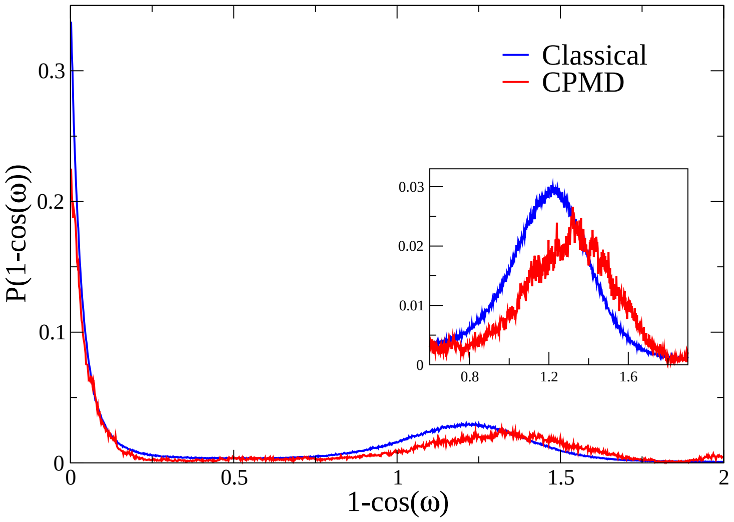

Figure 10.5: ω angular distribution functions obtained from the classical and CPMD simulations.

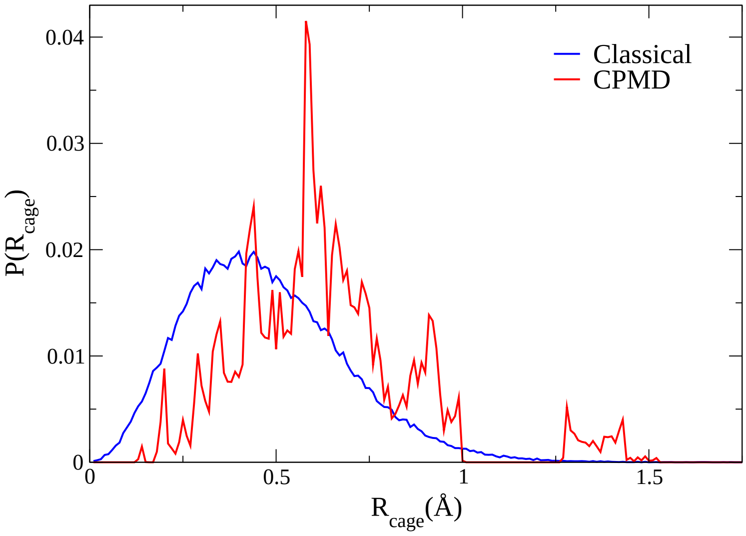

Figure 10.6: Distributions of the distance between the center of mass of the first hydration shell cage and the bromide ion (R_{cage}) obtained from the classical and CPMD simulations.

maximally localized Wannier centers are directly related to the positions of the electrons in the system, and, as a consequence, it is possible to compute the molecular (or atomic) dipole moment assuming that the electrons belong to the nearest neighbour atom. The peak of the anion dipole moment distribution is found at $\mu = 0.85D$ (see Figure 10.7), indicating that the bromide electron density is strongly influenced by the presence of the surrounding water molecules and that the inclusion of polarization effects is essential to provide a correct description of the anion-water interactions.

The flexible and unstructured layout of the bromide ion first solvation shell is reflected on its dynamical behavior. Several water exchange events between the first and the second hydration shell have been observed during the two simulations, and the rate of the water exchange processes has been evaluated by means of mean residence times of the water molecules in the first hydration shell. This property has been determined using both the Impey and direct method for the classical simulation and, due to the short simulation time, only with the direct method for the CPMD one. The problems encountered when trying to apply the Impey procedure for calculating the residence time from relatively short Molecular Dynamics simulations have been already pointed out in the literature [124], and are related to the difficulties of fitting the Impey survival function $n_{hyd}(t)$ to the required exponential form. On the other hand, it is always possible to count the number of exchange processes between water molecules belonging to the first and second coordination shell during a simulation, and then to determine the residence time by means of the direct method. Nevertheless, it is important to bear in

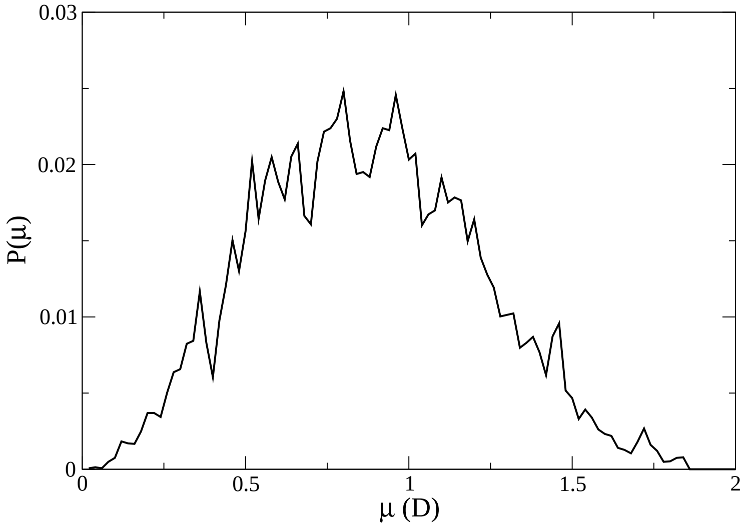

Figure 10.7: Distribution of the bromide dipole moment calculated from the CPMD simulation.

mind that also with the latter method the shorter the simulation time, and consequently the number of exchange processes, the higher the error on the calculated residence time. Table 10.1 lists the residence times of first shell water molecules obtained from the two trajectories for t* values of 0.0, 0.5 and 2.0 ps. A proper choice of the t* parameter is fundamental for systems with a very flexible first hydration shell, where the g(r)'s show only shallow minima and the definition of the first hydration shell can be somehow arbitrary, as in the present case. Our results show that the calculated residence times are strongly dependent on the choice of t* with both procedures, the dependence being stronger for the direct method. A t* values of 0.5 ps has been found to be the more appropriate choice [124], which corresponds to

	t*=0.0 ps			t*=0.5 ps			t*=2.0 ps		
Trajectory	N_{ex}	τ_I	τ_d	N_{ex}	τ_I	τ_d	N_{ex}	τ_I	τ_d
Classical	113862	2.4	0.7	28939	3.8	2.6	16305	5.8	4.7
CPMD	77	-	0.4	5	-	5.7	2	-	14.3

Table 10.1: Residence times (ps) of water molecules in the first coordination shell of the bromide ion, calculated using the Impey procedure (τ_I) and the direct method (τ_d) as a function of the t* parameter, for the classical and CPMD simulations. N_{ex} is the number of exchange events observed during the total simulation time.

the average lifetime of a water-water hydrogen bond. With this choice, the residence time calculated from the CPMD simulation is longer as compared to the classical one. However, the results obtained from both simulations are in the order of magnitude of the experimental determination [33]. Note that the residence time calculated from our CPMD simulation with a t* value of 2.0 ps approaches the value obtained by Raugei and Klein using the Impey method and the same t*, for the HBr aqueous solution (19 ps) [38]. The same authors suggested for the Br^- aqueous solution a slightly lower value of the residence time, as compared to the HBr one.

10.3 EXAFS data analysis

10.3.1 Methods

Br^- K-edge X-ray absorption spectra were recorded at the BM29 beamline of the ESRF [99]. The sample was a 0.1M RbBr aqueous solution kept in a cell with Kapton windows. The absorption coefficient was measured in transmission mode, and the monochromator was equipped with two flat Si(311) crystals for high-energy operation with excellent resolution. In order to reduce harmonic contamination, the crystals were kept slightly detuned with a feedback system. The incident and transmitted fluxes were monitored by ionization chambers filled with Kr gas. The storage ring was operating in 2/3 fill mode with a typical current of 200 mA after refill. Data points were collected for 1 s each, and three spectra were recorded and averaged.

The EXAFS data analysis was carried out using the same procedure as employed for the group 12 metal ions. All the details can be found in section 6.2.2. The values of the muffin-tin radii used for hydrogen and oxygen are always 0.2 Å and 0.9 Å, respectively, while a radius of 2.3 Å has been employed for bromine.

10.3.2 Results

As shown in section 10.2, the classical and CPMD simulations of the bromide ion in aqueous solution provide a different description of the hydration properties of the Br^- aqua ion. In particular, the Br-water first shell distances are shifted towards shorter distances in the classical case and the distributions of coordination numbers obtained from the two simulations are very different. In the classical Molecular Dynamics simulation the dominant species are indeed seven- and eight-fold hydration complexes, while only the six-fold complexes in the CPMD simulation.

In order to identify which method provides the most reliable structural results, we have compared the Molecular Dynamics results with the EXAFS experimental data. $\chi(k)$ theoretical signals have been calculated by means of equation 4.5, starting from the classical and CPMD Br-H and Br-O g(r)'s. The structural parameters derived from the simulations were kept fixed during the EXAFS analysis. In this way the first hydration shell structure ob-

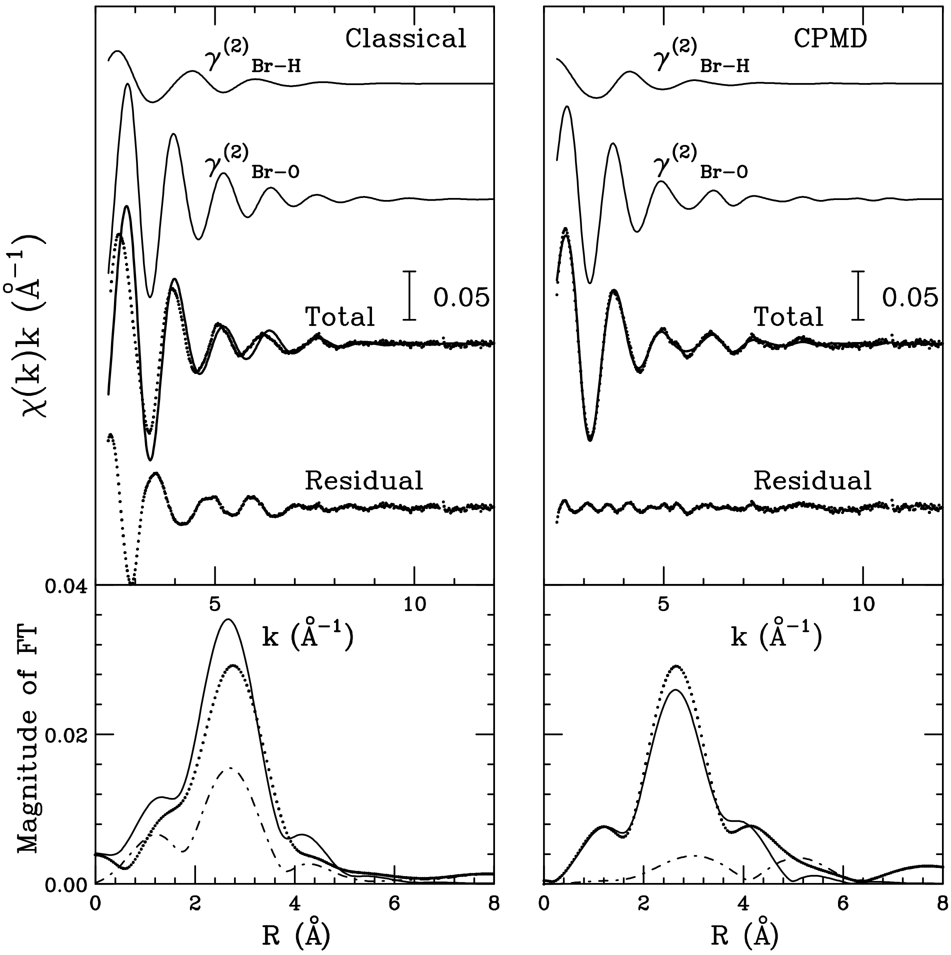

Figure 10.8: Upper panels: Comparison between the EXAFS theoretical signals (solid line) calculated from the classical and CPMD Br-O and Br-H g(r)'s and experimental data (dotted line). The residual signals (dotted line) are also shown. Lower panels: Nonphase-shifted corrected Fourier transforms of the experimental data (dotted line), of the theoretical signals (solid line), and of the residual curves (dot-dashed line).

tained from the simulations can be directly compared with experimental data and the validity of the theoretical framework used in the simulations can be assessed. In the upper panels of Figure 10.8, the comparison between the experimental signal and the theoretical curves is reported for the classical and CPMD simulations (left and right panels, respectively). The first two curves from the top are the Br-H and Br-O first shell contributions, while the reminder of the figures shows the total theoretical contribution compared with the experimental spectrum, and the resulting residuals. In the case of the classical Molecular Dynamics simulation the agreement between the calculated and experimental EXAFS spectra is quite poor and the presence of a leading frequency can be clearly identified in the residual curve. This behavior is due to the short value of the average Br-H and Br-O first shell distances obtained from the classical Molecular Dynamics simulation as compared to the experimental results. This finding is reinforced by the Fourier transform (FT) moduli of the EXAFS χ(k) theoretical, experimental and residual signals shown in the lower panels of Figure 10.8. The FT's have been calculated in the k-range 3.0-7.5 Å^{-1} with no phase shift correction applied. The theoretical first-neighbor peak is found to be shifted towards shorter distances than predicted by the experiment. Conversely, the theoretical $\chi(k)$ signal calculated from the CPMD g(r)'s match the experimental data very well and a good agreement is found also looking at the FT's. Therefore, the structural and dynamical information derived from the CPMD simulation is basically correct.

Even if the Br-O two body signal provides the most important contribution to the total $\chi(k)$, the hydrogen atoms in the first hydration shell provide a sizeable contribution especially in the k region below 6 Å^{-1}. The importance of the hydrogen contribution to the EXAFS spectra of metal cations in aqueous solution has been pointed out in previous works [7, 34]. To the best of our knowledge this is the first time that the hydrogen scattering is included in the EXAFS analysis of anion aqueous solutions. Note that a reliable calculation of the contribution associated with the hydrogen atoms in the first hydration shell relies on a proper description of the anion-water and water-water interaction in the Molecular Dynamics simulations. In this context use of advanced methods in the CPMD setup has been found to be essential to provide a correct description of the structural properties of the Br^- hydration shells.

Chapter 11

Summary and conclusions

In this work a detailed investigation of the structural and dynamic properties of Zn^{2+}, Cd^{2+}, Hg^{2+} and Br^- ions in aqueous solution has been carried out, combining XAS and Molecular Dynamics simulations. Cd^{2+}, Hg^{2+} and Br^- aqua ions have been studied at ambient conditions, while the Zn^{2+} hydration properties were investigated in conditions of very high pressures (up to 2.85 GPa).

For the group 12 metal cations, it has been further developed and applied a computational procedure for the generation of effective two-body ion-water potentials, to be used in the classical Molecular Dynamics simulations. In this methodology the effective ion-water potential is obtained by fitting the parameters of a suitable analytical function on an ab initio PES of the M^{2+}-H_2O system generated including many-body effects by means of the PCM.

As far as the Hg^{2+} aqua ion is concerned, the combined use of XAS experimental data, accurate ab initio calculations and classical Molecular Dynamics simulation allowed us to determine, in contrast to all the previously reported results, a quite unexpected sevenfold coordination of the hydrated Hg^{2+} complex in aqueous solution. The unusual hydration structure found for the Hg^{2+} ion prompted us to evaluate the effect of water-water interactions on the structural and dynamic results obtained, using two different water models, namely the SPC/E and TIP5P. Both simulations started from the generally accepted octahedral hydration structure, and only after a very long induction time (630 and 755 ps for the SPC/E and TIP5P water models, respectively), the first transition to a heptacoordinated cluster was observed. Thereafter, several transitions between stable hexa- and heptacoordinated complexes took place. As a consequence, a simulation time of 60 ns was necessary to properly define the first shell coordination number. It is important to stress that our approach, based on long enough classical simulations with potential not biased toward a specific first shell coordination number, is the only method able to determine the coordination geometry of a hydrated ion having a slow exchange dynamic. QM/MM or Car-Parrinello simulations, in fact, are computationally very expensive, and can model the aqueous solutions only up to a few ten or hundred picoseconds. Therefore, these techniques are not able to reproduce exchange events occurring in the nanosecond time scale, and, in this case, they can provide useful information only if the initial configuration is the correct one. The two water models used in the simulations provided quite similar first shell structural parameters, essentially describing the Hg^{2+} ion as being coordinated by seven water molecules. The simulation results have been compared with EXAFS experimental data which indicated that the structural and dynamical information derived from the two trajectories was basically correct. However, since in the case of disordered systems such as ionic solution the EXAFS technique is not able to provide unique information on the coordination number of the photoabsorber atom, in order to obtain a definite answer on the Hg^{2+} first shell coordination number, we have carried out also the analysis of XANES region of the spectrum. In the first step of our investigation, we have carried out a static fit analysis of the XANES experimental data that has shown that an hexahydrated cluster, even when a highly disordered model was used to fit the data, was not able to reproduce the XANES region of the spectrum. Conversely, a very good agreement with the experimental spectrum has been obtained when a sevenfold cluster with C_2 symmetry was considered. However, the best-fit model cluster obtained from the XANES data analysis shows a larger distance dispersion as compared to the MP2 and Molecular Dynamics theoretical results. To shed light on this point, we have decided to extend the study of Hg^{2+} aqua ion, using our Molecular Dynamics trajectories to interpret the XANES experimental data. We have thus calculated the XANES spectra using the structural information obtained from the Molecular Dynamics simulations without carrying out any minimization in the structural parameter space. It was possible to correctly reproduce the experimental spectrum only when a heptacoordinated first shell model as obtained from the Molecular Dynamics simulations was considered and when the second coordination shell was included in the calculation. The results of this dynamical XANES analysis demonstrated that the discrepancy we had found in our static fit analysis was mainly due to the lack of the second hydration shell contribution in the XANES calculations.

We have obtained quite unexpected results also in the case of the Cd^{2+} aqua ion. In fact both the SPC/E and TIP5P simulations started from a hexacoordinated first-shell configuration around the Cd^{2+} ion, and after an induction time of 230 and 155 ps for the SPC/E and TIP5P models, respectively, the system went to heptacoordination. In both cases, a very flexible Cd^{2+} ion first hydration shell was detected, which transits among coordination numbers of six, seven, and, for relatively short times, eight. However, while for the Hg^{2+} aqua ion a uniform picture emerged from the simulations carried out with different water models, for the Cd^{2+} ion we have obtained from the two trajectories a different behaviour: in the TIP5P case, the dominant species existing in solution is the heptacoordinated one, while the SPC/E simulation increases the stability of the hexacoordinated cluster. On the other hand, the structure of hexa- and heptahydrated complexes was the same regardless the water model used in the simulations. The EXAFS analysis results have shown that the structural and dynamical information derived from the two simulations was basically correct but also in this case to assess in an unambiguous way the ion coordination and to determine the effective percentages of the two hydration complexes we had to resort to analysis of the XANES starting from the Molecular Dynamics simulations. We calculated several averaged theoretical spectra with variable percentages of hexa- and heptacoordinated complexes, and by comparing them with the experimental spectrum we obtained a quite informative picture. Indeed the index of agreement between the theoretical and experimental spectrum has shown a regular trend and a minimum was obtained for a hexacoordinated cluster percentage of 20%, which corresponds to the result obtained from the TIP5P simulation. Thus, the XANES analysis provided a definitive experimental proof of the flexible nature of the Cd ion, unambiguously showing the presence of a dominant percentage of heptahydrated species. It is noteworthy to observe also in this case that it is been possible to suitably describe the hydration properties of the Cd^{2+} ion only by very long Molecular Dynamics simulations.

The results of the Molecular Dynamics simulations of the Zn^{2+} ion in high-density water have shown that the octahedral structure of the Zn^{2+} first hydration shell remain stable also under high pressure, while the Zn-O first shell distance is shortened and the hydrated complex becomes more disordered. The effect of the pressure changes is more evident on the structure of the second coordination shell which is compressed and becomes more disordered and unstructured with increasing pressure. Moreover, the increase of pressure has a strong effect on the dynamics of water, which has been investigated by calculating the reorentational correlation times of the water molecules surrounding the ion. The water mobility was found to increase, as a consequence of the rupture of the hydrogen bond network caused by the pressure increase. The first shell structural results obtained from the simulations have been experimentally validated by EXAFS spectroscopy. The EXAFS analysis confirmed that the Zn^{2+} first shell hydration number does not vary with pressure, while the first hydration shell is shifted towards the ion. Moreover, the higher mobility of the water molecules in the first coordination sphere under high pressure was highlighted by the trend of the Debye-Waller factors, that increase with pressure. The structural transformations occurring to water with increasing density were also investigated by Molecular Dynamics simulations. The results of this study have shown that solvent molecules are not significantly influenced by the presence of the solute and impact of the ion on the structure of water does not extend in an appreciable way beyond the first hydration shell. Conversely, substantial changes of the water structural properties were found with increasing pressure. These changes are caused by a collapse of the second coordination shell, as a consequence of the rupture of the hydrogen bond network. Radial and angular distribution function analyses have shown that the effect of pressure is to increase the number of interstitial water molecules, i.e. water molecules that are not hydrogen bonded to any of the first shell water molecules around a given central one. This study of the Zn^{2+} ion in high-density water paves the way for future investigations on the hydration structure under high pressure conditions of ions having more flexible hydration shells, such as the two ions that we have studied here at ambient conditions, namely Cd^{2+} and Hg^{2+}.

Finally, we have studied the hydration properties at ambient conditions of one halide, i.e. the bromide ion. In this case we have employed a completely different approach with respect to that used for the group 12 metal cations. In particular, we have resorted to Car-Parrinello Molecular Dynamics simulations, since they are able to describe the dynamics of the system in a self-consistent fashion, taking into account in a natural way polarization effects, which have a strong impact on the halide-water interactions. Moreover, in the case of the Br^- ion, the estimated residence time of the first shell water molecules is three orders of magnitude shorter (picosecond timescale) as compared to Cd^{2+} and Hg^{2+} (nanosecond timescale), so that sufficiently long ab initio Molecular Dynamics can be carried out to achieve a proper sampling of the phase space of the system. Besides the quantum mechanical Car-Parrinello Molecular Dynamics simulation, we have performed also a classical Molecular Dynamics simulation of Br^- and we have compared the results obtained from the two different approaches. The picture that has emerged from our simulations is that the pair potentials used in the classical approach are too rigid giving rise to a first hydration shell which is too tightly bound to the bromide ion. Indeed the classical Br-O and Br-H radial distribution functions are shifted towards shorter distances as compared to the CPMD ones and all the angular distribution function results have shown that the flexibility of the first coordination shell increases going from the classical to the CPMD simulation. In both trajectories the bromide ion transits among several coordination numbers, but while in the classical case the dominant species existing in solution were seven- and eight-fold hydration complexes, only six-fold complexes played a dominant role in the CPMD trajectory. The comparison of our theoretical results with EXAFS experimental data has allowed us to identify which method provides the most reliable structural results. In the case of the classical Molecular Dynamics simulation a quite poor agreement between the theoretical and experimental spectrum was found, while the spectrum calculated from the CPMD g(r)'s matched the experimental data very well, thus showing that the structural and dynamical information derived from the CPMD simulation was basically correct.

In conclusion, the application of both experimental and computational techniques used in this work paves the route for the systematic use of an integrated approach, with increased reliability, in the structural investigation of disordered systems. The characterization of the structural and dynamical properties of disordered systems is indeed a difficult task, that can hardly be obtained using a single technique. Conversely, the combined use of experimental and theoretical methods has allowed us to obtain reliable information on all the investigated systems. It is important to stress that the procedure which has been further developed and applied in this work to generate the ion-water pair potential can be extended to the study of other solutions, and the application of this methodology to investigate ion coordination in organic solvents will be the subject of future work.

As a last remark, we would like to point out that the procedure we have developed to analyze the XANES in combination with Molecular Dynamics simulations is absolutely general and can be applied to the study of liquid samples and biological media.

Bibliography

[1] D.T. Richens
The chemistry of aqua ions John Wiley & S. (1997).

[2] H. Othaki, T. Radnai
Chem. Rev. **93**, 1157, (1993).

[3] H.S. Frank, W.Y. Wen
Discuss. Faraday Soc. **24**, 133, (1957).

[4] R.W. Gurney
Ionic processes in solution McGraw Hill, New York (1953).

[5] A. Kuzmin, S. Obst, J. Purans
J. Phys.: Condens. Matter **9**, 10065, (1997).

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

[7] P. D'Angelo, V. Barone, G. Chillemi, N. Sanna, W. Meyer-Klauke, N.V. Pavel
J. Am. Chem. Soc. **124**, 1958, (2002).

[8] K.W. Jennette
Environ. Health Perspect. **40**, 233, (1981).

[9] E.C. Foulkes
Exper. Biol. Med. **223**, 234, (2000).

[10] P. D'Angelo, G. Chillemi, V. Barone, G. Mancini, N. Sanna, I. Persson
J. Phys. Chem. B **109**, 9178, (2005).

[11] G. Chillemi, V. Barone, P. D'Angelo, G. Mancini, I. Persson, N. Sanna
J. Phys. Chem. B **109**, 9186, (2005).

[12] C.C. Pye, M.R. Tomney, W.W. Rudolph
Can. J. Anal. Sci. Spect. **51**, 140, (2006).

[13] G. Johansson, M. Sandström
Acta Chem. Scand., Ser. A **32**, 109, (1978).

[14] M. Eigen
Pure Appl. Chem. **6**, 97, (1963).

[15] R. Åekesson, I. Persson, M. Sandström, U. Wahlgren
Inorg. Chem. **33**, 3715, (1994).

[16] C.S. Babu, C.J. Lim
J. Phys. Chem. A **110**, 691, (2006).

[17] C. Kritayakornupong, B. M. Rode
J. Chem. Phys. **118**, 5065, (2003).

[18] M. Sandström, I. Persson, S. Ahrland
Acta Chem. Scand. Ser. A **32**, 627, (1978).

[19] O. Sobolev, G.J. Cuello, G. Roman-Ross, N.T. Skipper, L. Charlet
J. Phys. Chem. A **111**, 5123, (2007).

[20] L.W. Flanagin, P.B. Balbuena, K.P. Johnsto, P.J. Rossky
J. Phys. Chem. **99**, 5196, (1995).

[21] D. Sebastiani, M. Parrinello
Chem. Phys. Chem. **3**, 675, (2002).

[22] T. Yamaguchi
J. Mol. Liq. **78**, 43, (1998).

[23] K. Bagchi, S. Balasubramanian, M.L. Klein
J. Chem. Phys. **107**, 8561, (1997).

[24] A.G. Kalinichev, Y.E. Gorbaty, A.V. Okhulkov
J. Mol. Liq. **82**, 57, (1999).

[25] A.K. Soper
Chem. Phys. **258**, 121, (2000).

[26] A.K. Soper, M.A. Ricci
Phys. Rev. Lett. **84**, 2881, (2000).

[27] E. Schwegler, G. Galli, F. Gygi
Phys. Rev. Lett. **84**, 2429, (2000).

[28] A. Saitta, F. Datchi
Phys. Rev. E **67**, 020201, (2003).

[29] T. Strässle, A.M. Saitta, Y. Le Godec, G. Hamel, S. Klotz, J.S. Loveday, R.J. Nelmes
Phys. Rev. Lett. **96**, 067801, (2006).

[30] A. Filipponi, S. De Panfilis, C. Oliva, M.A. Ricci, P. D'Angelo, D.T. Bowron
Phys. Rev. Lett. **91**, 165505, (2003).

[31] B. Winter, M. Faubel
Chem. Rev. **106**, 1176, (2006).

[32] D. Laage, J.T. Hynes
Proc. Natl. Acad. Sci. **104**, 11167, (2007).

[33] P. Bopp
The Physical Chemistry of Aqueous Solution
M.-C. Belliment-Funel and G. W. Neilson, Dordrecht (1987).

[34] P. D'Angelo, A. Di Nola, A. Filipponi, N.V. Pavel, D. Roccatano
J. Chem. Phys. **100**, 985, (1994).

[35] S.L. Wallen, B.J. Palmer, D.M. Pfund, J.L. Fulton, M. Newville, Y.J. Ma, E.A. Stern
J. Phys. Chem. A **101**, 9632, (1997).

[36] C.G. Elles, I.A. Shkrob, R.A. Crowell, D.A. Arms, E.C. Landahl
J. Chem. Phys. **128**, 061102, (2008).

[37] P.J. Merkling, R. Ayala, J.M. Martinez, R.R. Pappalardo, E.S. Marcos
J. Chem. Phys. **119**, 6647, (2003).

[38] S. Raugei, M. L. Klein
J. Chem. Phys. **116**, 196, (2002).

[39] C.D. Cappa, J.D. Smith, K.R. Wilson, B.M. Messer, M.K. Gilles, R.C. Cohen, R.J. Saykally
J. Phys. Chem. B **109**, 7046, (2005).

[40] W.J. Herhre, P.v.R. Schleyer, L. Radom, J.A. Pople
Ab-initio molecular orbital theory John Wyley & S. (1985).

[41] F. Jensen
Introduction to computational chemistry John Wyley & S. (1999).

[42] B.H. Bransden, C.J. Joachain
Physics of atoms and molecules Prentice Hall (2003).

[43] P. Hohenberg, W. Kohn
Phys. Rev. **136**, B864, (1964).

[44] L.H. Thomas
Proc. Camb. Phil. Soc. **23**, 542, (1927).

[45] E. Fermi
Z. Phys. **48**, 73, (1928).

[46] R.G. Parr, W. Yang
Density Functional Theory of atoms and molecules Oxford University Press (1995).

[47] J. Hutter
Introduction to ab initio Molecular Dynamics Lecture notes (2002).

[48] J. Tomasi, B. Mennucci, R. Cammi
Chem. Rev. **105**, 2999, (2005).

[49] M. Cossi, V. Barone, R. Cammi, J. Tomasi
Chem. Phys. Lett. **255**, 327, (1996).

[50] V. Barone, M. Cossi
J. Phys. Chem. A **102**, 1995, (1998).

[51] M. Cossi, V. Barone, B. Mennucci, J. Tomasi
Chem. Phys. Lett. **286**, 253, (1998).

[52] F. Floris, M. Persico, A. Tani, J. Tomasi
Chem. Phys. Lett. **199**, 518, (1992).

[53] M. P. Allen, D. Tildesley
Computer Simulations of Liquids Clarendon Press, Oxford (1987).

[54] D. van der Spoel, E. Lindahl, B. Hess, A.R. van Buuren, E. Apol, P.J. Meulenhoff, D.P. Tieleman T.M. Sijbers, K. A. Feenstra, R. van Drunen, H.J.C. Berendsen
Gromacs User Manual version 3.3 www.gromacs.org (2005).

[55] P.M. Morse
Phys. Rev. **34**, 57, (1929).

[56] L. Verlet
Phys. Rev. **34**, 1311, (1967).

[57] R.W. Hockney, S.P. Goel
J. Comp. Phys. **14**, 148, (1974).

[58] D. Frenkel, B. Smit
Understanding Molecular Simulations: From Algorithms to Applications Academic Press (2003).

[59] P.P. Ewald
Ann. Phys. **64**, 253, (1921).

[60] T. Darden, D. York, L. Pedersen
J. Chem. Phys. **98**, 10089, (1993).

[61] U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, L.G. Pedersen
J. Chem. Phys. **103**, 8577, (1995).

[62] B. Hess, H. Bekker, H.J.C. Berendsen, J.G.E.M. Fraaije
J. Comp. Chem. **18**, 1463, (1997).

[63] J.P. Ryckaert, G. Ciccotti, H.J.C. Berendsen
J. Comp. Phys. **23**, 327, (1977).

[64] H.J.C. Berendsen, J.P.M. Postma, A. Di Nola, J.R. Haak
J. Chem. Phys. **81**, 3684, (1984).

[65] S. Nosé
Mol. Phys. **52**, 255, (1984).

[66] W.G. Hoover
Phys. Rev. A **31**, 1695, (1985).

[67] M. Parrinello, A. Rahman
J. Appl. Phys. **52**, 7182, (1981).

[68] S. Nosé, M. L. Klein
Mol. Phys. **50**, 1055, (1983).

[69] R. Car, M. Parrinello
Phys. Rev. Lett. **55**, 2471, (1985).

[70] J.J. Rehr, R.C. Albers
Rev. Mod. Phys. **72**, 621, (2000).

[71] D.E. Sayers, E.A. Stern, F.W. Lytle.
Phys. Rev. Lett. **27**, 1204, (1971).

[72] E.A. Stern, D.E. Sayers, F.W. Lytle.
Phys. Rev. B **11**, 4836 (1975).

[73] C.R. Natoli, M. Benfatto
J. Phys. (Paris) Colloq. **47**, C8-11, (1986).

[74] M. Benfatto, C.R. Natoli
J. Phys. (Paris) Colloq. **48**, C9-1077, (1987).

[75] L. Hedin, B.I. Lundqvist
J. Phys. C **4**, 2064, (1971).

[76] A. Filipponi, A. Di Cicco, C.R. Natoli.
Phys. Rev. B **52**, 15122 (1995).

[77] M. Benfatto, S. Della Longa.
J. Synchrotron Radiat. **8**, 1087, (2001).

[78] J.E. Muller, O. Jepsen, J.W. Wilkins
Solid State Commun. **42**, 365, (1982).

[79] M. Benfatto, S. Della Longa, C.R. Natoli
J. Synchrotron Radiat. **10**, 51, (2003).

[80] P. D'Angelo, M. Benfatto, S. Della Longa, N.V. Pavel
Phys. Rev. B: Condens. Matter Mater. Phys. **66**, 064209, (2002).

[81] S. Della Longa, A. Arcovito, M. Girasole, J.L. Hazemann, M. Benfatto
Phys. Rev. Lett. **87**, 155501, (2001).

[82] F. Floris, M. Persico, A. Tani, J. Tomasi
Chem. Phys. Lett. **277**, 126, (1994).

[83] W.R. Wadt, P.J. Hay
J. Chem. Phys. **82**, 284, (1985).

[84] W.R. Wadt, P.J. Hay
J. Chem. Phys. **82**, 270, (1985).

[85] A. Wilson, T. van Mourik, T.H. Dunning Jr.
J. Mol. Struct. (Theochem) **388**, 339, (1996).

[86] F.P. Rotzinger
J. Phys. Chem. B **109**, 1510, (2005).

[87] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople
Gaussian 03, Revision B.01, Gaussian, Inc., Pittsburgh PA, (2003).

[88] *SAS/STAT User's Guide, Version 6.12* SAS Institute Inc., Cary (2000).

[89] J. Bard
Nonlinear Parameter Estimation Academic Press, New York (1974).

[90] H.J.C. Berendsen, J.R. Grigera, T.P. Straatsma
J. Chem. Phys. **91**, 6269, (1987).

[91] M.W. Mahoney, W.L. Jorgensen
J. Chem. Phys. **112**, 8910, (2000).

[92] H.J.C. Berendsen, D. van der Spoel, R. van Drunen
Comput. Phys. Commun. **91**, 43, (1995).

[93] G. Hummer, L.R. Pratt, A.E. Garcia
J. Phys. Chem. A **102**, 7885, (1998).

[94] A. Vishnyakov, A.P. Lyubartsev, A. Laaksonen
J. Phys. Chem. A. **105**, 1702, (2001).

[95] D.L. Bergman, L. Laaksonen, A. Laaksonen
J. Mol. Graph. Model. **15**, 301, (1997).

[96] R.W. Impey, P.A. Madden, I.R. McDonald
J. Phys. Chem. **87**, 5071, (1983).

[97] D. van der Spoel, P.J. van Mareen, H.J.C. Berendsen
J. Chem. Phys. **108**, 10220, (1998).

[98] P. Madden, D. Kivelson
Adv. Chem. Phys. **56**, 467, (1994).

[99] A. Filipponi, M. Borowski, D.T. Bowron, S. Ansell, S. De Panfilis, A. Di Cicco, J.P. Itie
Rev. Sci. Instrum. **71**, 2422, (2000).

[100] C. Hermes, E. Gilberg, M.H. Koch
Nucl. Instrum. Methods Phys. Res. **222**, 207, (1984).

[101] R.F. Pettifer, C. Hermes
J. Phys. Colloq. **47**, C8-127, (1986).

[102] A. Filipponi, A. Di Cicco
Phys. Rev. B, **52**, 15135, (1995).

[103] M.O. Krause, J.H. Oliver
J. Phys. Chem. Ref. Data **8**, 329, (1979).

[104] A. Filipponi
J. Phys. B: At. Mol. Opt. Phys. **33**, 2835, (2000).

[105] www.cern.ch/minuit.

[106] A.V. Egorov, A.V. Komolkin, A.P. Lyubartsev, A. Laaksonen
Theor. Chem. Acc. **115**, 170, (2006).

[107] D. R. Lide
CRC Handbook of Physics and Chemistry, 85th ed.
CRC press: Boca Raton, FL (2005).

[108] V. Barone
J. Chem. Phys. **120**, 3059, (2004).

[109] P. D'Angelo, O.M. Roscioni, G. Chillemi, S. Della Longa, M. Benfatto
J. Am. Chem. Soc. **128**, 1853, (2006).

[110] A. Kodre, J. Padežnik Gomilšek, A. Mihelič, I. Arčon
Radiat. Phys. Chem. **75**, 188, (2006).

[111] A. Botti, F. Bruni, S. Imberti, M.A. Ricci, A.K. Soper
J. Chem. Phys. **120**, 10154, (2004).

[112] R. Mancinelli, A. Botti, F. Bruni, M.A. Ricci, A.K. Soper
Phys. Chem. Chem. Phys. **9**, 2959, (2007).

[113] S.J. Stuart, B.J. Berne
J. Phys. Chem. **100**, 11934, (1996).

[114] L.X. Dang, B.C. Garret
J. Chem. Phys. **99**, 2972, (1993).

[115] A. Tongraar, B.M Rode
Phys. Chem. Chem. Phys. **5**, 357, (2003).

[116] T.P. Lybrand, I. Ghosh, J.A. McCammon
J. Am. Chem. Soc. **107**, 7793, (1985).

[117] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein
J. Chem. Phys. **79**, 926, (1983).

[118] J. Hutter, A. Alavi, T. Deutch, M. Bernasconi, S. Goedecker, D. Marx, M. Tuckerman, M. Parrinello
CPMD MPI für Festkörperforschung and IBM Zurich Research Laboratory: Stuttgart, (1995-1999).

[119] A.D. Becke
Phys. Rev. A **38**, 3098, (1988).

[120] C. Lee, W. Yang, R.G. Parr
Phys. Rev. B **37**, 785, (1988).

[121] N. Trouiller, J.L. Martins
Phys. Rev. B **43**, 1993, (1991).

[122] O.A. von Lilienfeld, I. Tavernelli, U. Rothlisberger, D. Sebastiani
Phys. Rev. Lett. **93**, 153004, (2004).

[123] I.C. Lin, A.P. Seitsonen, M.D. Coutinho-Neto, I. Tavernelli, U. Rothlisberger
J Phys Chem B. **113**, 1127, (2009).

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

[125] N. Marzari, D. Vanderbilt
Phys. Rev. B **56**, 12847, (1997).

[126] P.L. Silvestrelli, N. Marzari, D. Vanderbilt, M. Parrinello
Solid State Commun. **107**, 7, (1998).

Publications

1. Structural and Dynamical Properties of the Hg^{2+} aqua ion:
   A Molecular Dynamics Study.
   G. Mancini, N. Sanna, V. Barone, V. Migliorati, P. D'Angelo,
   G. Chillemi.
   J. Phys. Chem. B, **112** 4694, (2008).

2. Integrated experimental and theoretical approach for the structural characterization of Hg^{2+} aqueous solutions.
   P. D'Angelo, V. Migliorati, G. Mancini, V. Barone, G. Chillemi.
   J. Chem. Phys., **128** 84502, (2008).

3. A Coupled Molecular Dynamics and XANES Data Analysis Investigation of Aqueous Cadmium(II).
   P. D'Angelo, V. Migliorati, G. Mancini, G. Chillemi.
   J. Phys. Chem. A, **112** 11833, (2008).

4. Ion Hydration in high-density water.
   V. Migliorati, G. Chillemi, G. Mancini, A. Zitolo, S. Tatoli, A. Filipponi,
   P. D'Angelo.
   J. Phys.: Conf. Ser., **190** 012057, (2009).

5. On the hydration properties of the bromide aqua ion: the interplay of quantum and classical Molecular Dynamics, and X-ray absorption spectroscopy.
   V. Migliorati, L. Guidoni, P. D'Angelo.
   In preparation.

Acknowledgments

The first person I wish to thank is Paola D'Angelo, who has been much more than simply a supervisor for me. From a scientific perspective, it is impossible to list the infinite teachings she has imparted. Just from observing her at work, I have been learning how to organize the work, to write articles and to present the results (her fame in this field is well known to all who have met her). On top of that I have also found a real friend. It has been a great pleasure to work with her all these years. I have enjoyed every single moment of this period (even when, as it often happens in the research field, the results did not come - and it has happened frequently!). Next person who needs to be mentioned is Giordano Mancini for his invaluable help in the field of computational chemistry. Many of the things I am able to do today in this field I have learned from him. Sometimes I miss the period which we spent interpreting results and reasoning together when I had just started my thesis. It was fun for me but probably slightly less for him as he would get stressed because of my never-ending, insistent questioning. Thirdly I would like to thank Giovanni Chillemi, the great expert of Molecular Dynamics simulations of our research group. I have always felt at ease working with him and I have learned really a lot from our interesting discussions. I wish to thank Leonardo Guidoni, who has introduced me, about one year ago, to the wonderful world of Car-Parrinello and QM/MM MD simulations. This field has always had a particular appeal for me and I will always be grateful to Leonardo for his teachings. To my best friend and colleague Andrea Zitolo, with whom I have shared every single instant of this long adventure, a massive thank you. Apart from his help for downloading and installing codes, he has been like a brother to me. Getting used to working with others will be extremely difficult. One thing I am sure of is I will never find anyone else to laugh with the way we do. I would also like to thank Maria Montagna, my "physicist friend", with whom I have taken my first steps in the field of ab initio simulations. Despite only having met her last year we have built a strong relationship that feels has been there all my life. A special thanks to Prof. Nicolae Viorel Pavel for the help he has given me throughout the years and the chance to work in his group, and to Prof. Vincenzo Barone for his fundamental contribution to the development of the procedures used in this work to generate the effective pair potentials. I am grateful to Prof. Francesco Gianturco and to Prof. Ivano Tavernelli for their support. I also acknowledge Claudia Leggio and Simone Tatoli. Gabriele and Salvatore, my favourite students. A big thank you to Prof. Guido Gigli, the first to teach me quantum mechanics. And to my dear

friend Giada Penna, for what she knows.
I would like to thank my parents and my sister, to whom I owe everything.
My last thanks go to Massimo, for giving me happiness. Without him who knows where I would be today...